A new type of soliton with particle properties

U. Enz

Philips Research Laboratories, Eindhoven, The Netherlands
(Received 12 July 1976)

This paper describes stable field configurations of two scalar fields $\theta(x,y,z,t)$ and $\phi(x,y,z,t)$. The field configurations follow from a simple least action principle based on an energy density which is a function of θ, ϕ, and their first derivatives. The description is Lorentz-invariant. The structures are of a stringlike type and are characterized by several integers. It is shown, that the simplest closed strings, described by the integers $N=1$, $M=1$, $P\pm 1$, are stable. The structures $P=1$ and $P=-1$ are related by mirror symmetry. Three constants enter in the basic action principle: a length l, a constant E with the dimension of energy time length, and a dimensionless parameter γ. All properties of these field configurations have discrete values, which is a direct consequence of the nonlinearity of the basic expression for the energy density. An attempt is made to identify these structures with elementary particles, the electron and the positron in the simplest case $P=1$ and $P=-1$. To this aim, the total energy of the field structures is equated to the rest energy of the particles. The constants E, l, and γ are related to the fundamental physical constants h, m, e. The model proposed represents a classical field structure with quantized properties.

INTRODUCTION

This paper deals with an attempt to find equations describing stable, singularity-free field configurations in three-dimensional space. The field configurations will turn out to be such that high but finite field values and high energy densities occur in a limited spacial region, say inside a sphere of radius R, and fall to zero at distances $r \gg R$ from this region. The classical field introduced here has a simple geometrical meaning and is governed by nonlinear equations derived from a least action principle. In general, the field is not identical with known physical fields, but an identification will be attempted in certain limiting cases. The procedure is first to introduce this field, to investigate stable solutions, and then to identify its asymptotic parts with physical fields. The ultimate aim is to identify the structure described in this way with a stable elementary particle. In such a classical structural field theory, there is no room for the separate notions field and particle; the field is considered to be the fundamental entity, the particle being a localized distribution of the field. The motivation to search for such a theory was given long ago.[1-3]

The theory is based on the assumption of an energy density depending on two real scalar field variables θ and ϕ, and on their first derivatives with respect to the four-dimensional coordinates x, y, z, t. The fields θ and ϕ have the character of angular variables in an auxiliary Euclidean space of three dimensions u_i. These assumptions are a generalization of an earlier attempt[3] to obtain discrete particle properties in a theory dealing with a single field θ. The present paper first recalls the results of this simple case and then treats the generalization to two scalar fields, and finally deals with the physical interpretation.

SINGLE SCALAR FIELD

When dealing with a single field θ, the least action principle is introduced as follows

$$\delta W = 0,$$

$$W = \int \{K \sin^2\theta + A[(\nabla\theta)^2 - c^{-2}(\partial\theta/\partial t)^2]\}\, dx\, dy\, dz\, dt. \quad (1)$$

This expression was found by exploiting a close conceptual and formal analogy between moving domain walls in magnetic crystals and moving particles.[3,4] The field $\theta(x, y, z, t)$ is visualized as an angle in an auxiliary plane u_1, u_2 and describes the direction of a unit vector $\mathbf{n}$ in this plane. The term $K \sin^2\theta$ describes an anisotropy in the plane u_1, u_2 and has been chosen in this form for simplicity. Higher order terms $K_i \sin^{2i}\theta$ could be added. The constants A and K have the dimensions of energy per unit length and energy density respectively, and c is the velocity of light. The Euler equation corresponding to (1) reads:

$$\Box\theta = (K/2A)\sin 2\theta. \quad (2)$$

We first specialize to one spacelike coordinate, assuming $\theta = \theta(x,t)$. This case has been studied independently and from a different point of view by Perring and Skyrme,[5] and there has been much interest in recent years in its leading to a special class of soliton solutions.[6]

A solution of (2) is

$$\sin\theta = \pm\,(\cosh\bar{x}/\bar{x}_0)^{-1}, \quad (3)$$

with

$$\bar{x} = x - vt, \quad \bar{x}_0 = \pi(A/K)^{1/2}(1 - v^2/c^2)^{1/2}, \quad (4)$$

where v is a constant velocity $< c$. This solution represents a stable field configuration or an object with an internal structure moving with velocity v along the x axis. Inside a spacelike region of the order of magnitude $\bar{x}_0$ on the x axis the angle θ changes from $\theta = 0$ to $\theta = \pi$, describing a clockwise or counterclockwise rotation of the vector $\mathbf{n}$ by an angle $\pm\pi$. An invariant $\pm\pi$ of topological nature may therefore be ascribed to the structure (3). Most of the energy associated with (3) is concentrated in the region $\bar{x}_0$. In the rest system ($v = 0$) the structure is characterized by the discrete surface energy density

$$E_s = 4(AK)^{1/2} \quad (5)$$

corresponding to a discrete mass density

$$m_s = E_s/c^2, \quad (6)$$

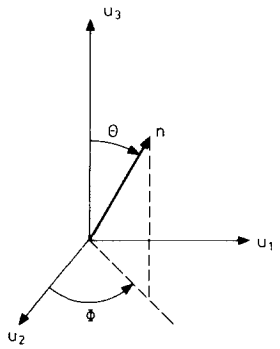


FIG. 1. Definition of the angular variables θ and ϕ as polar and azimuthal angles defining the direction of the unit vector $\mathbf{n}$ in the space u_1, u_2, u_3.

and by the length

$$l_0 = \pi(A/K)^{1/2}. \tag{7}$$

The solution can be visualized as an infinite planar sheet oriented parallel to the $y-z$ plane with mass density (6) and thickness (7). The occurrence of discrete quantities is a direct consequence of the nonlinearity of Eq. (2).

The structure discussed thus far may be considered as a "one-dimensional stable particle." This model particle exhibits a number of properties which are common to real particles: discrete mass, a length l_0 which can be considered as an elementary length, and an invariant $\pm\pi$. The description is Lorentz-invariant and free from singularities. It may be asked whether Eq. (2) has similar solutions describing localized stable fields in three-dimensional space. However, it has been shown by Derrick[7] that stable, time-independent solutions of (2) do not exist in three dimensions. This holds moreover for a large class of similar nonlinear equations.[7] On the other hand, Anderson and Derrick[2] showed that stable time-dependent (oscillating) solutions do exist for some related equations. This type of stability does not have the absolute nature of the topological stability of solution (3) but is *metastable* in the sense that it is destroyed by sufficiently large perturbations.[8] In the next section, we describe an approach that leads to *topologically stable* field configurations in three dimensions.

SOLUTION FOR TWO SCALAR FIELDS θ, ϕ

In this section we consider a unit vector $\mathbf{n}$ in an auxiliary Euclidian space of *three* dimensions u_i. The direction of $\mathbf{n}$ is determined by the polar angle θ and the azimuthal angle ϕ (Fig. 1). $\theta(x,y,z,t)$ and $\phi(x,y,z,t)$ are considered as two real scalar fields in physical space. In this way the physical space is mapped onto the sphere of radius 1 centered in the origin of the auxiliary space.

By analogy to (1) a least action principle is postulated:

$$\delta W = 0, \quad W = \int (K\sin^2\theta + AD_0 + ED_0^2)d^4x, \tag{8}$$

with the abbreviation

$$D_0 = [(\nabla\theta)^2 - c^{-2}\theta_t^2] + \sin^2\theta[(\nabla\phi)^2 - c^{-2}\phi_t^2] \tag{9}$$

being used. The operator ∇ stands for the gradient in the three spacelike coordinates. The constants K and A have the same meaning (and dimension) as before. The term AD_0 describes an energy density associated with the rate of change of the direction of $\mathbf{n}$. A quadratic term, ED_0^2, has been added, representing an additional energy density for high rates of change of θ and ϕ. The introduction of this term is the simplest way to assure stability, as will become clear in the next section. E is a constant with the dimension of energy times length. The constants K, A, E define a characteristic length

$$l = (E/K)^{1/4} \tag{10}$$

and a dimensionless constant

$$\gamma = EK/A^2. \tag{11}$$

Other lengths can be constructed with the aid of l and γ, e.g., the length l_0, (7), which is equal to $\pi\gamma^{-1/4}l$. If all lengths are expressed in units of l, Eq. (8) can be written in the dimensionless form

$$\delta W = 0, \quad W = (E/c)\int (\sin^2\theta + \gamma^{-1/2}D + D^2)d^4\xi, \tag{12}$$

where

$$D = l^2 D_0.$$

The assumptions (8) and (9) characterize the basic properties of space; it will be shown that they lead to stable field configurations of the type described in the introduction.

The Euler equations corresponding to the minimum principle (12) are two coupled, nonlinear partial differential equations of second order for the two scalar functions θ and ϕ. The equations read

$$\begin{aligned} &2(\gamma^{-1/2} + 2D)\Box\theta + 4(\nabla\theta\nabla D - \theta_\tau D_\tau) \\ &\quad - \{1 + (\gamma^{-1/2} + 2D)[(\nabla\phi)^2 - \phi_\tau^2]\}\sin(2\theta) = 0 \end{aligned} \tag{13}$$

and

$$\begin{aligned} &[(\gamma^{-1/2} + 2D)\Box\phi + 2(\nabla\phi\nabla D - \phi_\tau D_\tau)]\sin\theta \\ &\quad + 2[\gamma^{-1/2} + 2D)(\nabla\phi\nabla\theta - \phi_\tau\theta_\tau)]\cos\theta = 0, \end{aligned} \tag{14}$$

where $\tau = tc/l$.

No attempt is made to solve these equations in general form, but we show that simple solutions having cylindrical symmetry can be found. For this purpose, we introduce reduced cylindrical coordinates ρ, φ, ζ, τ with $x = l\rho\sin\varphi$, $y = l\rho\cos\varphi$, $z = l\zeta$ and write

$$D = (\theta_\rho^2 + \rho^{-2}\theta_\varphi^2 + \theta_\zeta^2 - \theta_\tau^2) + \sin^2\theta(\phi_\rho^2 + \rho^{-2}\phi_\varphi^2 + \phi_\zeta^2 - \phi_\tau^2). \tag{15}$$

We assume $\phi = \pm\frac{1}{2}L\varphi$ where L is an integer, and $\theta = \theta(\rho)$, i.e., that θ is a function of the coordinate ρ only. Uneven values of the "index" L are admissible because the basic action principle (8) remains unchanged if $\mathbf{n}$ is replaced by $-\mathbf{n}$. Here, however, we assume $L = 2$ for simplicity, so that we have $\phi = \varphi$. Equation (14) is satisfied by these special forms of the fields, and expression (15) reduces to

$$D = \theta_\rho^2 + \rho^{-2}\sin^2\theta. \tag{16}$$

From Eq. (13) we then find the following second order equation:

$$\begin{aligned} \theta_{\rho\rho} = (1 + 6\gamma^{1/2}\theta_\rho^2 + 2\gamma^{1/2}\rho^{-2}\sin^2\theta)^{-1}\{&\sin 2\theta(\tfrac{1}{2}\gamma^{1/2} + (\tfrac{1}{2} - \gamma^{1/2}\theta_\rho^2 \\ &+ \gamma^{1/2}\rho^{-2}\sin^2\theta)\rho^{-2}) - [1 + 2\gamma^{1/2}(\theta_\rho^2 - \rho^{-2}\sin^2\theta)]\rho^{-1}\theta_\rho\}. \end{aligned} \tag{17}$$

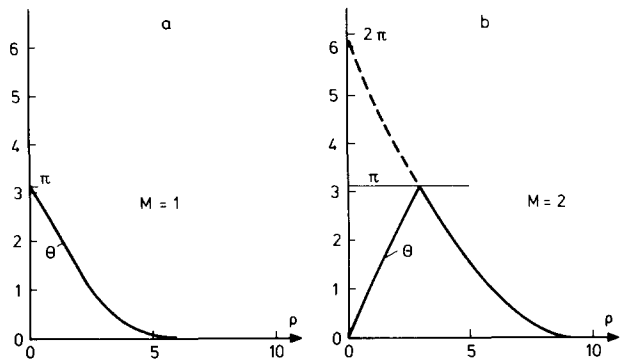


FIG. 2. Solution $\theta=\theta(\rho)$ for a stable field configuration with cylindrical symmetry. (a) $M=1$, (b) $M=2$. Dashed curve: numerical solution for $\theta(0)=2\pi$. Solid curve: solution with proper boundary condition $\theta(0)=0$. The solutions depend on the dimensionless parameter γ. In both examples $\gamma=1$.

This equation can be solved numerically. From Eq. (16) it can be seen that for divergence-free solutions the boundary condition for $\rho=0$ is given by $\sin\theta=0$ or $\theta=M\pi$, where M is an integer. We impose the further boundary condition $\theta=0$ for $\rho\to\infty$. Assuming $\gamma=1$, we find for $M=1$ the solution shown in Fig. 2(a). The first derivative in the origin is $\theta_\rho(0)=-0.9205$. This solution constitutes a stable field configuration which is infinitely extended in the z direction. The stability is of a topological type and is related to the fact that ϕ changes by 2π on any simple path one time encircling the z axis. The field $\theta(\rho)$ has the value π on the entire z axis and falls rapidly to zero with increasing ρ. The dimensionless energy density $\sin^2\theta+D+D^2$ decreases rapidly as well (Fig. 3) so that only the inside of a cylinder with a radius of, say, $\rho=5$ contributes appreciably to the integral (12). For static solutions with cylindrical symmetry, this integral reduces to

$$W=2\pi\sqrt{EK}\,(z_1-z_0)(t_1-t_0)I_1(\gamma) \tag{18}$$

with

$$I_1(\gamma)=\int_0^\infty(\sin^2\theta+\gamma^{-1/2}D+D^2)\rho\,d\rho, \tag{19}$$

where z_1-z_0 and t_1-t_0 are the integration intervals on the z and t axis respectively. $I_1(\gamma)$ is a dimensionless integral, a measure of the energy per unit length of the structure on the z axis. It still contains γ. In the present example, with $\gamma=1$ we find $I_1(\gamma)=12.9$. For $\gamma\ll1$ the integral approaches the asymptotic value $I_1(\gamma)=3.93/\gamma$.

All these properties lead to the intuitive picture that the solution of (17) can be considered as a stable cylinder or string with lateral dimensions of a few units of l. The orientation of the string axis in space is arbitrary due to the symmetry properties of D.

Apart from the simple solution $M=1$ described above, one can find numerical solutions for $M>1$. Definite integrals I_M are associated with these solutions. As θ is defined as a polar angle with $0\leq\theta\leq\pi$, the proper boundary condition is $\theta(0)=0$ for even M and $\theta(0)=\pi$ for uneven M. The solutions consist of M segments (or shells) $0\leq\theta(\rho)\leq\pi$ with alternating positive and negative slope θ_ρ. For segments with $\theta_\rho>0$ we have $\phi=\varphi+\pi$ and for segments with $\theta_\rho<0$ we have $\phi=\varphi$. Fig. 2(b) gives the solution for $M=2$ with $\theta_\rho(0)=1.302$.

At this point, an interesting property of the solutions $\theta(\rho)$ of Eq. (17) may be mentioned. This equation can be written in the form

$$\theta_{\rho\rho}+(1/\rho)\theta_\rho=f(\rho,\theta,\theta_\rho)=\epsilon(\rho) \tag{20}$$

and can be interpreted as Poisson's equation based on a density function $\epsilon(\rho)$. This density is uniquely defined by the solutions $\theta(\rho)$ of Eq. (17). As $\theta(\rho)$ decays exponentially for large ρ, it follows directly from Dirichlet's theorem that

$$\int_0^\infty\epsilon(\rho)\rho\,d\rho=0. \tag{21}$$

The numerical solution of (17) fulfills this condition.

The string can also be treated with the aid of a Ritz approximation. The exact solution of Fig. 2(a) ($M=1$) is then replaced by $\theta=\pi-a\rho$ in the interval $0\leq\rho\leq\pi a^{-1}$ and by $\theta=0$ for $\rho>\pi a^{-1}$. Equation (16) then reads

$$D=\rho^{-2}\sin^2 a\rho+a^2. \tag{22}$$

Using elementary integrals, we obtain for the integral I_1 (with $\gamma=1$)

$$I_{1R}(a)=\tfrac{1}{2}\pi^2+\tfrac{1}{2}\mathrm{Cin}2\pi+a^2(\tfrac{1}{2}\pi^2+\mathrm{Cin}4\pi)+\tfrac{1}{4}a^{-2}\pi^2, \tag{23}$$

which acquires a minimum value of $I_{1R}(a_0)=15.07$ for $a_0=0.744$. This value of the integral is not unreasonable compared with the exact result of 12.9. It gives an indication of the accuracy of a similar approximation which will be made below for a more complicated field structure.

Concluding this section, we remark that the introduction of two angular variables leads to at least one stable field configuration of a string like type. This result is to be compared with the case (1) of a single scalar θ, which leads to a stable planar structure of infinite extension. If three angular variables representing Eulerian angles in a four-dimensional space u_i are introduced, which are governed by an action principle analogous to (8), a localized field structure with spherical symmetry is obtained. In this structure $\theta=\pi$ is limited to a single point. Skyrme[9] was the first to describe a solution equivalent to this pointlike structure. The stringlike structure has some similarities with a disclination in a nematic liquid crystal.[10]

Summarizing, we observe that the introduction of one, two, or three angular variables leads to stable structures of planar, stringlike, and pointlike shape respec-

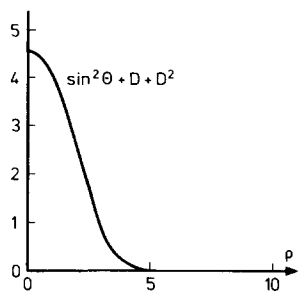


FIG. 3. Density $\sin^2\theta+D+D^2$ as a function of the reduced cylindrical coordinate ρ. $\gamma=1$.

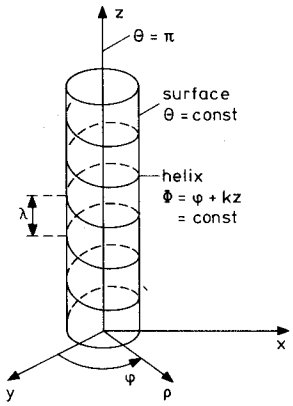


FIG. 4. Twisted string with a surface $\theta = \text{const}$ (cylinder) and a curve $\phi = \text{const}$ (helix) indicated.

tively. Our further considerations will be based on stringlike structures.

STABILITY OF CLOSED, TWISTED STRINGS

We first note that by putting

$$\phi = \varphi \mp kz/l \mp \omega tc/l \tag{24}$$

and by still assuming a time-independent function $\theta = \theta(\rho)$, a slightly more general solution than (16) can be found. A period λ on the z axis, defined by $k = 2\pi l/\lambda$, and a rotation of the vector $\mathbf{n}$ with constant dimensionless angular frequency ω are described in this way. Equation (14) is still satisfied by the above choice of ϕ and θ. Equation (15) then reads

$$D = \theta_\rho^2 + \sin^2\theta(\rho^{-2} + k^2 - \omega^2). \tag{25}$$

A differential equation analogous to (17) can be found by combining Eqs. (25) and (13), which has again stringlike solutions. Solutions with $k \neq 0$ are called twisted strings. In Fig. 4, a twisted string is sketched schematically. The axis $\theta = \pi$ coincides with the z axis, and the surfaces defined by $\theta = \text{const}$ are cylinders. A helix $\phi = \text{const}$ is also indicated.

If the same Ritz approximation as before ($\theta = \pi - a\rho$) is applied to the twisted string, we obtain for the integral (12)

$$W = \frac{E}{c} 2\pi(\zeta_1 - \zeta_0)(\tau_1 - \tau_0)\left[\frac{\pi^2}{2}\gamma^{-1/2} + \frac{\mathrm{Cin}2\pi}{2}\gamma^{-1/2} + \left(\frac{\pi^2}{2} + \mathrm{Cin}2\pi - \tfrac{1}{4}\mathrm{Cin}4\pi\right)(k^2 - \omega^2) + a^{-2}\left(\frac{\pi^2}{4} + \frac{\pi^2}{4}\gamma^{-1/2}(k^2 - \omega^2) + \frac{3\pi^2}{16}(k^2 - \omega^2)^2\right) + a^2\left(\frac{\pi^2}{2} + \mathrm{Cin}4\pi\right)\right] \tag{26}$$

with $\zeta = z/l$ and $\tau = ct/l$.

For large values of $(k^2 - \omega^2)$ the structure of the twisted string is modified as compared with the untwisted string in the sense that the lateral dimensions are decreased and the total minimum energy is increased.

At this point a transition is made to a more qualitative consideration. The string is now imagined as an object that can be flexed without losing its topological structure, and, for moderate curvature, without appreciable change of the energy per unit length. For the planar solution of Eq. (2) (Bloch walls) the corresponding transition is common practice.[11]

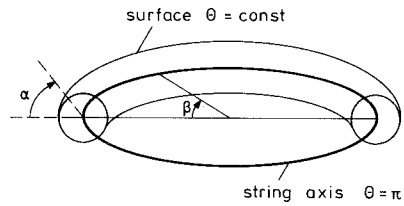


FIG. 5. Sketch of a closed string. A torus defined by $\theta = \text{const}$ is indicated.

Closed strings are now considered, i.e., strings of which the axis defined by $\theta = \pi$ forms a closed line. This is done with the aid of Fig. 5. The string axis is a circle of radius R. An angle β is introduced to define a point on this axis. $R\beta$ corresponds to the coordinate z of the straight string. A further angle α is introduced to define an angular coordinate corresponding to the angle φ of the straight string.

Two sets of structures can be defined by writing

$$\phi = N\beta \pm \alpha, \tag{27}$$

where $N = 2\pi R/\lambda$ is an integer indicating the number of full 2π twists on the circumference of the string. The condition $\theta = \text{const}$ defines surfaces topologically equivalent to a torus as indicated in Fig. 5. The two sets of structures (27) can be called right-handed and left-handed closed twisted strings. The spatial distribution of θ and ϕ for these structures follows in principle from Eqs. (13) and (14). However, the numerical solution of these equations for this three-dimensional field configuration is a mathematical problem which has not yet been solved.

In the following we present some considerations of a qualitative nature concerning the properties of closed strings.

(a) The two structures (27) are related by mirror symmetry, which can be described by an integer P taking the values $+1$ or -1.

(b) The closed string with the twist number $N = 0$, i.e., the untwisted closed string is unstable.

(c) A continuous line defined by $\theta = 0$ (or, more generally, $\sin\theta = 0$) passes through the closed string $N = 1$.

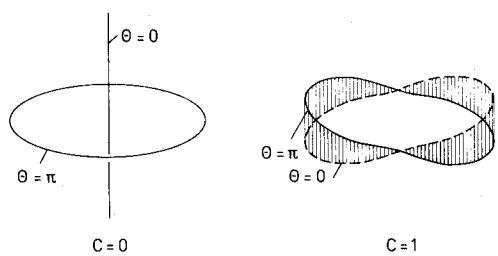


FIG. 6. The two structures denoted by $C = 0$ and $C = 1$. The line defined by $\theta = 0$ is a straight line ($C = 0$) or a closed curve ($C = 1$).

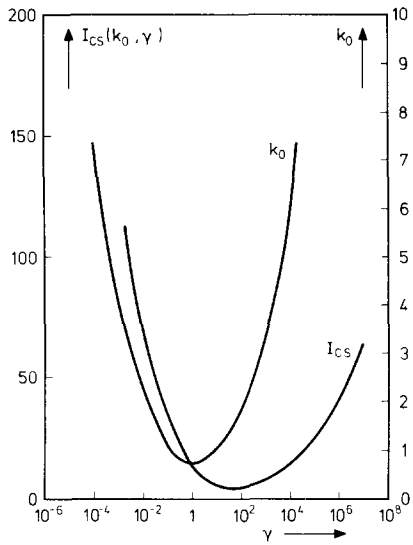


FIG. 7. Dimensionless integral $I_{cs}(k_0, \gamma)$, a measure for the minimum total energy of the closed string $L=2$, $N=M=1$, $C=0$ as a function of γ. The curve has a minimum of $I_{cs}=4.16$ for $\gamma \approx 34$ with a corresponding $k_0 \approx 1$. The curve $k_0(\gamma)$ is also plotted.

The field ϕ changes by 2π on any closed path encircling this line once. The presence of this line, which can be considered as a string $\theta=0$, follows from the property of single valuedness of the field ϕ in all points except where $\theta=0$. The string $\theta=0$ is therefore a stable entity of much the same kind as the closed string $\theta=\pi$ discussed up till now. In the structure denoted by $C=0$ in Fig. 6 this string $\theta=0$ is represented as a straight line.

(d) The following argument shows that the closed string $N=1$ is stable. Due to the property (c) $\nabla\theta$ is of the order of π/R inside the closed string. For small values of R the dominating term ED_0^2 contributes as π^4/R^4 to the energy density, whereas the volume is proportional to R^3. Therefore, for $R \to 0$, the total energy increases in proportion to $1/R$. For large values of R, on the other hand, the energy is proportional to R because, according to (26) the energy is proportional to the total length of the string. Hence, there is at least one radius R_0 defining a string of minimum energy. This argument is valid for static as well as for time dependent solutions θ. This makes it clear why it is necessary to introduce the term ED^2 in the action principle (8).

(e) In a crude way, the energy of the closed string can be calculated starting from equation (26) and assuming the circumference of the closed string to be $2\pi R = N\lambda$ and therefore

$$\zeta_1 - \zeta_0 = 2\pi N/k. \tag{28}$$

The additional terms following from the finite curvature of the string are neglected. For reason described below we assume

$$\omega = \gamma^{1/4}. \tag{29}$$

The action (26) for the closed string $N=1$ can therefore be written in the form

$$W_{cs} = (E/c)4\pi^2(\tau_1 - \tau_0)k^{-1}\{\gamma^{-1/2}C_1 + C_2(k^2 - \gamma^{1/2}) + a^{-2}[\gamma^{-1/2}k^2\pi^2/4 + \tfrac{3}{16}\pi^2(k^2 - \gamma^{1/2})^2] + a^2C_3\}, \tag{30}$$
$$= (E/c)4\pi^2(\tau_1 - \tau_0)I_{cs}$$

with

$$C_1 = \tfrac{1}{2}\pi^2 + \tfrac{1}{2}\mathrm{Cin}2\pi \approx 6.15,$$
$$C_2 = \tfrac{1}{2}\pi^2 + \mathrm{Cin}2\pi - \tfrac{1}{4}\mathrm{Cin}4\pi \approx 6.59, \tag{31}$$
$$C_3 = \tfrac{1}{2}\pi^2 + \mathrm{Cin}4\pi \approx 8.05,$$

where the dimensionless integral I_{cs} is a measure of the energy of the closed string. I_{cs} can be treated as a function of k and a and minimized with respect to these variables for any value of γ. The minima are characterized by $k=k_0(\gamma)$ and $a=a_0(\gamma)$. As a_0 and k_0 are found to be near to each other, we assume $a=k$. The resulting $I_{cs}(k_0, \gamma)$ is plotted in Fig. 7. The curve $I_{cs}(\gamma)$ shows a minimum for $\gamma \approx 34$; the corresponding value of k_0 (and of a_0) is 1.43. The minimum value of the integral is $I_{cs,\,min}=4.16$. It should be emphasized, however, that these results represent a rough approximation, which, moreover, neglects the effects of the finite curvature of the string. Therefore, the values of I_{cs} given in Fig. 7 are probably much to high. It may be mentioned here that the assumption $L=1$ leads to similar structures and values of I_{cs} which are slightly smaller than those given in Fig. 7.

(f) The action principle (8) is symmetric with respect to $\theta=\pi$ and $\theta=0$; an asymmetry enters only by virtue of the boundary condition $\theta \to 0$ for $r \to \infty$. This leads us to suspect that solutions may exist in which this symmetry dominates. In (c) the existence of a string $\theta=0$ passing through the closed string $\theta=\pi$ ($N=1$) has been discussed. The string $\theta=0$ is either infinite in length or closed in such a way as to interlock with $\theta=\pi$. This property can be described with a constant C assuming the values $C=0$ and $C=1$ respectively. The two structures are shown in Fig. 6. The structure with $C=1$ approaches the above symmetry more closely. It describes a closed Möbius strip with a 2π twist having strip boundaries defined by $\theta=0$ and $\theta=\pi$.

(g) So far, the discussion has been based on static functions ϵ and a time dependence of ϕ given by (24). However, a time dependence of θ is not excluded, and this would lead to a negative contribution to the value of $D(a)$ and therefore to the integral. The time dependence of θ possibly comes in through a rotation of the whole structure $C=1$ (Fig. 6). We are unable to estimate the difference in energy between the structures $C=0$ and $C=1$, but it is probable that for $C=1$ the integral I_{cs} is considerably lowered due to the negative contribution associated with θ_τ.

(h) If it is assumed that the structure $C=1$ is stable, it follows that outside of sphere of radius R_1 containing the whole structure of both strings $\theta=\pi$ and $\theta=0$, and where $ED^2 \ll AD$, Eq. (13) reduces to

$$\Box\theta \cong \theta(\gamma^{1/2} - \omega^2). \tag{32}$$

This occurs because in this outside region ϕ takes the form $\phi = \mathrm{const} + \omega t$. For

$$\omega = \gamma^{1/4} \tag{33}$$

TABLE I. Elements of three-dimensional field structures.

u_i	Auxiliary space of three dimensions
K	Constant describing anisotropy in u_i
A, E	"Exchange" constants
θ, ϕ	Scalar field variables, polar and azimuthal angle in u_i. $\theta = \theta(x, y, z, t)$, $\phi = \phi(x, y, z, t)$
l	Length parameter describing the structures.
γ	Dimensionless parameter
L	Integer indicating "index" of the straight string
M	Integer indicating "number of shells"
N	Integer indicating "number of twists"
C	Constant taking the values 0 or 1
P	Constant taking the values -1 or $+1$, describing the symmetry of the solution

(in units c/l) the asymptotic static solution is

$$\theta \propto 1/r. \tag{34}$$

The argument (21), therefore, does not apply here, and for $C=1$

$$\int \epsilon(r) r^2 \, dr \neq 0. \tag{35}$$

(i) For $N>1$ the stability is not obvious, a decay into simpler structures is not excluded topologically. Table I summarizes the various parameters of the structures described. The structures discussed in the last sections have not, to my knowledge, been described before. In these structures, a new synthesis between continuous field properties (θ, ϕ) and discrete properties (N, M, etc.) is realized.

INTERPRETATION

Thus far, we have mainly dealt with geometrical and topological properties of the model, except for the constants A, K, E, c which were introduced as physical constants from the beginning. In this section, a physical interpretation is attempted. The main result up till now is that stable structures or field configurations have been found which follow directly from the basic action principle. These structures represent entities of finite, discrete action or energy localized in a limited region of three-dimensional space. The characteristic properties of the model are: (a) no singularities occur, all field values and densities remain finite; (b) integers characterizing the structures (like N or M) occur in a natural way together with classical continuous fields; (c) the stability, at least for $N=M=1$ is of a topological type; (d) the description is Lorentz-invariant as becomes clear from Eq. (9).

These properties make it tempting to try to identify these structures with elementary particles. However, in general this would be a formidable task, and we limit ourselves to the discussion of some arguments that might contribute to such an identification. The idea is that a particular structure, specified by the integers L, N, M, C, P, corresponds to a particular free elementary particle in such a way that the total energy of the structure represents the rest energy of the particle. The energy is defined as the spacelike part of the basic integral.

With $\tau \equiv ct/l$ we may write according to Eq. (30):

$$W_{cs} = (t_1 - t_0)(E/l) 4\pi^2 I_{cs}(\gamma, L, N, M, C). \tag{36}$$

Here, we consider I_{cs} as the exact result of the integral, based on a time dependent field θ ($C=1$), and not as the approximative value given in Fig. 7 which is probably much too high. If we adopt the hypothesis that the two simplest stable structures $L=2$,[12] $N=M=1$, $C=1$, $P=\pm 1$ can be identified with the electron and the positron, we find

$$(E/l) 4\pi^2 I_{cs}(\gamma) = mc^2, \tag{37}$$

where m is the mass of the electron. If it is further assumed that the structure with $C=1$ is stable, and $\omega = \gamma^{1/4}$ [Eq. (33)], the field θ falls off as $\pi r_0/r$ for large distances r as compared with the string radius [Eq. (34)]. The energy density in this region is then given by

$$A\pi^2 r_0^2/r^4 \tag{38}$$

because the term containing D^2 can be neglected. Here, r_0 is a length which follows in principle from the exact solution of the minimum principle for $C=1$. The length r_0 is probably much smaller than R_0, the stable string radius. The energy density (38) is identified with the energy density $e^2 r^{-4}$ of the electrical field of the electron in the limit $r \to \infty$, and we find

$$e^2 = \pi^2 A r_0^2 = \pi^2 E \gamma^{-1/2} r_0^2/l^2, \tag{39}$$

where e is the elementary charge. This means that the charge is quantized. It may be further assumed that the circumference $2\pi R_0 = 2\pi k_0^{-1} l$ of the stable string is equal to the Compton wavelength, yielding

$$l = k_0 \hbar/mc. \tag{40}$$

Here again k_0 is considered to be the exact value of the equilibrium wavenumber corresponding to the proper minimum of I_{cs}. In principle, we can determine E, l, γ (or alternatively A, K, E) from Eqs. (37), (39), and (40). For E we find

$$E = \hbar c k_0 (4\pi^2 I_{cs})^{-1}. \tag{41}$$

However, as we are unable to calculate r_0, a further specification of the constants does not seem possible at this stage. On the other hand, it can be stated that the basic action principle is a well-posed mathematical problem. To find a more accurate solution than the one presented here should therefore be possible if sufficient computing facilities are available. We expect that k_0 and I_{cs} will turn out to be at least one order of magnitude smaller than the figures given here.

To conclude, some more general remarks may be made.

(1) A major property of the described field structure is that all its parameters are quantized, which is a direct consequence of the nonlinearity of the basic equation. The quantized properties are of two kinds: dimensionless quantum numbers (integers like N, M) and quantized physical properties (like mass and charge). We therefore believe that a further quantization according to the rules of quantum electrodynamics is unnecessary

and even meaningless. The model therefore represents a classical particle which quantized properties.

(2) The notion of exact localization of a particle has no meaning in this model; it should be replaced by saying that the major part of the energy is contained in a finite region of space.

(3) A mirror symmetry exists between the structures $N=M=C=1$, $P=1$, and $P=-1$ (electron and positron) so that the charge is described by CP. A closer examination shows that the ensemble of two structures $P=1$ and $P=-1$ is not topologically stable and may annihilate. This statement is not true for two structures both described by $P=1$ or $P=-1$.

(4) Structures with $N>1$ or $M>1$ have a higher energy and are possibly related to heavier particles.

(5) The interaction of two particles is not a free additional property, but is inherent in the model. Roughly speaking, the interaction is the consequence of the coexistence in space of the two structures, which extend through the whole of space.

(6) The dimensionless parameter γ describes a property of the vacuum and is not associated with a particular field structure. Nevertheless, the structures and the integrals I_{cs} depend in an essential way on γ, e.g., γ determines the relative contributions of the terms D and D^2 in the basic integral. From Eqs. (39) and (41) we find a relation between γ and the fine structure constant α:

$$\alpha = e^2/\hbar c = \tfrac{1}{4}\gamma^{-1/2}r_0^2 l^{-2} k_0 I_{cs}^{-1}. \tag{42}$$

(7) The described structure can be considered to be an extended oscillator of angular frequency $c/R_0 = mc^2/\hbar$ determined by (40). This "internal motion" may be related to the spin of the particle. Considerations on moving oscillators of this frequency were at the root of the original work of de Broglie,[13] leading to quantum mechanics. Therefore, we believe that a reconciliation of the present model with quantum mechanics may prove to be possible along de Broglie's lines of thinking concerning the causal interpretation of quantum mechanics.[13]

In conclusion, we may state that the proposed model exhibits a number of promising physical properties; its validity will depend on the exact values of the dimensionless constants and mass ratios.

Note added in proof: In deriving the asymptotic behavior (32) of θ, a term of the order of θ^3 has been neglected. This term is of the same order of magnitude as $\Delta\theta$, so that this neglect is not justified. If in the basic action principle (12) D^2 is replaced by D^3, the corresponding term in (32) turns out to be of the order of θ^5 so that θ then falls off as r_0/r asymptotically. This modification of (12), however, does not change the topological properties of the model nor the stability of the string $N=M=1$. The physical interpretation remains also unchanged. The asymptotic behavior of θ for the modified action principle will be treated in more detail in a future paper. One of the results is that, given $C=1$, the "charge" r_0 turns out to be independent of the details of the internal structure of the "particle."

[1] A. Einstein and L. Infeld, *Evolution der Physik* 284 (P. Zsolnay, Vienna, 1950).
[2] D. L. T. Anderson and G. H. Derrick, J. Math. Phys. **11**, 1336 (1970).
[3] U. Enz, Phys. Rev. **131**, 1392 (1963).
[4] U. Enz, Helv. Phys. Acta **37**, 245 (1964).
[5] J. K. Perring and T. H. R. Skyrme, Nucl. Phys. **31**, 550, 1962. For an extended reference list see A. C. Scott, Proc. IEEE **61**, 1443 (1973).
[6] N. J. Zabrusky and M. D. Kruskal, Phys. Rev. Lett. **15**, 240 (1965).
[7] G. H. Derrick, J. Math. Phys. **5**, 1252 (1964).
[8] D. L. T. Anderson, thesis (Univ. St. Andrews, Scotland, 1969).
[9] T. H. R. Skyrme, Nucl. Phys. **31**, 556 (1962); J. Math. Phys. **12**, 1735 (1971).
[10] P. G. de Gennes, *The Physics of Liquid Crystals* (Oxford U. P., Oxford, 1974).
[11] J. C. Slonczewski, Intern. J. Magn. **2**, 85 (1972).
[12] A possible alternative is $L=1$, which leads to a similar structure with a slightly lower value of I_{cs}.
[13] L. de Broglie, *Nonlinear Wave Mechanics* (Elsevier, New York, 1960).

Convergence of lattice approximations and infinite volume limit in the $(\lambda\phi^4-\sigma\phi^2-\mu\phi)_3$ field theory

Yong Moon Park*

Department of Theoretical Physics, University of Bielefeld, Germany
(Received 15 March 1976)

By unified method we prove the convergence of the lattice approximation of the $(\lambda\phi^4-\sigma\phi^2-\psi\phi)_3$ field model with periodic, Dirichlet and Neumann boundary conditions in a finite box. This then allows us to take the inifinite volume limit of the Dirichlet states by the Nelson's monotonicity argument. The model under consideration satisfies all the Wightman axioms except possibly the uniqueness of vacuum for $\mu = 0$ and the mass gap.

I. INTRODUCTION

This is a continuation of our previous works[1,2] in proving the convergence of the lattice approximation in the $(\lambda\phi^4 - \sigma\phi^2 - \mu\phi)_3$ quantum field model with various classical boundary conditions in a finite box. With a unified method we prove the convergence of the lattice approximation of the model with periodic, Dirichlet and Neumann boundary conditions. We then use our results together with uniform bounds of the Schwinger functions[3,4] and the Nelson's monotonicity argument[5,6] to take the infinite volume limit of the Dirichlet states of the model under consideration. The resulting theory satisfies all the Wightman axioms except possibly the uniqueness of vacuum for $\mu = 0$ and the mass gap.

During the last few years there has been much progress in the construction of the $(\lambda\phi^4 - \sigma\phi^2 - \mu\phi)_3$ quantum field theory. The main results have been the proofs of existence and semi-boundedness of the volume cutoff Hamiltonian,[7] convergence of the lattice and momentum cutoff Schwinger functions with free and periodic boundary conditions as the cutoffs are removed[1,2,8] and convergence of the infinite volume limit for the weakly coupled model.[9,10] Recently, Park[3] and Seiler and Simon[4] have established uniform bounds of the volume cutoff Schwinger functions of the models with free and periodic boundary conditions. Fröhlich[11] has constructed the infinite volume theory for the strongly coupled model with weakly coupled boundary condition.

In the $P(\phi)_2$ field theory the control of the boundary conditions provided great flexibility in the study of the infinite volume limit.[5,6,12–15] The main purpose of this work is the completion of our program in proving the convergence of the lattice approximation in the $(\lambda\phi^4 - \sigma\phi^2 - \mu\phi)_3$ quantum field model with various classical boundary conditions in which we are interested. Then one may develop the $(\lambda\phi^4 - \sigma\phi^2 - \mu\phi)_3$ field theory parallel to the $P(\phi)_2$ model.

The method we will use in proving the convergence of the lattice approximation for Dirichlet (D) and Neumann (N) boundary conditions is essentially same as that for free (F) and periodic (P) boundary conditions developed in Refs. 1, 2. Since the proof is very complicated (even if it is not so hard) and since the over-all structure of the proof is same as that for P. b. c., we feel that it is convenient to reproduce the proof for P. b. c. This makes the paper better organized and thus more readable. That is, we will prove the convergence of the lattice approximation for D, N, and P boundary conditions simultaneously.

We now discuss briefly our main ideas of the proof. Let Λ be a box in R^3 and let δ be the lattice spacing parameter. Let $C^X_\Lambda(m\delta, n\delta)$ be the covariance of the lattice free fields $\phi_\delta(n)$ with X (= D, N, P) boundary condition. By the standard eigenvector expansion[13] one may write

$$C^X_\Lambda(m\delta, n\delta) = \sum_{k\in \tilde{T}^X_\delta} f^X_k(m\delta)\bar{f}^X_k(n\delta)\mu_\delta(k)^{-2}. \tag{1.1}$$

See Sec. 2 for the details. Let $S^X_{\Lambda,\delta}(f_1,\ldots,f_n)$ be the corresponding lattice cutoff Schwinger functions of the model under consideration. A direct proof of the convergence of $S^X_{\Lambda,\delta}$ as $\delta \to 0$ seems to be very difficult (in fact, we are still not able to produce a direct proof). Therefore, we will introduce the lattice and momentum cutoff Schwinger functions $S^X_{\Lambda,\delta,\kappa}(f_1,\ldots,f_n)$. See Sec. 3 for the definition. It is easy to show that $S^X_{\Lambda,\delta,\kappa} \to S^X_{\Lambda,\delta}$ as $\kappa \to 1$ and $S^X_{\Lambda,\delta,\kappa} \to S^X_{\Lambda,\kappa}$ as $\delta \to 0$. Hence, if we prove that

$$S^X_{\Lambda,\delta,\kappa}(f_1,\ldots,f_n) \underset{\kappa\to 1}{\to} S^X_{\Lambda,\delta}(f_1,\ldots,f_n) \text{ uniformly in } \delta, \tag{1.2}$$

the convergence of the lattice approximation follows from the standard 3ϵ argument. As in Refs. 1, 2 we shall use the method of the inductive expansion developed by Glimm and Jaffe[7] and modified by Feldman[8] to establish (1.2). To prove (1.2) for $X = \mathrm{D}, \mathrm{N}$ by the inductive expansion method,[1,2,7,8] one must check the following:

(a) isolation of localization factors,

(b) bounds of the kernels of Feynman graphs by majorizing functions,

(c) estimations of small graphs including the mass renormalization cancellations,

(d) that the partition function is nonvanishing.

Since we do not have translation invariance for D, N boundary conditions, we need to modify the method used in Refs. 1, 2, 7, 8 to handle (a) for $X = \mathrm{D}, \mathrm{N}$. Roughly speaking, we will be able to isolate the distance factors from the following inspection: Let $C^X_{\Lambda,\kappa}(m\delta, n\delta)$ be the lattice and momentum cutoff covariance (defined in Sec. 3) and let $|x-y|_X$ denote ordinary Euclidean distance for $X = \mathrm{D}, \mathrm{N}, \mathrm{F}$ and periodic distance for $X = \mathrm{P}$. Then

it follows that

$$|C^X_{\Lambda,\kappa}(m\delta, n\delta)| \sim |(m-n)\delta|_X^{-j} \text{ for any } j \in N^+$$

uniformly in κ, δ, where $m\delta, n\delta \in \Lambda$. This means that the momentum and lattice cutoff covariance decays polynomially. This fact is enough to isolate the distance factors. See Step 3 in the proof of Theorem 4.1, Sec. 4. In Sec. 3 we establish basic estimates to take care of (b). We shall give the estimates (c) in the Appendix. To prove the nonvanishing of the partition function (d), we shall apply the method of the mass shift transformation used in the previous work[2] together with the nonvanishing of the partition function for sufficiently large mass m_0^2. Since the boundary terms for D, N, P boundary conditions are independent of the mass, one may apply the mass shift transformation. See proof of Theorem 2.1 in Sec. 3.

We remark that as a by-product we also establish the convergence of the momentum cutoff Schwinger functions in the $(\lambda\phi^4 - \sigma\phi^2 - \mu\phi)_3$ model with D, N, P boundary conditions in a finite box (Corollary 4.3). To be more precise, if $X = $ D, N, P, the results of this paper may be summarized by a diagram:

$$\begin{array}{ccccc} S^X_{\Lambda,\delta}(f_1,\ldots,f_n) & \xrightarrow{\delta\to 0} & S^X_{\Lambda}(f_1,\ldots,f_n) & \xrightarrow[|\Lambda|\to\infty]{X=D} & S^D_{\infty}(f_1,\ldots,f_n) \\ \uparrow{\scriptstyle \kappa\to 1} & & \uparrow{\scriptstyle \kappa\to 1} & & \\ S^X_{\Lambda,\delta,\kappa}(f_1,\ldots,f_n & \xrightarrow{\delta\to 0} & S^X_{\Lambda,\kappa}(f_1,\ldots,f_n) & & \end{array}$$

The organization of the paper is as follows: In Sec. 2 we introduce the notation and definitions on the lattice fields and the lattice cutoff Schwinger functions with X (= D, N, P) boundary conditions. We then state our main theorems, namely, convergence of the lattice approximation and the infinite volume limit of the Dirichlet states in the $(\lambda\phi^4 - \sigma\phi_2 - \mu\phi)_3$ field model. We also give the proof of the existence of the infinite volume limit of the Dirichlet states. In Sec. 3 we introduce the Schwinger functions with a joint lattice and momentum cutoff and establish basic estimates which we will use later on. Assuming the convergence of the inductive expansion as $\kappa \to 1$ uniformly in δ, we establish the convergence of the lattice approximation. In Sec. 4 we prove the uniform convergence of the inductive expansion. In the Appendix we establish estimates for small graph including the mass renormalization cancellations.

Apology: After having established the convergence of the lattice approximation of the models with free and periodic boundary conditions more than one year ago,[1,2] the author felt that the result was not as fully utilized as it deserved mainly because the result for Dirichlet boundary condition was lacking. We hoped that someone would study the problem for Dirichlet boundary condition, since all the techniques we need contained in Refs. 1, 2, 7, 8 except the control of the isolation of the distance factors. We are very reluctant to come back to this subject. We tried hard to simplify this paper to improve its readibility. In some parts, we skip detailed proofs and give only a bare description of the proofs because of notational complications. We apologize for that. But we hope that the underlying idea is apparent to the reader.

Note: Recently Feldman and Osterwalder[16] have announced the infinite volume limit of the Dirichlet states of the model we consider. Their argument is also based on the monotonicity of the Dirichlet Schwinger functions, but they did not provide any description of the proof showing the convergence of the Dirichlet Schwinger functions. Since the preprint has not yet appeared, we produce this paper.

II. NOTATION, DEFINITION AND MAIN RESULTS

In this section we first introduce the lattice fields and the lattice cutoff Schwinger functions (Euclidean Green's functions) of the $(\lambda\phi^4 - \sigma\phi^2 - \mu\phi)_3$ quantum field model with D, P, N boundary conditions in a finite box. Throughout this paper we assume that $\lambda \in R^+$, $\sigma, \mu \in R$. We then state our main results. We specialize for the case where Λ is a box, say $(-l_1/2, l_1/2)\times(-l_2/2, l_2/2)\times(-l_3/2, l_3/2)$. Let δ be the lattice spacing parameter for the lattice $L_\delta = \{n\delta \mid n = (n^{(1)}, n^{(2)}, n^{(3)}) \in Z^3\}$. We denote $\Lambda_\delta = \Lambda \cap L_\delta$ the set of lattice points within Λ and $\partial\Lambda_\delta$ denotes the points in Λ_δ which have the nearest neighbors outside Λ_δ. Following Ref. 13, we assume that l_1, l_2, and l_3 are odd multiple of δ for P and N boundary conditions and even multiple of δ for Dirichlet boundary condition (if Λ is a cube with length of side l, this means that $\delta = l/(2j+1)$ for P., N. b. c. and $\delta = l/j$ for D. b. c., $j \in N^+$). We introduce the lattice fields $\phi_\delta(n)$ as the real Gaussian random process indexed by the lattice in Λ_δ with mean zero and covariance given by[2,13]

$$C^X_\Lambda(m\delta, n\delta) = \sum_{k\in T^X_\delta} f^X_k(m\delta)\bar{f}^X_k(n\delta)\mu_\delta(k)^{-2}, \tag{2.1}$$

where

$$\begin{aligned} &\mu_\delta(k)^{+2} = \delta^{-2}\Big(6 - 2\sum_{i=1}^{3}\cos(\delta k^{(i)})\Big) + m_0^2, \\ &T^X_\delta = T^X \cap [-\pi/\delta, \pi/\delta], \\ &T^X = \begin{cases} \dfrac{2\pi}{l_1}Z \times \dfrac{2\pi}{l_2}Z \times \dfrac{2\pi}{l_3}Z, & X = \text{P}, \\ \dfrac{\pi}{l_1}Z \times \dfrac{\pi}{l_2}Z \times \dfrac{\pi}{l_3}Z, & X = \text{D, N}, \end{cases} \end{aligned} \tag{2.2}$$

and

$$\begin{aligned} &f^{\text{P}}_k(x) = (l_1 l_2 l_3)^{-1/2}\exp(ik\cdot x), \\ &f^{\text{N}}_k(x) = (l_1 l_2 l_3)^{-1/2}\prod_{i=1}^{3} g^{(i)}_{k^{(i)}}(x^{(i)}), \\ &g^{(i)}_{k^{(i)}}(x^{(i)}) = \begin{cases} \cos(k^{(i)}x^{(i)}) & \text{if } l_i k^{(i)}/\pi \text{ is even}, \\ \sin(k^{(i)}x^{(i)}) & \text{if } l_i k^{(i)}/\pi \text{ is odd}, \end{cases} \\ &f^{\text{D}}_k(x) = (l_1 l_2 l_3)^{-1/2}\prod_{i=1}^{3} h^{(i)}_{k^{(i)}}(x^{(i)}), \\ &h^{(i)}_{k^{(i)}}(x^{(i)}) = \begin{cases} \sin(k^{(i)}x^{(i)}) & \text{if } l_i k^{(i)}/\pi \text{ is even}, \\ \cos(k^{(i)}x^{(i)}) & \text{if } l_i k^{(i)}/\pi \text{ is odd}. \end{cases} \end{aligned} \tag{2.3}$$

We note that the functions $\{\delta^{3/2} f^X_k(m\delta)\}_{k\in T^X_\delta}$ satisfy the completeness condition[13]:

$$\delta^3 \sum_{k\in T^X_\delta} f^X_k(m\delta)\bar{f}^X_k(n\delta) = \delta_{m,n} \tag{2.4}$$

for $m\delta, n\delta \in \Lambda_\delta$. It is convenient to introduce the covariance for the free (F) boundary condition given by

$$C^{\text{F}}(m\delta, n\delta) = (2\pi)^{-3}\int_{-\pi/\delta}^{\pi/\delta}\exp[ik\cdot(m-n)\delta]\mu_\delta(k)^{-2}d^3k. \tag{2.5}$$

Let $d\mu^X_{\Lambda,\delta}$ be the underlying Gaussian measures on

$\mathcal{S}'(\Lambda)$ with X. b. c. We introduce the smeared free fields and Wick product of free fields by

$$\phi_\delta(f) = \delta^3 \sum_{n\delta \in \Lambda_\delta} \phi_\delta(n) f(n\delta),$$
$$:\phi_\delta^m:_X(g) = \delta^3 \sum_{n\delta \in \Lambda_\delta} :\phi_\delta(n)^m:_X g(n\delta) \tag{2.6}$$

for $f, g \in L^\infty(\Lambda)$, where $:\ :_X$ is the Wick ordering with respect to the Gaussian measure $d\mu^X_{\Lambda,\delta}$.

We now introduce the lattice cutoff interacting action for the $(\lambda\phi^4 - \sigma\phi^2 - \mu\phi)_3$ field model with X (= D, N, P) boundary condition. We write $\langle\cdot\rangle^{(X)}_{\Lambda,\delta}$ for $\int \cdot\, d\mu^X_{\Lambda,\delta}$. The interacting action is given by

$$V^{(X)}_{\Lambda,\delta} = V^{(X)}_{I,\Lambda,\delta} + V^{(X)}_{C,\Lambda,\delta},$$
$$V^{(X)}_{I,\Lambda,\delta} = :(\lambda\phi^4_\delta - \sigma\phi^2_\delta - \mu\phi_\delta):_X(\chi_\Lambda), \tag{2.7}$$
$$V^{(X)}_{C,\Lambda,\delta} = \tfrac{1}{2}\lambda^2\langle(:\phi^4_\delta:_X(\chi_\Lambda))^2\rangle^{(X)}_{\Lambda,\delta} - \tfrac{1}{6}\lambda^3\langle(:\phi^4_\delta:_X(\chi_\Lambda))^3\rangle^{(X)}_{\Lambda,\delta}$$
$$+ \tfrac{1}{2}\lambda^2\delta m^2_\delta :\phi^2_\delta:_X(\chi_\Lambda),$$
$$\delta m^2_\delta = (4^2\times 6)\delta^3 \sum_{n\delta\in\Lambda_\delta} [C^F(m\delta, n\delta)]^3,$$

where C^F is the covariance for the free boundary condition defined in (2.5) and χ_Λ is the characteristic function of Λ. One may easily check that the above definition of δm^2_δ is equivalent to that of Ref. 1 (at most, up to finite mass renormalization). See the details in the Appendix.

Remark 2.1: We hold the coefficient δm^2_δ of the mass counter term fixed. We always use the δm^2_δ appropriate to free boundary condition.[1] We do this for convenience in using the conditioning theory and monotonicity argument for the Dirichlet states. A more appropriate coefficient of the mass renormalization counter term for X. b. c. would be

$$\delta m^{(X)}_{\Lambda,\delta} = (4^2\times 6)\sum_{n\delta\in\Lambda_\delta} [C^X_\Lambda(m\delta, n\delta)]^3. \tag{2.8}$$

It is then easy to check that for $X = $ D, N, P and $|\Lambda| > 1$

$$|\delta m^2_\delta - \delta m^{(X)}_{\Lambda,\delta}| \le \text{const}, \quad \text{uniformly in } \Lambda \text{ and } \delta. \tag{2.9}$$

We will give the proof of (2.9) in the Appendix.

The partition function and the Schwinger functions are defined by

$$Z^X_{\Lambda,\delta} = \langle\exp(-V^X_{\Lambda,\delta})\rangle^{(X)}_{\Lambda,\delta}, \tag{2.10}$$

$$S^X_{\Lambda,\delta}(f_1,\ldots,f_n) = (Z^X_{\Lambda,\delta})^{-1}\langle\phi_\delta(f_1)\cdots\phi_\delta(f_n)\exp(-V^X_{\Lambda,\delta})\rangle^{(X)}_{\Lambda,\delta}.$$

We now give the main results of this paper:

Theorem 2.1: For $\lambda > 0$, $\sigma, \mu \in R$ in (2.7) and $X = $ D, N, P:

(a) (The convergence of the lattice approximation.) The limits

$$\lim_{\delta\to 0} Z^X_{\Lambda,\delta} = Z^X_\Lambda,$$
$$\lim_{\delta\to 0} S^X_{\Lambda,\delta}(f_1,\ldots,f_n) = S^X_\Lambda(f_1,\ldots,f_n)$$

exist for $f_i \in \mathcal{S}(\Lambda)$.

(b) There exist constants $K_1(\lambda,\sigma,\mu)$ and $K_2(\lambda,\sigma,\mu,\Lambda)$ and a suitable Schwartz space norm $|\cdot|$ such that

$$0 < Z^X_\Lambda \le \exp(K_1|\Lambda|),$$
$$|S^X_\Lambda(f_1,\ldots,f_n)| \le K_2(\Lambda) n! \prod_{i=1}^n |f_i|,$$

where $|\Lambda|$ is the volume of Λ.

(c) For $\tau \in C$ and $f \in \mathcal{S}(\Lambda)$, the limit

$$\lim_{\delta\to 0}\langle\exp(\tau\phi_\delta(f) - V^X_{\Lambda,\delta})\rangle^{(X)}_{\Lambda,\delta} = Z^X_\Lambda(\tau f)$$

exists and is analytic in τ.

Remark 2.2: In Refs. 3, 17 we have established the following:

$$\exp(-a|\Lambda|) \le Z^{F,P} \le \exp(b|\Lambda|),$$
$$|S^P_\Lambda(f_1,\ldots,f_n)| \le K^n n! \prod_{i=1}^n |f_i|, \tag{2.11}$$

where a, b, and K are constants independent of Λ. Similar results also hold for free boundary condition.[4,17] Hence the result of Theorem 2.1 (b) is much weaker than (2.11) for $X = $ P.

We postpone the proof of Theorem 2.1 at the later part of this paper. As consequences of Theorem 2.1 we have the following result:

Corollary 2.2: Let λ, σ, and μ be as in Theorem 2.1 and let $X = $ D, N, P.

(a) There exists a unique measure dq^X_Λ for each $X = $ D, N, P on $\mathcal{S}'(\Lambda)$ such that

$$S^X_\Lambda(f_1,\ldots,f_n) = \int \phi(f_1)\cdots\phi(f_n)\, dq^X_\Lambda$$

for $f_i \in \mathcal{S}(\Lambda)$.

(b) (Lee–Yang theorem and correlation inequalities.) The Lee–Yang theorem[5,15] and the correlation inequalities[5,15] whi h hold for the $(\lambda\phi^4 - \sigma\phi^2 - \mu\phi)_2$ field theory also hold for the $(\lambda\phi^4 - \sigma\phi^2 - \mu\phi)_3$ field theory.

Proof: (a) This follows from Theorem 2.1 (b), Minlos' theorem and the method used in Ref. 8.

(b) This is a consequence of Theorem 2.1 (a) and (c). ∎

Now we use the above result and the uniform bounds of the Schwinger functions with P(F) b. c.[3,4] to take the infinite volume limit of the Dirichlet states.

Theorem 2.3: The infinite volume limit of the Dirichlet Schwinger functions

$$S^D_\infty(f_1,\ldots,f_n) = \lim_{|\Lambda|\to\infty} S^D_\Lambda(f_1,\ldots,f_n)$$

exists for $f_i \in \mathcal{S}(R^3)$. The Schwinger functions $\{S^D_\infty(f_1,\ldots,f_n) | n = 0, 1, 2, \cdots\}$ satisfy the axioms of Osterwalder and Schrader[18] with the possible exception of clustering. These are moments of an unique measure dq^D on $\mathcal{S}'(R^3)$. The theory satisfies all the Wightman axioms with the possible exception of the uniqueness of vacuum for $\mu = 0$ and the mass gap.

Proof: Since we have fixed the coefficient δm^2_δ of the mass renormalization counterterm independent of boundary conditions, it follows from Refs. 5, 6 that for $f_i > 0$

$$S^{D}_{\Lambda,\delta}(f_1,\ldots,f_n) \le S^{P,F}_{\Lambda,\delta}(f_1,\ldots,f_n),$$
$$S^{D}_{\Lambda,\delta}(f_1,\ldots,f_n) \le S^{D}_{\Lambda',\delta}(f_1,\ldots,f_n) \quad \text{if } \Lambda \subset \Lambda'.$$

From the convergence of the lattice approximation (Theorem 2.1 (a) and the result in Refs. 1, 2) it follows that for $f_i > 0$

$$S^{D}_{\Lambda}(f_1,\ldots,f_n) \le S^{P,F}_{\Lambda}(f_1,\ldots,f_n) \le K^n n! \prod_{i=1}^{n} |f_i|,$$
$$S^{D}_{\Lambda}(f_1,\ldots,f_n) \le S^{D}_{\Lambda'}(f_1,\ldots,f_n), \quad \Lambda \subset \Lambda'. \tag{2.12}$$

Here we have used (2.11) (or the result of Ref. 15). The first part of the theorem follows from (2.12) together with the method used in the $(\lambda\phi^4 - \sigma\phi^2 - \mu\phi)_2$ field theory.[5,6,14,15] By the Osterwalder and Schrader's reconstruction theorem, we only need to prove the uniqueness of vacuum for $\mu \neq 0$. This follows from correlation inequalities, Lee–Yang theorem, and the method used in Ref. 15. This completes the proof of the theorem. ∎

The rest of this paper is devoted to the proof of Theorem 2.1. We shall employ the method used in Refs. 1, 2 with some modifications, which are necessary to control the localization factors.

III. REDUCTION OF THE PROBLEM

In this section we first introduce the lattice and momentum cutoff Schwinger functions and then reduce the proof of Theorem 2.1 to the uniform convergence (with respect to δ) of the lattice and momentum cutoff Schwinger functions as $\kappa \to 1$ (Propositions 3.2–3.3). At the end of the section we also establish basic estimates which we will use in the following sections. Following Glimm and Jaffe,[7] we introduce a momentum cutoff function of the form

$$\kappa(k) = \prod_{i=1}^{3} [\eta(k^{(i)}/\beta^{(i)}) - \eta(k^{(i)}/\alpha^{(i)})],$$
$$\alpha^{(i)} < \beta^{(i)}, \tag{3.1}$$
$$\alpha^{(i)}, \beta^{(i)} \in \{M_0 = 0, M_j = M_1^{(1+\nu)^{j-1}} \text{ if } j \ge 1\},$$

where $M_1 > 1$ and $\nu > 0$ are constants given in Ref. 7, and η is a fixed $C_0^\infty(R)$ function satisfying

$$\eta(x) = 1 \text{ for } |x| \le \tfrac{1}{2}, \quad 0 < \eta(x) < 1 \text{ for } \tfrac{1}{2} < |x| < 2$$
$$\eta(x) = 0 \text{ for } |x| \ge 2 \text{ and } \eta(x) = \eta(-x)$$

and by convention $\eta(k/0) = 0$. We introduce

$$k_\delta^{(i)} = 2\delta^{-1}\sin(\delta k^{(i)}/2). \tag{3.2}$$

We define the lattice and momentum cutoff free fields by

$$\phi^X_{\delta,\kappa}(n) = \delta^3 \sum_{m\delta \in \Lambda_\delta} \phi_\delta(m\delta) \sum_{k \in T^X_\delta} f^X_k(m\delta)\bar{f}^X_k(n\delta)\kappa(k_\delta). \tag{3.3}$$

Notice that the momentum cutoff function $\kappa(k_\delta)$ depends on δ.

Direct computation yields

$$\langle \phi^X_{\delta,\kappa}(m)\phi^X_{\delta,\kappa}(n)\rangle^{(X)}_{\Lambda,\delta} = C^X_{\Lambda,\kappa}(m\delta, n\delta),$$
$$C^X_{\Lambda,\kappa}(m\delta, n\delta) = \sum_{k \in T^X_\delta} f^X_k(m\delta)\bar{f}^X_k(n\delta)\mu_\delta(k)^{-2}\kappa(k_\delta)^2. \tag{3.4}$$

The lattice and momentum cutoff partition function and "unnormalized" (un) Schwinger functions are defined by

$$Z^X_{\Lambda,\delta,\kappa} = \langle \exp(-V^X_{\Lambda,\delta,\kappa})\rangle^{(X)}_{\Lambda,\delta},$$
$$S^{X,\mathrm{un}}_{\Lambda,\delta,\kappa}(f_1,\ldots,f_n) = \langle \phi_\delta(f_1)\cdots\phi_\delta(f_n)\exp(-V^X_{\Lambda,\delta,\kappa})\rangle^{(X)}_{\Lambda,\delta}, \tag{3.5}$$

where the triple cutoff interacting measure $V^X_{\Lambda,\delta,\kappa}$ is obtained from $V^X_{\Lambda,\delta}$ in (2.7) by replacing $\phi_\delta(n)$ and $C^X_\Lambda(m\delta, n\delta)$ by $\phi^X_{\delta,\kappa}$ and $C^X_{\Lambda,\kappa}(m\delta, n\delta)$ respectively. We then have

Lemma 3.1: Let $\lambda > 0$, $\sigma, \mu \in R$, and let $X = \mathrm{D}, \mathrm{N}, \mathrm{P}$.

(a) For $\delta > 0$ and $f_i \in \mathcal{S}(\Lambda)$

$$\lim_{\kappa \to 1} Z^X_{\Lambda,\delta,\kappa} = Z^X_{\Lambda,\delta},$$
$$\lim_{\kappa \to 1} S^{X,\mathrm{un}}_{\Lambda,\delta,\kappa}(f_1,\ldots,f_n) = S^{X,\mathrm{un}}_{\Lambda,\delta}(f_1,\ldots,f_n).$$

(b) Let κ have compact support. Then for $f_i \in \mathcal{S}(\Lambda)$

$$\lim_{\delta \to 0} Z^X_{\Lambda,\delta,\kappa} = Z^X_{\Lambda,\kappa},$$
$$\lim_{\delta \to 0} S^{X,\mathrm{un}}_{\Lambda,\delta,\kappa}(f_1,\ldots,f_n) = S^{X,\mathrm{un}}_{\Lambda,\kappa}(f_1,\ldots,f_n)$$

exist.

Proof: Under the assumptions, all the objects we deal with are well defined by virtue of the momentum cutoffs κ and δ. For example, it is easy to check that $V^X_{\Lambda,\delta,\kappa} \in L^P(\mathcal{S}'(\Lambda), d\mu^X_{\Lambda,\delta})$ and

$$\exp(-V^X_{\Lambda,\delta,\kappa}) \le \mathrm{const}(\delta,\kappa),$$

where $\mathrm{const}(\delta,\kappa)$ is finite if either $\delta > 0$ or else κ has compact support. The proof of the lemma follows as in Refs. 5, 13. ∎

For convenience we write

$$Z^X_{\Lambda,\kappa} = Z^X_{\Lambda,\delta=0,\kappa},$$
$$S^{X,\mathrm{un}}_{\Lambda,\kappa}(f_1,\ldots,f_n) = S^{X,\mathrm{un}}_{\Lambda,\delta=0,\kappa}(f_1,\ldots,f_n). \tag{3.6}$$

As in Refs. 1, 2 we reduce the proof of Theorem 2.1 to the following propositions:

Proposition 3.2: Let $X = \mathrm{D}, \mathrm{N}, \mathrm{P}$ and $f_i \in \mathcal{S}(\Lambda)$, $i = 1,\ldots,n$.

(a) There are constants $K_1(\lambda,\sigma,\mu)$ and $K_2(\lambda,\sigma,\mu)$ independent of Λ, and a Schwartz space norm $|\cdot|$ such that

$$Z^X_{\Lambda,\kappa} \le \exp(K_1|\Lambda|)$$
$$S^{X,\mathrm{un}}_{\Lambda,\kappa}(f_1,\ldots,f_n) \le n!\left(\prod_{i=1}^{n}|f_i|\right)\exp(K_2|\Lambda|).$$

(b) Let $\delta \ge 0$. As $\kappa \to 1$,

$$Z^X_{\Lambda,\delta,\kappa} \to Z^X_{\Lambda,\delta} \quad \text{uniformly in } \delta,$$
$$S^{X,\mathrm{un}}_{\Lambda,\delta,\kappa}(f_1,\ldots,f_n) \to S^{X,\mathrm{un}}_{\Lambda,\delta}(f_1,\ldots,f_n) \quad \text{uniformly in } \delta.$$

(c) For $\tau \in C$ and $f \in \mathcal{S}(\Lambda)$ we write $Z^X_{\Lambda,\delta,\kappa}(\tau f) = \langle \exp(\tau\phi_\delta(f) - V^X_{\Lambda,\delta,\kappa})\rangle^{(X)}_{\Lambda,\delta}$. Then, as $\kappa \to 1$,

$$Z^X_{\Lambda,\delta,\kappa}(\tau f) \to Z^X_{\Lambda,\delta}(\tau f) \text{ uniformly in } \delta.$$

The convergence is uniform in τ for a compact subset of C.

Proposition 3.3: Let m_0^2 be sufficiently large (depending on Λ). Then

$$\lim_{\kappa \to 1} X^X_{\Lambda,\kappa} = Z^X_\Lambda > 0.$$

The proofs of Propositions 3.2 and 3.3 are delayed to the next section. Using the above results, we now prove Theorem 2.1.

Proof of Theorem 2.1: We first note that Lemma 3.1, Proposition 3.2 (b)—(c), and a 3ϵ argument yield

$$Z^X_{\Lambda,\delta} \xrightarrow[\delta\to 0]{} Z^X_\Lambda, \quad Z^X_{\Lambda,\delta} \xrightarrow[\delta\to 0]{} Z^X_\Lambda(\tau f),$$
$$S^{X,\mathrm{un}}_\Lambda(f_1,\ldots,f_n) \xrightarrow[\delta\to 0]{} S^{X,\mathrm{un}}_\Lambda(f_1,\ldots,f_n). \tag{3.7}$$

If we prove that

$$Z^X_\Lambda > 0 \tag{3.8}$$

for all $m_0^2 > 0$, Theorem 2.1 follows from Proposition 3.2 (a), (3.7), and (3.8).

We now prove (3.8) by using Proposition 3.3, the method of the mass shift transformation, and the Jensen's inequality. Identifying $\phi_\delta(n) = q_n$ and following Guerra, Rosen, and Simon.[5,13] One may write

$$Z^X_{\Lambda,\delta} = \mathrm{const} \int \prod^n dq_n \exp[-\tfrac{1}{2} q \cdot A^X_\Lambda \cdot q - \sum_n V^X_{\Lambda,\delta}(q_n)], \tag{3.9}$$

where the constant is the normalization factor for free measure, $\sum_n V^X_{\Lambda,\delta}(q_n)$ is the polynomial obtained by replacing $\phi_\delta(n)$ with q_n in the definition of $V^X_{\Lambda,\delta}$, and A^X_Λ is the matrix corresponding to the $(-\Delta^X_\delta + m_0^2)$. For details we refer to Ref. 13, Sec. IX. Let $Z^X_{\Lambda,\delta}(m_0^2)$, $V^X_{\Lambda,\delta}(m_0^2)$, and $\langle\cdot\rangle^{(X)}_{\Lambda,\delta,m_0^2}$ denote the partition function, the interacting action, and the Gaussian expectation with respect to the mass m_0^2 respectively. With this notation one may write

$$Z^X_{\Lambda,\delta}(m^2) = \langle \exp(-V^X_{\Lambda,\delta}(m^2))\rangle^{(X)}_{\Lambda,\delta,m^2}.$$

From Proposition 3.3 it follows that $Z^X_\Lambda(m^2) > 0$ for sufficiently large m^2. Also the expression (3.9) yields

$$Z^X_{\Lambda,\delta}(m^2 - b^2) = \langle \exp[-V^X_{\Lambda,\delta}(m^2 - b^2) + \tfrac{1}{2} b^2 :\phi^2_\delta:_{m^2}(\chi_\Lambda)\rangle^{(X)}_{\Lambda,\delta,m^2} \times \{\langle \exp[\tfrac{1}{2} b^2 :\phi^2_\delta:_{m^2}(\chi_\Lambda)\rangle^{(X)}_{\Lambda,\delta,m^2}\}^{-1} \tag{3.10}$$

for $m^2 > b^2$ and $X = \mathrm{D}, \mathrm{N}, \mathrm{P}$, where $:\ :_{m^2}$ means the Wick ordering corresponding to the mass m. The above is the "mass shift transformation." It is not hard to check that

$$\langle \exp[\tfrac{1}{2} b^2 :\phi^2_\delta:_{m^2}(\chi_\Lambda)\rangle^{(X)}_{\Lambda,\delta,m^2} \geq \exp(-d|\Lambda|), \tag{3.11}$$

where d is a constant independent of Λ, δ. We denote C^X_{Λ,m_0^2} by the covariance corresponding to the mass m_0 and

$$\delta C^X_{\Lambda,m_1^2,m_2^2} = C^X_{\Lambda,m_1^2} - C^X_{\Lambda,m_2^2}.$$

Using the formula for the change of the Wick ordering (see also Ref. 2, Lemma 3.7) one may check that there is a constant $c^X(\Lambda,\delta)$ such that

$$V^X_{\Lambda,\delta}(m^2 - b^2) = V^X_{\Lambda,\delta}(m^2) + 6 \sum_{n\delta\in\Lambda_\delta} [\delta C^X_{\Lambda,m^2,m^2-b^2}(n\delta, n\delta)] :\phi^2_\delta:_{m^2} + c^X(\Lambda,\delta), \tag{3.12}$$

where $|c^X(\Lambda,\delta)| \leq \mathrm{const}(\Lambda)$ uniformly in δ. The proof of (3.12) is elementary. We only remark that two divergent scalar terms appear during the change of interacting action from $V^X_{\Lambda,\delta}(m^2 - b^2)$ to $V^X_{\Lambda,\delta}(m^2)$ coming from the mass and vacuum renormalization counter terms. These are cancelled out explicitly and leave a regular scalar term. We denote the second term in the right-hand side of (3.12) by $:\phi^2_\delta:_{m^2}(\delta C^X_\Lambda)$. From Theorem 4.1 in the next section it follows that

$$|\langle [:\phi^2_\delta:_{m^2}(\delta C^X_\Lambda)] \exp[-V^X_{\Lambda,\delta}(m^2)]\rangle^{(X)}_{\Lambda,\delta,m^2}| \leq \mathrm{const}(\Lambda) \tag{3.13}$$

uniformly in δ. The above uniform bound is a consequence of smoothness of $\delta C^X_\Lambda(n\delta, n\delta)$ and Theorem 4.1. We do not produce the explicit proof of (3.13) and leave it to the reader. We remark that $\delta C^P_\Lambda(n\delta, n\delta)$ is a constant bounded uniformly in δ. Let $m^2 > b^2$ and let m^2 be sufficiently large so that $Z^X_\Lambda(m^2) > 0$. Then the relation (3.10)—(3.12) and Jensen's inequality yield

$$\begin{aligned} Z^X_{\Lambda,\delta}(m^2 - b^2)[Z^X_{\Lambda,\delta}(m^2)]^{-1} &\geq \exp[-\mathrm{const}(\Lambda)]\langle \exp[-V^X_{\Lambda,\delta}(m^2) + \tfrac{1}{2} b^2 :\phi^2_\delta:_{m^2}(\chi_\Lambda) \\ &\quad + :\phi^2_\delta:_{m^2}(\delta C^X_\Lambda)]\rangle^{(X)}_{\Lambda,\delta,m^2} [Z^X_{\Lambda,\delta}(m^2)]^{-1} \\ &\geq \exp[-\mathrm{const}(\Lambda)] \exp\{\langle (\tfrac{1}{2} b^2 :\phi^2_\delta:_{m^2}(\chi_\Lambda) + :\phi^2_\delta:_{m^2}(\delta C^X_\Lambda)) \\ &\quad \times \exp[-V^X_{\Lambda,\delta}(m^2)]\rangle^{(X)}_{\Lambda,\delta,m^2} [Z^X_{\Lambda,\delta}(m^2)]^{-1}\}. \end{aligned} \tag{3.14}$$

Since $Z^X_\Lambda(m^2) > 0$ and since $Z^X_{\Lambda,\delta} \to Z^X_\Lambda$ as $\delta \to 0$, it follows that there exist a constant $\epsilon(\Lambda) > 0$ such that

$$Z^X_{\Lambda,\delta}(m^2) \geq \epsilon(\Lambda) > 0$$

for sufficiently small δ. We use the above result and (3.13) [and the analogous result for $:\phi^2_\delta:_{m^2}(\chi_\Lambda)$] to bound (3.14) from below by $\exp[-\mathrm{const}(\Lambda)]$. Hence we have

$$Z^X_{\Lambda,\delta}(m^2 - b^2) \geq \exp[-\mathrm{const}(\Lambda)] Z^X_{\Lambda,\delta}(m^2).$$

By taking the limit as $\delta \to 0$ and using the fact that $Z^X_{\Lambda,\delta}(m^2) > 0$, we have completed the proof of (3.8) and so Theorem 2.1. ∎

As explained briefly in the Introduction, the proof of Proposition 3.2—3.3 is based on the inductive expansion developed by Glimm and Jaffe,[7] and modified by Feldman.[8] To control the expansion, one needs basic estimates. The rest of this section is devoted to establish a technical lemma, which summarizes these basic estimates. We begin by introducing more notation. For $f \in L^\infty(\Lambda)$ we write

$$\tilde{f}_\delta(k) = (2\pi)^{-3/2} \delta^3 \sum_{n\delta\in\Lambda_\delta} f(n\delta) \exp(-ik\cdot n\delta), \quad k \in T^X_\delta. \tag{3.15}$$

The above is a discrete version of the Fourier transformation. During the inductive expansion we will divide the box Λ into the union of cubes.[7] To prevent double counts of the lattice points on the boundary of cubes, we consider half open cubes of the form

$$\Delta = \{x \mid x \in \Lambda,\ x^{(i)} \in [a^{(i)}, b^{(i)}),\ |a^{(i)} - b^{(i)}| = d,\ i = 1,2,3\}.$$

In the rest of this paper we use the following convention: *The center of a cube* Δ *refers to the nearest lattice side* $n\delta \in \Delta$ *from the geometric center of* Δ. If there are several nearest lattice points, we pick up one of them as the center of Δ. This definition is identical to the usual definition when $\delta = 0$. As in Refs. 1, 2 we also write

$$\begin{aligned} &\mu(k)^2 = k^2 + m_0^2, \quad \mu(k^{(i)})^2 = (k^{(i)})^2 + m_0^2, \\ &\mu_\delta(k^{(i)})^2 = \delta^{-2}[2 - \cos(\delta k^{(i)})] + m_0^2, \quad i = 1,2,3, \\ &\tilde{\chi}_{\Delta,\delta}(k) = (2\pi)^{-3/2} \delta^3 \sum_{n\delta\in\Delta} \exp(-ik\cdot n\delta), \end{aligned} \tag{3.16}$$

$$F_{\Delta,\delta}(k)=\prod_{i=1}^{3}[|\Delta|^{1/3}\mu_\delta(k^{(i)})+1]^{-1},$$

where $|\Delta|$ is the volume of Δ. Finally we introduce a discrete version of derivatives with respect to $k\in T_\delta^X$ variables:

$$\begin{aligned}(D_\Lambda^X)^m&=\prod_{i=1}^{3}(D_\Lambda^{X,i})^{m^{(i)}},\quad m=(m^{(1)},m^{(2)},m^{(3)}),\\(D_\Lambda^{D,1}f)(k)&=(l_1/2\pi)[f(k^{(1)}+\pi/l_1,k^{(2)},k^{(3)})\\&\quad-f(k^{(1)}-\pi/l_1,k^{(2)},k^{(3)})],\\(D_\Lambda^{D,N,1}f)(k)&=(l_1/\pi)[f(k^{(1)}+\pi/2l_1,k^{(2)},k^{(3)})\\&\quad-f(k^{(1)}-\pi/2l_1,k^{(2)},k^{(3)})].\end{aligned}\tag{3.17}$$

$D_\Lambda^{X,2}$ and $D_\Lambda^{X,3}$ are defined in a similar manner. The following is the result corresponding to Ref. 1, Lemmas 2.1—2.3.

Lemma 3.4: Assume $k\in T_\delta^X$, $X=D,N,P$.

(a) For each $k\in R^3$, $\mu_\delta(k)\to\mu(k)$ as $\delta\to 0$.

(b) $\mu_\delta(k)^{-1}\le(\pi/2)\mu(k)^{-1}$.

(c) For $f\in\mathcal{S}(\Lambda)$, $|\tilde f_\delta(k)|\le \mathrm{const}(f)\mu(k)^{-2}$.

(d) $|(D_\Lambda^X)^m\mu_\delta(k)^{-2}|\le O(1)\mu(k)^{-2-|m|}$.

(e) We assume that $|\Delta|\le 1$ and the center of Δ is at the origin:

$$|(D_\Lambda^X)^m\tilde\chi_{\Delta,\delta}(k)|\le O(1)|\Delta|^{1+|m|/3}F_{\Delta,\delta}(k).$$

(f) Let κ be the momentum cutoff function defined in (3.1)—(3.3):

$$|(D_\Lambda^X)^m\kappa(k_\delta)|\le O(1)\min\{(\alpha^{(i)})^{-|m|}\,|\,i=1,2,3\}\chi_\kappa,$$

where χ_κ is the characteristic function of the support of $\kappa(k)$.

Proof: (a) Obvious. (b) This follows from the following inequality[5]:

$$2y^2\ge 1-\cos y\ge 2\pi^{-2}y^2\quad\text{if } y\in[-\pi,\pi].\tag{3.18}$$

(c) From the definitions in (2.2) and (3.15) it follows that

$$|\mu_\delta(k)\tilde f_\delta(k)|\le O(1)\delta\sum_{n\delta\in\Lambda_\delta}\Big|6f(n\delta)-\sum_{|n'-n|=1}f(n'\delta)\Big|.$$

We use the method employed in proving Lemma 2.2 in Ref. 1 to bound the above by $\mathrm{const}(f)$, where $\mathrm{const}(f)$ depends on the volume of the support of f. The lemma follows from the part (b).

(d)—(e) The proof follows by replacing D^m in the proof of Lemma 2.3, Ref. 1, by $(D_\Lambda^X)^m$ in (3.17) and adapting a method similar to that of Ref. 1.

(f) This follows from a straightforward computation and (3.18). We leave the detailed proof to the reader. ∎

IV. CONVERGENCE OF THE INDUCTIVE EXPANSION

In this section we prove Proposition 3.2 and Proposition 3.3 and thus complete the proof of Theorem 2.1. The proof is based on the method of inductive expansion developed by Glimm and Jaffe[7] and modified by Feldman.[8] We organize the proof close to that of Refs. 7, 8. We already have used the method successfully to establish the convergence of the lattice approximation of the $\lambda\phi_3^4$ field model with F and P boundary conditions.[1,2] The difficulty for $X=D,N$ is that we do not have the translation invariance of the free Gaussian measure $d\mu_\Lambda^{D,N}$. Hence it is much more complicated to isolate the localization factors for $X=D,N$ than for $X=F,P$. We must introduce a modification to the method of Glimm and Jaffe[7] so that we can overcome this difficulty.

As in Refs. 7, 8 the notation G may refer, depending on context, to the topological graph G, the function $G(q)$ on $\mathcal{S}'(\Lambda)$, or the kernel $G(k_i)$. Following Ref. 8, we introduce a norm of the graph G: For $\gamma>2\alpha>0$

$$\|G^{(X)}\|_{1,\gamma,\alpha}=\sup_{\rho}\sup_{C}\|\rho_\alpha^e C M^\gamma|G^{(X)}|\,\|_{H.S.}.$$

Here ρ_α^e, C, M^γ and $|\cdot|$ are "operators" that modify the graph G and its kernel. See Ref. 8 for detailed discussion. In our case the notation may differ slightly from that of Ref. 8. Let $C_{\Lambda,\delta}^X$ be the operator defined by its kernel in (2.1) and let $C_{\Lambda,\delta}^F$ be the corresponding operator for F boundary condition. We then have that[13] for $\delta\ge 0$

$$0<C_{\Lambda,\delta}^X\le cC_{\Lambda,\delta}^F,\tag{4.1}$$

where the constant c is independent of Λ,δ for $|\Lambda|\ge 1$. The above result has been proved in Ref. 13 for two-dimensional space—time and $\delta=0$. The same method yields (4.1). It then follows that

$$\begin{aligned}\|\phi^{(X)}(f)\|_{1,\gamma,\alpha}&=K\|(C_\Lambda^X)^{1/2-\gamma}f\|_2\\&\le K'\|(C_\Lambda^F)^{1/2-\gamma}f\|_2\equiv|f|_\gamma,\\\|\phi_\delta^{(X)}(f)\|_{1,\gamma,\alpha}&=K\|(C_{\Lambda,\delta}^X)^{1/2-\gamma}f\|_{l^2}\\&\le K'\|(C_{\Lambda,\delta}^F)^{1/2-\gamma}f\|_{l^2}\equiv|f_\delta|_\gamma.\end{aligned}$$

Notice that $|f|_\gamma$ and $|f_\delta|_\gamma$ are finite for $f\in C^\infty(\Lambda)$. From the method used in Ref. 7 it also follows that

$$\begin{aligned}\Big\|\prod_{i=1}^{n}\phi^{(X)}(f_i)\Big\|_{1,\gamma,\alpha}&\le n!\prod_{i=1}^{n}|f_i|_\gamma,\\\Big\|\prod_{i=1}^{n}\phi_\delta^{(X)}(f_i)\Big\|_{1,\gamma,\alpha}&\le n!\prod_{i=1}^{n}|f_{i,\delta}|_\gamma.\end{aligned}\tag{4.2}$$

For the detailed derivation we refer to Ref. 7.

Proposition 3.2 and Proposition 3.3 will follow as corollaries of the following results:

Theorem 4.1: Assume G is a graph having N external legs. Then there is a constant $K_1(\lambda,\sigma,\mu,\gamma,\alpha)$ independent of Λ, δ, and κ such that

$$|\langle G\exp(-V_{\Lambda,\delta,\kappa}^X)\rangle_{\Lambda,\delta}^{(X)}|\le N^N\|G\|_{1,\gamma,\alpha}\exp(K_1|\Lambda|)$$

for $X=D,N,P$.

Theorem 4.2: Let $\|G\|_{1,\gamma,\alpha}<\infty$ for some $0<\gamma<\gamma_0$. Then, as $\kappa\to 1$,

$$\lim_{\kappa\to 1}\langle G\exp(-V_{\Lambda,\delta,\kappa}^X)\rangle_{\Lambda,\delta}^{(X)}=[G]_{\Lambda,\delta}^{(X)}$$

exists for all $\delta\ge 0$. The convergence is uniform with respect to δ.

We postpone the proofs of Theorem 4.1 and Theorem

4.2. We now prove Proposition 3.2 and Proposition 3.3.

Proof of Proposition 3.2: (a) and (b). This follows from Theorem 4.1, Theorem 4,2, and (4.2) by setting $G=\prod_{i=1}^{n}\phi_\delta(f_i)$ (and $G=1$).

(c) This follows from the method which we will use to prove Theorems 4.1—4.2 and the method used in Ref. 8. Feldman[8] has shown a result similar to that of Proposition 3.2 (c) with $\delta=0$ and $X=\mathrm{F}$ by modifying the method used in proving Theorems 2—3 of Ref. 8. analogous procedure gives us the proof.

Proof of Proposition 3.3: This follows from Theorem 4.1—4.2 together with the method used in proving Theorem 3.5 of Ref. 2. Intuitively, one expects that $Z_\Lambda^X\to 1$ as $m_0^2\to\infty$, since $C_\Lambda^X(x,y)\to 0$ as $m_0^2\to\infty$. This proves the proposition.

Corollary 4.3: Let $S^X_{\Lambda,\kappa}(f_1,\ldots,f_n)$ be the momentum cutoff Schwinger function. Then the limit

$$\lim_{\kappa\to 1} S^X_{\Lambda,\kappa}(f_1,\ldots,f_n)=S^X_\Lambda(f_1,\ldots,f_n)$$

exists for $f_i\in\mathcal{S}(\Lambda)$, $X=\mathrm{D},\mathrm{N},\mathrm{P}$.

Proof: We set $\delta=0$ and $G=\prod_{i=1}^{n}\phi(f)$ (and $G=1$). Then the corollary follows from (4.2), Theorem 4.2, and the fact that $Z_\Lambda^X>0$ [Theorem 2.1 (b)]. ∎

Before proving Theorems 4.1—4.2, we describe briefly the structure of the proof which is parallel to that in Refs. 7, 8. Roughly speaking, the authors of the cited papers have proved the result corresponding to Theorem 4.1 and Theorem 4.2 for $\langle G\exp(-V^{\mathrm{F}}_{\Lambda,\kappa})\rangle^{(\mathrm{F})}$ by expanding it by a so-called "inductive construction."[7] Each inductive step consists of three main steps: (a) a high momentum (P—C) expansion, (b) a low momentum (Wick) expansion, (c) combinatoric estimates. Then the problem was reduced to estimates of the elementary integrals labeled by Feynman graphs. The combinatoric estimates have been used to bound the number of the terms in a sum of Feynman graphs. After isolating localization factors the kernels of graphs in each term have been bounded by majorizing functions. For the details we refer to Refs. 7, 8. As in Refs. 1, 2 we will replace $\langle G\exp(-V^{\mathrm{F}}_{\Lambda,\kappa})\rangle^{(\mathrm{F})}$ by $\langle G\exp(-V^X_{\Lambda,\delta,\kappa})\rangle^{(X)}_{\Lambda,\delta}$ and will perform exactly the same inductive expansion as in Refs. 7, 8. We will also assign combinatoric factors similar to those of Ref. 7 for combinatoric estimates. We will have to introduce a modification to Ref. 7 (and Refs. 1, 2) to isolate localization factors because of lack of translation invariance for $X=\mathrm{D},\mathrm{N}$. The rest of the proof will be similar to that in Refs. 1, 2, 7, 8.

We need more notations and some observations. We define

$$C^X_{\Lambda,\kappa,\kappa'}(m\delta,n\delta)\equiv\langle\phi_{\delta,\kappa}(m\delta)\phi_{\delta,\kappa'}(n\delta)\rangle^{(X)}_{\Lambda,\delta}=\sum_{k\in T^X_\delta} f^X_k(m\delta)\bar{f}^X_\delta(n\delta)u_\delta(\kappa,\kappa';k),\qquad(4.3)$$

$$u_\delta(\kappa,\kappa';k)=\mu_\delta(k)^{-2}\kappa(k_\delta)\kappa'(k_\delta).$$

It is easy to check that

$$C^{\mathrm{P}}_{\Lambda,\kappa,\kappa'}(m\delta,n\delta)=\frac{1}{|\Lambda|}\sum_{k\in T^{\mathrm{P}}_\delta} u_\delta(\kappa,\kappa';k)\exp[-ik\cdot[m-n)\delta],$$

$$C^X_{\Lambda,\kappa,\kappa'}(m\delta,n\delta)=\frac{1}{|\Lambda|}\sum_{k\in T^X_\delta} u_\delta(\kappa,\kappa';k)\prod_{i=1}^{3}\tfrac{1}{2}\{\exp[-ik^{(i)}(m^{(i)}-n^{(i)})\delta]+\epsilon_X\exp[-ik^{(i)}(m^{(i)}\delta+n^{(i)}\delta+l_i)]\},\qquad(4.4)$$

$X=\mathrm{D},\mathrm{N}$, $\epsilon_{\mathrm{N}}=1$, $\epsilon_{\mathrm{D}}=-1$.

We write, for $f\in L^\infty(\Lambda)$,

$$\tilde{f}^X_\delta(k_1,\ldots,k_m)=|\Lambda|^{m/2}\delta^3\prod_{n\delta\in\Lambda_\delta} f(n\delta)\prod_{i=1}^{m} f^X_{k_i}(n\delta).\qquad(4.5)$$

For $X=\mathrm{P}$ it follows from (3.15) and (4.5) that

$$\tilde{f}^{\mathrm{P}}_\delta(k_1,\ldots,k_m)=(2\pi)^{3/2}\tilde{f}_\delta(k_1+\cdots+k_m).$$

We next consider the momentum cutoff function $\kappa(k_\delta)$ more closely. Following Feldman,[8] we define λ, u, and U—the maximum lower cutoff, the minimum upper cutoff, and the maximum upper cutoff of a group of legs—by

$$\lambda=\max_{i,l}\{2,\alpha_l^{(i)}\},\quad u=\min_{i,l}\{\beta_l^{(i)}\},\quad U=\max_{i,l}\{\beta_l^{(i)}\},\qquad(4.6)$$

where $\eta_{\beta,\alpha}(k^{(i)}_{l,\delta})=\eta(k^{(i)}_{l,\delta}/\beta_l^{(i)})-\eta(k^{(i)}_{l,\delta}/\alpha_l^{(i)})$ is the momentum cutoff function in the ith space—time direction for the leg l. Because of the lattice cutoff we may assume that

$$\beta_l^{(i)}\le 2\pi/\delta\quad\text{for all } i \text{ and } l.\qquad(4.7)$$

Let

$$\mathrm{supp}\eta_{\beta,\alpha}(k^{(i)}_{l,\delta})\cap[-\pi/\delta,\pi/\delta]=[-\beta^{(i)}_{l,\delta},-\alpha^{(i)}_{l,\delta}]\cup[\alpha^{(i)}_{l,\delta},\beta^{(i)}_{l,\delta}].$$

We denote λ_δ, u_δ, and U_δ as the δ-dependent maximum lower cutoff, minimum upper cutoff, and maximum upper cutoff of a group of legs obtained by replacing $\alpha_l^{(i)},\beta_l^{(i)}$ by $\alpha^{(i)}_{l,\delta},\beta^{(i)}_{l,\delta}$, respectively, in (4.6). From (3.18) and the definition of k_δ in (3.2) it follows that

$$|\lambda/\lambda_\delta|^{\pm 1}\le O(1),\quad |u/u_\delta|^{\pm 1}\le O(1),\quad |U/U_\delta|^{\pm 1}\le O(1).\qquad(4.8)$$

Finally, from (4.3) it also follows that

$$|(C^X_{\Lambda,\kappa,\kappa'})(n\delta,n\delta)|\le O(1)U.\qquad(4.9)$$

We are ready to prove Theorem 4.1 and Theorem 4.2. As we stated before, we will only give a sketch of the proof and leave the details to the reader.

Sketch of Proof of Theorem 4.1: We only consider the proof of the theorem for the $\lambda\phi^4_3$ field model (i.e., the case of $\sigma=\mu=0$). Since the term $\sigma\phi^2+\mu\phi$ does not introduce any divergent counter terms, we only need to change certain combinatoric factors slightly in the following proof to account for this additional term. This can be done easily. See also the argument given in Ref. 11. We follow the same steps as in the proof of Theorem 3.1, Ref. 1.

Step 1: *The inductive expansion*: As a consequence of (4.8) and (4.9) one may employ the expansion of Glimm and Jaffe[7] and Feldman[8] to obtain

$$|\langle G''\exp(-V^X_{\Lambda,\delta,\kappa})\rangle^{(X)}_{\Lambda,\delta}|\le\Big|\sum_G I^X(G)\Big|,\qquad(4.10)$$

where $I^X(G)$ is the elementary integrals labeled by the Feynman graph G. We remark that in dealing with Wick vertices we have used a polynomial in the fields similar

to that in Sec. 3.2 of Ref. 7. This can be justified as a consequence of (4.9). We note that $I^X(G)$'s can be obtained "essentially" from those of Refs. 7, 8 by replacing

$$\left.\begin{array}{l}\int d^3k \\ \mu(k)^{-1} \\ (g\chi_\Delta)^\sim(k_1+\cdots+k_m) \\ \delta m_\kappa^2 \\ \kappa_S(k)\end{array}\right\} \text{ by } \left\{\begin{array}{l}\left(\frac{(2\pi)^3}{|\Lambda|}\right)\sum_{k\in T_\delta^P},\ \left(\frac{\pi^2}{|\Lambda|}\right)\sum_{k\in T_\delta^{D,N}} \\ \mu_\delta(k)^{-1} \\ (\tilde{\chi}_\Delta^X)(k_1,\ldots,k_m) \\ \delta m_{\delta,\kappa}^2 \\ \kappa_S(k_\delta).\end{array}\right. \tag{4.11}$$

The estimate leading to Theorem 4.1 is completed by first using the method of combinatoric factors to bound the number of terms in the sum of graph G and then bounding the size of each term.

Step 2: *The combinatoric factors*: The combinatoric bounds given in Ref. 8, Lemma 4.1, apply equally well to our case. This follows as a consequence of (4.8). See also the argument given in Step 2 of the Proof of Theorem 3.1 of Ref. 2. Using the combinatoric estimates, we have that

$$\left|\sum_G I^X(G)\right| \le \sup_G c(G)\left|I^X(G)\right|, \tag{4.12}$$

where $c(G)$ is the combinatoric coefficient given above.[7,8]

Step 3: *Localization factors*: This step is one of the most complicated parts of the proof. As explained before we need a modification to the method used in Refs. 1, 2, 7, 8 to deal with D, N boundary conditions. Since we do not have translation invariance of $d\mu^X_{\Lambda,\delta}$, $X=D,N$, we cannot translate each cube in $I^X(G)$ by a simple method. See Ref. 7, Sec. 5.2. To demonstrate our method for isolation of the distance factors from $I^X(G)$, we first consider the simplest graph given by

$$\begin{aligned} I^X_{\Delta,\Delta'}(G) &= \frac{1}{|\Lambda|}\sum_{k\in T_\delta^X}\{(\tilde{\chi}^X_\Delta)(k)(\bar{\tilde{\chi}}^X_{\Delta'})(k)\}u_\delta(\kappa,\kappa';k) \\ &= \delta^6 \sum_{\substack{n\delta\in\Delta \\ n'\delta\in\Delta'}} C^X_{\Lambda,\kappa,\kappa'}(n\delta,n'\delta), \end{aligned} \tag{4.13}$$

where Δ and Δ' are the cubes centered at $p\delta$ and $p'\delta, p, p'\in Z^3$, respectively. Let Δ_0 and Δ_0' be the cubes obtained by translating Δ and Δ' to the origin. We write

$$\begin{aligned} \tilde{\chi}_\Delta(k) &= \prod_{i=1}^3 \tilde{\chi}^{(i)}_\Delta(k^{(i)}), \\ \tilde{\chi}^{(i)}_\Delta(k^{(i)}) &= \delta \sum_{\substack{n^{(i)}\delta \\ n\delta\in\Delta}} \exp(-ik^{(i)}n^{(i)}\delta). \end{aligned} \tag{4.14}$$

We then have the following result:

Lemma 4.4: Let $I^X_{\Delta,\Delta'}$ be given by (4.13). For $X=\mathrm{D},\mathrm{N},\mathrm{P}$ and $\gamma=0,2,4,\cdots$,

$$\begin{aligned} &\left|\left|(p\delta-p'\delta)\right|^\gamma_X I^X_{\Delta,\Delta'}(G)\right| \\ &\le O(1)\max_{\pm 1}\left\{(\tfrac{1}{2})_X\frac{1}{|\Lambda|}\sum_{k\in T_\delta^X}\left|\left(\sum_{i=1}^3 (D^{X,i}_\Lambda)^2\right)^{\gamma/2}u_\delta(\kappa,\kappa';k)\right.\right. \\ &\left.\left.\times\tilde{\chi}_{\Delta_0}(\pm k)\tilde{\chi}_{\Delta_0'}(\pm k)\right|\right\}, \end{aligned}$$

where $\tilde{\chi}_\Delta(\pm k)=\prod_{i=1}^3\tilde{\chi}^{(i)}_\Delta((-1)^{n_i}k^{(i)})$ for $n=(n_1,n_2,n_3)\in Z^3$, $|n_i|\le 1$, and $(1/2)_P=1$ and $(1/2)_{D,N}=1/2$.

Remark 4.1: One may easily check that the expression in the lemma is bounded by $O(1)$ as a consequence of Lemma 3.4. Hence, the above result implies the uniform polynomial decay of the lattice and momentum cutoff covariance.

Proof: We first consider the lemma for $X=\mathrm{P}$ and $\gamma=2$. Translating Δ and Δ' to the origin, we get

$$\begin{aligned} I^P_{\Delta,\Delta'} &= \frac{(2\pi)^3}{|\Lambda|}\sum_{k\in T_\delta^P}u_\delta(\kappa,\kappa';k)\tilde{\chi}_{\Delta_0}(k)\tilde{\chi}_{\Delta_0'}(-k) \\ &\quad\times\exp[-ik\cdot(p-p')\delta]. \end{aligned} \tag{4.15}$$

For $|x^{(i)}|_P\le l_i/2$ and $|x^{(i)}|_{D,N}\le l_i$ we define

$$\begin{aligned} |x^{(i)}|^2_{P,\delta} &\equiv \tfrac{1}{2}\pi^2(l_i/2\pi)^2\{2-2\cos[(2\pi/l_i)x^{(i)}]\}, \\ |x^{(i)}|^2_{D,N,\delta} &\equiv \tfrac{1}{2}\pi^2(l_i/\pi)^2\{2-2\cos[(\pi/l_i)x^{(i)}]\}, \\ |x|^2_{X,\delta} &\equiv \sum_{i=1}^3 |x^{(i)}|^2_{X,\delta}. \end{aligned} \tag{4.16}$$

From (3.18) it follows that

$$|x^{(i)}|^2_X \le |x^{(i)}|^2_{X,\delta}, \quad X=\mathrm{D},\mathrm{N},\mathrm{P}, \tag{4.17}$$

if $|x^{(i)}|_P\le l_i/2$ and $|x^{(i)}|_{D,N}\le l_i$. We also note that

$$|n^{(i)}\delta|^2_{X,\delta}\exp(ik^{(i)}n^{(i)}\delta) = \tfrac{1}{2}\pi^2(D^{X,i}_\Lambda)^2\exp(ik^{(i)}n^{(i)}\delta), \quad i=1,2,3. \tag{4.18}$$

We now use (4.15), (4.17), and (4.18) and the discrete version of the integration by parts to obtain the lemma for $X=\mathrm{P}$ and $\gamma=2$. Here we also used the periodicity of $u_\delta(\kappa,\kappa';k)$ (period $2\pi/\delta$). The lemma for general γ follows from a method similar to that for $\gamma=2$. This proves the lemma for $X=\mathrm{P}$.

We next establish the lemma for $X=\mathrm{D}$ and $\gamma=2$. We substitute (4.4) into (4.13). We then translate Δ and Δ' to the origin to obtain

$$\begin{aligned} I^D_{\Delta,\Delta'} &= \frac{1}{|\Lambda|}\sum_{k\in T_\delta^D}u_\delta(\kappa,\kappa';k)\ \delta^6\Bigg\{\sum_{\substack{n\delta\in\Delta_0 \\ n'\delta\in\Delta_0}}\prod_{i=1}^3 \\ &\quad\times\tfrac{1}{2}\{\exp[-ik^{(i)}(n^{(i)}-n'^{(i)})\delta+ik^{(i)}(p^{(i)}-p'^{(i)}\delta] \\ &\quad-\exp[-ik^{(i)}(n^{(i)}+n'^{(i)})\delta+ik^{(i)}(p^{(i)}\delta+p'^{(i)}\delta+l_i)]\}. \end{aligned} \tag{4.19}$$

Let A be the set of all subsets of $\{1,2,3\}$. By expanding the above we have

$$\begin{aligned} I^D_{\Delta,\Delta'} &= \sum_{a\in A}(\pm 1)\left\{(\tfrac{1}{2})^3\frac{(2\pi)^2}{|\Lambda|}\sum_{k\in T_\delta^D}u_\delta(\kappa,\kappa';k)\right. \\ &\quad\times\left[\prod_{i=1}^3\tilde{\chi}^{(i)}_{\Delta_0}(k^{(i)})\tilde{\chi}^{(i)}_{\Delta_0'}(+k^{(i)})\exp\left(i\sum_{i\in a}k^{(i)}(p^{(i)}-p'^{(i)})\delta\right.\right. \\ &\quad\left.\left.\left.+\sum_{i\in\tilde{a}}k^{(i)}(p^{(i)}\delta+p'^{(i)}\delta+l_i)\right)\right]\right\} \\ &\equiv \sum_{a\in A} I^{D,a}_{\Delta,\Delta'}, \end{aligned} \tag{4.20}$$

where $\tilde{a}$ is the complement of $a\in A$ and (+) and (−) signs depend on a. An elementary calculation yields

$$|x^{(i)}-y^{(i)}| \le |x^{(i)}+y^{(i)}+l_i| \quad \text{for } |x^{(i)}|, |y^{(i)}|\le l_i/2. \tag{4.21}$$

We consider each term in (4.20) separately. From (4.17) and (4.21) it follows that

$$\left| |p\delta - p'\delta|^2_{\mathrm{D}} I^{\mathrm{D},a}_{\Delta,\Delta'} \right| \leq \left| \left(\sum_{i\in a} |(p^{(i)} - p'^{(i)})\delta|^2_{\mathrm{D},\delta} + \sum_{i\in\tilde{a}} |(p^{(i)} + p'^{(i)})\delta + l_i|^2_{\mathrm{D},\delta} \right) I^{\mathrm{D},a}_{\Delta,\Delta'} \right| . \tag{4.22}$$

We now use (4.18) and the argument below (4.18) to bound the above by

$$O(1)(\tfrac{1}{2})^3 \frac{\pi^2}{|\Lambda|} \sum_{k\in T^{\mathrm{D}}_\delta} \left| \left(\sum_{i=1}^{3} (D^{\mathrm{D},i}_\Lambda)^2 u_\delta(\kappa,\kappa';k) \times \prod_{i=1}^{3} \tilde{\chi}^{(i)}_{\Delta'_0}(k^{(i)}) \tilde{\chi}_{\Delta'_0}(\pm k^{(i)}) \right) \right| . \tag{4.23}$$

Since the above bound holds for each $a\in A$ and since there are 2^3 terms in (4.20), we proved the lemma for $X=\mathrm{D}$ and $\gamma=2$. The lemma for the other cases follows from a straightforward modification of the method used above. This completes the proof of the lemma. ∎

We now generalize the method used above to isolate the distance factors from more general graphs in (4.12). We begin by introducing more notation. We define

$$Z^m_{\leq 1} \equiv \{(n_1,\ldots,n_m) \,|\, n_j = 0,1,\ 1\leq j\leq m\},$$
$$Z^{3m}_{\leq 1} \equiv Z^m_{\leq 1}\times Z^m_{\leq 1}\times Z^m_{\leq 1}. \tag{4.24}$$

For given $a\in Z^{3m}_{\leq 1}$ $[a=(a^{(1)},a^{(2)},a^{(3)})]$ we define an operator P_a by

$$(P_a\tilde{\chi}_\Delta)(k_1+\cdots+k_m) = \prod_{i=1}^{3} (P_{a^{(i)}}\tilde{\chi}^{(i)}_\Delta)(k_1+\cdots+k_m),$$
$$(P_{a^{(i)}}\tilde{\chi}^{(i)}_\Delta)(k^{(i)}_1+\cdots+k^{(i)}_m) = \tilde{\chi}^{(i)}_\Delta((-1)^{n_1}k^{(i)}_1+\cdots+(-1)^{n_m}k^{(i)}_m), \tag{4.25}$$
$$a^{(i)} = (n_1,\ldots,n_m)\in Z^m_{\leq 1}.$$

We extend the above definition to the function $F_{\Delta,\delta}(k)$ defined in (3.16):

$$(P_aF_{\Delta,\delta})(k_1+\cdots+k_m) = \prod_{i=1}^{3} (P_{a^{(i)}}F^{(i)}_{\Delta,\delta})(k^{(i)}_1+\cdots+k^{(i)}_m). \tag{4.26}$$

Here we have written $F_{\Delta,\delta}=\Pi_{i=1}^3 F^{(i)}_{\Delta,\delta}$. Eventually, we will assign the operator P_a to the kernel of each vertex in a given Feynman graph G. Let $w(k_l)$ be the kernel of the graph G consisting of the kernels $\{w_i(k_{l,i}) \,|\, i=1,\ldots,q\}$ of the vertices in G. One may write

$$w(k_l) = \prod_{i=1}^{q} w_i(k_{l,i}).$$

Let M be the total number of legs in G. We assign an operator P_a, $a\in Z^{3M}_{\leq 1}$, to the kernel $w(k_l)$ in the following manner:

$$P_aw(k_l) = \prod_{i=1}^{q} (P_{a_i}w_i)(k_{l,i}),$$

where $(P_{a_i}w_i)(k_{l,i})$ is obtained from $w_i(k_{l,i})$ by replacing $\tilde{\chi}_\Delta(k_1+\cdots+k_m)$ in $w_i(k_{l,i})$ by $(P_{a_i}\tilde{\chi}_\Delta)(k_{l,i})$. Hence, for given P_a, $a\in Z^{3M}_{\leq 1}$, a_i's are defined implicitly and vise versa. We are now ready to control localization factors. Following Glimm and Jaffe,[7] we first divide cubes in G'', P, C, W-vertices such that the conditions in Ref. 7, Sec. 5.2, are satisfied. We then obtain

$$I^X(G) = \sum_\alpha I^X(G_\alpha), \tag{4.27}$$

where G_α is the graph corresponding to that of Refs. 7, 8. Hence, one may write

$$I^X(G_\alpha) = (1/|\Lambda|)^M \sum_{k's\in T^X_\delta} w^X_\alpha(k), \tag{4.28}$$

where M is the total number of k's in the summation. Let d^X_i be the scaled distance analogous to that defined in Ref. 7. We prove the main result in this step.

Proposition 4.5: For $X=\mathrm{D}$, N, P we have

$$I^X(G_\alpha) \leq \left(\prod_{\substack{\text{lines } l \\ \text{connect to} \\ P,\,C,\,W\ (\text{or } G'')\ \text{vertices}}} [\alpha_1 (d^X_l)^{\gamma m}] \right) \max_{a's} \{\bar{I}^X_a(G)\},$$
$$\bar{I}^X_a(G) = \left((\tfrac{1}{2})_X \frac{1}{|\Lambda|}\right)^m \sum_{k's\in T^X_\delta} \overline{P_aw}(k),$$
$$(\tfrac{1}{2})_P = 1, \quad (\tfrac{1}{2})_{\mathrm{D,N}} = \tfrac{1}{2},$$

where, with exception of the kernels of the mass renormalization cancellation diagrams, the kernel $\overline{P_aw}(k)$ is obtained from the corresponding majorizing kernel $\bar{w}(k)$ of Ref. 1, (3.15), by replacing each function $F_{\delta,\Delta}(k_1+\cdots+k_m)$ in $\bar{w}(k)$ of Ref. 1 by $(P_{a'}F_{\Delta,\delta})(k_1+\cdots+k_m)$, where $a'\in Z^{3m}_{\leq 1}$ is implicitly defined by a given $a\in Z^{6m}_{\leq 1}$. The kernels for the mass renormalization graphs are given by (4.34) in the following proof.

Remark 4.2: For $X=\mathrm{P}$ there is only one term on the right-hand side of the proposition, namely, the term corresponding to $P_a=1$ (see Ref. 2). One may expect that the term $\bar{I}^X_a(G)$ for $P_a=1$ is dominating the others corresponding to the case of $P_a\neq 1$. We will exploit this observation in the Appendix.

Proof: The complete proof of the proposition is very complicated mainly because of notational complication. In principle, the proposition will follow from the method used in proving Lemma 4.4 together with Lemma 3.4. For $X=\mathrm{P}$ the result follows as in the proof of Lemma 4.4 (see also the step 3 of Ref. 2, Theorem 3.1). We give the proof for $X=\mathrm{D}$. A similar method gives us the proof for $X=\mathrm{N}$.

Let us assume that the graph G_α does not contain W vertices and mass renormalization diagrams. We write $I^X(G_\alpha)$ in the lattice space expression (i.e., in the configuration space). For instance, see (4.13). We then have

$$I^{\mathrm{D}}(G_\alpha) = \delta^{3M'} \sum_{\substack{n_i\delta\in\Delta_i \\ i=1,\ldots M'}} \prod_{\substack{\text{lines } l \\ \text{connecting} \\ \text{between vertices}}} (C^{\mathrm{D}}_{\Lambda,\kappa,\kappa'})(\hat{n}_l\delta,\hat{n}'_l\delta), \tag{4.29}$$

where $\begin{cases} \hat{n}_l\neq\hat{n}'_{l_2} & \hat{n}_l\delta,\hat{n}'_l\delta\in\{n_i\delta \,|\, i=1,\ldots,M'\}, \\ \text{each } n_i\delta\in\Delta_i & \text{for some } \Delta_i, \\ \text{each } \hat{n}'_l\delta\in\Delta_j & \text{for some } \Delta_j. \end{cases}$

In the above expression each $C^{\mathrm{D}}_{\Lambda,\kappa,\kappa'}(\hat{n}_l\delta,\hat{n}'_l\delta)$ represents a line l connecting two lattice points localized in two cubes, say, Δ_i and Δ_j $(\Delta_i\neq\Delta_j)$, and M' is the number of vertices in the graph G_α. We now follow the procedure used in the proof of Lemma 4.4. We first substitute (4.4) into (4.29) and then translate each cube Δ_i, $i=1,\ldots,M'$, to the origin. One then has

$$I^{D}(G_\alpha)=\delta^{3M'}\sum_{\substack{n_l\delta\in\Delta_{i,0}\\ i=1,\dots,M'}}\prod_{\text{lines } l}\Bigg(\sum_{k_l\in T_\delta^D}u_\delta(\kappa,\kappa';k_l)$$
$$\times\frac{1}{|\Lambda|}\prod_{i=1}^{3}\tfrac{1}{2}(\exp\{-ik_l^{(i)}[(\hat n_l^{(i)}-\hat n_l'^{(i)})\delta$$
$$+(p_l^{(i)}-p_l'^{(i)})\delta]\}-\exp\{-ik_l^{(i)}[(\hat n_l^{(i)}$$
$$+\hat n_l'^{(i)})\delta+(p_l^{(i)}\delta+p_l'^{(i)}\delta+l_i)]\}\Bigg), \tag{4.30}$$

where $\Delta_{i,0}$ is the cube of volume $|\Delta_i|$ centered at the origin, and $p_l\delta$ and $p_l'\delta$ are the centers of the cubes connected by the line l. We expand (4.30); we then have 2^{3M_l} terms, where M_l is the total number of the lines in G_α. Each term has a multiplication factor $(1/2)^{3M_l}$. By changing the order of summations in (4.30) we may write

$$I^{D}(G_\alpha)=\sum_{a\in Z_{\le 1}^{6M_l}}[(\pm 1)|\Lambda|^{-M}2^{-6m_l}\sum_{k_l's\in T_\delta^D}\prod_{\text{vertices}}(P_a\tilde w)(k_l)], \tag{4.31}$$

where

$$\Bigg(\prod_{\text{vertices}}(P_a\tilde w)(k_l)\Bigg)$$
$$=\Bigg(\prod_{\text{lines}}u_\delta(\kappa,\kappa';k_l)\Bigg)\prod_{i=1}^{3}\Bigg[\Bigg\{\prod_{\text{vertices}}(P_a\tilde\chi_\Delta^{(i)})(k_{1,l}^{(i)}+\cdots+k_{m,l}^{(i)})\Bigg\}$$
$$\times\prod_{\text{lines }l}\begin{cases}\exp[ik_l^{(i)}(p_l^{(i)}-p_l'^{(i)})\delta] & \text{if } n_l^{(i)}=0\\ \exp[ik_l^{(i)}(p_l^{(i)}\delta+p_l'^{(i)}\delta+l_i)] & \text{if } n_l^{(i)}=1\end{cases}\Bigg].$$

Here $\{k_{1,l},\dots,k_{m,l}\}\subset\{k_l\,|\,l=1,\dots,M_l\}$ and n_l is a component of $a\in Z_{\le 1}^{3M_l}$. Notice that each component of line l has a factor either $\exp[ik_l^{(i)}(p_l^{(i)}-p_l'^{(i)})\delta]$ or else $\exp[ik_l^{(i)}(p_l^{(i)}\delta+p_l'^{(i)}\delta+l_i)]$ depending on $a\in Z_{\le 1}^{3M_l}$. We now employ the method in the proof of Lemma 4.4 to pull out the distance factor from the above exponential functions. See the step used in (4.20)–(4.23). The proposition follows from Lemma 3.2 and the method used in Ref. 7.

For a graph G_α containing W vertices, we follow a similar procedure. We need to be careful to handle W vertices. The kernel of an individual W vertex has the form [(6.2.2)–(6.2.4) of Ref. 7]

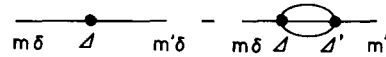
$$w(k_1,\dots,k_l)=(\mu_{1,\delta}\cdots\mu_{l,\delta})^{-1}v_\delta(k_1,\dots,k_l),$$
$$v_\delta(k_1,\dots,k_l)=c\sum_{J_1\subset I_1}\prod_{\substack{i=1\\ \{1,\dots,m\}=J_1}}^{m}\frac{\tilde\chi_\Delta^{D}(k_i)}{\tilde\chi_\Delta(0)}$$
$$\times\ \tilde\chi_\Delta^{D}(k_{m+1},\dots,k_l)\eta(k_\delta), \tag{4.32}$$

where η is the product of momentum cutoff functions. See Ref. 7 for the detailed discussion. We follow a procedure similar to that used before. In the procedure we consider (4.32) as the kernel of one W vertex. That is, in the final step in which (4.31) was derived, the sign of each $k_j^{(i)}$, $j=1,\dots,l$, $i=1,2,3$, is the same in each term in (4.32):

$$(P_av_\delta)(k_1,\dots,k_l)=c\sum_{J_1\subset I_1}\prod_{i=1}^{m}\frac{\tilde\chi_\Delta(-1)^{n_1}k_i}{\tilde\chi_\Delta(0)}\ \tilde\chi_\Delta(-1)^{n_{m+1}}k_{m+1}$$
$$+\cdots+(-1)^{n_l}k_l), \tag{4.33}$$

where $a=(n_1,\dots,n_l)$, $n_i=(n_i^{(1)},n_i^{(2)},n_i^{(3)})$. This allows us to exhibit the low momentum cancellation explicitly as in Ref. 7. The rest follows as in Ref. 7.

Finally, we consider the mass renormalization cancellation diagrams. The majorizing function for a given diagram is very complicated because of our lack of the translation invariance of covariances for $X=D,N$. We will derive it carefully. The diagram under consideration corresponds to the following expression in the lattice space:



$$=\Bigg\{\tfrac{1}{2}\lambda^2\delta m^2_{\delta,\kappa,\kappa'}\sum_{\substack{n\delta\in\Delta\\ n'\delta\in\Delta'}}\delta_{\Delta,\Delta'}-\tfrac{1}{2}\lambda^2 4^2\cdot 6$$
$$\times\sum_{\substack{n\delta\in\Delta\\ n'\delta\in\Delta'}}\Bigg(\prod_{j=1}^{3}C^{D}_{\Lambda,\kappa_j,\kappa_j'}(n\delta,n'\delta)\Bigg)\Bigg\}$$
$$\times C^{D}_{\Lambda,\kappa_4,\kappa_4'}(m\delta,n\delta)C^{D}_{\Lambda,\kappa_5,\kappa_5'}(n'\delta,m'\delta), \tag{4.34}$$

where $\kappa=\prod_{i=1}^{3}\kappa_i$ and $\kappa'=\prod_{i=1}^{3}\kappa_i'$. We use (4.4) for $C^{D}_{\Lambda,\kappa_j,\kappa_j'}$, $j=4,5$, in the above expression. We do not decompose $C^{D}_{\Lambda,\kappa_j,\kappa_j'}$ for $j=1,2,3$. We follow the procedure used before. After isolating localization factors, the majorizing function of the kernel of the mass renormalization cancellation diagram has the form

$$\Bigg|D^{m}_{\Lambda,k_4}D^{m'}_{\Lambda,k_5}\Bigg[(u_{4,\delta}{}^{-1}u_{5,\delta}{}^{-1}\kappa_S(k_{4,\delta})\kappa_S(k_{5,\delta}))$$
$$\times\sum_{\substack{n\delta\in\Delta\\ n'\delta\in\Delta'}}\tfrac{1}{2}\lambda^2\Bigg\{\delta m^2_{\delta,\kappa,\kappa'}\delta_{\Delta,\Delta'}-4^2\cdot 6$$
$$\times\Bigg(\prod_{j=1}^{3}C^{X}_{\Lambda,\kappa_j,\kappa_j'}((n+p)\delta,(n'+p')\delta)\Bigg)\Bigg\}$$
$$\times\Bigg(\prod_{i=1}^{3}\exp(\pm ik_4^{(i)}n^{(i)}\delta)\Bigg)\Bigg(\prod_{i=1}^{3}\exp(\pm ik_5^{(i)}n'^{(i)})\Bigg)\Bigg]\Bigg|, \tag{4.35}$$

where $X=D$, $p\delta$ and $p'\delta$ are the centers of Δ and Δ' respectively and the $(\pm)$ signs in the exponential functions are determined by a given $a\in Z_{\le 1}^{6M_l}$ in the proposition. The detailed derivation of (4.34) is left to the reader. This completes the proof of the proposition.

Step 4: *Estimates of* $\tilde I_a^X(G)$'s: We assert that, for each $a\in Z_{\le 1}^{6M_l}$ and $X=P,D,N$, $\tilde I_a^X(G)$ is bounded above by a product of factors given by those of Ref. 1, Step 4 in the proof of Theorem 3.2 of Ref. 2 (equivalently those of Lemma 5.1, Ref. 8). The theorem now follows from Proposition 4.5, the above assertion, and the method in the proof of Theorem 3, Ref. 8.

We prove our assertion. We use the method of decomposing big graphs into small graphs to estimate $\tilde I_a^X(G)$ for each $a\in Z_{\le 1}^{6M_l}$ and $X=D,N,P$. We adapt the decomposition process of Refs. 1, 7, 8. As in Refs. 1, 7, 8, it is sufficient to establish the bounds corresponding to Ref. 1, Propositions A.1, 1–5 for small graphs [for each $a\in Z_{\le 1}^{6M_l}$ and $X(=D,N,P)$]. The estimates of the small graphs corresponding to those of Ref. 1 are given in the Appendix. See Propositions A.1–2 in the Appendix. This proves our assertion and completes the proof of the theorem.

Proof of Theorem 4.2: We employ the method used in

the proof of Theorem 3 of Ref. 8 (also Theorem 3.4 of Ref. 1) together with Theorem 4.1 and the method of its proof. The proof then follows as in Refs. 8. We leave again the detailed proof to the reader.

APPENDIX: ESTIMATES OF SMALL GRAPHS AND THE MASS RENORMALIZATION CANCELLATIONS

In this appendix we establish the basic estimate for small graphs including the mass renormalization cancellations and obtain the same upper bounds as those of Propositions A.1—A.5 of Ref. 1 (equivalently Propositions 5.3.3—5.3.6 of Ref. 7 and Theorem 6 of Ref. 8). We will only sketch the proof. For the detailed structure of the proofs we refer the reader to Refs. 1, 7, 8. We begin by proving the following:

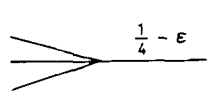
Proposition A.1:

For each $a \in Z_{\leq 1}^{6M^1}$ and X (=D, N, P), the bounds of the small graphs corresponding to Proposition A.1.1 and Propositions A.1.3—A.1.5 of Ref. 1 hold in our cases.

Proof: We will reduce the problem to that of Ref. 1. From Proposition 4.5 it follows that if one replaces

$$\left.\begin{aligned}&\int_{-\pi/6}^{\pi/6} d^3k\\ &F_{\delta,\Delta}(k_1+\cdots+k_m)\end{aligned}\right\} \text{ by } \begin{cases}(\tfrac{1}{2})_X^3 \dfrac{(2\pi)^3}{|\Lambda|}\displaystyle\sum_{k\in T_\delta^X}\\ (P_a F_{\Delta,\delta})(k_1+\cdots+k_m)\end{cases} \tag{A1}$$

in the expressions for small graphs in Ref. 1, one obtains the corresponding expressions in our cases. To give the general structure of the proof, we prove the proposition corresponding to Proposition A.1.1 (a) of Ref. 1. The majorizing kernel of the graph

$A_a^{(X)} =$ 

is given by

$$v_a^X(k_1,\ldots,k_4) = |\Delta|(u_1u_2u_3)^{-1}u_4^{-3/4+6}(P_aF_{\Delta,\delta})(k_1+\cdots+k_4)\prod_{i=1}^4 \chi_{\kappa S,i}.$$

See Proposition 4.5 and Sec. 6.2 of Ref. 7 for the details. Hence,

$$\|A_a^{(X)}\|_{3,1} = \|(A_a^{(X)})^*(A_a^{(X)})\|_{1,1}^{1/2} \leq \|(A_a^{(X)})^*(A_a^{(X)})\|_{H.S.}^{1/2}$$

$$\leq \left[(\tfrac{1}{2})_X^6\left(\frac{(2\pi)^3}{|\Lambda|}\right)^2 \sum_{k_1,k_5\in T_\delta^X} |\tilde{v}_a^X(k_1,k_5)|^2\right]^{1/2}, \tag{A2}$$

$$|\tilde{v}_a^X(k_1,k_5)| \leq (\tfrac{1}{2})_X^9\left(\frac{(2\pi)^3}{|\Lambda|}\right)^3 \sum_{\substack{k_i\in T_\delta^X\\ i=2,3,4}} |v_a^X(k_2,k_3,k_4,k_1)| \times |v_a^X(k_2,k_3,k_4,k_5)|.$$

Since $u(k)$ and $\kappa_S(k)$ [and also $u_\delta(k)$ and $\kappa_s(k_\delta)$] are even functions of k, we can set $P_a=1$ in the estimate of (A2) for all $a \in Z_{\leq 1}^{6M^1}$. We also note that since

$$[(|k^{(i)}|+1)/(|k^{(i)}\pm\pi/l_i|+1)]^{\pm 1} \leq O(1), \tag{A3}$$

we can replace

$$(\tfrac{1}{2})_X^3\left(\frac{(2\pi)^3}{|\Lambda|}\right)\sum_{k\in T_\delta^X} \text{ by } \int_{-\pi/\delta}^{\pi/\delta} d^3k \tag{A4}$$

to get the upper bound of (A2). Hence, $\|A_a^{(X)}\|_{3,1}$ in (A2) is bounded by that of Ref. 1 for free boundary condition [(A.3.5)—(A.3.7) of Ref. 1]. As in Ref. 1, $\|A_a^{(X)}\|_{3,1}$ is bounded by

$$O(1)|\Delta|^{\epsilon'}(\log v(2))^{1/2} \quad \text{for } \epsilon'>0.$$

This proves the proposition for the graph under consideration.

To prove the proposition for the remaining cases, we only need to combine the method used above together with that used in Ref. 1. In the most cases, the estimations are reduced to estimates of operator norms of the majorizing kernel of a single vertex (or Hilbert—Schmidt norm of the kernel). During a certain stage of the estimation we can set $P_a=1$ and use the replacement in (5.4) by the same reasoning as before. One may check explicitly that for a given graph the upper bound for $P_a \neq 1$ is bounded by that for $P_a=1$. Then the problems are reduced to those for free boundary condition in Ref. 1. The proof follows as in Ref. 1. ∎

We finally consider the cancellation of the mass renormalization. We state the result.

Proposition A.2: (The mass renormalization cancellation): Let $w_a^X(k_4,k_5)$ be the kernel of the mass renormalization cancellation diagrams given in (4.34). Then the result corresponding to that of Proposition A.1.2 of Ref. 1 holds for all $a \in z_{\leq 1}^{6M^1}$ and X:

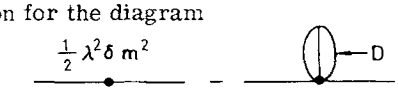
$$\|w_a^X\| \leq O(1)\,(|\Delta|\,|\Delta'|)^\epsilon \lambda^{-\epsilon d^{-n}}$$

for some $\epsilon>0$. Here d is the scaled distance from Δ to Δ'.

Proof: We will reduce the problem to that of free boundary condition treated in Ref. 1. Then the result will follow as in Ref. 1. We introduce

$$\delta m_{\Lambda,\delta,\kappa,\kappa'}^{(X)}(m\delta) = 4^2\cdot 6\,\delta^3 \sum_{n\delta\in\Lambda_\delta}\prod_{i=1}^3 [C_{\Lambda,\kappa_i,\kappa_i'}^X(m\delta,n\delta)]. \tag{A5}$$

The above is the coefficient of the mass renormalization counterterm appropriate to X boundary condition. We will prove the proposition for $X=$D. The proposition for $X=$P, N follows from a straightforward modification of the method for $X=$D. We first show the result in the proposition for the diagram



That is, we prove

$$\left\|\left(\frac{\pi}{|\Lambda|}\right)^2 \sum_{k_4,k_5\in T_\delta^D}\sum_{m\delta\in\Delta_0}\left[[\delta m_{\delta,\kappa,\kappa'}^2 - \delta m_{\Lambda,\delta,\kappa,\kappa'}^{(D)}((m+p)\delta)]\right.\right.$$
$$\times D_{\Lambda,k_4}^m D_{\Lambda,k_5}^{m'}\left(\prod_{i=4}^5 u_\delta(k_i)^{-1}\kappa_S(k_{i,\delta})\right)$$
$$\left.\left.\times\prod_{i=1}^3 \exp[\pm i(k_4^{(i)}\pm k_5^{(i)})mb]\right]\right\| \leq O(1)|\Delta|^{2\epsilon}\lambda^{-\epsilon}, \tag{A6}$$

where Δ_0 is a cube of volume $|\Delta|\leq 1$ centered at the origin. We substitute (4.4) into (A5) to obtain

$$\delta m^{(D)}_{\Lambda,\kappa,\kappa'}(m\delta)=\left(\frac{1}{|\Lambda|}\right)^3\sum_{\substack{k_i\in T^D_\delta\\ i=1,2,3}} v(k_1,k_2,k_3)$$

$$\times\prod_{i=1}^{3} I^{(i)}(k_1^{(i)},k_2^{(i)},k_3^{(i)};m\delta), \tag{A7}$$

where

$$v(k_1,k_2,k_3)=\left(\prod_{i=1}^{3} u_\delta(k_i)^{-2}\kappa_{S,i}(k_{i,\delta})\kappa'_{S,i}(k_{i,\delta})\right), \tag{A8}$$

$$I^{(i)}=\delta^3\sum_{\substack{n^{(i)}\delta\\ n\delta\in\Lambda_\delta}}\left[\prod_{j=1}^{3}\tfrac{1}{2}\{\exp(-ik_j^{(i)}m^{(i)}\delta)-\exp(ik_j^{(i)}m^{(i)}\delta+il_i)\}\right.$$

$$\left.\times\exp(ik_j^{(i)}n^{(i)}\delta)\right]$$

$$=4^2\cdot 6\cdot(\tfrac{1}{2})^9\delta_{k_1^{(i)}+k_2^{(i)}+k_3^{(i)},0}\left[\left(\prod_{i=1}^{3}[1-\exp(ik_j^{(i)}l_i)]\right)\right.$$

$$+\left(\sum_{j=1}^{3}\exp(ik_j^{(i)}l_i)\right)-\{\exp[i(k_1^{(i)}+k_2^{(i)})l_i]+\text{cyclic}\}$$

$$+\{\exp[-i(k_1^{(i)}+k_2^{(i)})m^{(i)}\delta+ik_3^{(i)}(m^{(i)}\delta+l_i)]+\text{cyclic}\}$$

$$+\{\exp[-ik_1^{(i)}m^{(i)}\delta+ik_2^{(i)}(m^{(i)}\delta+l_i)+ik_3^{(i)}(m^{(i)}\delta+l_i)]$$

$$\left.+\text{cyclic}\}\right]$$

$$=\sum_{j=1}^{5} I_j^{(i)}(k_1^{(i)},k_2^{(i)},k_3^{(i)};\ m^{(i)}\delta).$$

We substitute (A8) into (A7). We then obtain 15 terms from (A7). We write

$$\delta m^{(D)}_{\Lambda,\delta,\kappa_S,\kappa'_S}(m\delta)=\sum_{\substack{j_1,j_2',j_3''=1\\ j=(j_1,j_2',j_3'')}}^{5}\delta m^{(D),j}_{\Lambda,\delta,\kappa_S,\kappa'_S}(m\delta), \tag{A9}$$

where $\delta m^{(D),j}$, $j=(j_1,j_2',j_3'')$ is the term corresponding to $I_j^{(1)}I_{j'}^{(2)}I_{j''}^{(3)}$, $j_1,j_2',j_3''=1,\ldots,5$. The contribution of $\delta m^{(D),j}$, $j\neq(1,1,1)$ (the term containing at least one $I_j^{(i)}$, $j=2,\ldots,5$), to (A6) is very small. In fact, it is bounded by $O(1)l_i^{-1}|\Delta|^{2\epsilon}\lambda^{-\epsilon}d^{-\epsilon}$. The main reason is that each term $I_j^{(i)}$, $j=2,\ldots,5$, contains at least one factor $\exp(ik_j^{(i)}l_i)$, $i=1,2,3$. We multiply a factor $(l_i)^{-2}|l_i|^2_{D,\delta}\geq O(1)$ to $\delta m^{(D),j}$ and use the technique employed in proving Proposition 4.5. In this way one may pull out a convergence factor from each $\exp(ik_j^{(i)}l_i)$. The proof is easy, and we leave it to the reader. Hence, it is sufficient to consider the contribution coming from $\delta m^{(D),j}$, $j=(1,1,1)$, i.e., the term obtained solely from the $I_1^{(i)}$, $i=1,2,3$, in (A8). We note that

$$\delta m^{(D),(1,1,1)}_{\Lambda,\delta,\kappa_S,\kappa'_S}=4^2\cdot 6\left(\frac{1}{|\Lambda|}\right)^3\sum_{\substack{k_j\in T^D_\delta\\ (\pi/l_i)k_j^{(i)}=\text{odd}\\ i,j=1,2,3}} v(k_1,k_2,k_3)\delta_{k_1+k_2+k_3,0},$$

$$\delta m^2_{\delta,\kappa_S,\kappa'_S}=4^2\cdot 6\cdot(2\pi)^{-9}\int_{-\pi/\delta}^{\pi/\delta}d^3k_1\,d^3k_2\,d^3k_3$$

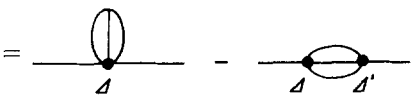
$$\times v(k_1,k_2,k_3)\delta(k_1+k_2+k_3).$$

The only difference between $\delta m^{(D),(1,1,1)}$ and δm^2 is that of the discrete integration (summation) and the continuous integration. It is easy to check that

$$|\delta m^2_{\delta,\kappa_S,\kappa'_S}-\delta m^{(D),(1,1,1)}_{\Lambda,\delta,\kappa_S,\kappa'_S}|\leq O(1)\lambda^{-\epsilon}.$$

We replace $\delta m^{(D)}$ by $\delta m^{(D),(1,1,1)}$ in (A6) and use the above result. Then the bound (A6) follows from Lemma 3.2. This proves (A6).

As a consequence of (A6) and (4.34), the following result is sufficient to establish the proposition: The Hilbert–Schmidt norm of the kernel of the diagram

$w_{\Delta,\Delta'}(k_4,k_5)=$ 

$$=\left|D^m_{\Lambda,k_4}D^{m'}_{\Lambda,k_5}\left[\left(\prod_{i=4}^{5}u^{-1}_{\delta,i}\kappa_{S,i}\right)\sum_{\substack{n\delta\in\Delta_0\\ n'\delta\in\Delta'_0}}\frac{\lambda^2}{2}\cdot 4^2\cdot 6\right.\right.$$

$$\times\left\{\delta_{\Delta_0,\Delta'_0}\delta_{n,n'}\left(\sum_{m\delta\in\Lambda_\delta}\prod_{j=1}^{3}C^D_{\Lambda,\kappa_j,\kappa'_j}(m\delta,(n'+p')\delta)\right)\right.$$

$$\left.-\prod_{j=1}^{3}C^D_{\Lambda,\kappa_j,\kappa'_j}((n+p)\delta,(n'+p')\delta)\right\}$$

$$\left.\times\left(\prod_{i=1}^{3}\exp(\pm ik_4^{(i)}n\delta)\right)\left(\prod_{i=1}^{3}\exp(\pm ik_S^{(i)}n'\delta)\right)\right]\Bigg| \tag{A10}$$

is bounded by

$$||w_{\Delta,\Delta'}||_{\text{H.S.}}\leq O(1)\ (|\Delta|\ |\Delta'|)^{-\epsilon}\lambda^{-\epsilon}d^{-n}. \tag{A11}$$

The proof of the bound in (A11) follows from the method used in Appendix A.4 of Ref. 1, a method similar to that in the Step 3 of the proof of Proposition 4.5, and Lemma 3.4. We do not produce the detailed proof of (A11); we only give a description of the proof. We first substitute (4.4) into (A10) and expand it. We then have 2^9 terms. After the summation over the lattice space Λ_δ, each term has a form similar to that of (A4.1) of Ref. 1. In principle one may obtain the expression of each term in (A10) from that of (A4.1) of Ref. 1 by replacing

$$\left.\begin{array}{l}\displaystyle\int_{-\pi/\delta}^{\pi/\delta}d^3k\\ (\chi_{\Delta_0})\tilde{\delta}(k_1+k_2+k_3+k_4)\\ (\chi_{\Delta'_0})_{\tilde{\delta}}(-k_1-k_2-k_3-k_5)\end{array}\right\}\text{ by }\left\{\begin{array}{l}(\pi^3/|\Lambda|)\displaystyle\sum_{k\in T^D_\delta}\\ (P_a\tilde{\chi}_{\Delta_0})(k_1+k_2+k_3+k_4)\\ (P_{a'}\tilde{\chi}_{\Delta'_0})(-k_1-k_2-k_3-k_5).\end{array}\right. \tag{A12}$$

when $(P_a\tilde{\chi}_\Delta)=\tilde{\chi}_\Delta(\pm(k_1+k_3+k_3)\pm k_4)$ and $P_{a'}\tilde{\chi}_{\Delta'}=\tilde{\chi}_{\Delta'}(\pm(k_1+k_2+k_3)\pm k_5)$. Otherwise there is no divergence. We use Lemma 3.2 and follow a method similar to that used in Appendix A.4 of Ref. 1 to bound each term in (A10). The rest of the proof follows from a straightforward modification of the method in Ref. 1. This proves the lemma for $X=D$. The proof of the proposition for $X=N,P$ follows in a similar manner. This completes the proof. ∎

*On leave from the Department of Mathematics, Yonsei University, Seoul, Korea.

[1]Y. Park, J. Math. Phys. **16**, 1065 (1975).
[2]Y. Park, J. Math. Phys. **16**, 2183 (1975).
[3]Y. Park, J. Math. Phys. **17**, 1143 (1976).

[4]E. Seiler and B. Simon, "Nelson's Symmetry and All That in the Yukawa$_2$ and $\lambda\phi_3^4$ Field Theories," Princeton preprint.
[5]F. Guerra, R. Rosen, and B. Simon, Ann. Math. **101**, 111 (1975).
[6]E. Nelson, *Probability Theory and Euclidean Field Theory*, Springer Lecture Notes in Physics (Springer, New York, 1973).
[7]J. Glimm and A. Jaffee, Fortschr. Physik **21**, 327 (1973).
[8]J. Feldman, Comm. Math. Phys. **37**, 93 (1974).
[9]J. Feldman and K. Osterwalder, "The Wightman Axioms and Mass Gap for Weakly Coupled $(\lambda\phi^4)_3$ Quantum Field Theories," Harvard preprint (1975).
[10]J. Magnen and R. Seneor, "Infinite Volume Limit of the $\lambda\phi_3^4$ Model," preprint (1975).
[11]J. Fröhlich, "Existence and Analyticity in the Bare Parameter of the $(\lambda(\phi\cdot\phi)^2 - \sigma\varphi_1^2 - \mu\phi_1)_3$-Quantum Field Models, I," manuscript.
[12]J. Glimm, A. Jaffee, and T. Spencer, *The Particle Structure of the Weakly Coupled $P(\phi)_2$ Model and Other Applications of High Temperature Expansion*, Springer Lecture Notes in Physics (Springer, New York, 1973).
[13]F. Guerra, R. Rosen, and B. Simon, "Boundary Conditions in the $P(\phi)_2$ Euclidean Quantum Field Theory," preprint (1975).
[14]F. Guerra, R. Rosen, and B. Simon, Comm. Math. Phys. **41**, 19 (1975).
[15]B. Simon, *The $P(\phi)_2$ Euclidean (Quantum) Field Theory* (Princeton U.P., Princeton, N.J., 1974).
[16]J. Feldman and K. Osterwalder, "The Construction of $\lambda\phi_3^4$ Quantum Field Models," Marseille Conference (1975).
[17]Y. Park, J. Math. Phys. **17**, 1073 (1976).
[18]K. Osterwalder and R. Schrader, Comm. Math. Phys. **31**, 83 (1973); **42**, 281 (1975).

Equilibrium properties of fluids in the semiclassical limit

S. K. Sinha* and Y. Singh

Department of Physics, Banaras Hindu University, Varanasi-221005, India
*and *Department of Physics, L. S. College, Bihar University, Muzaffarpur, India*
(Received 12 August 1975)

The problem of calculating the equilibrium properties of dense fluids in the semiclassical limit when the quantum effects are small is studied. Expressions are given for the pressure, free energy, and the radial distribution function in terms of the properties and correlation functions of the classical system and s-body "modified" Mayer functions $f^{s}_{1,2,\ldots,s}$. It is shown that the correct radial distribution function of a fluid in the semiclassical limit is generated from the classical radial distribution function if we replace in turn each f^0 bond ($f^0_{12} = e^{-\beta\phi(1,2)} - 1$) by an effective f^{eff} bond, where $f^{\text{eff}} = f^0 + (1+f^0)f^{\text{II}} + (1+f^0)(1+f^{\text{II}})\mathcal{L}$ and where $\mathcal{L}$ is subset of the line-irreducible graphs each of which contain one f^{III} bond. The effective pair bond correct to the second order in thermal wavelength λ ($= \{2\pi\hbar^2\beta/m\}^{1/2}$) for a fluid of hard spheres is calculated for $\lambda/d = 0.1$, and 0.2 at reduced densities $\rho^* = 0.3$ and 0.6. The most striking effect of the quantum mechanics on the structure of a hard-sphere fluid is found at and near the point of contact of the hard spheres.

I. INTRODUCTION

In recent years considerable progress has been made in understanding the structural properties of classical fluids, made of spherical or nearly spherical molecules, thanks to the molecular dynamics and Monte Carlo methods,[1] and to very ingenious theories like the scaled particle,[2] the Percus–Yevick theory,[3] the hypernetted chain,[4] and various perturbation schemes.[5–11] However, when dealing with the fluids in which deviation takes place at a microscopic level from classical law, our theoretical understanding is far from satisfactory.[12]

In the semiclassical limit, when quantum effects are small and can be treated as a correction to the classical system, the usual way of studying the properties of fluids is to expand them in powers of Planck's constant h.[13–20] The first term of this series is a classical value and other terms arise due to quantum effects. In the Wigner–Kirkwood (WK) method,[13,15] expansion is done in powers of the kinetic energy operator $\hbar^2\nabla^2$, which leads to a series in powers of $\hbar^2$. Since ∇^2 operates on the potential energy term, the WK method fails in cases where the intermolecular potential is a nonanalytic function of distance. Such systems are dealt with by the Hemmer–Jancovici (HJ) method[19,20] in which expansion is done in terms of the Ursell function and which leads to a series in powers of $\hbar$. Recently, in a series of papers, Singh and Ram[21,22] and others[23,24] have investigated the effect of quantum mechanics on the structural and thermodynamic properties of fluids. Expressions for the first quantum corrections were derived using the WK method for the analytic potential case and the HJ method for the nonanalytic potential case. The fluid of hard spheres has been treated for the thermodynamic properties to the second quantum correction term by Gibson.[25] It is found that the contributions of the higher-order terms (corrections) increase with the density. At liquid densities, one has to consider several terms of the series even at sufficiently high temperatures. This suggests developing a theory or method which should enable us to sum the series.

This paper is concerned with the evaluation of the thermodynamic properties and low-order correlation functions of dense fluids in the semiclassical limit. The major emphasis is on the development of a general method applicable to all types of fluids and with suitable approximations to all order of the quantum corrections. We use the grand canonical ensemble and functional differentiation technique to derive the required results. These results are given in terms of the properties and correlation functions of the classical system. Then it is shown how the different quantum correction terms for different fluids can be derived from these expressions.

In Sec. III, we define an effective pair potential (functions of temperature and density) that will be found if the properties of a fluid, in which quantum corrections are present, are interpreted on the assumption that such corrections are absent. Usually, the examination of different properties leads to different effective pair potentials. Here we shall confine ourselves to that effective pair potential which generates the correct radial distribution function in the semiclassical limit.

Expansion of the grand partition function and the s-body correlation function are developed in Sec. II in terms of diagrams[26] (or graphs). A diagram is a collection of circles (or vertices) and bonds connecting the vertices. There are two types of circles, white and black. Each white circle has a label and position associated with it, but black vertices are unlabeled. Each bond has associated with it a function of positions. The value of a diagram is defined in terms of these functions and an integration over the positions which can be assigned to each black circle. According to convention,[26] a factor which is determined by the topological structure of the diagram is associated with the value of the diagram. For a detailed discussion of this aspect of the problem, readers are referred to the articles by Stell,[26(c)] and Morita and Hiroike.[26(b)]

The exchange effects, which are introduced by the Bose–Einstein or Fermi–Dirac statistics, have been

neglected in this paper. In the case of dilute gas, it has been found that the exchange terms decrease rapidly with increasing temperature and are negligible in the temperature region we are concerned with.[27] However, for dense fluids the situation is not so clear and it is possible that the exchange effects do form a significant part of the total correction at intermediate temperatures. We propose to examine this effect in a future publication.

II. GENERAL FORMULATION

The grand paritition function and the grand canonical s-particle density distribution function are defined by

$$\Xi(V,\beta,Z)=\sum_{N\geq 0}\frac{1}{N!}\int\cdots\int\exp\left[\sum_{i=1}^{N}\gamma(i)\right]\times W_N(1,2,\ldots,N)\prod_{i=1}^{N}di \tag{1}$$

and

$$n_s(1,2,\ldots,s/z)=\Xi^{-1}\sum_{N\geq s}\frac{1}{(N-s)!}\int\cdots\int\exp\left[\sum_{i=1}^{N}\gamma(i)\right]\times W_N(1,2,\ldots,N)\prod_{i=s+1}^{N}di, \tag{2}$$

where $\gamma(i)=l_n z(i)=-3l_n\lambda+\beta\mu-\beta\phi(i)$, the fugacity $z=\lambda^{-3}\exp(+\beta\mu)$ (μ is the chemical potential) and

$$W_N(1,2,\ldots,N)=\lambda^{3N}\langle 1,2,\ldots,N|\bar{e}^{\beta\hat{H}}|1,2,\ldots,N\rangle. \tag{3}$$

Here $\hat{H}_N$ is the Hamiltonian operator of a system of N identical particles, each of mass m put in a container of volume V. Under the assumptions that the total potential energy of interaction is pairwise additive and that the molecules are spherically symmetric, one can write for the Hamiltonian

$$H_N=\sum_{i=1}^{N}\frac{p_i^2}{2m}+\sum_{i=1}^{N}\phi(i)+\sum_{1\leq i<j\leq N}\phi(i,j), \tag{4}$$

where $\phi(i)$ is the potential energy of a particle i situated at $\bar{r}_i$ due to external forces, and $\phi(i,j)$ is the pair potential energy between particles i and j.

In the semiclassical limit one has

$$W_N=W_N^c W_N^m, \tag{5}$$

where W_N^c is the classical value of W_N and is defined as

$$W_N^c=\exp\left[-\beta\sum_{i<j}\phi(i,j)\right]. \tag{6}$$

In the WK expansion, W_N^m is expressed as

$$W_N^m=1+\frac{\hbar^2}{m}\sum_{i=1}^{N}\theta_2(i)+\cdots, \tag{7}$$

where

$$\theta_2(i)=-\frac{\beta^2}{12}\left[\nabla_i^2\sum_{i<j}\phi(i,j)-\frac{\beta}{2}\left(\nabla_i\sum_{i<j}\phi(i,j)\right)^2\right], \tag{8}$$

and in the Hemmer–Jancovici expansion

$$W_N^m=1+\sum_{i<j}U_2^m(i,j)+\sum_{i<j<k}U_3^m(i,j,k)+\cdots, \tag{9}$$

where

$$\begin{aligned}&U_1^m(1)=W_1^m(1)=1,\\ &U_2^m(1,2)=W_2^m(1,2)=1,\\ &U_3^m(1,2,3)=W_3^m(1,2,3)-W_2^m(1,2)\\ &\qquad\qquad-W_2^m(1,3)-W_2^m(2,3)+2,\end{aligned} \tag{10}$$

and so on. Here $U_l^m(1,2,\ldots,l)$ is a "modified" Ursell cluster function of l distinguishable particles. From (10) it is clear that (i) the product property of W_l^m's implies the cluster property of the U_l^m's and vice versa, and (ii) U_l^m can be found from a solution of the quantum-mechanical l-body problem. Equation (9) represents the expansion of W_N^m in a more general way and reduces to (7) if the potential function is differentiable.

We find it more convenient to work in the grand canonical ensemble with what we call the "modified s-particle Mayer function," $f_{12\ldots s}^s$ being defined so that it reduces to zero if any one of the particles $1,2,\ldots,s$ becomes remote from the others and

$$W_N^m(1,2,\ldots,N)=\prod_{i<j}^{N}(1+f_{ij}^{II})\prod_{i<j<k}^{N}(1+f_{ijk}^{III})\cdots. \tag{11}$$

Comparing (11) with (9) we find that

$$\begin{aligned}&U_1^m(1)=1,\\ &U_2^m(1,2)=f_{12}^{II},\\ &U_3^m(1,2,3)=f_{12}^{II}f_{13}^{II}+f_{12}^{II}f_{23}^{II}+f_{13}^{II}f_{23}^{II}\\ &\qquad\qquad+f_{12}^{II}f_{13}^{II}f_{23}^{II}+f_{123}^{III},\\ &\cdots\cdots\cdots\cdots\end{aligned} \tag{12}$$

$$U_l^m(1,2,\ldots,l)=\sum_{C_l}\prod_{C_l}(f_{ij}^{II}f_{ijk}^{III}\cdots), \tag{13}$$

where the sum of products is carried out over all connected graphs C of l labeled points.

The s-particle cluster function which arises in a natural fashion in the study of a statistical mechanical system is defined by a functional derivative of $\ln\Xi$[26(c)]:

$$X_s(1,2,\ldots,s)=\prod_{1\leq i\leq s}\exp[\gamma(i)]\,\frac{\delta^s\ln\Xi}{\prod_{1\leq i\leq s}\exp[\gamma(i)]}. \tag{14}$$

The other definition of the cluster functions is obtained from their relationship with the s-particle distribution function,

$$n_S(1,2,\ldots,s)=\sum_{(\sum\alpha p_\alpha=s)}X_\alpha(i_1 i_2,\ldots,i_{p_\alpha}), \tag{15}$$

where the sum of the products is carried out over all possible divisions of the s-particles with the condition that $\sum\alpha p_\alpha=S$. By successive variational differentiation of $X_1(1)$ we find

$$\begin{aligned}&\frac{\delta X_1(1)}{\delta\gamma(l)}=X_2(1,l)+X_1(1)\delta(1-l),\\ &\cdots\cdots\cdots\cdots\\ &\frac{\delta X_S(1,2,\ldots,s)}{\delta\gamma(l)}=X_{s+1}(1,2,\ldots,s;l)\\ &\qquad\qquad+X_S(1,2,\ldots,s)\sum_{i=1}\delta(i-l).\end{aligned} \tag{16}$$

Substitution of (11) in (1) and (2), respectively, leads to

$$\Xi(v,\beta,z)=\sum_{N\geq 0}^{\infty}\frac{1}{N!}\int\cdots\int\exp\Big[\sum_{i=1}\gamma(i)-\beta\sum_{i<j}^{N}\phi(i,j)\Big]\times\Big[\prod_{i<j}^{N}(1+f_{ij}^{II})\prod_{i<j<k}^{N}(1+f_{ijk}^{III})\cdots\Big]\prod_{i=1}^{N}di, \tag{17}$$

$$n_s(1,2,\ldots,s/z)=\Xi^{-1}\sum_{N\geq s}\frac{1}{(N-s)!}\times\int\cdots\int\exp\Big[\sum_{i=1}^{N}\gamma(i)-\beta\sum_{i<j}^{N}\phi(i,j)\Big]\times\Big[\prod_{i<j}^{N}(1+f_{ij}^{II})\prod_{i<j<k}^{N}(1+f_{ijk}^{III})\cdots\Big]\prod_{i=s+1}^{N}di. \tag{18}$$

For future use we break Ξ and n_s into two parts:

$$\Xi=\Xi_0+\Xi^1,$$

and

$$n_s=n_s^0+n_s^1, \tag{19}$$

where

$$\Xi_0(V,\beta,z)=\sum_{N\geq 0}\frac{1}{N!}\int\cdots\int\exp\Big[\gamma(i)-\beta\sum_{i<j}\phi(i,j)\Big]\prod_{i=1}^{N}di, \tag{20}$$

$$\Xi^1(v,\beta,z)=\Xi_0(v,\beta,z)\times\Big[\frac{1}{2!}\int\int n_2^0(1,2)f_{12}^{II}\,d1\,d2+\cdots\Big], \tag{21}$$

and

$$n_s^0(1,2,\ldots,s/z)=\Xi_0^{-1}\sum_{N\geq s}\frac{1}{(N-s)!}\int\cdots\int\exp\Big[\sum_{i=1}^{N}\gamma(i)-\beta\sum_{i<j}\phi(i,j)\Big]\prod_{i=s+1}^{N}di. \tag{22}$$

n_s^1 will be defined later on [see Eq. (28) below].

A. Cluster expansion of Ξ, n_S and x_l at constant fugacity z

The expansion of Ξ at constant z is obtained from (17) in terms of the composite graphs, with f^{II}, f^{III}, $\cdots$ bonds (represented in graphs by lines, shaded triangles, ...) and X_l^0 polyhedron containing $l-1$ lines (represented by a vertex, dotted line, triangle, ...) where X_l^0 is a cluster function of l labeled points of reference system. That is

$$X_l^0(1,2,\ldots,l)=\prod_{i=1}^{l}\exp[\gamma(i)]\frac{\delta^l\ln\Xi_0}{\prod_{i=1}^{l}\delta\exp[\gamma(i)]} \tag{23}$$

so that

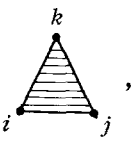
$$n_1^0(1)=X_1^0(1),\quad n_2^0(1,2)=X_2^0(1,2)+X_1^0(1)X^0(2),$$

$$n_3^0(1,2,3)=X_3^0(1,2,3)+X_2^0(1,2)X_1^0(3)+X_2^0(1,3)X_1^0(2)+X_2^0(2,3)X_1^0(3)+X_1^0(1)X^0(2)X_1^0(3), \tag{24}$$

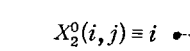
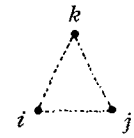
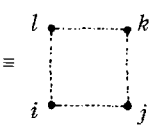
$$\cdots\cdots\cdots\cdots$$

$$n_s^0(1,2,\ldots,s/z)=\sum\prod X^0(i_1,i_2,\ldots,i_p)\quad(\sum\alpha p_\alpha=s).$$

For graphical representation the notation is

$f_{ij}^{II}\equiv$, $f_{ijk}^{III}\equiv$  ,

and

$X_1^0(i)\equiv$  , $X_2^0(i,j)\equiv$  , $X_3^0(i,j,k)\equiv$  , $X_4^0(i,j,k,l)\equiv$  .

Thus,

$\Xi(z)=\Xi(z)$ [1 + sum of all distinct (topologically) composite graphs with no labeled vertices, at most one f^{II}-bond connecting any two vertices and/or one f^{III}-bond connecting any three vertices, with X_l^0 polyhedron $(l\geq 1)$]. (25)

Using Lemma 3 of Morita and Hiroike[26(b)] we get the following prescription from (25) for the graphical expansion of $\ln\Xi$:

$$\ln\Xi(z)=\ln\Xi_0(z)+\xi(z)$$

where

$\xi(z)$ = sum of all distinct connected composite graphs with no labeled vertices, at most one f^{II}-bond connecting any two vertices and/or one f^{III} bond connecting any three vertices, with X_l^0 polyhedron $(l \geq 1)$. (26)

From (14), (16), and (26), we derive the graphical expansion of X_l:

$X_l(1,2,\ldots,l/z) = X_l^0(1,2,\ldots,l/z)$ + [sum of all the distinct connected composite graphs, composed of l white vertices labeled $1,2,\ldots,l$, respectively, some or no unlabeled vertices, at most one f^{II}-bond connecting any two vertices and/or one f^{III}-bond connecting any three vertices, with X_l^0 polyhedron $(l \geq 1)$]. (27)

Using relation (15), we obtain the graphical expansion of $n_s(1,2,\ldots,s/z)$,

$$n_s(1,2,\ldots,s/z) = n_s^0(1,2,\ldots,s/z) + n_s^1(1,2,\ldots,s/z)$$

where

$n_s^1(1,2,\ldots,s/z)$ = [sum of all distinct graphs composed of s white vertices labeled as $1,2,\ldots,s$, respectively, some or no unlabeled vertices, at most one f^{II}-bond connecting any two points and/or one f^{III}-bond connecting any three vertices with X_l^0 polyhedron $(l \geq 1)$] (Each composite graph contains at least one labeled vertex.) (28)

Equations (26)—(28) provide the starting point for further topological reduction in the following section.

B. Expansion of $\ln\Xi$, and n_s at constant ρ

Since the number density ρ is easily associated with physical measurements other than fugacity, we consider in this section the problem of expressing $\ln\Xi$ and n_s as functionals of ρ and n_s^0. This is done with the help of the functional Taylor expansion.

Let $L(z)$ be any function, such as $\ln\Xi(z)$ or $n_s(1,2,\ldots,s/z)$, and letting $L_0(z)$ be its value for a reference (classical) system, then [see Eq. (19)]

$$L(z) = L_0(z) + L_1(z). \tag{29}$$

Taking the functional Taylor expansion of the right-hand side of (29) about z_0, we have

$$L(z) = L_0(z_0) + L_1(z_0) + \int \Delta z(\bar{r}_1') \left(\frac{\partial L_0(z)}{\partial z(\bar{r}_1')}\right)_{z=z_0} d\bar{r}_1' + \int \Delta z(\bar{r}_1') \left(\frac{\partial L_1(z)}{\partial z(\bar{r}_1')}\right)_{z=z_0} d\bar{r}_1'$$
$$+ \frac{1}{2}\int\int \Delta z(\bar{r}_1')\Delta z(\bar{r}_2') \left(\frac{\partial^2 L_0(z)}{\partial z(\bar{r}_1')\partial z(\bar{r}_2')}\right)_{z=z_0} d\bar{r}_1' d\bar{r}_2' + \cdots. \tag{30}$$

Here z_0 is the fugacity of the reference system and

$$\Delta z = z - z_0.$$

We then transform

$$\frac{\partial L_0(z)}{\partial z(\bar{r}_1')} = \int \frac{\partial L_0(z)}{\partial n_1^0(\bar{r}_1/z)} \frac{\partial n_1^0(\bar{r}_1/z)}{\partial z(\bar{r}_1')} d\bar{r}_1,$$

$$\frac{\partial^2 L_0(z)}{\partial z(\bar{r}_1')\partial z(\bar{r}_2')} = \int\int \frac{\partial^2 L_0(z)}{\partial n_1^0(\bar{r}_1/z)\partial n_1^0(\bar{r}_2/z)} \frac{\partial n_1^0(\bar{r}_1/z)}{\partial z(r_1')} \frac{\partial n_1^0(\bar{r}_2/z)}{\partial z(\bar{r}_2')} d\bar{r}_1 d\bar{r}_2 \tag{31}$$
$$+ \int \frac{\partial L_0(z)}{\partial n_1^0(\bar{r}_1/z)} \frac{\partial^2 n_1^0(\bar{r}_1/z)}{\partial z(\bar{r}_1')\partial z(\bar{r}_2')} d\bar{r}_1,$$

so that (30) can be written as

$$L(z) = L_0(z_0) + L_1(z_0) + \int\int \Delta z(\bar{r}_1') \left(\frac{\partial L_0(z)}{\partial n_1^0(\bar{r}_1/z)} \frac{\partial n_1^0(\bar{r}_1/z)}{\partial z(\bar{r}_1')}\right)_{z=z_0} d\bar{r}_1' d\bar{r}_1$$
$$+ \int\int \Delta z(\bar{r}_1') \left(\frac{\partial L_1(z)}{\partial n_1^0(\bar{r}_1/z)} \frac{\partial n_1^0(\bar{r}_1/z)}{\partial z(\bar{r}_1')}\right)_{z=z_0} d\bar{r}_1' d\bar{r}_1 + \frac{1}{2}\int\int \Delta z(\bar{r}_1')\Delta z(\bar{r}_2') \left(\frac{\partial^2 L_0(z)}{\partial n_1^0(\bar{r}_1/z)\partial n_1^0(\bar{r}_2/z)} \frac{\partial n_1^0(\bar{r}_1/z)}{\partial z(\bar{r}_1')}\right.$$
$$\left. \times \frac{\partial n_1^0(\bar{r}_2/z)}{\partial z(\bar{r}_2')}\right)_{z=z_0} d\bar{r}_1' d\bar{r}_2' d\bar{r}_1 d\bar{r}_2 + \frac{1}{2}\int\int \Delta z(\bar{r}_1')\Delta z(\bar{r}_2') \left(\frac{\partial L_0(z)}{\partial n_1^0(\bar{r}_1/z)} \frac{\partial^2 n_1^0(\bar{r}_1/z)}{\partial z(\bar{r}_1')\partial z(\bar{r}_2')}\right)_{z=z_0} d\bar{r}_1' d\bar{r}_2' d\bar{r}_1 + \cdots. \tag{32}$$

Using the relation

$$n_1(\bar{r}_1/z) = z(\bar{r}_1) \frac{\delta \ln\Xi(z)}{\delta z(\bar{r}_1)}$$

from (26) we get

$$n_1(\bar{r}_1/z)=n_1^0(\bar{r}_1/z)+z(\bar{r}_1)\frac{\delta\xi(z)}{\delta z(\bar{r}_1)}\ . \tag{33}$$

Expanding rhs of (33) about z_0, we obtain

$$n_1(\bar{r}_1/z)=n_1^0(\bar{r}_1/z_0)+\left(z(\bar{r}_1)\frac{\delta\xi(z)}{\delta z(\bar{r}_1)}\right)_{z=z_0}+\int\Delta z(\bar{r}_1')\left(\frac{\partial n^0(\bar{r}_1/z)}{\partial z(\bar{r}_1')}\right)_{z=z_0}d\bar{r}_1+\int\Delta z(\bar{r}_1')\left[\frac{\partial}{\partial z(\bar{r}_1')}\left(z(\bar{r}_1)\frac{\partial\xi(z)}{\partial z(\bar{r}_1)}\right)\right]_{z=z_0}d\bar{r}_1$$
$$+\frac{1}{2}\int\int\Delta z(\bar{r}_1')\Delta z(\bar{r}_2')\left(\frac{\partial^2 n^0(\bar{r}_1/z)}{\partial z(\bar{r}_1')\partial z(\bar{r}_2')}\right)_{z=z_0}d\bar{r}_1'\,d\bar{r}_2'+\cdots. \tag{34}$$

At constant density we have[21]

$$n_1(\bar{r}_1/z)=n_1^0(\bar{r}_1/z_0)=\rho(\bar{r}_1). \tag{35}$$

Then each order of correction in (34) is separately zero. Thus

$$\int\Delta z(\bar{r}_1')\left(\frac{\partial n_1^0(\bar{r}_1/z)}{\partial z(\bar{r}_1')}\right)_{z=z_0}d\bar{r}_1'=-z_0(\bar{r}_1)\frac{\partial\xi(z_0)}{\partial z_0(\bar{r}_1)}\ ,\ \frac{1}{2}\int\Delta z(\bar{r}_2')\left(\frac{\partial^2 n_1^0(\bar{r}_1/z)}{\partial z(\bar{r}_1')\partial z(\bar{r}_2')}\right)_{z=z_0}d\bar{r}_2'$$
$$=-\left[\frac{\partial}{\partial z_0(\bar{r}_1')}\left(z_0(\bar{r}_1)\frac{\partial\xi(z_0)}{\partial z_0(\bar{r}_1)}\right)\right], \tag{36}$$

and so on. Substitution of (36) into (32) leads to

$$L(z)=L_0(z_0)+L_1(z_0)-\int z_0(\bar{r}_1)\frac{\partial\xi(z_0)}{\partial z_0(\bar{r}_1)}\frac{\partial L_0(z_0)}{\partial\rho(\bar{r}_1)}d\bar{r}-\int z_0(\bar{r}_1)\frac{\partial L_1(z_0)}{\partial\rho(\bar{r}_1)}\frac{\partial\xi(z_0)}{\partial z_0(\bar{r}_1)}\frac{(z_0)}{z_0(r_1)}d\bar{r}_1+\frac{1}{2}\int z_0(\bar{r}_1)$$
$$\times z_0(\bar{r}_2)\frac{\partial\xi(z_0)}{\partial z_0(\bar{r}_1)}\frac{\partial\xi(z_0)}{\partial z_0(\bar{r}_2)}\frac{\partial^2 L_0(z_0)}{\partial\rho(\bar{r}_1)\partial\rho(\bar{r}_2)}d\bar{r}_1\,d\bar{r}_2+\int z_0(\bar{r}_2)\frac{\partial\xi(z_0)}{\partial z_0(\bar{r}_2)}\frac{\partial}{\partial\rho(\bar{r}_2)}\left(z_0(\bar{r}_1)\frac{\partial\xi(z_0)}{\partial z_0(\bar{r}_1)}\right)\frac{\partial L_0(z_0)}{\partial\rho(\bar{r}_1)}d\bar{r}_1\,d\bar{r}_2+\cdots \tag{37}$$

$$=L_0(z_0)+L_1(z_0)-\int z_0(\bar{r}_1)\frac{\partial\xi(z_0)}{\partial z_0(\bar{r}_1)}\frac{\partial L_0(z_0)}{\partial\rho(\bar{r}_1)}d\bar{r}_1-\int z_0(\bar{r}_1)\frac{\partial\xi(z_0)}{\partial z_0(\bar{r}_1)}\frac{\partial L_1(z_0)}{\partial\rho(\bar{r}_1)}d\bar{r}_1$$
$$+\frac{1}{2}\int\frac{\partial}{\partial\rho(\bar{r}_2)}\left(z_0(\bar{r}_1)z_0(\bar{r}_2)\frac{\partial\xi(z_0)}{\partial z_0(\bar{r}_1)}\frac{\partial\xi(z_0)}{\partial z_0(\bar{r}_2)}\frac{\partial L_0(z)}{\partial\rho(\bar{r}_1)}\right)d\bar{r}_1\,d\bar{r}_2+\cdots \tag{38}$$

For a uniform system

$$\rho(\bar{r}_1)=\rho(\bar{r}_2),\quad z_0(\bar{r}_1)=z_0(\bar{r}_2)$$

so that

$$L(z)=L_0(z_0)+L_1(z_0)-z_0\frac{\partial\xi(z_0)}{\partial z_0}\frac{\partial}{\partial\rho}[L_0(z_0)+L_1(z_0)]+\frac{1}{2}\frac{\partial}{\partial\rho}\left[\left(z_0\frac{\partial\xi(z_0)}{\partial z_0}\right)^2\frac{\partial L_0(z_0)}{\partial\rho}\right]+\cdots. \tag{39}$$

a. Thermodynamic properties

With the help of (39), we get the expression for $\ln\Xi$,

$$\ln\Xi(z)=\ln\Xi_0(z_0)+\xi(z_0)-\rho\frac{\partial\xi(z_0)}{\partial\rho}-\frac{1}{\beta}\rho\left(\frac{\partial\rho}{\partial p^c}\right)\left(\frac{\partial\xi(z_0)}{\partial\rho}\right)^2+\frac{1}{2}\frac{\partial}{\partial\rho}\left[\frac{1}{\beta}\rho^2\left(\frac{\partial\rho}{\partial p^c}\right)\left(\frac{\partial\xi(z_0)}{\partial\rho}\right)^2\right]+\cdots$$
$$=\ln\Xi_0(z_0)+\xi(z_0)-\rho\frac{\partial\xi(z_0)}{\partial\rho}+\frac{\rho^2}{2\beta}\frac{\partial}{\partial\rho}\left[\left(\frac{\partial\rho}{\partial p^c}\right)\left(\frac{\partial\xi(z_0)}{\partial\rho}\right)^2\right]. \tag{40}$$

Here use has been made of the relations

$$\rho=z_0\frac{\partial\ln\Xi_0}{\partial z_0}\ \text{ and }\ z_0\frac{\partial\rho}{\partial z_0}=\rho\frac{\partial\rho}{\partial(\beta P_c)}\ . \tag{41}$$

The pressure P of the system is given by

$$\beta P=V^{-1}\ln\Xi. \tag{42}$$

Substitution of (26) and (40) in (42) leads to

$$\beta P=\beta P_c-\frac{1}{2!}\rho^2\int\left(g_2^0(1,2)+\rho\,\frac{\partial g_2^0(1,2)}{\partial\rho}\right)U_2^m(1,2)\,d2-\frac{1}{3!}\rho^3\int\int\left(2g_3^0(1,2,3)+\rho\,\frac{\partial g_3^0(1,2,3)}{\partial\rho}\right)$$

$$\times U_3^m(1,2,3)\,d2\,d3-\frac{1}{8}\rho^4\int\int\int\left[3\left[g_4^0(1,2,3,4)-g_2^0(1,2)\,g_2^0(3,4)\right]+\left(\rho\,\frac{\partial g_4^0(1,2,3,4)}{\partial\rho}\right)\right.$$

$$\left.-2\rho\,g_2^0(1,2)\,\frac{\partial g_2^0(3,4)}{\partial\rho}\right]U_2^m(1,2)\,U_2^m(3,4)\,d2\,d3\,d4+\frac{1}{2\beta}\rho^2\,\frac{\partial}{\partial\rho}\left\{\left(\frac{\partial\rho}{\partial p^c}\right)\left[\frac{1}{2!}\,\frac{\partial}{\partial\rho}\left(\rho^2\int g_2^0(1,2)\right.\right.\right.$$

$$\left.\left.\left.\times U_2^m(1,2)\,d2+\frac{1}{3!}\rho^3\int g_3^0(1,2,3)\,U_3^m(1,2,3)\,d2\,d3\right)\right]^2\right\}+\cdots \tag{43}$$

Integrating with density (43) gives an expression for the Helmholtz free energy per particle

$$\beta a=\beta a^c-\tfrac{1}{2}\rho\int U_2^m(1,2)\,g_2^0(1,2)\,d2-\tfrac{1}{6}\rho^2\int\int U_3^m(1,2,3)\,g_3^0(1,2,3)\,d2\,d3-\tfrac{1}{8}\rho^3\int\int\int U_2^m(1,2)\,U_2^m(3,4)$$

$$\times\left[g_4^0(1,2,3,4)-g_2^0(1,2)\,g_2^0(3,4)\right]d2\,d3\,d4+\frac{\rho^2}{8\beta}\left(\frac{\partial\rho}{\partial p^c}\right)\beta\left(\int U_2^m(1,2)\,\frac{\partial}{\partial\rho}[\rho^2 g_2^0(1,2)]\,d2\right)^2+\cdots. \tag{44}$$

Other thermodynamic properties can be obtained from (43) and (44).

b. Radial distribution function

For the two-particle distribution function we find from (39) and (28) that

$$n_2\left(\frac{1,2}{z}\right)=n_2^0\left(\frac{1,2}{z_0}\right)+n_2'\left(\frac{1,2}{z_0}\right)-\frac{1}{\beta}\rho\left(\frac{\partial\rho}{\partial p^c}\right)\frac{\partial\xi(z_0)}{\partial\rho}\,\frac{\partial}{\partial p}\left[n_2^0\left(\frac{1,2}{z_0}\right)+n_2'\left(\frac{1,2}{z_0}\right)\right]$$

$$+\frac{1}{2\beta}\,\frac{\partial}{\partial\rho}\left\{\left(\rho\,\frac{\partial\rho}{\partial p^c}\right)^2\frac{\partial}{\partial\rho}\left[n_2^0\left(\frac{1,2}{z_0}\right)\right]\left(\frac{\partial}{\partial\rho}\,\xi(z_0)\right)^2\right\}. \tag{45}$$

Hence the radial distribution function of the system in the semiclassical limit is

$$g_2(1,2)=g_2^0(1,2)+g_2^0(1,2)U_2^m(1,2)+2\rho\int g_3^0(1,2,3)U_2^m(1,3)\,d3+\tfrac{1}{2}\rho^2\int\left[g_4^0(1,2,3,4)-g_2^0(1,2)\,g_2^0(3,4)\right]$$

$$\times U_2^m(3,4)\,d3\,d4-\frac{1}{2\beta\rho}\left(\frac{\partial\rho}{\partial p^c}\right)\beta\left(\frac{\partial}{\partial\rho}\left[\rho^2 g_2^0(1,2)\right]\right)\left(\frac{\partial}{\partial\rho}\left[\rho^2\int g_2^0(3,4)U_2^m(3,4)\,d3\,d4\right]\right)$$

$$+\rho\int U_3^m(1,2,3)\,g_3^0(1,2,3)\,d3+\rho^2\int g_4^0(1,2,3,4)\,U_3^m(1,3,4)\,d3\,d4+\tfrac{1}{2}\rho^2\int\left[g_4^0(1,2,3,4)\right.$$

$$\left.-g_2^0(1,2)\,g_2^0(3,4)\right]U_2^m(1,2)\,U_2^m(3,4)\,d3\,d4+\rho^2\int\left[g_4^0(1,2,3,4)-g_2^0(1,3)\,g_2^0(2,4)\right.$$

$$\left.\times U_2^m(1,3)\,U_2^m(2,4)\,d3\,d4+\rho^3\int\left[g_5^0(1,\ldots,5)-g_3^0(1,2,3)\,g_2^0(4,5)\right]U_2^m(1,3)U_2^m(4,5)d3\,d4\,d5\right.$$

$$+\tfrac{1}{8}\rho^3\int\left[g_5^0(1,\ldots,5)-g_2^0(1,2)\,g_3^0(3,4,5)\right]U_3^m(3,4,5)\,d3\,d4\,d5-\frac{1}{6\rho\beta}\left(\frac{\partial\rho}{\partial p^c}\right)\beta$$

$$\times\left(\frac{\partial}{\partial\rho}\left[\rho^2 g_2^0(1,2)\right]\right)\left[\frac{\partial}{\partial\rho}\left(\rho^3\int g_3^0(1,2,3)\,U_3^m(1,2,3)\,d2\,d3\right)\right]+\cdots. \tag{46}$$

Following exactly a similar method one can derive the expressions for the higher-order correlation functions.

We make here the following observations:

1. For a fluid of hard spheres the leading contribution from U_l^m to an integral in (43), (44), or (46) is of order λ^{l-1}. Therefore, the first quantum correction is obtained from those terms which are linear in U_2^m and

$$U_2^m\approx\exp(-\Omega^2) \tag{47}$$

where $\Omega=(2\pi)^{1/2}(d/\lambda)(r/d-1)$; d is the diameter of a sphere. For the second quantum correction, one has to consider all those terms which involve U_2^m [here taken equal to $(1/\sqrt{2})(\lambda/d)\Omega^2\,\mathrm{erfc}\,\Omega$, where $\mathrm{erfc}\,x=(2/\sqrt{\pi})\int_x^\infty\times\exp(-t^2)\,dt$], $U_3^m(1,2,3)$, and $U_2^m(1,3)\,U_2^m(2,3)$, and so on.

Since at the point of contact of spheres 1 and 2

$$U_3^m(1,2,3)\approx-U_2^m(1,3)-U_2^m(2,3) \tag{48}$$

and

$$U_2^m(1,2)\approx-1,$$

a contribution to the first quantum correction to the radial distribution function at $r=d$ also comes from those terms which contain $U_3^m(1,2,3)$ and $U_2^m(1,2)U_2^m(1,3)$.[25] From this consideration one finds that at the contact point,

$$g_2(d)=0. \tag{49}$$

2. If we put

$$U_2^m(1,2)=-\frac{1}{12}\,\frac{\hbar^2}{m}\,\beta\{2\nabla_{12}^2\phi(1,2)-\beta[\nabla_{12}\phi(1,2)]^2\}$$

and

$$U_3^m(1,2,3)=\frac{\hbar^2\beta^2}{12m}[\nabla_1\phi(1,2)\cdot\nabla_1\phi(1,3)$$

$$+\nabla_2\phi(1,2)\cdot\nabla_2\phi(2,3)$$

$$+\nabla_3\phi(3,1)\cdot\nabla_3\phi(3,2)].$$

From the terms written explicitly in (46), we get [except the three terms which involve $U_2^m(i,j)\,U_2^m(k,l)$ where $i,j,k,l=1,2,3,4$ and $i\neq j$, $k\neq l$] the first quantum correction to the radial distribution function for a fluid whose particles interact with a pair potential which is an analytic function of distance.[21,23,24] Higher-order quantum corrections (which are not given here because of their length) will be obtained by considering a greater number of terms in (46) and by the proper choice of U_2^m and U_3^m.

Exact evaluation of the integrals appearing in (43), (44), or (46) is not possible because the values of classical correlation functions $g_3^0(1,2,3)$, $g_4^0(1,2,3,4),\cdots$ are not known. It may, however, be noted that in the perturbation theory of classical fluids such integrals have been evaluated either by using simplified superposition approximation[28] or the Barker–Honderson discrete summation method.[29] But, any such attempts here can have only limited success for reasons discussed in the Introduction. In the following section, we propose a computationally convenient method through the introduction of an effective pair potential which generates the correct radial distribution function in the semiclassical limit.

III. THE EFFECTIVE PAIR POTENTIAL FOR THE PAIR DISTRIBUTION FUNCTION

Here we consider the density expansion of (46). The s-body distribution function for the reference system is given as

$$g_s^0(1,2,\ldots,s)=\exp\left[-\sum_{i<j}^{s}\phi(i,j)\right]\times\left[1+\sum_{l=1}^{\infty}a_l(1,\ldots,s)\rho^l\right],\tag{50}$$

where the coefficient $a_l(1,\ldots,s)$ is the cluster integral involving s base points (white circles) and l field points (black circles). Now we define the two-body Mayer function for reference systems,

$$f_{12}^0=\exp[-\beta\phi(1,2)]-1.$$

In terms of graphs, (50) can be written as[26]

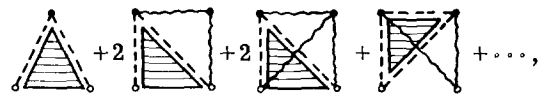
$g_s^0(1,2,\ldots,s)=$ sum of distinct simple linear graphs consisting of white 1-circles labeled $1,2,\ldots,s$, respectively, some or no black circles, and some or no f^0 bonds, such that there is a path from each black circle to each white circle and the graphs are free of articulation circles, i.e., the graphs are root-connected and one-irreducible. (51)

If the deviations from the classical behavior are small so that the contribution of all the diagrams (i) involving more than one f^{III} bond and (ii) involving f^{IV} or higher bonds are negligibly small, then we have to consider only those graphs in (46) with one f^{III} bond or none. Under this condition we find that all the diagrams of (46) can be generated if we replace in turn each f^0 bond in (51) by an effective f^{eff} bond where

$$f^{eff}=f^0+(1+f^0)f^{II}+(1+f^0)(1+f^{II})L.\tag{52}$$

L is a subset of the elementary graphs each of which contains one f^{III} link. That is

L_{12}



(53)

where the dotted line indicates a $(1+f^0)$ bond and the curly line an f^0 bond. Casonova *et al.*[30] have used the same subset of graphs for defining an effective pair potential which generates the exact two-body correlation function in the presence of three-body forces. We have found[31] that L_{12} can be approximated by the following integral equation:

$$L_{12}=\rho\frac{1}{g_2^0(1,2)}\int g_3^0(1,2,3)f_{123}^{III}\,d3.\tag{54}$$

From (12) we find that (54) can be written as

$$L_{12}=\frac{\rho}{g_2^0(1,2)}g_3^0(1,2,3)[U_3^m(1,2,3)-2U_2^m(1,2)U_2^m(1,3)-U_2^m(1,3)U_2^m(2,3)-U_2^m(1,2)U_2^m(1,3)U_2^m(2,3)]\,d3\tag{55}$$

IV. EFFECTIVE PAIR POTENTIAL AND A FLUID OF HARD-SPHERES

In the fluid of hard spheres, the leading contribution to L_{12} which is of order λ^2, comes from cluster formations in which two distances of a triangle formed by three particles 1, 2, 3 lie within d and $d+\lambda$, and the third distance is greater than $d+\lambda$. If $r_{12}>d+\lambda$, L_{12} can be evaluated following the method of Jancovici.[32] The result is

$$L_{12}=\frac{2\pi d^2}{r_{12}}\lambda^2\rho[g_2^0(d)]^2\left\{\left[\left(\frac{\pi}{12\theta}-\frac{\theta}{12\pi}\right)\times\sin\theta-\frac{1}{2\pi}\cos\theta\right]-\frac{1}{8}\right\},\tag{56}$$

where

$$\theta=\pi-\cos^{-1}(\tfrac{1}{2}\cos\alpha)$$

and

$$\cos\alpha=1-\frac{r_{12}^2}{2d^2}.$$

For the configuration in which r_{12} and r_{13} (or r_{23}) lie within the interval $(d,d+\lambda)$ and r_{23} (or r_{13}) is greater than $d+\lambda$, we find that

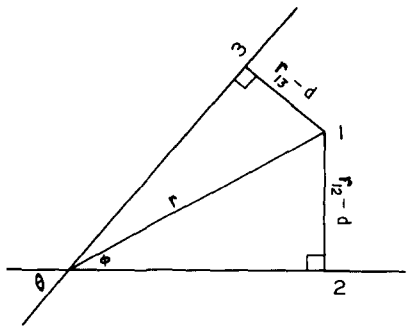


FIG. 1. Position of three spheres in a configuration in which $r_{23} \geq d+\lambda$.

$$\mathcal{L}_{12} = \frac{4\pi}{r_{12}} \rho\, dg_2^0(d) \int_d^\infty \int_d^{2d} [W_3^m(r_{12}, r_{13}, r_{23}) - 1 - U_2^m(1,2) - U_2^m(1,3) - U_2^m(1,2)U_2^m(1,3)] \times g_2^0(r_{23}) r_{23}\, dr_{23}\, dr_{13}. \tag{57}$$

Jancovici[32] has evaluated W_3^m for a configuration in which two short distances on which integration has to be done are $r_{13} \sim d$ and $r_{23} \sim d$ and the long distance which has to be kept fixed is $r_{12} \geq d+\lambda$. Here we have a different situation; one distance on which integration has to be performed is a long distance. However, for the calculation of W_3^m, we adopt the method of Jancovici with proper changes so as to suit our problem. Thus, in the high temperature limit $(\lambda/d \to 0)$, $W_3^m(r_{12}r_{13}r_{23})$ is found to be proportional to the probability of one particle of mass $m/2$ moving in a plane wedge (Fig. 1), the summit angle of which is $\theta = \pi - \cos^{-1}(\frac{1}{2}\cos\alpha)$, where $\cos\alpha = 1 - r_{23}^2/2d^2$; $r_{12}-d$ and $r_{13}-d$ are the distances of the particle 1 to the edges, on which the wavefunctions of the system of three-particles are constrained to vanish. In terms of polar coordinates (r, θ, ϕ) chosen as in Fig. 1, the three distances can be written as

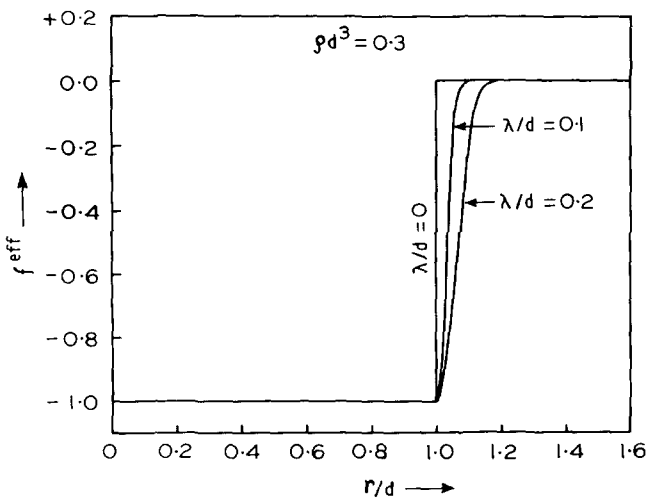


FIG. 2. f^{eff} as a function of interparticle separation at the reduced density $\rho^* = 0.3$.

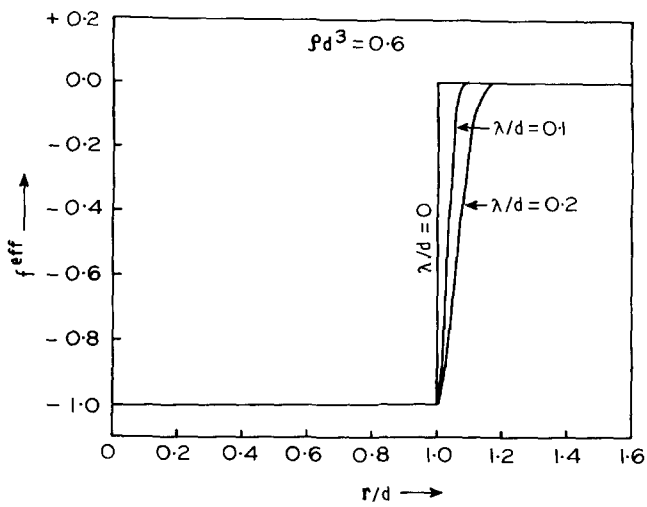


FIG. 3. f^{eff} as a function of interparticle separation at the reduced density $\rho^* = 0.6$.

$$r_{12} = d + r\sin\phi,$$
$$r_{13} = d + r\sin(\theta - \phi), \tag{58}$$
$$r_{23} = d(2 + 4\cos\theta)^{1/2}$$

and

$$W_3^m(r,\theta,\phi) = \frac{4}{\theta} \sum_{m=1}^{\infty} \exp\left(-\frac{\pi r^2}{\lambda^2}\right) \times I_{n\pi/\theta}(\pi r^2/\lambda^2) \sin^2(n\pi\phi/\theta) \tag{59}$$

where I is a Bessel function of imaginary argument.

Equation (57) is solved numerically and the results obtained from (52) for $\lambda/d = 0.1$ and 0.2 are given in Figs. 2 and 3, respectively, at reduced densities ρ^* $(=\rho d^3) = 0.3$ and 0.6. We find that the contribution of term involving $\mathcal{L}$ is very small compared to other terms.

From the behavior of f^{eff} it follows that the effect of the quantum mechanics is to enhance and make soft the hard core diameter. Due to this, a substantial change in the structure of a hard-sphere fluid is found in the neighborhood of the point of contact. The effective pair potential is obtained from f^{eff} by the relation

$$-\beta\phi^{\text{eff}}(1,2) = \ln(1 + f^{\text{eff}}).$$

We find that the softness of the hard core depends very weakly on the density but quite strongly on the temperature.

[1]W.W. Wood, F.R. Parker, and J.D. Jacobson, Nuovo Cimento Suppl. **9**, 133 (1958); B.J. Alder and T.E. Wainwright, *ibid.* **9**, 116 (1958).
[2]H. Reiss, H.L. Frisch, and J.L. Lebowitz, J. Chem. Phys. **31**, 369 (1959).
[3]J.K. Percus and G.J. Yevick, Phys. Rev. 110, 1 (1958); E. Thiele, J. Chem. Phys. **39**, 474 (1963); M.S. Wertheim, Phys. Rev. Lett. **10**, 321 (1963).
[4]J.M.J. van Lecuven, J. Grovessveld, and J. de Boer, Physica **25**, 792 (1959); T. Morita and K. Hiroike, Prog.

Theor. Phys. **23**, 1003 (1960); E.J. Meeron, J. Math. Phys. **1**, 192 (1960); G.S. Rushbrooke, Physica **26**, 259 (1960); M.S. Green, J. Chem. Phys. **33**, 1403 (1960); L. Verlet, Nuovo Cimento **18**, 77 (1960).
[5]J.L. Lebowitz, G. Stell, and S. Baer, J. Math. Phys. **6**, 1282 (1965).
[6]J.A. Barker and D. Henderson, J. Chem. Phys. **47**, 2856 4714 (1967).
[7]G.A. Mansoori and F.B. Confield, J. Chem. Phys. **51**, 4958 (1968).
[8]J. Rasaiah and G. Stell, Molec. Phys. **18**, 249 (1970).
[9]J.D. Weeks, D. Chandler, and H.C. Anderson, J. Chem. Phys. **51**, 4958 (1971).
[10]H.C. Anderson and D. Chandler, J. Chem. Phys. **57**, 1918 (1972).
[11]L. Verlet and J.J. Weis, Phys. Rev. A **5**, 939 (1972).
[12]E. Feenberg, *Theory of Quantum Fluids* (Academic, New York, 1969).
[13]E. Wigner, Phys. Rev. **40**, 749 (1932).
[14]G.E. Uhlenbeck and L. Groper, Phys. Rev. **41**, 79 (1932).
[15]J.G. Kirkwood, Phys. Rev. **44**, 31 (1933).
[16]K. Husmi, Proc. Phys. Math. Soc. Jpn. Ser. 3 **22**, 264 (1940).
[17]J.E. Mayer and W. Band, J. Chem. Phys. **15**, 141 (1947).
[18]H.S. Green, J. Chem. Phys. **19**, 955 (1947).
[19]P.C. Hemmer, Phys. Lett. A **27**, 377 (1968).
[20]B. Jancovici, Phys. Rev. **178**, 295 (1969).
[21]Y. Singh and J. Ram, Molec, Phys. **25**, 145 (1973); **28**, 197 (1974).
[22]Y. Singh, Molec. Phys. **27**, 1687 (1974); J. Ram and Y. Singh, Molec. Phys. (to be published).
[23]G. Nienhuis, J. Math. Phys. **11**, 239 (1970).
[24]W.G. Gibson, Molec. Phys. **28**, 793 (1974).
[25]W.G. Gibson, preprint (1975).
[26](a) G.E. Uhlenbeck and G.W. Ford in *Studies in Statistical Mechanics*, edited by J. de Boer and G.E. Uhlenbeck (North-Holland, Amsterdam, 1962), Vol. I; (b) T. Morita and K. Hiroike, Prog. Theor. Phys. **25**, 537 (1961); (c) G. Stell in *The equilibrium theory of classical fluids*, edited by H.L. Frisch and J.L. Lebowitz (Benjamin, New York, 1964), and references given therein.
[27]J.O. Hirschfelder, C.F. Curtiss, and R.B. Bird, *Molecular theory of gases and liquids* (Wiley, New York, 1964), Chap. 6; J.J. D' Arrunda and R.N. Hill, Phys. Rev. A **1**, 1791 (1970); S.Y. Larsen, J.E. Kilpatrick, E.H. Leeb, and H.F. Jordon, Phys. Rev. **140**, A129 (1965).
[28]W.R. Smith, D. Henderson, and J.A. Barker, J. Chem. Phys. **53**, 508 (1970).
[29]J.A. Barker and D. Henderson, Ann. Rev. Phys. Chem. **23**, 439 (1972).
[30]G. Gasanova, R.J. Dulla, D.A. Jonah, J.S. Rowlinson, and G. Saville, Molec. Phys. **18**, 589 (1970).
[31]J. Ram and Y. Singh, J. Chem. Phys. (to be published).
[32]B. Jancovici, Phys. Rev. **184**, 119 (1969); B. Jancovici and S.P. Merkuriev, preprint (1975).
[33]W.G. Gibson and S.G. Byrnes, Phys. Rev. A **11**, 270 (1975).

Half-space analysis basic to the time-dependent BGK model in the kinetic theory of gases

C. E. Siewert and E. E. Burniston

Departments of Nuclear Engineering and Mathematics, North Carolina State University, Raleigh, North Carolina 27607

(Received 12 July 1976; revised manuscript received 7 September 1976)

The elementary solutions of the linearized time-dependent BGK equation are shown to have, for the case of no discrete eigenvalues, the half-range expansion property necessary for half-space analysis. Also the partial indices corresponding to the basic matrix Riemann problem encountered are shown, for the general case, to be nonnegative, as required for the half-space analysis.

I. INTRODUCTION

The time-dependent BGK model in the kinetic theory of gases can be linearized and expressed in the manner

$$\left(\frac{\partial}{\partial t}+c_x\frac{\partial}{\partial x}+1\right)h(x,\mathbf{c},t)=(\pi)^{-3/2}\int_{E_3}h(x,\mathbf{c}',t)[1+2\mathbf{c}\cdot\mathbf{c}'+\tfrac{2}{3}(c'^2-\tfrac{3}{2})(c^2-\tfrac{3}{2})]e^{-c'^2}d^3c', \tag{1}$$

where $h(x,\mathbf{c},t)$ represents the perturbation of the distribution from the Maxwellian distribution, $\mathbf{c}$, with components c_x, c_y, and c_z and magnitude c, is the velocity, t is the time, and x is the space variable. In the manner of Cercignani,[1] we find that Eq. (1) can be decomposed, by taking moments, into a set of two coupled equations plus three uncoupled equations. Since the uncoupled equations have been discussed in considerable detail,[1] we consider here only the coupled equations,

$$\left(\frac{\partial}{\partial t}+\mu\frac{\partial}{\partial x}+1\right)\Psi(x,\mu,t)=(\pi)^{-1/2}\int_{-\infty}^{\infty}[\mathbf{Q}(\mu)\tilde{\mathbf{Q}}(\mu')+\mathbf{P}(\mu)\tilde{\mathbf{P}}(\mu')]\Psi(x,\mu',t)\,e^{-\mu'^2}d\mu'. \tag{2}$$

Here, the elements of the two-vector $\Psi(x,\mu,t)$ are related[1] to the density and temperature of the gas, and x, μ, and t represent, respectively, the position, velocity component, and time, in dimensionless units. In addition,

$$\mathbf{Q}(\mu)=\begin{bmatrix}(\frac{2}{3})^{1/2}(\mu^2-\frac{1}{2}) & 1\\ (\frac{2}{3})^{1/2} & 0\end{bmatrix} \tag{3a}$$

and

$$\mathbf{P}(\mu)=(2)^{1/2}\mu\begin{bmatrix}1 & 0\\ 0 & 0\end{bmatrix} \tag{3b}$$

We note that the time-independent version of Eq. (2) has been studied extensively by Kriese, Chang, and Siewert.[2]

Since we wish ultimately to solve initial and boundary-value problems relevant to Eq. (2), we first will establish the required elementary solutions.

II. ELEMENTARY SOLUTIONS

We seek solutions of Eq. (2) of the form

$$\Psi(x,\mu,t)=\exp(st)\Phi(\nu,\mu;s)\exp[-(s+1)x/\nu], \tag{4}$$

where, in general, s is complex, but $s\neq-1$, and ν is to be determined. Equation (4) can be entered into Eq. (2) to yield, after some elementary analysis has been carried out,

$$(\nu-\mu)\Phi(\nu,\mu;s)=\omega\nu\mathbf{Q}(\mu)[\mathbf{I}+\gamma\nu\mu\mathbf{D}]\mathbf{M}(\nu;s), \tag{5}$$

where

$$\gamma=2s/(s+1), \tag{6a}$$

$$(\pi)^{1/2}\omega=1/(s+1), \tag{6b}$$

$$\mathbf{D}=\begin{bmatrix}0 & 0\\ 0 & 1\end{bmatrix}, \tag{7}$$

and the normalization vector is given by

$$\mathbf{M}(\nu;s)=\int_{-\infty}^{\infty}\tilde{\mathbf{Q}}(\mu)\Phi(\nu,\mu;s)\,e^{-\mu^2}d\mu. \tag{8}$$

Since the velocity component $\mu\in(-\infty,\infty)$, we can solve Eq. (5) for $\nu\in(-\infty,\infty)$ by writing

$$\Phi(\nu,\mu;s)=\omega\left[\nu P\nu\left(\frac{1}{\nu-\mu}\right)+\hat{\lambda}(\nu)\,\delta(\nu-\mu)\right]\mathbf{Q}(\mu)\times(\mathbf{I}+\gamma\nu\mu\,\mathbf{D})\mathbf{M}(\nu;s). \tag{9}$$

Here $P\nu(1/x)$ denotes the Cauchy principal-value distribution, and $\delta(x)$ represents the Dirac delta distribution. We note that Eq. (9) is a generalization of the "singular eigenfunction" introduced in 1960 by Case[3] and discussed extensively in the text by Case and Zweifel.[4] In Eq. (9) the function $\hat{\lambda}(\nu)$ is considered, at this point, "arbitrary"; however, if we multiply Eq. (9) by $\tilde{\mathbf{Q}}(\mu)\exp(-\mu^2)$ and integrate over μ, we find

$$[\lambda(\nu;s)-\hat{\lambda}(\nu)\Psi(\nu;s)]\mathbf{M}(\nu;s)=0, \tag{10}$$

where

$$\Psi(\mu;s)=\omega e^{-\mu^2}\tilde{\mathbf{Q}}(\mu)\mathbf{Q}(\mu)(\mathbf{I}+\gamma\mu^2\mathbf{D}), \tag{11}$$

and

$$\lambda(\mu;s)=\mathbf{I}+\mu P\int_{-\infty}^{\infty}\Psi(\nu;s)\frac{d\nu}{\nu-\mu}. \tag{12}$$

From Eq. (10) we deduce that $\det[\lambda(\nu;s)-\hat{\lambda}(\nu)\Psi(\nu;s)]=0$ and hence that there exist two $\hat{\lambda}$'s, i.e., $\hat{\lambda}_1(\nu)$ and $\hat{\lambda}_2(\nu)$. We thus write our so-called continuum solutions as

$$\Phi_\alpha(\nu,\mu;s)=\omega\left[\nu P\nu\left(\frac{1}{\nu-\mu}\right)+\hat{\lambda}_\alpha(\nu)\,\delta(\nu-\mu)\right]\times\mathbf{Q}(\mu)(\mathbf{I}+\gamma\nu\mu\,\mathbf{D})\,\mathbf{M}_\alpha(\nu;s),\quad \alpha=1\text{ and }2. \tag{13}$$

In regard to the discrete spectrum, we consider now $\nu \notin (-\infty, \infty)$ and write

$$\Phi(\pm\nu_\alpha, \mu; s) = \omega\left(\frac{\nu_\alpha}{\nu_\alpha \mp \mu}\right)\mathbf{Q}(\mu)(\mathbf{I} \pm \gamma\nu_\alpha\mu\mathbf{D})\mathbf{M}(\nu_\alpha; s), \tag{14}$$

where

$$\Lambda(\pm\nu_\alpha; s)\mathbf{M}(\nu_\alpha; s) = \mathbf{0}, \tag{15}$$

$$\Lambda(z; s) = \mathbf{I} + z\int_{-\infty}^{\infty} \Psi(\mu; s)\frac{d\mu}{\mu - z}, \tag{16}$$

and ν_α is used to denote each of the "positive" zeros of $\Lambda(z; s) = \det\boldsymbol{\Lambda}(z; s)$.

As we have discussed in a previous paper,[5] the dispersion function can be written as

$$\begin{aligned}\Lambda(z; s) = \frac{1}{(s+1)^3}\{&\tfrac{1}{3}s^2z^2 + (s+1)(s-\tfrac{1}{3})(s+\tfrac{1}{2}) \\ &+ [\tfrac{2}{3}s^2z^4 + \tfrac{1}{3}z^2(4s^2-1) + \tfrac{1}{2}(s+1)(\tfrac{11}{3}s+1)]\Lambda(z) \\ &+ \tfrac{2}{3}(s+1+2s^2)\Lambda^2(z)\},\end{aligned} \tag{17}$$

where

$$\Lambda(z) = 1 + \frac{1}{\sqrt{\pi}}z\int_{-\infty}^{\infty} e^{-\mu^2}\frac{d\mu}{\mu - z}. \tag{18}$$

We have shown[5] that $\Lambda(z; s)$ has $\kappa(s)$ pairs of zeros, where $\kappa(s)$ can be either 0, 1, 2, or 3 when s is contained respectively in S_0, S_1, S_2, or S_3, as previously defined.[5]

Having established the required elementary solutions of Eq. (2), we now formally write our general solution (with s as a parameter) as

$$\begin{aligned}\Psi(x, \mu, t; s) = e^{st}\Bigg\{&\sum_{\alpha=1}^{\kappa(s)}[A(\nu_\alpha)\Phi(\nu_\alpha, \mu; s)\exp[-(s+1)x/\nu_\alpha] \\ &+ A(-\nu_\alpha)\Phi(-\nu_\alpha, \mu; s)\exp[(s+1)x/\nu_\alpha]] \\ &+ \int_{-\infty}^{\infty}\sum_{\alpha=1}^{2}A_\alpha(\nu)\Phi_\alpha(\nu, \mu; s) \\ &\times \exp[-(s+1)x/\nu]d\nu\Bigg\}.\end{aligned} \tag{19}$$

Here $A(\pm\nu_\alpha)$, $A_1(\nu)$, and $A_2(\nu)$ are the expansion coefficients to be determined from the boundary and initial conditions. If we let $\mathbf{A}(\nu)$ denote the expansion vector

$$\mathbf{A}(\nu) = A_1(\nu)\mathbf{M}_1(\nu; s) + A_2(\nu)\mathbf{M}_2(\nu; s), \tag{20}$$

then Eq. (19) can be written as

$$\begin{aligned}\Psi(x, \mu, t; s) = e^{st}\Bigg\{&\sum_{\alpha=1}^{\kappa(s)}[A(\nu_\alpha)\Phi(\nu_\alpha, \mu; s)\exp[-(s+1)x/\nu_\alpha] \\ &+ A(-\nu_\alpha)\Phi(-\nu_\alpha, \mu; s)\exp[(s+1)x/\nu_\alpha]] \\ &+ \int_{-\infty}^{\infty}\Phi(\nu, \mu; s)\mathbf{A}(\nu)\exp[-(s+1)x/\nu]d\nu\Bigg\},\end{aligned} \tag{21}$$

where the continuum *matrix* is

$$\begin{aligned}\Phi(\nu, \mu; s) = &\omega\nu Pv\left(\frac{1}{\nu - \mu}\right)\mathbf{Q}(\mu)(\mathbf{I} + \gamma\nu\mu\mathbf{D}) \\ &+ \delta(\nu - \mu)e^{\nu^2}\tilde{\mathbf{Q}}^{-1}(\nu)\boldsymbol{\lambda}(\nu; s).\end{aligned} \tag{22}$$

III. HALF-RANGE ANALYSIS

We wish now to show that the "eigenvectors" established in the previous section have an important property that allows us to write

$$\mathbf{I}(\mu) = \sum_{\alpha=1}^{\kappa(s)}A(\nu_\alpha)\Phi(\nu_\alpha, \mu; s) + \int_0^\infty \Phi(\nu, \mu; s)\mathbf{A}(\nu)d\nu, \quad \mu \in (0, \infty), \tag{23}$$

where $\mathbf{I}(\mu)$ is an arbitrary two-vector which is Hölder continuous on any bounded interval of the positive real axis and further satisfies

$$|I_\alpha(\mu)| < Ce^{\mu}, \quad \alpha = 1, 2, \quad \mu \in (0, \infty), \tag{24}$$

where C is a positive constant.

Equation (23) is the statement equivalent to Case's[3] half-range completeness theorem for the one-speed neutron problem and clearly will be required when we wish to solve explicitly a typical half-space problem.

In order to illustrate explicitly the analysis required to prove Eq. (23), we consider currently only those values of $s \in S_0$, so we can allow $\kappa(s)$ to be zero. By introducing the sectionally analytic vector function

$$\mathbf{N}(z) = \frac{1}{2\pi i}\int_0^\infty \nu(\mathbf{I} + \gamma\nu z\mathbf{D})\mathbf{A}(\nu)\frac{d\nu}{\nu - z}, \tag{25}$$

with limiting values

$$\begin{aligned}\mathbf{N}^{\pm}(t) = &\frac{1}{2\pi i}P\int_0^\infty \nu(\mathbf{I} + \gamma\nu t\mathbf{D})\mathbf{A}(\nu)\frac{d\nu}{\nu - t} \\ &\pm \tfrac{1}{2}t(\mathbf{I} + \gamma t^2\mathbf{D})\mathbf{A}(t),\end{aligned} \tag{26}$$

we can express the equation

$$\mathbf{I}(\mu) = \int_0^\infty \Phi(\nu, \mu; s)\mathbf{A}(\nu)d\nu, \quad \mu \in (0, \infty) \text{ and } s \in S_0, \tag{27}$$

in the form

$$\begin{aligned}\mu\boldsymbol{\Pi}(\mu)\tilde{\mathbf{Q}}(\mu)e^{-\mu^2}\mathbf{I}(\mu) = &\boldsymbol{\Omega}^+(\mu; s)\boldsymbol{\Pi}^{-1}(-\mu)\mathbf{N}^+(\mu) \\ &- \boldsymbol{\Omega}^-(\mu; s)\boldsymbol{\Pi}^{-1}(-\mu)\mathbf{N}^-(\mu).\end{aligned} \tag{28}$$

Here we have introduced

$$\boldsymbol{\Pi}(z) = \mathbf{I} - \left(\frac{z}{z_1}\right)\mathbf{D}, \quad (\gamma)^{1/2}z_1 = i, \tag{29}$$

and

$$\boldsymbol{\Omega}^{\pm}(\mu; s) = \boldsymbol{\Pi}(\mu)[\boldsymbol{\lambda}(\mu; s) \pm \pi i\mu\boldsymbol{\Psi}(\mu; s)]\boldsymbol{\Pi}^{-1}(\mu). \tag{30}$$

It can easily be shown that the matrix

$$\boldsymbol{\Omega}(z; s) = \boldsymbol{\Pi}(z)\boldsymbol{\Lambda}(z; s)\boldsymbol{\Pi}^{-1}(z) \tag{31}$$

has the limiting values given by Eq. (30) and can be written as

$$\boldsymbol{\Omega}(z; s) = \mathbf{I} + z\int_{-\infty}^{\infty}\hat{\boldsymbol{\Psi}}(\mu; s)\frac{d\mu}{\mu - z}, \tag{32}$$

where

$$\hat{\boldsymbol{\Psi}}(\mu; s) = \boldsymbol{\Pi}(\mu)\boldsymbol{\Psi}(\mu; s)\boldsymbol{\Pi}^{-1}(\mu). \tag{33}$$

The $\boldsymbol{\Omega}(z; s)$ matrix has other important properties that we will soon require,

$$\overline{\boldsymbol{\Omega}(\bar{z}; \bar{s})} \equiv \bar{\boldsymbol{\Omega}}(z; s) = \boldsymbol{\Omega}(-z; s) = \tilde{\boldsymbol{\Omega}}(z; s). \tag{34}$$

If now we let $\mathbf{X}(z; s)$ denote a canonical solution, of ordered normal form at infinity,[5,6] to the matrix Riemann problem defined by

$$\mathbf{X}^+(\mu;s)=\mathbf{G}(\mu;s)\,\mathbf{X}^-(\mu;s),\quad \mu\in[0,\infty), \tag{35}$$

where

$$\mathbf{G}(\mu;s)=\tilde{\mathbf{\Omega}}^+(\mu;s)[\tilde{\mathbf{\Omega}}^-(\mu;s)]^{-1}, \tag{36}$$

then we can write the solution to Eq. (28) as

$$\tilde{\mathbf{X}}(z;s)\,\mathbf{\Pi}^{-1}(-z)\,\mathbf{N}(z)=\frac{1}{2\pi i}\left(\int_0^\infty \mathbf{\Gamma}(\mu)\frac{d\mu}{\mu-z}+R(z)\right), \tag{37}$$

where

$$\mathbf{\Gamma}(\mu)=\mu\tilde{\mathbf{X}}^+(\mu;s)\,[\mathbf{\Omega}^+(\mu;s)]^{-1}\,\mathbf{\Pi}(\mu)\tilde{\mathbf{Q}}(\mu)\,e^{-\mu^2}\mathbf{I}(\mu). \tag{38}$$

In Eq. (37) we use $\mathbf{R}(z)$ to denote a vector of rational functions. At this point we wish to make use of the fact (proved in the next section) that the partial indices, κ_1 and κ_2, basic to the Riemann problem defined by Eqs. (35) and (36), are nonnegative. Since we are considering here the case $\kappa=0$, then clearly $\kappa_1=\kappa_2=\kappa=0$, and thus we can normalize our canonical solution by taking

$$\mathbf{X}(\infty;s)=\mathbf{I},\quad \kappa=0. \tag{39}$$

On investigating Eq. (37) for large z and noting that $\mathbf{R}(z)$ can be singular only at $z=-z_1$, we conclude, after examining the form of Eq. (25) for large $|z|$, that

$$\mathbf{R}(z)=(z_1+z)^{-1}\mathbf{R}, \tag{40}$$

where the constant vector $\mathbf{R}$ can be expressed as

$$\mathbf{R}=\int_0^\infty \mathbf{\Gamma}(\mu)\,d\mu-(\mathbf{I}-\mathbf{D})\int_0^\infty \nu\mathbf{A}(\nu)\,d\nu-\gamma z_1\mathbf{D}\int_0^\infty \mathbf{A}(\nu)\nu^2\,d\nu. \tag{41}$$

Thus we can now write Eq. (37) as

$$\mathbf{\Pi}^{-1}(-z)\,\mathbf{N}(z)=\tilde{\mathbf{X}}^{-1}(z;s)\frac{1}{2\pi i}\left[\int_0^\infty \mathbf{\Gamma}(\mu)\frac{d\mu}{\mu-z}+\frac{1}{z+z_1}\,\mathbf{R}\right]. \tag{42}$$

If now we notice that Eq. (26) yields

$$\mathbf{N}^+(t)-\mathbf{N}^-(t)=t\mathbf{\Pi}(t)\,\mathbf{\Pi}(-t)\,\mathbf{A}(t), \tag{43}$$

we can obtain from Eq. (42) the expression

$$t\mathbf{\Pi}(t)\,\mathbf{A}(t)=\int_0^\infty\left(\mathbf{U}(t)\frac{P}{\mu-t}+\mathbf{V}(t)\,\delta(\mu-t)\right)\mathbf{X}(-t;s)\,\mathbf{\Omega}(\infty;s)\times\mathbf{\Gamma}(\mu)\,d\mu+\mathbf{U}(t)\,\mathbf{X}(-t;s)\,\mathbf{\Omega}(\infty;s)\frac{\mathbf{R}}{z_1+t}. \tag{44}$$

In developing Eq. (44) we have used the fact that the $\mathbf{X}$ matrix factors $\mathbf{\Omega}(z;s)$ in the following manner:

$$\tilde{\mathbf{\Omega}}(z;s)=\mathbf{X}(z;s)\,\mathbf{\Omega}(\infty;s)\,\tilde{\mathbf{X}}(-z;s),\quad \kappa=0. \tag{45}$$

Here we have defined

$$2\pi i\mathbf{U}(t)=[\mathbf{\Omega}^+(t;s)]^{-1}-[\mathbf{\Omega}^-(t;s)]^{-1} \tag{46a}$$

and

$$2\mathbf{V}(t)=[\mathbf{\Omega}^+(t;s)]^{-1}+[\mathbf{\Omega}^-(t;s)]^{-1}. \tag{46b}$$

We can find from Eq. (44) the moments of $\mathbf{A}(\nu)$ required in Eq. (41) to establish $\mathbf{R}$. After using the integral representation

$$\mathbf{\Omega}^{-1}(z;s)\,\mathbf{X}(-z;s)=\mathbf{\Omega}^{-1}(\infty;s)+\int_0^\infty \mathbf{U}(t)\,\mathbf{X}(-t;s)\frac{dt}{t-z}$$

$$\kappa=0, \tag{47}$$

to help simplify our result, we find

$$\mathbf{R}=-2z_1[\mathbf{I}+\tilde{\mathbf{X}}(-z_1;s)\,\mathbf{D}\tilde{\mathbf{X}}^{-1}(z_1;s)]^{-1}\times\tilde{\mathbf{X}}(-z_1;s)\,\mathbf{D}\tilde{\mathbf{X}}^{-1}(z_1;s)\int_0^\infty \mathbf{\Gamma}(\mu)\frac{d\mu}{\mu-z_1}. \tag{48}$$

Since $\mathbf{R}$ is now established explicitly, we consider the proof that $\mathbf{I}(\mu)$ can be represented as given by Eq. (23) completed for the case $\kappa=0$.

IV. PARTIAL INDICES

As mentioned in the previous section, we require in our proof of "half-range completeness" the knowledge that the partial indices basic to the matrix Riemann problem defined by Eqs. (35) and (36) are nonnegative. The proof that we will develop here is similar to the one given previously[7] for a problem relating to the scattering of polarized light; however, because $\mathbf{\Omega}(z;s)$ is not symmetric and because the problem contains a complex parameter s, some additional work is required. We consider in this section the total index κ to be 0, 1, 2, or 3.

First of all, we note that

$$\mathbf{\Phi}(z;s)\equiv\tilde{\mathbf{\Omega}}(z;s)\,\tilde{\mathbf{X}}^{-1}(-z;s) \tag{49}$$

is a solution of the Riemann problem defined by

$$\mathbf{\Phi}^+(\mu;s)=\mathbf{G}(\mu;s)\,\mathbf{\Phi}^-(\mu;s),\quad \mu\in[0,\infty), \tag{50}$$

where $\mathbf{G}(\mu;s)$ is given by Eq. (36), and thus[6] $\mathbf{\Phi}(z;s)$ can be expressed as

$$\mathbf{\Phi}(z;s)=\mathbf{X}(z;s)\,\mathbf{P}(z), \tag{51}$$

where $\mathbf{P}(z)$ is a matrix of polynomials. It is clear that Eqs. (49) and (51) yield the factorization

$$\tilde{\mathbf{\Omega}}(z;s)=\mathbf{X}(z;s)\,\mathbf{P}(z)\,\tilde{\mathbf{X}}(-z;s). \tag{52}$$

We note that by definition[6] a canonical solution of ordered normal form at infinity is such that

$$\lim_{|z|\to\infty}\mathbf{X}(z;s)\begin{bmatrix}z^{\kappa_1}&0\\0&z^{\kappa_2}\end{bmatrix}=\mathbf{K},\ \det\mathbf{K}\neq 0, \tag{53}$$

where $\kappa_1\leq\kappa_2$ and κ_2 are the partial indices and $\kappa_1+\kappa_2=\kappa$. If we use Eq. (53) in Eq. (52), as $|z|\to\infty$, we can readily deduce that $\kappa_1\geq 0$ unless $P_{11}(z)\equiv 0$. Thus to show that $\kappa_2\geq\kappa_1\geq 0$, we need to prove that $P_{11}(z)\neq 0$.

If we now change s to $\bar{s}$ in Eqs. (35) and (36) and take the complex conjugate of the resulting equations, we can use Eq. (34) to deduce that

$$\bar{\tilde{\mathbf{X}}}^-(\mu;s)=\mathbf{G}_*(\mu;s)\bar{\tilde{\mathbf{X}}}^-(\mu;s),\quad \mu\in[0,\infty), \tag{54}$$

where

$$\mathbf{G}_*(\mu;s)=\overline{[\mathbf{G}(\mu;\bar{s})]}^{-1}=\mathbf{\Omega}^+(\mu;s)[\mathbf{\Omega}^-(\mu;s)]^{-1} \tag{55}$$

and we have defined

$$\bar{\tilde{\mathbf{X}}}(z;s)=\overline{\mathbf{X}(\bar{z};\bar{s})}. \tag{56}$$

Using the fact that $\mathbf{W}(z;s)=\mathbf{\Pi}(z)\,\mathbf{\Pi}(-z)\,\mathbf{\Lambda}(z;s)$ is symmetric, we can deduce a convenient relationship,

$$\mathbf{B}(\mu)\,\mathbf{G}_*(\mu;s)=\mathbf{G}(\mu;s)\,\mathbf{B}(\mu), \tag{57}$$

where

$$\mathbf{B}(\mu)=\mathbf{\Pi}^{-1}(\mu)\,\mathbf{\Pi}(-\mu). \tag{58}$$

It is not difficult now to deduce that

$$\overline{\overline{\mathbf{X}}}(z;s)=\mathbf{B}^{-1}(z)\mathbf{X}(z;s)\mathcal{R}(z), \tag{59}$$

where $\mathcal{R}(z)$ is a matrix of rational functions. Further, we observe from Eq. (59) that $\mathcal{R}(z)$ must be of the form

$$\mathcal{R}(z)=\left(\frac{1}{z_1-z}\right)\hat{P}(z), \tag{60}$$

where $\hat{P}(z)$ has polynomial elements. If we let

$$\mathbf{X}(z;s)\to\mathbf{K}(s)\begin{bmatrix}z^{-\kappa_1} & 0\\ 0 & z^{-\kappa_2}\end{bmatrix}, \quad |z|\to\infty, \tag{61}$$

then we can find the general form of $\hat{P}(z)$ by investigating Eqs. (59) and (60) as $|z|\to\infty$:

$$\hat{P}(z)\to z\begin{bmatrix}T_{11} & T_{12}\,z^{\kappa_1-\kappa_2}\\ T_{21}\,z^{\kappa_2-\kappa_1} & T_{22}\end{bmatrix}, \quad |z|\to\infty, \tag{62}$$

where

$$\mathbf{T}=\mathbf{K}^{-1}(s)\begin{bmatrix}-1 & 0\\ 0 & 1\end{bmatrix}\overline{\mathbf{K}(\bar{s})}. \tag{63}$$

We assume here that $\kappa_1\neq\kappa_2$, for otherwise no proof that κ_1 and κ_2 are nonnegative is required, and thus we can consider

$$\mathbf{K}(s)=\begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix} \tag{64a}$$

or

$$\mathbf{K}(s)=\begin{bmatrix}1 & b(s)\\ 0 & 1\end{bmatrix} \tag{64b}$$

and deduce that the most general form of $\hat{P}(z)$ is

$$\hat{P}(z)=\pm\begin{bmatrix}\hat{P}_{11}+z & 0\\ \hat{P}_{21}(z) & \hat{P}_{22}-z\end{bmatrix}, \tag{65}$$

where $\hat{P}_{11}$ and $\hat{P}_{22}$ are constants. If now we use Eqs. (65) and (60) and evaluate Eq. (59) at $z=0$, we obtain

$$\overline{X_{12}(0;\bar{s})}=\pm\left(\frac{1}{z_1}\right)\hat{P}_{22}X_{12}(0;s) \tag{66a}$$

and

$$\overline{X_{22}(0;\bar{s})}=\pm\left(\frac{1}{z_1}\right)\hat{P}_{22}X_{22}(0;s). \tag{66b}$$

Equations (66) allow us to prove the required statement that the polynomial $P_{11}(z)$ appearing in Eq. (52) is not identically zero. Since $\Omega(0;s)=\mathbf{I}$, we can solve Eq. (52) to obtain

$$P_{11}(0)=X^{-2}(0;s)\,[X_{12}^2(0;s)+X_{22}^2(0;s)], \tag{67}$$

where $X(0;s)=\det\mathbf{X}(0;s)$. To allow $P_{11}(0)=0$ yields

$$X_{22}(0;s)=\pm iX_{12}(0;s) \tag{68}$$

which contradicts Eqs. (66). Since $P_{11}(0)\neq 0$ it follows that the partial indices basic to the Riemann problem defined by Eqs. (35) and (36) are nonnegative.

V. THE H MATRIX

If we go back to Eq. (52) and use the normalization $\mathbf{X}(\infty;s)=\mathbf{I}$, $\kappa=0$, we can write

$$\tilde{\Omega}(z;s)=\mathbf{X}(z;s)\,\mathbf{X}^{-1}(0;s)\,\tilde{\mathbf{X}}^{-1}(0;s)\,\tilde{\mathbf{X}}(-z;s). \tag{69}$$

Therefore, if we define the H matrix by

$$\mathbf{H}(z;s)=\tilde{\mathbf{X}}^{1}(-z;s)\,\tilde{\mathbf{X}}(0;s), \tag{70}$$

then a factorization of $\Omega(z;s)$ becomes

$$\tilde{\Omega}(z;s)=\tilde{\mathbf{H}}^{-1}(-z;s)\,\mathbf{H}^{-1}(z;s). \tag{71}$$

Since ultimately we wish to express all of our results in terms of the convenient H matrix, we can use

$$\mathbf{H}^{-1}(z;s)=\mathbf{H}^{-1}(\infty;s)+\frac{1}{2\pi i}\times\int_{-\infty}^{0}\tilde{\mathbf{H}}(-t;s)\,[\tilde{\Omega}^{+}(t;s)-\tilde{\Omega}^{-}(t;s)]\frac{dt}{t-z} \tag{72}$$

or

$$\mathbf{H}^{-1}(z;s)=\mathbf{I}-z\int_{0}^{\infty}\tilde{\mathbf{H}}(t;s)\,\hat{\Psi}(t;s)\frac{dt}{t+z} \tag{73}$$

to compute $\mathbf{H}(z;s)$ for $z\notin(0,\infty)$ after we have solved

$$\mathbf{H}^{-1}(\mu;s)=\mathbf{I}-\mu\int_{0}^{\infty}\tilde{\mathbf{H}}(t;s)\,\hat{\Psi}(t;s)\frac{dt}{t+\mu}, \quad \mu\in[0,\infty), \tag{74}$$

iteratively. It is clear that Eq. (74) has a solution since we know that $\mathbf{X}(z;s)$ exists and the subsequent definition of $\mathbf{H}(z;s)$ in terms of $\mathbf{X}(z;s)$; however, the recent work of Zweifel and co-workers[8,9] could prove very useful for showing that an iterative solution of Eq. (74) converges to the desired result.

VI. SOUND-WAVE PROPAGATION

It is evident that we can readily solve half-space problems based on Eq. (2) subject to a free-surface boundary condition of the form

$$\Psi(0,\mu,t)=\exp(st)\,\mathbf{F}(\mu), \quad \mu\geq 0, \tag{75}$$

and a specified condition as $x\to\infty$. Here, we consider $\mathbf{F}(\mu)$ to be, in general, an arbitrary Hölder function. For example, for sound-wave propagation in a half-space defined by $\Psi(x,\mu,t)\to 0$ as $x\to\infty$, and

$$\Psi(0,\mu,t)=\exp(i\omega t)\,\mathbf{F}(\mu), \quad \mu>0, \tag{76}$$

we simply let $s=i\omega$ and write the desired solution as

$$\Psi(x,\mu,t)=\exp(i\omega t)\left[\sum_{\alpha=1}^{\kappa}A(\nu_\alpha)\,\Phi(\nu_\alpha,\mu;i\omega)\exp[-(i\omega+1)\,x/\nu_\alpha]+\int_0^\infty\Phi(\nu,\mu;i\omega)A(\nu)\exp[-(i\omega+1)x/\nu]\,d\nu\right]. \tag{77}$$

If we constrain Eq. (77) to meet Eq. (76), we get

$$\mathbf{F}(\mu)=\sum_{\alpha=1}^{\kappa}A(\nu_\alpha)\,\Phi(\nu_\alpha,\mu;i\omega)+\int_0^\infty\Phi(\nu,\mu;i\omega)\,\mathbf{A}(\nu)\,d\nu, \quad \mu\geq 0. \tag{78}$$

The solution of Eq. (78) is given in Sec. III for the case $\kappa=0$; we note from our previous work[5] that $\omega>2.14517\cdots\Rightarrow\kappa=0$. We are confident that explicit solutions of Eq. (78) for a general index will soon be forthcoming.

ACKNOWLEDGMENT

The authors are grateful to Dr. J.R. Thomas, Jr. for some helpful discussions regarding this work.

[1]C. Cercignani, *Mathematical Methods in Kinetic Theory* (Plenum, New York, 1969).
[2]J.T. Kriese, T.S. Chang, and C.E. Siewert, Int. J. Eng. Sci. 12, 441 (1974).
[3]K.M. Case, Ann. Phys. 9, 1 (1960).
[4]K.M. Case and P.F. Zweifel, *Linear Transport Theory* (Addison-Wesley, Reading, Massachusetts, 1967).
[5]C.E. Siewert, E.E. Burniston, and J.R. Thomas, Jr., Phys. Fluids 16, 1532 (1973).
[6]N.I. Muskhelishvili, *Singular Integral Equations* (Noordhoff, Groningen, The Netherlands, 1953).
[7]C.E. Siewert and E.E. Burniston, Astrophys. J. 174, 629 (1972).
[8]R.L. Bowden and P.F. Zweifel, Astrophys. J. "A Banach Space Analysis of the Chandrasekhar *H*-Equation" (in press).
[9]R.L. Bowden, P.F. Zweifel, and R. Menikoff, J. Math. Phys. 17, 1722 (1976).

The lattice of verifiable propositions of the spin-1 system*

Bror O. Hultgren, III

Department of Physics, Simmons College, Boston, Massachusetts 02115

Abner Shimony

Departments of Philosophy and Physics, Boston University, Boston, Massachusetts 02215
(Received 25 March 1976)

A lattice of verifiable propositions L_V for a spin-1 system is constructed by admitting only propositions which correspond to appropriate Stern–Gerlach filters. L_V is a complete, orthocomplemented, weakly modular lattice, and it satisfies the first part of the atomicity axiom of Jauch and Piron, but not the second part (the covering law) nor related axioms of Zierler and MacLaren. Doubt is therefore thrown upon the program of recovering the Hilbert space formulation of quantum mechanics from empirically justified axioms. The class of admissible states on L_V is exhaustively characterized, and it is shown that there exist some nonquantal states but none that are dispersion free.

I. INTRODUCTION

In spite of the innumerable confirmations of quantum mechanics, it is still far from clear how completely the standard formulation of the theory, namely, the Hilbert space formulation, is justified empirically. Conceptual difficulties, such as the problem of measurement, can be taken as indications that the Hilbert space formulation of quantum mechanics is an excessive extrapolation beyond the empirical evidence. The discovery of superselection rules,[1] moreover, shows that certain features of the Hilbert space formulation can be modified without loss, and indeed with increase of explanatory power. Understandably, therefore, much of the recent work in the foundations of quantum mechanics has been devoted to attempts to formulate the theory in a stepwise axiomatic manner,[2–8] with hopes of exhibiting the empirical support of each axiom. We are rather skeptical on general methodological grounds about the likelihood that such programs will yield definitive results, because of the difficulty of assessing the empirical consequences of individual axioms. Furthermore, it seems probable to us that the solutions to the problems of the foundations of quantum mechanics will be inseparable from new discoveries in such areas as elementary particle theory and space–time theory. We nevertheless feel that a careful examination of the various reformulations of quantum mechanics which have recently been proposed can be illuminating, and at the least can yield interesting negative results concerning some of the programs in the foundations of quantum mechanics.

The purpose of this paper is to develop an instrument which will be useful in the enterprise of examining reformulations of quantum mechanics. The instrument is the set L_V, consisting of those propositions which may reasonably be regarded to be empirically testable concerning a spin-1 system (with nonzero rest mass). A precise characterization of L_V will be given in Sec. IV. It suffices for the present to say that a proposition is in L_V if its truth value can be determined by means of a suitable spin measuring device, such as a Stern–Gerlach apparatus. If the standard formulation of quantum mechanics (summarized in Sec. II and Sec. III) is applied to the spin-1 system, there is a projection operator (or equivalently, a closed linear subspace of a three-dimensional Hilbert space) corresponding to each proposition in L_V, but there are also projection operators which do not correspond to members of L_V and which in fact do not seem to correspond in any natural way to testable propositions. It is not our intention, in paying special attention to L_V, to insist that every term in an acceptable physical theory be susceptible of an operational interpretation.[9] However, the availability of an operational interpretation of a term does provide *prima facie* evidence that there is an element of physical reality correlated with the term, and in the absence of other evidence about the physical significance of the projection operators in the Hilbert space formulation it is a good working hypothesis that those operators corresponding to members of L_V have a definite physical status which the others do not have.

In Sec. V we establish exhaustively the structure of L_V. In principle this could be done quite directly from experiments, using reasonable interpolations and inductive generalizations. Our procedure, however, will be to draw upon the quantum mechanical predictions for the spin-1 system, relying upon the assumption that the quantum mechanical predictions concerning actual observations of spin are all correct. Thus, even though we are engaged in a critical investigation of the Hilbert space formulation of quantum mechanics, we may make free use of the consequences of this formulation in any domain in which it has been successful. There is no inconsistency in this procedure, for one can suspect that some parts of a formalism lack physical content and nevertheless believe that the formalism is correct whenever it is physically significant.

Once the structure of L_V is determined we check in Sec. VI whether the axioms proposed in several alternative formulations of quantum mechanics are valid in L_V. We are particularly interested in the axioms of Jauch and Piron,[4,5] because of the remarkable mathematical work of Piron,[4] showing that the Hilbert space formulation of quantum mechanics can be essentially recovered (with small modifications, such as allowance for superselection rules) from these axioms. We find that all the axioms of Jauch and Piron are satisfied by L_V except one, the "covering law." The failure of this axiom for a system as simple, well-understood, and unimpeachably physical as L_V throws doubt upon the

validity of the axiom and makes it appear that the axiom is motivated by a goal, the recovery of the Hilbert space formulation, rather than by physical evidence. We also show that two of the axioms of Zierler[8] and one of MacLaren,[7] both with programs similar to that of Jauch and Piron, fail to hold in L_V. These negative results constitute, in our opinion, a serious obstacle to the program of recovery the Hilbert space formulation of quantum mechanics from empirically well justified principles.

In Sec. VII we study the states (in the sense of σ-additive measures) defined on L_V. We prove a theorem which essentially exhibits all the possible states on L_V. We also prove in a new way a result previously obtained by Kochen and Specker[10] and by Belinfante,[11] that there exists no dispersion-free state on L_V. This result is philosophically very significant. It might be conjectured that the nonexistence of a dispersion-free state on the standard quantum mechanical lattice of propositions for the spin-system is due to the occurrence of propositions having no physical significance; but such a conjecture would obviously not be true of L_V, because of the undeniably physical character of all of its propositions. Thus, restricting attention to propositions which have definite physical significance does not suffice to save one important class of hidden-variable theories.

II. THE CONCEPTS OF PROPOSITION AND STATE

The fundamental concept in our investigation is a *proposition* concerning a physical system. The precise explanation of this concept cannot be given without solving some of the deep problems of the foundations of quantum mechanics (including the measurement problem, which is essentially the problem of determining exactly when and how a potentiality is realized). We nevertheless can convey the intended meaning in a preliminary manner, sufficient for the structural investigation of this paper. A proposition is—with two convenient exceptions—a bivalent potentiality of the system, one realization of it being identified as the truth of the proposition and the other as its falsity. The two exceptions are the "impossible" proposition ϕ, which can only be false, and the "necessary" proposition 1, which can only be true. There is no *a priori* assurance that in an arbitrary state of the system a given proposition has a definite truth value, and indeed according to the usual interpretation of quantum mechanics every proposition, except ϕ and 1, is unrealized in some states of the system. We do not take the facts about the constitution of a system, such as its mass or charge or composition out of more elementary parts, to be propositions, although one could consider them to be associated in a many—one manner with the necessary proposition.

Another way of conveying the intended meaning of "proposition" is to say that an ideal test for a proposition would be an experiment with only two possible outcomes, such that one outcome is a sufficient condition for correctly saying that the proposition is true and the other outcome is a sufficient condition for correctly saying that it is false. Such a bivalent experiment can be schematically represented by a filter, through which the system can either pass or not pass, passage being a sufficient condition for the truth of the proposition at the moment of completion of the experiment and nonpassage for its falsity at that moment.

Several cautionary remarks are essential to prevent misunderstanding at this point.

(i) It is a great idealization to suppose that a definite bivalent experiment can be performed upon a system regardless of the state prior to the experiment. A minimum, though usually far from sufficient, condition in practice is that the system enter the forward aperture of the filter; but this condition implies approximate localization and would not be satisfied, for example, if the system were in a state of nearly exact linear momentum. (The condition of approximate localization imposed upon the initial state is fortunately not troublesome in the investigations of the present paper, since propositions about spin are compatible with those about position.)

(ii) Even if we disregard the first difficulty, we still would not be justified in associating a proposition with a specific experimental procedure, but only with an equivalence class of bivalent procedures, such that if any one yields the outcome "true" (or "false") regarding the proposition in question, so will any other. The equivalence class of procedures is an "open" class, in the sense that a physical apparatus as yet uninvented could conceivably interact with the system in a manner adequate for a bivalent experimental test of the proposition of interest.

(iii) In actual experimental situations it is an idealization to say that one outcome is a sufficient condition for the truth of the proposition and the other outcome for its falsity. Almost always there is a nonnegligible probability of an erroneous correlation of experimental outcomes with truth values of the propositions, not only because of technical difficulties or perturbations, but sometimes also for reasons of principle.[12–14]

Remarks (i), (ii), and (iii) are strong reasons against an operationalist interpretation either of specific propositions or of the concept of proposition (see Fig. 9, pp. 371—3, 408—9, and 425). Nevertheless, these cautionary remarks do not preclude laboratory experiments which yield realizations of propositions, and indeed in enough situations to provide weighty evidence regarding relations among propositions and hence regarding the structure of the set of propositions.

We shall suppose that every proposition p has a unique orthocomplement p', which is always realized when p is realized, but in such a way that when realized p and p' have opposite truth values. Let us consider any one of the equivalence class of yes-or-no experiments associated with p, but idealized so as to be performable whatever the initial state of the system may be. Then p' is that proposition which is realized by the same experiment, but such that a sufficient condition for its truth (respectively, falsity) is the outcome which is sufficient for the falsity (respectively, truth) of p. In this way, the bivalence of an arbitrary experimental realization of p can be used to indicate a sense in which p and p' are exhaustive. It must be emphasized, however, that in spite of the exhaustiveness of p and p',

there are many states in which both p and p' are unrealized potentialities (unless p is either ϕ or 1).

The fundamental relation between propositions is *implication*. A necessary and sufficient condition for p to imply q, symbolized by $p \leq q$, is that in every state in which p is true, q is also true. A bivalent experimental test of p cannot be considered without further information to be a test of q, since q may not be realized if p is false.

We shall assume the following axioms concerning the set of propositions (henceforth designated by L), the operation of orthocomplementation, and the relation of implication:

P (partial ordering): (a) $a \leq a$ for all $a \in L$,

(b) if $a \leq b$ and $b \leq a$, then $a = b$,

(c) if $a \leq b$ and $b \leq c$, then $a \leq c$.

B (boundedness from below): There exists a proposition ϕ in L such that $\phi \leq a$ for all $a \in L$.

O (orthocomplementation): For every $a \in L$ there exists a unique proposition $a' \in L$ such that

(a) $(a')' = a$,

(b) for any $x \in L$, if $x \leq a$ and $x \leq a'$, then $x = \phi$,

(c) if $a \leq b$, then $b' \leq a'$.

We regard these axioms as analytic, in the sense of being implicit in the concepts of proposition, orthocomplementation, and implication, and experimental evidence for them is not needed.

Some derivative concepts can now be introduced.

If $\{a_i \mid i$ belonging to an index set $I\}$ is a subset of L (assumed for the moment to be partially ordered, but with no assumption of boundedness from below or of orthocomplementation), and if there exists a proposition x such that

(i) $x \leq a_i$ for all $i \in I$,

(ii) $y \leq a_i$ for all $i \in I$ implies $y \leq x$,

then x is *the greatest lower bound* or *g.l.b.* or the *generalized conjunction* of the $\{a_i\}$, and it is denoted by $\wedge_{i \in I} a_i$. If I consists of two indices 1 and 2, then a more convenient notation for the g.l.b. is $a_1 \wedge a_2$.

If a partially ordered set of propositions L is such that for every pair a, b their g.l.b. exists, then the set is a *lattice*. If any subset $\{a_i \mid i \in I\}$ has a g.l.b., then L is a *complete lattice*. If any denumerable subset $\{a_i \mid i = 1, 2, \ldots\}$ has a g.l.b., then L is a *σ-lattice*. (The statement of these definitions does not presuppose that L is bounded from below or orthocomplemented.)

If $\{a_i \mid i \in I\}$ is a subset of L such that $\wedge_{i \in I} a_i'$ exists, then $(\wedge_{i \in I} a_i')'$ is *the least upper bound* or *l.u.b.* or *generalized disjunction* of the $\{a_i\}$, and it is denoted by $\vee_{i \in I} a_i$. If I consists of two indices 1 and 2, then a more convenient notation for the l.u.b. is $a_1 \vee a_2$. Caution is needed to avoid misconceptions suggested by the name "generalized disjunction." The concept as defined permits the possibility—which indeed obtains in quantum mechanics and is contrary to the ordinary notion of disjunction—that $\vee_{i \in I} a_i$ is realized and true while not a single one of the a_i is realized.

If $a \leq b'$, then a and b are *disjoint*. (Note that disjointness is *not* defined as $a \wedge b = \phi$.)

The proposition 1 is defined as ϕ'. [It follows from axioms B and O(b) that for all $a \in L$, $a \leq 1$, so that the existence of the necessary proposition need not be separately postulated.]

In our informal discussion of the concept of proposition we spoke several times of the state of the system, without explaining the locution. We are primarily interested in pure states, which can roughly be characterized as maximal specifications of the system. In classical physics a maximal specification essentially consists in the simultaneous assignment of truth values to all propositions concerning the system. The structure of the set of propositions postulated by quantum mechanics precludes the existence of pure states in the classical sense (results of Gleason[15] and others) and motivates a probabilistic conception of state. (If the logical difference between certainty and 100% probability is set aside, then the classical conception of pure state is subsumed under the probabilistic conception.) We therefore adopt Mackey's conception of a state[3]: namely, a real-valued function defined on the set of propositions L such that

(i) $m(a) \geq 0$, for all $a \in L$,

(ii) $m(1) = 1$,

(iii) if $\{a_i \mid i = 1, 2, \cdots\}$ is a finite or denumerable subset of L such that $a_i \leq a_j'$ for $i \neq j$, then (if $\vee a_i$ exists) $m(\vee a_i) = \sum m(a_i)$.

Mackey's conception seems to catch as much of the classical conception of a probability measure on the set of propositions as could be expected, given the physical evidence that not all propositions are realizable together. In particular, according to condition (iii) complete additivity of m holds for a denumerable set of pairwise disjoint propositions. Although his conception might be weakened or generalized, it so elegantly combines conservatism regarding probability theory with recognition of the exigencies of microphysics that departures from it would be reasonable only if they led to substantial theoretical extension or clarification. States in Mackey's sense can be classified as mixed or pure by the following criterion: m is a *mixed* state if there exist two distinct states m_1 and m_2 and two positive real numbers c_1 and c_2 with $c_1 + c_2 = 1$, such that for all $a \in L$ $m(a) = c_1 m_1(a) + c_2 m_2(a)$. [Distinctness of m_1 and m_2 is equivalent to the existence of some $b \in L$ such that $m_1(b) = m_2(b)$.] If no such decomposition of m exists, then m is *pure*.

The criterion of purity has the virtue of not referring to simultaneous assignment of truth values to propositions.

We shall have little occasion in this paper to use the concept of *observable* (as it is commonly called, though the usage is misleading, since most instances of the concept are not associated with actual procedures of observation or even with Gedanken experiments). It will suffice to say that the concept can reasonaly be defined in terms of the set of propositions (see Ref. 3, p. 63, and Ref. 5, pp. 97–9).

III. THE HILBERT SPACE FORMULATION OF QUANTUM MECHANICS

One needs to add a single strong axiom H to the axioms P, B, and O of the preceding section in order to obtain the standard formulation of the kinematics of nonrelativistic quantum mechanics (not including, however, some of the principles concerning composite systems, notably symmetrization and antisymmetrization). Moreover, with very modest additional assumptions about temporal evolution the dynamics of quantum mechanics (the time-dependent Schrödinger equation) is also derivable from axioms P, B, O, and H (Ref. 3, pp. 81–3). Axiom H is as follows:

Axiom H: The partially ordered, orthocomplemented set of propositions (now to be called L_H) has the structure of the lattice $\mathcal{L}$ of closed linear subspaces of a separable complex Hilbert space $\mathcal{H}$. Specifically, there is a 1–1 mapping $f: L_H \to \mathcal{L}$ (onto) such that

(i) $f(a') = (f(a))^{\perp}$ (where $\perp$ denotes orthogonal complementation in $\mathcal{H}$), and

(ii) $a \leq b$ iff $f(a) \subseteq f(b)$.

[Note that since $f(a)$ and $f(b)$ both belong to $\mathcal{L}$, the set theoretical inclusion $f(a) \subseteq f(b)$ implies that $f(a)$ is a subspace of $f(b)$.]

Adjoining axiom H to the preceding axioms has important consequences, which can easily be demonstrated, for all the derivative concepts defined in Sec. II. The impossible proposition corresponds to the empty subspace of $\mathcal{H}$: $f(\phi) = 0$. The necessary proposition corresponds to $\mathcal{H}$ itself: $f(1) = \mathcal{H}$, the generalized conjunction[16] $a \wedge_H b$ corresponds to the intersection of the subspaces matched with a and b: $f(a \wedge_H b) = f(a) \cap f(b)$. Moreover, since $\mathcal{L}$ is a σ lattice so is L_H, and $f(\wedge_{i=1H} a_i) = \cap_{i=1} f(a_i)$. The generalized disjunction of a and b corresponds to the subspace spanned by $f(a)$ and $f(b)$, and more generally $f(\vee_{i=1H} a_i)$ equals the subspace spanned by $\cup_{i=1} f(a_i)$. That generalized disjunction does not correspond to set theoretical union is one of the crucial ways in which L_H is nonclassical. One consequence is that distributivity does not generally hold in L_H. Moreover, modularity does not hold if $\mathcal{H}$ is infinite dimensional (Ref. 5, p. 85). The following principles hold, whether $\mathcal{H}$ is finite or infinite dimensional:

Weak modularity: if $x \leq z$, then $x = z \wedge_H (z' \vee_H x)$.

Atomicity: 1. (existence of atoms) For every $x \in L_H$ there exists an atom p such that $p \leq x$. (p is an *atom* iff for all $q \in L$, $q \leq p$ iff $q = \phi$ or $q = p$.)

2. (covering law) If q is an atom, then $a \leq x \leq a \vee_H q \Rightarrow x = a$ or $x = a \vee_H q$. It is easy to see that the atoms in L_H correspond under the mapping f to one-dimensional subspaces, or rays, of $\mathcal{H}$.

The mapping f permits the construction of a set of states on L_H, which will be recognized as the usual pure states of quantum mechanics. Let ψ be a normalized vector in $\mathcal{H}$, and let P_a be the projection operator associated with the subspace $f(a)$ which corresponds to the proposition a. Then we define the measure m_ψ on L_H as follows: for all $a \in L_H$, $m_\psi(a) = (\psi, P_a \psi)$. It is easily checked that m_ψ satisfies Mackey's conditions for being a state. A deep theorem of Gleason[15] asserts that (with an almost trivial exception) all states on L_H are convex combinations of states of this type. Specifically, if L_H satisfies axioms P, B, O, and H, and $\dim(\mathcal{H})$ is greater than two, and if m is a state on L_H, then there exists a denumerable set of normalized vectors $\{\psi_i\} \subseteq \mathcal{H}$ and a set of positive real numbers $\{w_i\}$ such that $\sum w_i = 1$ and $m(a) = \sum w_i m_{\psi_i}(a)$ for all $a \in L_H$.

For future reference it is useful to note that Gleason's theorem holds whether $\mathcal{H}$ is a real or a complex Hilbert space, i.e., whether the scalars are the real or the complex numbers.

It is evident that Axiom H is a very strong assumption about the structure of the propositions of a physical system. Piron has demonstrated, however, that much of the content of Axiom H is contained in the assumptions that the set of propositions is an orthocomplemented σ-lattice (so that axioms P, B, and O hold *a fortiori*) and satisfies the conditions of weak modularity and atomicity. The exact content of Piron's theorem is rather complicated to state, and we refer to the original publication[4] and to the careful statement by Varadarajan.[17] Our concern in the present paper is with the validity of the assumptions of Piron's theorem, and our main result in Sec. V is that the second part of the atomicity condition (the "covering law") does not hold in the lattice L_V of verifiable propositions of the spin-1 system.

As preparation for the discussion of the lattice L_V it will be valuable to write down some details about the lattice L_H for the spin-1 system. We shall restrict our attention to the spin properties of the system, and abstract from properties defined in terms of position and linear momentum. The quantum mechanical formalism facilitates such a restriction of attention, since the complete Hilbert space $\mathcal{H}$ appropriate to the spin-1 system is $\mathcal{H} = L^2(E^3) \otimes \mathcal{H}_3$, which is the tensor product of the Hilbert space $L^2(E^3)$ of (equivalence classes of) square integrable functions on three-dimensional Euclidean space, with the three-dimensional complex Hilbert space $\mathcal{H}_3$. Formally, the spin-1 system can be treated as a composite system, of which one component is spinless and the other lacks the properties defined in terms of position and linear momentum. It is the second component in this formal decomposition which we shall henceforth refer to by the expression "the spin-1 system".

The propositions of L_H, in the case of the spin-1 system, correspond to linear subspaces of $\mathcal{H}_3$. (The condition of closure on these subspaces mentioned in Axiom H is automatically satisfied, because of the finite dimensionality of $\mathcal{H}_3$). The atoms of L_H correspond to one-dimensional linear subspaces, or rays, of $\mathcal{H}_3$. We shall use the notation $\langle \phi \rangle$ to designate the ray spanned by the vector ϕ. A direct physical interpretation can be given to a ray spanned by a vector ψ such that $(\mathbf{s} \cdot \hat{n})\psi = \lambda\psi$, where $\mathbf{s} \cdot \hat{n}$ is the spin operator in the direction $\hat{n}$ ($\hat{n}$ a unit vector in Euclidean three-space), and λ is 1 or 0 or -1, which are the three possible eigenvalues of $\mathbf{s} \cdot \hat{n}$ in units of $\hbar$. The proposition in L_H corresponding to this ray will be designated by $a(\hat{n}, \lambda)$, and the ray itself can be designated, in a manner which

does not single out any of its vectors, as $\langle \hat{n}, \lambda \rangle$,

$$f(a(\hat{n}, \lambda)) = \langle \hat{n}, \lambda \rangle = \langle \psi \rangle.$$

The intuitive meaning of $a(\hat{n}, \lambda)$ is the proposition that the spin of the system in the direction $\hat{n}$ is λ (in units of $\hbar$). There are also some propositions in L_H corresponding to two-dimensional subspaces of H_3 which have clear intuitive content: namely $(a(\hat{n}, \lambda))'$, hereafter designated as $b(\hat{n}, \lambda)$, which corresponds under the mapping f to $\langle \hat{n}, \lambda \rangle^{\perp}$. The content of $b(\hat{n}, \lambda)$ is that the spin of the systen in the direction $\hat{n}$ is (in units of $\hbar$) unequal to λ.

It is of the greatest importance for our subsequent discussion to notice that not every ray of H_3 is spanned by an eigenvector of $\mathbf{s} \cdot \hat{n}$ for any direction $\hat{n}$. (In the case of the spin-$\frac{1}{2}$ system, by contrast, every vector is an eigenvector of $\mathbf{s} \cdot \hat{n}$ for some $\hat{n}$—a fact which makes the spin-$\frac{1}{2}$ system unsuitable for the purposes of this paper.) Our assertion is proved by considering the effect of the standard rotation matrix

$$D^2(\alpha, \beta, 0) = \begin{pmatrix} \frac{1}{2}(1+\cos\beta)e^{-i\alpha} & -(1/\sqrt{2})\sin\beta\, e^{-i\alpha} & \frac{1}{2}(1-\cos\beta)e^{-i\alpha} \\ (1/\sqrt{2})\sin\beta & \cos\beta & -(1/\sqrt{2})\sin\beta \\ \frac{1}{2}(1-\cos\beta)e^{i\alpha} & (1/\sqrt{2})\sin\beta\, e^{i\alpha} & \frac{1}{2}(1+\cos\beta)\, e^{i\alpha} \end{pmatrix}$$

upon the column vectors

$$\psi(\hat{z}, 1) = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \psi(\hat{z}, 0) = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad \psi(\hat{z}, -1) = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

(α, β, and 0 are the Euler angles of a rotation.) It is easily seen that if $\hat{n}$ is obtained from $\hat{z}$ by the rotation $(\alpha, \beta, 0)$, then

$$\psi(\hat{n}, \lambda) \equiv D^1(\alpha, \beta, 0)\psi(\hat{z}, \lambda)$$

is an eigenvector of $\mathbf{s} \cdot \hat{n}$, and because of nondegeneracy any eigenvector of $\mathbf{s} \cdot \hat{n}$ is a scalar multiple of $\psi(\hat{n}, \lambda)$. [In particular, for any γ, $D^1(\alpha, \beta, \gamma)\psi(\hat{z}, \lambda)$ is a scalar multiple of $\psi(\hat{n}, \lambda)$, and therefore there is no loss of generality in taking the third Euler angle to be 0.] The vectors $\psi(\hat{n}, 1)$, $\psi(\hat{n}, 0)$, and $\psi(\hat{n}, -1)$ are respectively the first, second, and third columns of the matrix $D^1(\alpha, \beta, 0)$. It is evident upon inspection that there are vectors which are not scalar multiples of any of these columns, for any $\hat{n}$, e.g.,

$$\varphi = \begin{pmatrix} 1/\sqrt{6} \\ 1/\sqrt{3} \\ 1/\sqrt{2} \end{pmatrix}.$$

Nevertheless, there is a proposition in L_H corresponding to φ, namely $f^{-1}(\langle \varphi \rangle)$, and from the standpoint of Axiom H there is no difference in physical status between this proposition and $a(\hat{n}, \lambda)$.

IV. THE VERIFIABLE PROPOSITIONS OF THE SPIN-1 SYSTEM

Since spin is both measurable and theoretically well-understood, and since each spin component of the (massive) spin-1 system is known to have a spectrum consisting of the three points $\hbar$, 0, and $-\hbar$, it is clear that the propositions concerning this system of the form "the spin in direction $\hat{n}$ is λ ($\lambda = 1, 0$, or -1 in units of $\hbar$)" are physically significant.

A strict operationalist might contend that the evidence for the quantization of spin does not suffice to endow every proposition $a(\hat{n}, \lambda)$ with physical significance. The usual method of determining the value of a component of spin is that of Stern and Gerlach, in which a beam of particles passes through an inhomogeneous magnetic field perpendicular to the propagation direction. Consequently, the operationalist might object that once the direction of beam propagation is fixed, the only physically significant $a(\hat{n}, \lambda)$ are those such that $\hat{n}$ is in the direction of a field $\mathbf{H}$ capable of splitting the beam—a direction which cannot in practice be far from perpendicular to the propagation direction even if the Stern—Gerlach procedure is generalized, and in principle cannot be along the propagation direction.

We can give two answers to this objection. First, as discussed in Sec. II, a strictly operationalist understanding of the propositions concerning a physical system is implausible on several grounds. Consequently, if for some $\hat{n}$ propositions of the form $a(\hat{n}, 1)$, $a(\hat{n}, 0)$, and $a(\hat{n}, -1)$ are physically significant, as indicated by the practicality of Stern—Gerlach experiments, and if space is isotropic, as a great variety of considerations indicate, then there is no reason to deny physical significance to any $a(\hat{n}, \lambda)$. Second, we can actually go a long way toward satisfying the demands of the operationalist concerning the entire set $a(\hat{n}, \lambda)$, though at the price of imposing some experimental complications. If ions rather than neutral atoms are used as the particles for the experiment, the direction of the beam may be adiabatically changed (slowly enough to make spin flips improbable),[18] so that the final beam direction is perpendicular to any preassigned axis $\hat{n}$ along which one desires to measure the spin. In this way, even when the initial beam direction is specified, there is no proposition $a(\hat{n}, \lambda)$ which could not be tested by a Stern—Gerlach measurement. More precisely, a filter appropriate for testing $a(\hat{n}, \lambda)$ is prepared by blocking the two output channels of the Stern—Gerlach apparatus corresponding to spin values unequal to λ, so that the particle will pass through the filter only if its spin component in the direction $\hat{n}$ has the value λ (in units of $\hbar$). (See Fig. 1.)

By opening two channels and blocking one on a suitably oriented Stern—Gerlach apparatus, one prepares a filter appropriate for testing $b(\hat{n}, \lambda)$. (See Fig. 2.) Filters for the impossible proposition ϕ and the necessary proposition 1 are prepared, respectively, by closing all channels and by opening all channels of an arbitrarily oriented Stern—Gerlach apparatus (the operational con-

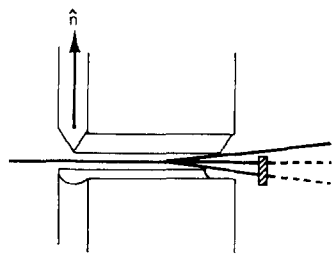


FIG. 1. Filter for the proposition $a(\hat{n}, 1)$.

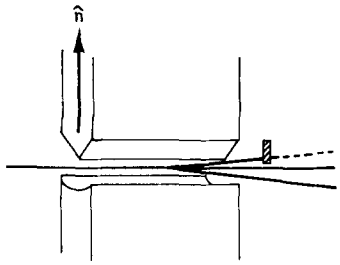


FIG. 2. Filter for the proposition $b(\hat{n}, 1)$.

straint, of course, being that the beam has been guided so as to enter the forward aperture of the filter). As discussed in Sec. II, the operation of orthocomplementation can be understood in a natural manner by reference to filters, even if one does not adopt an operationalist point of view. Specifically, if u is a verifiable proposition and $\hat{u}$ is a Stern–Gerlach apparatus adapted by closing some channels and keeping others open so as to be a filter for u, then u' is the proposition corresponding to the $\hat{u}'$ filter which is obtained by closing the open channels of $\hat{u}$ and opening its closed channels. Obviously, $b(\hat{n}, \lambda) = (a(\hat{n}, \lambda))'$ and $\phi' = 1$. The relation of implication between verifiable propositions can be similarly understood, following the general discussion of Sec. II. We shall designate by L_V the set of verifiable propositions, together with the operation of orthocomplementation and the relation of implication which have been indicated. The structure of L_V will be fully stated in Sec. V.

We cannot dismiss *a priori* the possibility that there are other physically significant propositions concerning the spin-1 system than the members of L_V, and it is even possible that reasonable experimental procedures can be devised for testing these propositions, in which case our usage of the term "verifiable" would be too narrow. Proposals for classifying a set of propositions larger than L_V as physically significant will have to be examined on their merits when they are set forth. There have been, apparently, very few such proposals for systems of any kind. In this section we shall discuss one due to Jauch, and two other proposals will be considered in Appendices B and C.

Jauch (Ref. 5, p. 75) essentially proposes the following method for extending any set S_0 of propositions with unequivocal physical significance to a larger set S which derivately acquires physical significance. Let $\hat{S}_0$ be the set of (equivalence classes of) filters corresponding to members of S_0. Then a composite device can be constructed by connecting in series finitely or denumerably many filters (with replicas allowed) from $\hat{S}_0$. Let $\hat{S}$ be the set of (equivalence classes of) devices so constructible and satisfying sufficient conditions to function as filters. Passage or nonpassage through a filter in $\hat{S}$ is a bivalent physical test in an idealized sense. Hence $\hat{S}$ determines a set of propositions S which may be considered to be physically significant.

If Jauch's procedure is legitimate, then it seems possible to extend L_V. Appendix A shows that for all atoms $a \in (L_H - L_V)$ there exist projection operators A_1 and A_2 on H_3 which correspond to propositions a_1 and a_2 in L_V, such that A corresponding to a is the limit (in the sense of uniform convergence) of the sequence $A_1A_2A_1A_2A_1\cdots$. If $\hat{a}_1$ and $\hat{a}_2$ are filters associated with a_1 and a_2, then the composite device $\hat{a}_1\hat{a}_2\hat{a}_1\hat{a}_2\hat{a}_1\cdots$ seems to be the kind of admissible filter envisaged in Jauch's proposal, and it is reasonably associated with the proposition a.[19]

Clearly, Jauch's proposal admits highly nonoperational procedures as tests for propositions. We do not condemn it on this account, since we have previously expressed skepticism about operationalist programs. What we do find disturbing about his proposal is the absence of clearly articulated and adequately justified rules governing the admission of filters with infinitely many components. Even more troublesome is the fact that in the case of systems other than the spin-1 system there are very plausible constructions of filters with a denumerable infinity of components which generate propositions not contained in the Hilbert space lattice, as we shall now show.

To see the difficulty, consider a spinless particle restricted to one dimension, and therefore quantum mechanically describable in the Hilbert space of square-integrable functions of a single variable $L^2(E)$. Let $[c_i, d_i]$ be a nested sequence of intervals on the real line ($c_i \leq c_{i+1}$, $\lim c_i = e$, $d_i \geq d_{i+1}$, $\lim d_i = e$) the intersection of which consists of exactly the one point e. Let s_i be the proposition that the particle is located in the ith interval, and let S_i be the corresponding projection operator on $L^2(E)$. There is no objection in principle to imagining a filter $\hat{s}_i$ corresponding to each s_i. Now consider the filter with infinitely many components $s = \hat{s}_1\hat{s}_2\hat{s}_3\cdots$. (Note that there is no need to resort to a complicated interleafing of the $\hat{s}_i$ in order to obtain a filter, as in Ref. 19, since the propositions s_i are compatible with one another.) Quantum mechanics predicts that any particle which passes through a device consisting of $\hat{s}_i, \ldots, \hat{s}_n$ in series will pass through any replica of one of these filters. It is reasonable then to extrapolate to the idealized filter $\hat{s}$ and assert $\hat{s} \leq \hat{s}_i$, for $i = 1, 2, \cdots$. The proposition s tested by $\hat{s}$ is intuitively the conjunction of the s_i, and its content is that the system is located at the point e. By contrast, the g.l.b. of the projection operators S_i on H is the null projection operator, which corresponds to the impossible proposition rather than to the proposition s. There is, of course, no projection operator on H corresponding to a proposition about the point location of particle. Thus, if Jauch's admission of filters with infinitely many components is intended to enrich the set of operationally defined propositions, it does so too well, for it appears to lead to propositions which are not represented in the Hilbert space formulation of quantum mechanics.

We do not wish to draw the conclusion that nothing can be salvaged from Jauch's ingenious proposal. It is possible that a reasonable set of operations upon physically unimpeachable filters may indeed generate a set of ideal filters with precisely the structure supposed by the Hilbert space formulation of quantum mechanics. We only contend that this program has not been achieved and is not likely to be easy. In the absence of satisfactory methods for treating filters with infinitely many

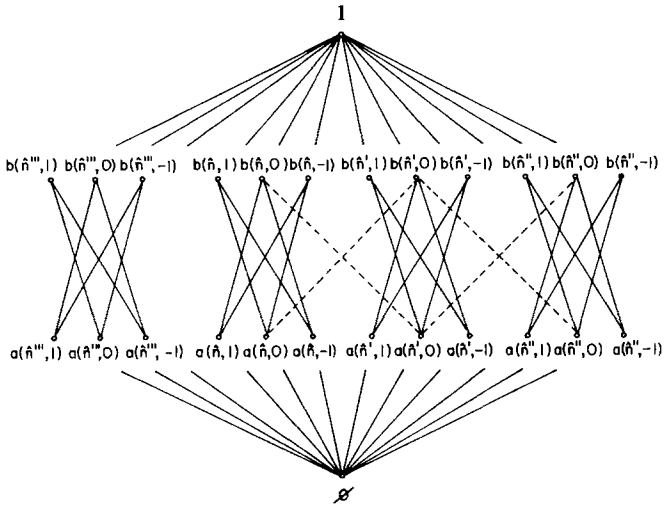


FIG. 3. A representative sample of propositions of L_V and their relations. A line (solid or dotted) indicates that the proposition represented by the lower dot implies the proposition represented by the higher dot. The absence of a line between two dots indicates that no implication relation holds between the corresponding propositions. The direction $\hat{n}'$ is orthogonal both to $\hat{n}$ and to $\hat{n}''$, but $\hat{n}$ is not parallel, antiparallel, or orthogonal to $\hat{n}''$; and $\hat{n}'''$ is not parallel or antiparallel ororthogonal to any of the other directions.

components, we can continue to explore the consequences of the working hypothesis of this paper: That those propositions of L_H which belong to L_V have a physical significance which other propositions in L_H do not possess.

V. THE STRUCTURE OF L_V

Since the orthocomplement of each verifiable proposition is immediately determined, the structure of L_V is completely known when one knows which propositions are related by the implication relation. The only interesting instances of the relation are those involving $a(\hat{n},\lambda)$ for some $\hat{n},\lambda$, and $b(\hat{n}',\lambda')$ for some $\hat{n}',\lambda'$, because it is evident that

$$\emptyset \leq x \quad \text{and} \quad x \leq 1 \quad \text{for all } x \in L_V, \tag{1}$$

$$a(\hat{n},\lambda) \leq a(\hat{n}',\lambda') \quad \text{iff } \hat{n}=\hat{n}' \text{ and } \lambda=\lambda', \text{ or } \hat{n}=-\hat{n}' \text{ and } \lambda=-\lambda', \tag{2}$$

$$b(\hat{n},\lambda) \leq b(\hat{n}',\lambda') \quad \text{iff } \hat{n}=\hat{n}' \text{ and } \lambda=\lambda', \text{ or } \hat{n}=-\hat{n}' \text{ and } \lambda=-\lambda'. \tag{3}$$

The implication $a(\hat{n},\lambda) \leq b(\hat{n}',\lambda')$ holds iff $\langle\hat{n},\lambda\rangle \subseteq \langle n',\lambda'\rangle^{\perp}$, or equivalently, iff the rays $\langle\hat{n},\lambda\rangle$ and $\langle\hat{n}',\lambda'\rangle$ are mutually orthogonal. The complete set of possibilities for obtaining mutual orthogonality can be found by first taking $\hat{n}'$ to be $\hat{z}$ and afterwards performing the appropriate rotations to achieve generality. Using the columns of $\hat{D}(\alpha,\beta,0)$ as the expressions for $\psi(\hat{n},1)$, $\psi(\hat{n},0)$, and $\psi(\hat{n},-1)$ respectively, we see that $\langle\hat{z},0\rangle$ is orthogonal to $\langle\hat{n},\lambda\rangle$ iff $\lambda=0$ and $\cos\beta=0$ (i.e., $\hat{n}$ is perpendicular to $\hat{z}$), and that $\langle\hat{z},\pm1\rangle$ is orthogonal to $\langle\hat{n},\lambda\rangle$ if $\lambda=\mp1$ and $\beta=0$ (i.e., $\hat{n}=z$) or $\lambda=\pm1$ and $\beta=\pi$ (i.e., $\hat{n}=-\hat{z}$). Generalizing, we obtain the following: $a(\hat{n},\lambda) \leq b(\hat{n}',\lambda')$ iff either

$$\hat{n}=\hat{n}' \quad \text{and} \quad \lambda\neq\lambda' \tag{4a}$$

$$\hat{n}=-\hat{n}' \quad \text{and} \quad \lambda\neq-\lambda', \tag{4b}$$

or

$$\hat{n}\perp\hat{n}' \quad \text{and} \quad \lambda=\lambda'=0. \tag{4c}$$

All instances of implication in L_V are comprised in Eqs. (1)—(4). We illustrate the implications in L_V in Fig. 3, using solid lines to indicate implications of types (1)—(3), (4a), and (4b), and dotted lines to indicate implications of type (4c).

It is evident that axioms P, B, and O hold of L_V. Furthermore, an examination of Fig. 3 makes it clear that any two propositions in L_V have a greatest lower bound, so that L_V is an orthocomplemented lattice. (Indeed, L_V is a complete lattice, but this is a fact that we shall not use.) Since $x\wedge_V y$ and $x\vee_V y$ exist for all $x,y\in L_V$, it will be very useful to compile an essentially exhaustive list of them (Tables I and II), and this can be done almost immediately from (1), (2), (3), and (4) or from inspection of Fig. 3. The most interesting entries in Tables I and II are those such that $x\wedge_V y \neq x\wedge_H y$ or $x\vee_V y\neq x\vee_H y$. In several cases $b(\hat{n},\lambda)\wedge_V b(\hat{n}',\lambda')=\emptyset$, whereas $b(\hat{n},\lambda)\wedge_H b(\hat{n}',\lambda')=f^{-1}(\langle\hat{n},\lambda\rangle^{\perp}\cap\langle\hat{n}',\lambda'\rangle^{\perp})\neq\emptyset$, since the intersection of two two-dimensional subspaces of a three-dimensional Hilbert space is a subspace of dimension at least one (a ray), but this ray may not be spanned by an eigenvector of $\mathbf{s}\cdot\hat{n}$ for any $\hat{n}$ and hence may not correspond to a verifiable proposition. If so, then the g.l.b. of $b(\hat{n},\lambda)$ and $b(\hat{n}',\lambda')$ in L_V is $\emptyset$. Likewise, there are several cases

TABLE I. Generalized conjunctions.

$\hat{n}'=\hat{n}$

$a(\hat{n},\lambda)\wedge_V a(\hat{n},\lambda')=a(\hat{n},\lambda)$ if $\lambda'=\lambda$,
$=\emptyset$ if $\lambda'\neq\lambda$.

$a(\hat{n},\lambda)\wedge_V b(\hat{n},\lambda')=a(\hat{n},\lambda)$ if $\lambda'\neq\lambda$,
$=\emptyset$ if $\lambda'=\lambda$.

$b(\hat{n},\lambda)\wedge_V b(\hat{n},\lambda')=b(\hat{n},\lambda)$ if $\lambda'=\lambda$,
$=a(\hat{n},\lambda'')$ if $\lambda'\neq\lambda$, where $\lambda''\neq\lambda$, $\lambda''\neq\lambda'$.

$\hat{n}'=-\hat{n}$

$a(\hat{n},\lambda)\wedge_V a(-\hat{n},\lambda')=a(\hat{n},\lambda)$ if $\lambda'=-\lambda$
$=\emptyset$ if $\lambda'\neq-\lambda$.

$a(\hat{n},\lambda)\wedge_V b(-\hat{n},\lambda')=a(\hat{n},\lambda)$ if $\lambda'\neq-\lambda$
$=\emptyset$ if $\lambda'=-\lambda$.

$b(\hat{n},\lambda)\wedge_V b(-\hat{n},\lambda')=b(\hat{n},\lambda)$ if $\lambda'=-\lambda$
$=a(\hat{n}'',\lambda'')$ if $\lambda'\neq-\lambda$ where $\lambda''\neq\lambda$, $\lambda''\neq-\lambda$.

$\hat{n}'\perp\hat{n}$

$a(\hat{n},\lambda)\wedge_V a(\hat{n},\lambda')=\emptyset$.

$a(\hat{n},\lambda)\wedge_V b(\hat{n}',\lambda')=a(\hat{n},\lambda)$ if $\lambda=0$ and $\lambda'=0$,
$=\emptyset$ if $\lambda\neq0$ and $\lambda'\neq0$.

$b(\hat{n},\lambda)\wedge_V b(\hat{n},\lambda')=a(\hat{n}'',0)$ if $\lambda'=\lambda=0$, where $n'\perp n''\perp n$
$=\emptyset$ if neither λ' nor $\lambda=0$.

If $\hat{n}$ and $\hat{n}'$ are neither parallel, antiparallel, nor perpendicular:
$b(\hat{n},0)\wedge_V b(\hat{n}',0)=a(\hat{n}'',0)$, where $n\perp n''\perp n'$.

$a(\hat{n},\lambda)\wedge_V b(\hat{n}',\lambda')=\emptyset$, $a(\hat{n},\lambda)\wedge_V a(\hat{n}',\lambda')=\emptyset$.

$b(\hat{n},\lambda)\wedge_V b(\hat{n}',\lambda')=\psi$ if $\lambda'\neq0$ or $\lambda\neq0$.

TABLE II. Generalized disjunctions.

$\hat{n}' = \hat{n}$	
$b(\hat{n},\lambda)\vee_V b(\hat{n},\lambda') = b(\hat{n},\lambda)$	if $\lambda' = \lambda$,
$= 1$,	if $\lambda' \neq \lambda$.
$b(\hat{n},\lambda)\vee_V a(\hat{n},\lambda') = b(\hat{n},\lambda)$,	if $\lambda' \neq \lambda$,
$= 1$,	if $\lambda' = \lambda$.
$a(\hat{n},\lambda)\vee_V a(\hat{n},\lambda') = a(\hat{n},\lambda)$,	if $\lambda' = \lambda$,
$= b(\hat{n},\lambda'')$,	if $\lambda' \neq \lambda$, where $\lambda'' \neq \lambda$, $\lambda'' \neq \lambda'$.
$\hat{n}' = -\hat{n}$	
$b(\hat{n},\lambda)\vee_V b(-\hat{n},\lambda') = b(\hat{n},\lambda)$,	if $\lambda' = -\lambda$,
$= 1$,	if $\lambda' \neq -\lambda$.
$b(\hat{n},\lambda)\vee_V a(-\hat{n},\lambda') = b(n,\lambda)$,	if $\lambda' \neq -\lambda$,
$= 1$,	if $\lambda' = -\lambda$.
$a(\hat{n},\lambda)\vee_V a(-\hat{n},\lambda') = a(n,\lambda)$,	if $\lambda' = -\lambda$,
$= b(n,\lambda'')$,	if $\lambda' \neq -\lambda$, where $\lambda'' \neq \lambda$, $\lambda'' \neq -\lambda'$
$\hat{n}' \perp \hat{n}$	
$b(\hat{n},\lambda)\vee_V b(\hat{n}',\lambda') = 1$.	
$b(\hat{n},\lambda)\vee_V a(\hat{n}',\lambda') = b(\hat{n},\lambda)$,	if $\lambda' = \lambda = 0$,
$= 1$,	otherwise.
$a(\hat{n},\lambda)\vee_V a(n',\lambda') = b(\hat{n}'',0)$,	if $\lambda = \lambda' = 0$, where $\hat{n} \perp \hat{n}'' \perp \hat{n}'$,
$= 1$,	otherwise.
If $\hat{n}$ and n' are neither parallel, antiparallel, nor perpendicular:	
$a(\hat{n},\lambda)\vee_V a(\hat{n}',\lambda') = 1$,	if $\lambda' \neq 0$ or $\lambda \neq 0$.
$a(\hat{n},0)\vee_V a(\hat{n}',0) = b(\hat{n}'',0)$,	where $\hat{n} \perp \hat{n}'' \perp \hat{n}'$.
$a(\hat{n},\lambda)\vee_V b(\hat{n}',\lambda') = 1$,	$b(\hat{n},\lambda)\vee_V b(\hat{n}',\lambda') = 1$.

in which $a(\hat{n},\lambda)\vee_V a(\hat{n}',\lambda') = 1$, whereas $a(\hat{n},\lambda)\vee_H a(\hat{n}',\lambda')$ can never be 1 since the subspace spanned by two rays has dimension of at most two. But again this subspace may not correspond to a verifiable proposition, and if so, the l.u.b. of $a(\hat{n},\lambda)$ and $a(\hat{n}',\lambda')$ in L_V is 1.

VI. THE VALIDITY OF VARIOUS AXIOMS IN L_V

In this section we use L_V as an instrument for examining several crucial axioms proposed in programs which aim at establishing the Hilbert space formulation on a firm basis. We find that all of Piron's axioms are satisfied by L_V except the second part of the atomicity axiom (the "covering law"). We do not examine other axiomatizations in detail, but we do show that two of Zierler's axioms and one axiom of MacLaren fail to hold in L_V. Also we find that modularity (which is assumed in the pioneering work of Birkhoff and von Neumann, though not in recent axiomizations) does not hold in L_V.

That L_V is an orthocomplemented σ lattice has already been pointed out in Sec. V. Consequently, the only axioms of Piron which must be checked for L_V are weak modularity and atomicity.

L_V is weakly modular if for all $x, z \in L_V$ such that $x \leq z$, $x = z \wedge_V (z' \vee_V x)$. Evidently, the equation holds if either x or z is ϕ or 1. If both x and z are of the form $a(\hat{n},\lambda)$, or both x and z are of the form $b(\hat{n},\lambda)$, then $x = z$, and again the equation holds. Hence, the only interesting cases are those in which $x = a(\hat{n},\lambda)$ and $z = b(\hat{n},\lambda')$. But by Sec. V, $a(\hat{n},\lambda) \leq b(\hat{n}',\lambda')$ iff (i) $\hat{n} = \hat{n}'$ and $\lambda \neq \lambda'$, or (ii) $\hat{n} = -\hat{n}'$ and $\lambda \neq -\lambda'$, or (iii) $\hat{n} \perp \hat{n}'$ and $\lambda = \lambda' = 0$. Using Tables I and II one sees that the equation holds in each of these cases.

The first part of Piron's axiom of atomicity asserts that for every $x \in L_V$ except ϕ there exists an atom p such that $p \leq x$. The validity of this assertion is evident from Eqs. (1)—(4) of Sec. V or from Fig. 3.

If L_V satisfies the second part of the axiom of atomicity (the "covering law"), then for every atom q of L_V the following holds: If $x \leq y \leq x \vee_V q$, then either $y = x$ or $y = x \vee_V q$. A counterexample is provided by choosing $q = a(\hat{n}, 1)$, $x = a(\hat{n}', -1)$, $y = b(\hat{n}', 1)$, with $\hat{n} \neq \pm\hat{n}'$. Then by Table II $x \vee_V q = 1$, so that y equals neither x nor $x \vee_V q$. The breakdown of the covering law in this case is clearly due to the fact that the l.u.b. of x and q in L_H is a proposition not associated with spin in any direction, and therefore their l.u.b. in L_V does not equal the l.u.b. in L_H.

Two axioms of Zierler[8] which we shall now consider use the concept of a *finite* proposition, i.e., a proposition which is a generalized disjunction of a finite number of atoms. The *dimension* of a proposition x is the l.u.b. of the number of implications in chains (consisting of distinct elements) of the form $\phi \leq x_1 \leq x_2 \leq \cdots \leq x$. It is clear from Sec. V that every proposition in L_V is finite, the dimension of $a(\hat{n},\lambda)$ being 1, of $b(\hat{n},\lambda)$ being 2, and of the necessary proposition being 3. The two axioms in question are the following:

1. If b, c, and d are elements of the sublattice of elements $\leq a$ (where a is finite), and if $d \leq c$ and $b \wedge c = \phi$, then $(d \vee b) \wedge c = d$.

2. If a and b are finite elements of the same dimension, then the sublattices L_a and L_b, which consist respectively of elements $\leq a$ and elements $\leq b$, are isomorphic.

That the first of these fails to hold in L_V is shown by taking the operations $\vee$ and $\wedge$ to be $\vee_V$ and $\wedge_V$, and letting $d = a(\hat{n},\lambda)$, $b = a(\hat{n}',\lambda')$, $c = b(\hat{n},\lambda'')$, where $\hat{n} \neq \pm\hat{n}$, $\hat{n}$ is not orthogonal to $\hat{n}'$, $\lambda \neq 0$, and $\lambda'' \neq \lambda$. Then $(d \vee_V b) = 1$, so that $(d \vee_V b) \wedge_V c = c$. That the second fails can be seen by taking a and b to be $b(\hat{n}, 0)$ and $b(\hat{n}, 1)$ respectively and looking at Fig. 3.

MacLaren[7] proposes the *semimodularity* axion, which can be stated as follows:

If x and y are propositions such that for all z with $x \leq z$, $(x \vee y) \wedge z = x \vee (y \wedge z)$, then for all z such that $y \leq z$, $(y \vee z) \wedge z = y \vee (x \wedge z)$.

A counterexample in L_V is provided by taking $x = b(\hat{n},\lambda)$ and $y = a(\hat{n}'',\lambda')$, where $\hat{n} \neq \pm\hat{n}''$ and $\hat{n}$ is not orthogonal to $\hat{n}''$. Then $x \vee_V y = 1$, so that $(x \vee_V y) \wedge_V z = z$. The only values of z such that $x \leq z$ are $b(\hat{n},\lambda)$ and 1, and for each of these alternatives $x \vee_V (y \wedge_V z)$ is seen to be z. Thus, the antecedent of the semimodularity axiom is satisfied for this choice of x and y. However, the consequent is not satisfied, since $y \leq z$ holds when z is chosen to be $b(\hat{n}'',\lambda'')$ where $\lambda'' \neq \lambda'$. Then $(y \vee_V x) \wedge_V z = 1 \wedge_V z = b(\hat{n}'',\lambda'')$, while $y \vee_V (x \wedge_V z) = a(\hat{n}'',\lambda') \vee_V \phi = a(\hat{n}'',\lambda')$.

A lattice is *modular* if $x \vee (y \wedge z) = (x \vee y) \wedge z$ whenever $x \leq z$. That L_V is not modular is seen by taking x

$= a(\hat{n}, 1)$, $y = a(\hat{n}', 1)$, $z = b(\hat{n}, 0)$, with $\hat{n} \neq \pm \hat{n}'$. As seen from Fig. 1, $a(\hat{n}, 1) \leq b(\hat{n}, 0)$. From Tables I and II we have

$$a(\hat{n}, 1) \vee_V [a(\hat{n}', 1) \wedge_V b(\hat{n}, 0)] = a(\hat{n}, 1) \vee_V \phi = a(\hat{n}, 1),$$

whereas

$$[a(\hat{n}, 1) \vee_V a(\hat{n}', 1)] \wedge_V b(\hat{n}, 0) = 1 \wedge_V b(\hat{n}, 0) = b(\hat{n}, 0).$$

The failure of modularity for L_V is an interesting curiosity, since L_V can be embedded in L_H, which is modular. (Modularity does not hold for the lattice of closed linear subspaces of an infinite dimensional Hilbert space, but it does hold if the Hilbert space is finite dimensional. See Ref. 5, p. 85.)

VII. STATES ON L_V

It is well known that if H has dimension greater than or equal to three, then there are no dispersion-free states on L_H, or equivalently, no states m such that $m(a)$ is either 0 or 1 for each $a \in L_H$ (Gleason,[15] Kochen and Specker,[10] Bell,[20] Belinfante[11]). This mathematical fact precludes interpreting the usual quantum mechanical states as probability distributions over a space of classically pure states, for these latter would have to be dispersion-free. However, someone might conjecture that L_V—which is the "physically significant" part of L_H—does admit dispersion-free states, and that L_H does not admit them just because it is laden with nonphysical elements. The conjecture is false, and its falsity is philosophically significant, for it shows that one of the most important nonclassical features of quantum mechanical states cannot be blamed upon the admission of nonphysical elements into the lattice of propositions.

The nonexistence of dispersion-free states on L_V can be read immediately from the arguments given by Kochen and Specker, and Belinfante concerning L_H. In this section we shall present a new proof of the nonexistence of dispersion-free states on L_V, by taking Gleason's theorem as applied to a real Hilbert space as our starting point. An advantage of this procedure is that it enables us to exhibit all the possible states on L_V.

Consider the orthocomplemented sublattice L_0 of L_V generated by all the propositions of the form $a(\hat{n}, 0)$ for arbitrary $\hat{n}$, i.e., closed under the operations of orthocomplementation and g.l.b. Tables I and II show that L_0 consists only of $\{a(\hat{n}, 0)\}$, $\{b(\hat{n}, 0)\}$, ϕ, and 1. We shall now show that L_0 is isomorphic to the lattice of subspaces of a real three-dimensional Hilbert space. We do this by constructing a concrete realization $H(R)$ of the space, consisting of all real multiples of the column vectors

$$\psi(\hat{n}, 0) = \begin{bmatrix} -2^{-1/2} \sin\beta \exp(-i\alpha) \\ \cos\beta \\ 2^{-1/2} \sin\beta \exp(i\alpha) \end{bmatrix}.$$

$H(R)$ is closed under real linear combinations, as can be shown by checking that $c\psi(\hat{n}, 0) + d\psi(\hat{n}', 0)$ is a real multiple of $\psi(\hat{n}'', 0)$ for some direction $\hat{n}''$. Because of rotational invariance it suffices to check this fact by choosing the polar axis perpendicular to both $\hat{n}$ and $\hat{n}'$, so that $\beta = \beta' = \pi/2$. Then

$$c\psi(n, 0) + d\psi(n', 0) = \begin{bmatrix} -2^{-1/2}[c \exp(-i\alpha) + d \exp(-i\alpha')] \\ 0 \\ 2^{-1/2}[c \exp(i\alpha) + d \exp(i\alpha')] \end{bmatrix}$$

and since the third row in this column vector is the complex conjugate of the negative of the first row

$$c\psi(\hat{n}, 0) + d\psi(\hat{n}', 0) = k \begin{bmatrix} -2^{-1/2} \exp(-i\alpha'') \\ 0 \\ 2^{1/2} \exp(i\alpha'') \end{bmatrix}$$

for some real k and α''.

If the inner product of two vectors of $H(R)$ is taken to be the matrix product of the Hermitian adjoint of the first with the second, then it can be checked that the conditions for an inner product of a real Hilbert space are satisfied. It is clear that $H(R)$ is three-dimensional. The correspondence between the atomic propositions $\{a(\hat{n}, 0)\}$ of L_0 and the rays of $H(R)$ is now obvious, and it then follows that the lattice L_0 is isomorphic to the lattice of subspaces of $H(R)$. If a dispersion-free state existed on L_V, it would automatically define a dispersion-free state on the sublattice L_0, and by the isomorphism just exhibited one would also be defined on the lattice of subspaces of a three-dimensional real Hilbert space. Since that would be in conflict with Gleason's theorem (see Sec. III), we conclude that there is no dispersion-free state on L_V.

Any state on L_V must be an extension of one of the class $\mathcal{M}_R$ of states which are definable on L_0 in accordance with Gleason's theorem, i.e., $\bar{m} \in \mathcal{M}_R$ if for all $a \in L_0$

$$\bar{m}(a) = \sum w_i (\psi_i, P_a \psi_i),$$

where the ψ_i are normalized vectors in $H(R)$, the w_i are positive real numbers summing to 1, and P_a is the projection operator on $H(R)$ corresponding to the proposition a according to the isomorphism indicated in the preceding paragraph. The following theorem gives an exhaustive compilation of the states on L_V.

Theorem: For m to be a state on L_V it is necessary and sufficient that there exists a state $\bar{m} \in \mathcal{M}_R$ and a nonnegative function $\alpha(\hat{n})$ such that

(i) $m(x) = \bar{m}(x)$ if $x \in L_0$,

(ii) $\alpha(\hat{n}) \leq 1 - m(a(\hat{n}, 0))$,

(iii) $\alpha(-\hat{n}) = 1 - m(a(\hat{n}, 0)) - \alpha(\hat{n})$,

(iv) $m(a(\hat{n}, 1)) = \alpha(\hat{n})$,

(v) $m(a(\hat{n}, -1)) = \alpha(-\hat{n})$,

(vi) $m(b(\hat{n}, \pm 1)) = 1 - m(a(\hat{n}, \pm 1))$.

To prove this theorem note that the only relations among the $a(\hat{n}, \pm 1)$, the $b(\hat{n}, \pm 1)$, and L_0 which impose constraints upon the function $m(x)$ are the following (as can be seen by examining Tables II and III):

$$a(\hat{n}, 1) = a(-\hat{n}, -1),$$

TABLE III. Disjointness relations in L_V.

$\hat{n}'=\hat{n}$
$a(\hat{n},\lambda)\perp a(\hat{n},\lambda')$, iff $\lambda\neq\lambda'$.
$a(\hat{n},\lambda)\perp b(\hat{n}',\lambda')$, iff $\lambda=\lambda'$.
$\hat{n}'=-\hat{n}$
$a(\hat{n},\lambda)\perp a(\hat{n}',\lambda')$, iff $\lambda\neq-\lambda'$.
$a(\hat{n},\lambda)\perp b(\hat{n}',\lambda')$, iff $\lambda=-\lambda'$.
$\hat{n}'\perp\hat{n}$
$a(\hat{n},\lambda)\perp a(\hat{n}',\lambda')$, iff $\lambda=\lambda'=0$.
Otherwise, $a(\hat{n},\lambda)$ is not disjoint to $a(\hat{n},',\lambda)$ or $b(\hat{n}',\lambda)$.

$$b(\hat{n},1)=b(-\hat{n},-1),$$

$$a(\hat{n},\lambda)\perp a(\hat{n},\lambda') \text{ for } \lambda\neq\lambda', \text{ and } a(n,\lambda)\vee_V a(\hat{n},\lambda')=b(\hat{n},\lambda'') \text{ if } \lambda,\ \lambda',\ \lambda'' \text{ are all unequal},$$

$$a(\hat{n},\lambda)\perp b(\hat{n},\lambda), \text{ and } a(\hat{n},\lambda)\vee_V b(\hat{n},\lambda)=1.$$

Hence, by the general conditions on states, if m is a state on L_V, then

$$m(a(\hat{n},1))=m(a(-\hat{n},-1)),$$

$$m(b(\hat{n},1))=m(b(-\hat{n},-1)),$$

$$m(b(\hat{n},\lambda))=m(a(\hat{n},\lambda'))+m(a(\hat{n},\lambda'')) \text{ for } \lambda,\ \lambda',\ \lambda'' \text{ unequal},$$

$$m(a(\hat{n},\lambda))+m(b(\hat{n},\lambda))=m(a(\hat{n},1))+m(a(\hat{n},0))+m(a(\hat{n},-1))=1.$$

These constraints and the previously noted constraint concerning states on L_0, as well as the general conditions on states, are clearly satisfied if conditions (i)—(vi) are fulfilled, thus proving the theorem.

All the quantum mechanical states on L_V (i.e., states of the form $m(x)=\sum w_i(\phi_i,P_x\phi_i)$, where the ϕ_i belong to the complex Hilbert space H_3, and P_x is a projection operator on this space) are included in the compilation given in the foregoing theorem. However, in this compilation there exist some nonquantum mechanical states. For example, let

$$\bar{m}(x)=(\psi,P_x\psi),$$

where $\mathbf{s}\cdot\hat{z}\psi=0$, and let $\alpha(\hat{n})=1-\bar{m}(a(\hat{n},0)$ for $\hat{n}=\hat{x}$ and $\hat{n}=\hat{y}$, but otherwise let $\alpha(\hat{n})$ have any value allowed by (ii) and (iii). Then

$$m(a(\hat{x},0))=m(a(\hat{y},0))=0,$$

and by (iv) of the theorem,

$$m(a(\hat{x},1))=m(a(\hat{y},1))=1.$$

But no pure quantum mechanical states assign the value 1 to two distinct atomic propositions, and hence no convex combination of quantum mechanical states would do so. Therefore m is not a quantum mechanical state, although it is a well-defined state on L_V when $\alpha(\hat{n})$ is fully specified.

VIII. CONCLUSIONS

The concept of a verifiable proposition of the spin-1 system has been examined, and it has been argued that the members of the lattice L_V have a preferred status which other propositions concerning the spin-1 system lack. The structure of L_V has been exhaustively investigated.

The results concerning L_V are not decisive evidence that the Hilbert space formulation of quantum mechanics cannot be fully justified. Nevertheless, the fact that the lattice of verifiable propositions of a specific physical system fails to satisfy some of the crucial axioms proposed in programs which aim at recovering the Hilbert space formulation does constitute a serious challenge to these programs. An obligation is placed upon advocates of such programs to exhibit the physical significance of propositions concerning the spin-1 system which are not elements of L_V, or at least to show that there are great mathematical advantages in extending L_V by admitting "ideal" elements. The possibility remains open that a structure somewhat different from the standard one would be adequate for deriving the well-known physically significant consequences of standard quantum theory, and may even be superior for the purpose of handling hitherto unsolved problems.

The fact that L_V does not admit a dispersion-free state is an interesting contribution to the evidence which has been gathered recently against hidden-variable theories.

APPENDIX A

Let ξ be a vector in H_3, but suppose that $\langle\xi\rangle$ does not belong to L_V. A composite filter $\hat{a}_1\hat{a}_2\hat{a}_1\hat{a}_2\hat{a}_1\cdots$ will be associated with $\langle\xi\rangle$ in accordance with Jauch's proposal if ξ lies in the intersection of the two subspaces $f(b(\hat{n},\lambda))$, $f(b(\hat{n}',\lambda'))$. Then $\hat{a}_1$ and $\hat{a}_2$ can be taken to be the filters associated respectively with the propositions $b(\hat{n},\lambda)$ and $b(\hat{n}',\lambda')$ of L_V. Clearly, then, a necessary and sufficient condition for a composite filter of the indicated kind to be constructible for $\langle\xi\rangle$ is that ξ be orthogonal to two nonparallel vectors $\psi(\hat{n},\lambda)$, $\psi(\hat{n},\lambda')$. We shall now show that for any ξ it is possible to find a vector $\psi(\hat{n},0)$ and a vector $\psi(\hat{n}',1)$ orthogonal to ξ.

We first write

$$\xi=\begin{vmatrix} d_1+ie_1\\ d_2+ie_2\\ d_3+ie_3\end{vmatrix} \text{ and } \sqrt{2}\psi(\hat{n},0)=\begin{vmatrix} -\sin\beta\exp(-i\alpha)\\ \sqrt{2}\cos\beta\\ \sin\beta\exp(i\alpha)\end{vmatrix}=\begin{vmatrix} -x+iy\\ \sqrt{2}z\\ x+iy\end{vmatrix}.$$

If ξ is orthogonal to $\psi(\hat{n},0)$, then the real and imaginary parts of $(\xi,\psi(\hat{n},0))$ must both vanish, imposing two conditions on the real vector (x,y,z), which can always be simultaneously satisfied, and except in degenerate cases essentially only in one way.

In order to investigate orthogonality to $\psi(\hat{n},1)$ we rewrite

$$2\psi(\hat{n},1)=\begin{pmatrix}(1+\cos\beta)\exp(-i\alpha)\\ \sqrt{2}\sin\beta\\ (1-\cos\beta)\exp(i\alpha)\end{pmatrix}$$

and multiply ξ by a scalar to yield a vector of the form

$$\xi' = \begin{pmatrix} c_1 \exp[i(\kappa - \delta)] \\ c_2 \\ c_3 \exp[i(\kappa + \delta)] \end{pmatrix},$$

where each c_i is real and nonnegative. If either c_1 or c_3 is 0, then ξ is orthogonal to either $\begin{pmatrix}1\\0\\0\end{pmatrix}$ or $\begin{pmatrix}0\\0\\1\end{pmatrix}$, and the required orthogonality is achieved. Hence, in the following argument both c_1 and c_3 will be assumed nonzero. $\psi(n, 1)$ is orthogonal to ξ' (and hence to ξ) iff

$$0 = c_1(1 + \cos\beta)\exp[i(\gamma + \kappa)] + c_2\sqrt{2}\sin\beta + c_3(1 - \cos\beta)\exp[i(-\gamma + \kappa)], \tag{A1}$$

where $\gamma = \alpha - \delta$. Take the real and imaginary parts of (A1),

$$0 = c_1(1 + \cos\beta)\cos(\gamma + \kappa) + \sqrt{2}c_2\sin\beta + c_3(1 - \cos\beta)\cos(-\gamma + \kappa), \tag{A2}$$

and

$$0 = c_1(1 + \cos\beta)\sin(\gamma + \kappa) + c_3(1 - \cos\beta)\sin(-\gamma + \kappa). \tag{A3}$$

Hence,

$$x \equiv \cos\beta = \frac{c_1\sin(\gamma + \kappa) + c_3\sin(-\gamma + \kappa)}{-c_1\sin(\gamma + \kappa) + c_3\sin(-\gamma + \kappa)}. \tag{A4}$$

Since c_1 and c_3 are nonzero, a sufficient condition for $-1 \leq \cos\beta \leq 1$ is that $\sin(\gamma + \kappa)$ and $\sin(-\gamma + \kappa)$ have opposite signs, which will be the case if the direction of γ is closer to the y axis than is the direction of κ, or equivalently if $\cos^2\gamma$ is less than $\cos^2\kappa$. We shall now show that except in special cases, which will be investigated separately, the equation resulting from the substitution of Eq. (A4) into Eq. (A2) has a solution γ for which $\cos^2\gamma < \cos^2\kappa$, so that β is a real angle.

From Eqs. (A2) and A4) and the assumptions made so far concerning c_1, c_3, γ, and κ, one obtains by a straightforward calculation

$$0 = 2c_1c_3u^2 - (2c_1c_3 + c_2^2)u + c_2^2w, \tag{A5}$$

where $u = \cos^2\gamma$ and $w = \cos^2\kappa$. Hence

$$u = \tfrac{1}{2} + r \pm \tfrac{1}{2}[1 + 4r(1 - 2w) + 4r^2]^{1/2}, \tag{A6}$$

where $r = c_2^2/4c_1c_3$. Since $0 \leq w \leq 1$, we have

$$r - \tfrac{1}{2} \leq \tfrac{1}{2}[1 + 4r(1 - 2w) + 4r^2]^{1/2} \leq r + \tfrac{1}{2}.$$

Hence, choosing the negative sign in Eq. (A6) we have $0 \leq u \leq 1$, which is a necessary and sufficient condition for γ to be a real angle. In order to check whether $\cos^2\gamma$ is less than $\cos^2\kappa$, we fix w and seek an r which maximizes u. It is easily shown that $du/dr = 0$ implies that w equals either 0 or 1. Hence, except for these two extreme values of w, u is monotonic in r, and therefore to check whether u is less than w we need only look at $r = 0$ and $r \to \infty$. At $r = 0$, Eq. (A6) (with the negative sign) yields $u = 0$, so that u is less than w unless $w = 0$. For $r \to \infty$,

$$u = \tfrac{1}{2} + r - r[1 + (1 - 2w)/r + 1/4r^2]^{1/2} \approx w - (w - w^2)/2r,$$

by keeping only terms to first order in $1/r$. Therefore, if w does not equal 0, u is less than w for large but finite r. (The case of infinite r has been excluded, since we have assumed that c_1 and c_3 are nonzero.) We have established, therefore, that $\cos^2\gamma$ is less than $\cos^2\kappa$ except if w equals 0 or 1, which we now investigate separately.

If w equals 0, then κ is $\pi/2$ or $3\pi/2$. If the former, let γ be $\pi/2$; if the latter, let γ be $-\pi/2$. In either case Eq. (A3) is satisfied and Eq. (A2) becomes

$$0 = [(c_1 - c_3)^2 - 2c_2^2] + 2(c_1^2 - c_3^2)\cos\beta + [(c_1 + c_3)^2 + 2c_2^2]\times\cos^2\beta.$$

With the proper choice of the root we obtain

$$\frac{c_3^2 - c_1^2 + 2c_2^2}{c_1^2 + 2c_1c_3 + c_3^2 + 2c_2^2} \leq \cos\beta \leq \frac{c_3^2 - c_1^2 + 2c_1c_3 + 2c_2^2}{c_1^2 + 2c_1c_3 + c_3^2 + 2c_2^2},$$

so that $-1 \leq \cos\beta \leq 1$. Hence both Eq. (A2) and Eq. (A3) are satisfied by some real angles β and γ.

If w equals 1, we distinguish two cases, $2c_1c_3 \leq c_2^2$ and $2c_1c_3 > c_2^2$. In the former case let γ be 0. Then Eq. (A3) is satisfied, and Eq. (A2) yields (for proper choice of the sign of the radical)

$$\cos\beta = \frac{c_3^2 - c_1^2 + 2[(c_2^2 - c_1c_3)^2 - (c_1c_3)^2]^{1/2}}{c_3^2 - 2c_1c_3 + c_1^2 + 2c_2^2}.$$

Therefore,

$$\frac{c_3^2 - c_1^2}{c_3^2 - 2c_1c_3 + c_1^2 + 2c_2^2} \leq \cos\beta \leq \frac{c_3^2 - c_1^2 + 2(c_2^2 - c_1c_3)}{c_3^2 - 2c_1c_3 + c_1^2 + 2c_2^2},$$

so that $-1 \leq \cos\beta \leq 1$. If, however, $2c_1c_3$ is greater than c_2^2 we return to Eq. (A5) but take the positive sign of the radical. Then

$$u = \tfrac{1}{2} + r + \tfrac{1}{2}(2r - 1) = 2r = 2c_2^2/4c_1c_3 < 1.$$

Hence, $0 \leq \cos^2\gamma < 1 = \cos^2\kappa$, and therefore real angles β and γ can be found so as to satisfy Eqs. (A2) and (A3).

All cases have now been covered, and the proof is complete.

APPENDIX B

One may try to use a suggestion of Stein (Ref. 9, p. 390) for the purpose of finding physically significant propositions outside L_V. Let x be a proposition in L_V and define $x(t)$ as the proposition which is realized at time 0 by virtue of realizing x at time t, the truth value of the former being the same as that of the latter in case of realization. Clearly, the expression "$x(t)$" is elliptical, since its content depends upon the dynamics of the system. We shall restrict our attention to the only physically realistic Hamiltonian of which we are aware for the nonrelativistic spin-1 system, $H = -\mu\mathbf{B}\cdot\mathbf{s}$, where $\mathbf{B}$ is in general a time-dependent magnetic induction. With this choice of the Hamiltonian it can be shown that $x \in L_V$ implies $x(t) \in L_V$, so that no extension of the class of physically significant propositions has been accomplished.

It suffices to give the proof for an infinitesimal time interval Δt (which the advantage that $\mathbf{B}$ can be taken to be effectively constant), since the result for finite t follows by iteration and passing to the limit. If P_x is the projection operator on the Hilbert space corresponding

to x, then the Hilbert space corresponding to x, then the projection operator corresponding to $x(t)$ is

$$P_{x(\Delta t)} = \exp(i\mu \mathbf{B}\cdot\mathbf{s}\Delta t/\hbar) P_x \exp(-i\mu \mathbf{B}\cdot\mathbf{s}\Delta t/\hbar).$$

With no loss of generality $\mathbf{B}$ can be taken in the z direction, so that

$$P_{x(\Delta t)} = \exp(iQs_z) P_x \exp(-iQs_z),$$

where Q is defined as $\mu B\Delta t/h$. If x is $a(\hat{n}, 1)$, with $\hat{n}$ given polar coordinates θ, ϕ, then

$$\exp(iQs_z)\begin{pmatrix} \tfrac{1}{2}(1+\cos\theta)\exp(-i\phi) \\ 2^{-1/2}\sin\theta \\ \tfrac{1}{2}(1-\cos\theta)\exp(i\phi) \end{pmatrix}$$

$$= \begin{pmatrix} \exp(iQ\hbar) & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \exp(-iQ\hbar) \end{pmatrix}\begin{pmatrix} \tfrac{1}{2}(1+\cos\theta)\exp(-i\phi) \\ 2^{-1/2}\sin\theta \\ \tfrac{1}{2}(1-\cos\theta)\exp(i\phi) \end{pmatrix}$$

$$= \begin{pmatrix} \tfrac{1}{2}(1+\cos\theta)\exp[-i(\phi - Q\hbar)] \\ 2^{-1/2}\sin\theta \\ \tfrac{1}{2}(1-\cos\theta)\exp[i(\phi - Q\hbar)] \end{pmatrix}$$

is an eigenvector of $P_{x(\Delta t)}$ and it also has the form of an eigenvector (with eigenvalue 1) of the spin operator in some direction $\hat{n}'$. Consequently, $x(\Delta t)$ equals $a(\hat{n}', 1)$ and hence belongs to L_V. A similar argument holds for $x = a(\hat{n}, -1)$ or $a(\hat{n}, 0)$, and no further argument is needed for $x = b(\hat{n}, \lambda)$.

APPENDIX C

In 1969, Jauch and Piron[6] published a formulation of quantum logic with differs radically from their earlier work. Two features of the new formulation are especially relevant to our argument: (a) They propose a new definition of conjunction, which may permit the extension of L_V to a larger set of physically significant propositions; (b) They claim to prove the covering law.

In order to make judgments on both of these points we must summarize their salient definitions and (since there are some obscurities) make some comments on how they should be construed.

(i) A yes—no experiment α is said to be *true* if a performance of α will necessarily give the result yes.

(ii) A partial ordering $<$ is defined on the class of yes—no experiments as follows: If α true $\Rightarrow \beta$ true, then $\alpha < \beta$.

(iii) The *proposition* $\{\alpha\}$ determined by the yes—no experiment α is $\{\beta \mid \beta < \alpha \text{ and } \alpha < \beta\}$.

(iv) If for each i in an index set I, α_i is a yes—no experiment, then $\Pi\,\alpha_i$ is the experiment which consists of randomly choosing and measuring one of the α_i with the result yes (no) being ascribed to $\Pi\,\alpha_i$ if the result yes (no) is obtained for α_i. If a_i is the proposition given by $a_i = \{\alpha_i\}$, then $\cap\, a_i$ is the proposition $\{\Pi\,\alpha_i\}$. (If $i = 1, 2$, we can write $\cap\, a_i$ as $a_1 \cap a_2$.)

(v) The state of a system is the set S of all true propositions of the system.

Definition (i) is clearly elliptical, and some phrase like "when the system is in state S" ought to be inserted between the words "true" and "if." Without such a qualification the predicate truth would be reserved only for necessary yes—no experiments, which surely is not the authors' intention. Instead of the phrase "when the system is in state S" one might suggest "when the system is prepared in manner X," but we suppose that the difference is probably not very important.

If the elliptical definition (i) is expanded as we have just suggested, there are important consequences for definitions (ii) and (iii): Specifically, whether or not the relation $<$ holds depends not only upon the experiments α and β, but also upon the extension of the set $\mathcal{S}$ of states. The reason is that the phrase "α true $\Rightarrow \beta$ true" is elliptical and means something like "in every state in which α is true β is also true." It follows that the concept of a proposition, defined in (iii), also depends upon the extension of $\mathcal{S}$, since a proposition is defined as an equivalence class of yes—no experiments symmetrically related by $<$. As an example, consider $a = \{\alpha\}$, a suppose that there is no state in which the measurement of α is certain to yield yes. In that case we have $\alpha < \phi$, where ϕ is the impossible experiment, and a then equals the impossible proposition. We see, therefore, that the new quantum logic of Jauch and Piron differs radically from their own earlier formulation, in which the structure of the lattice of propositions was in principle specified without reference to the set of possible states of the system; in fact, as we saw in Sec. II, the set of states was explicitly defined in terms of the lattice of propositions. We do not wish to condemn the new procedure of Jauch and Piron, since it is legitimate and often very fruitful to axiomatize several concepts in tandem. What is unsatisfactory in their procedure is first that they do not seem to recognize that the concept of state is involved in their definition of implication, and second that their characterization of the set of states consists of nothing more than their definition (v), which is insufficient to fix the structure of either the set of states or the set of propositions.

We now inquire whether definition (iv) of $\cap\, a_i$ provides a reasonable extension of the set of physically significant propositions. Let $b(\hat{n}, \lambda)$ and $b(\hat{n}', \lambda')$ be two propositions in L_V such that

$$\phi = b(n, \lambda) \wedge_V b(n', \lambda') \neq b(n, \lambda) \wedge_H b(n', \lambda').$$

What is the content of the proposition $b(\hat{n}, \lambda) \cap b(\hat{n}', \lambda')$, in the sense of definition (iv)? It is the proposition which is true in those and only those states in which $b(\hat{n}, \lambda)$ and $b(\hat{n}', \lambda')$ are both true. Now suppose that the set of states contains the state which is represented in the Hilbert space formulation by the one-dimensional intersection of the two two-dimensional subspaces $f(b(\hat{n}, \lambda))$ and $f(b(\hat{n}, \lambda'))$. (This is a reasonable supposition to ascribe to Jauch and Piron even though, as noted, they say little about the set of states.) Then $b(\hat{n}, \lambda) \cap b(\hat{n}', \lambda')$ is not identical to the impossible proposition.

Prima facie, then, we have obtained a physically significant proposition outside of L_V by performing an operation upon members of L_V. But has anything significant been achieved thereby? The only way that we can check whether $b(\hat{n},\lambda)\cap b(\hat{n}',\lambda')$ is true of a physical system at time t is to equip oneself with an ensemble of systems which are somehow guaranteed to be identical with the system of interest (perhaps because of identical preparation), and then to check $b(\hat{n},\lambda)$ on an arbitrarily large subensemble and $b(\hat{n}',\lambda')$ on another arbitrarily large subensemble. The two subensembles must be disjoint, because of the non-corealizability of $b(\hat{n},\lambda)$ and $b(\hat{n}',\lambda')$. There appears to be no way of checking $b(\hat{n},\lambda)\cap b(\hat{n}',\lambda')$ other than by investigating the features of an ensemble. But the ensemble in question can be characterized perfectly well in terms of the lattice L_V, and therefore only a verbal maneuver seems to have been performed in attributing physical significance to the proposition $b(\hat{n},\lambda)\cap b(\hat{n}',\lambda')$.

We shall not discuss the covering law in detail, because the argument of Jauch and Piron is rather lengthy and has to be made even lengthier in order to achieve rigor.[21] It suffices for our purpose to note that their argument depends crucially upon the following assumption: For any state S of a system and any proposition a, there exists an ideal measurement of the first kind of a. (A measurement of a is of the first kind if an answer "yes" implies that a is true immediately after the measurement. A measurement of a is ideal if and only if any proposition which is both corealizable with a and true prior to the measurement is true immediately after the measurement.) The assumption of Jauch and Piron is hard to judge without much more information about the physically allowable states and the physically possible measurements. It appears to us to be at least as strong an assumption as the covering law itself, and as difficult to judge *a priori*.

One judgment which we can make confidently is that ideal measurements of the first kind cannot be accomplished by measuring propositions which are conjunctions in the sense discussed in the preceding paragraph. Suppose that a particular system of interest is in a state S in which neither $q\equiv b(\hat{n},\lambda)\cap b(\hat{n}',\lambda')$ nor its orthocomplement q' is true. Measurement of q consists, as we have seen, in carrying out experimental tests of $b(\hat{n},\lambda)$ and $b(\hat{n}',\lambda')$ in subensembles selected from an ensemble of systems in state S; in one or both of these subensembles the answers obtained will be a mixture of yes and no. This procedure obviously does not affect the system of interest, and in particular it does not throw that system into a new state in which either q is true or q' is true. In fact, the procedure does not have this effect for *any* of the systems of the original ensemble, whether chosen to be tested in one of the two ways or not. In short, even if the new conception of conjunction of Jauch and Piron is accepted, it does not seem to advance their proposal to establish the covering law via the performability of ideal measurements of the first kind.

We conclude with a brief philosophical comment on the new conception of a state proposed by Jauch and Piron: that a state is the set of all true propositions (in their sense of "true") of the system. Since neither potentiality nor probability enters explicitly into this conception, it is much closer to the concept of state in classical physics than to the usual quantum mechanical conception. Only the circumstance that in standard quantum mechanics there is a one—one correspondence between pure states and atomic propositions—both being represented by rays—prevents outright discrepancy between the two conceptions, but it would be a mistake to rely heavily upon this circumstance, because (as argued in the present paper) the physical significance of many atomic propositions is doubtful. (Also see Stein's argument in Ref. 9, p. 431, that states cannot properly be regarded as subject to yes-or-no tests.) Even if it turns out that quantum mechanics can formally be cast into the new formulation of Jauch and Piron, we believe that their conception of state would nevertheless disguise one of the profound philosophical implications of quantum mechanics: that potentialities constitute an essential aspect of what a physical system is. (On this point we agree with Heisenberg, Ref. 22, p. 53.) Finally, Jauch himself recognizes[23] that the new conception of state is problematic when applied to one part of a system consisting of several correlated parts, as in the example of Einstein, Podolsky, and Rosen.

ACKNOWLEDGMENTS

We wish to thank Professor J. Stachel, Professor F.J. Belinfante, Professor M.A. Horne, and Dr. J. Leveille for helpful discussions.

*Based in part on the Ph.D. thesis of one of us (B.O.H.), Boston University, 1974; the work of one of us (A.S.) was supported in part by a grant from the National Science Foundation.

[1]G.C. Wick, A.S. Wightman, and E. Wigner, Phys. Rev. **88**, 101 (1952).
[2]G. Birkhoff and J. von Neumann, Ann. Math. **37**, 823 (1936).
[3]G.W. Mackey, *Mathematical Foundations of Quantum Mechanics* (Benjamin, New York, 1963).
[4]C. Piron, Helv. Phys. Acta **37**, 439 (1964).
[5]J. Jauch, *Foundations of Quantum Mechanics* (Addison-Wesley, Reading, Mass., 1968).
[6]J. Jauch and C. Piron, Helv. Phys. Acta **42**, 842 (1969).
[7]M. MacLaren, Argonne Nat. Lab. Pub. ANL-7065 (1965).
[8]N. Zierler, Pac. J. Math. **11**, 1151 (1961).
[9]H. Stein, in *Paradigms and Paradoxes*, edited by R.G. Colodny (University of Pittsburgh Press, Pittsburgh, 1972).
[10]S. Kochen and E. Specker, J. Math. Mech. **17**, 331 (1967).
[11]F.J. Belinfante, *A Survey of Hidden Variable Theories* (Pergamon, New York, 1973).
[12]E. Wigner, Z. Phys. **133**, 101 (1952).
[13]H. Araki and M. Yanase, Phys. Rev. **120**, 666 (1961).
[14]H. Stein and A. Shimony, in *Foundations of Quantum Mechanics*, edited by B. d'Espagnat (Academic, New York, 1971), p. 56.

[15]A.M. Gleason, J. Rat. Mech. Anal. **6**, 885 (1957).
[16]We shall use the notations $\wedge_H$ and $\vee_H$ for g.l.b. and l.u.b. in L_H, in order to contrast with the notations $\wedge_V$ and $\vee_V$ in the lattice of verifiable propositions. It is not necessary to introduce differentiating subscripts on the implication and orthocomplementation symbols, because, as will be seen, implication and orthocomplementation are simply the corresponding concepts in L_H restricted to the members of L_V.
[17]V.S. Varadarajan, *Geometry of Quantum Theory* (Van Nostrand, Princeton, N.J., 1968), Vol. I.
[18]C.W. Drake, in *Methods of Experimental Physics*, edited by L. Marton (Academic, New York, 1967), Vol. IVb, Sec. 4.3.
[19]A. Shimony, Found. Phys. **1**, 325 (1971).
[20]J.S. Bell, Rev. Mod. Phys. **33**, 447 (1966).
[21]W. Ochs, Commun. Math. Phys. **25**, 245 (1972).
[22]W. Heisenberg, *Physics and Philosophy* (Harper, New York, 1958).
[23]J. Jauch, in *Foundations of Quantum Mechanics*, edited by B. d'Espagnat (Academic, New York, 1971), p. 31—2.

Asymptotic approximations, with error estimates, of the scattering matrix for quantal Coulomb excitation by means of a nonlinear (Riccati) matrix differential equation

C. H. Rasmussen

Mathematics Department, University of Connecticut, Storrs, Connecticut 06268

G. H. Rawitscher

Physics Department, University of Connecticut, Storrs, Connecticut 06268
(Received 17 February 1976)

A scattering matrix function is defined, which obeys a nonlinear (Riccati) matrix differential equation, containing two coupling potential matrices U and W, which are slowly vanishing, and which are mildly oscillatory and rapidly oscillatory, respectively. The scattering matrix is the limiting value of this scattering function. The equation is first transformed to separate the effects of U and W, thereby yielding separate equations in each. The long range effects of U and W are included in approximations for the scattering matrix, errors are assessed, and a prescription is outlined for the numerical computation of these approximations. In the case where the effect of W is entirely neglected beyond a certain point, the approximation obtained by Alder and Pauli [Nucl. Phys. **128**, 193 (1969)] is recovered. An assessment of the error in this approximation is obtained.

1. INTRODUCTION

Numerical solution of the coupled radial equations which occur in the partial wave description of Coulomb excitation is very cumbersome in practice because of the slow decrease wth distance of the coupling potentials. This problem already occurs in the distorted wave Born approximation to the solution, in that DWBA integrals whose integrands are oscillatory, have to be carried out to prohibitively large distances. When coupling between the various inelastic channels is to be included to all orders, the difficulty is compounded, and inclusion of more than five or six channels becomes prohibitive using the conventional algorithms for solving coupled equations. A similar situation is encountered in atomic physics, where polarization potentials between atoms persist to large distances. Approximate methods of solution have been devised in both the nuclear[1-4] and in the atomic[5,6] cases. These methods exploit the fact that at large distances the coupling potentials become very small and either asymptotic expansions of the radial functions in powers of $(1/r)$ are presented, or else iterative solutions in orders of the coupling potentials are devised. The approximate solutions thus obtained are then matched to the solution obtained by the conventional numerical solution of the coupled equations carried out from the origin to the matching point. None of these studies, however, systematically explore errors incurred due to neglecting the higher order terms. In the present paper we investigate the long range behavior of the solutions of coupled radial equations in the presence of coupling potentials which decrease like the sum of inverse powers of the radial distance of r, $r^{-\lambda-1}$ with $\lambda \geq 1$. The purpose is to examine the leading terms in the solution, and the order of magnitude of the associated corrections so as to guide future numerical approximation methods. In particular, a scattering matrix $S(r)$ is defined as a function of r, whose limit at infinity is the scattering matrix $\mathcal{S}$, from which the scattering cross sections are derived. Approximations to $\mathcal{S}$ are obtained and examined by integrating a Riccati equation for $S(r)$. The zeroth order is identical to an approximate solution first proposed by Alder and Pauli, and the higher order terms provide a method for obtaining the asymptotic behavior of the corrections.

The first order correction terms provide a practical method for improving upon the method of Alder and Pauli, as is shown in a companion paper.[7] In that paper, methods for computing the correction terms are proposed and numerical examples are given.

Solutions of the coupled equations involving asymptotic series in powers of $1/r$ have been employed.[3,6] However, as used, the applicability of this method depends on the requirement that the differences between the wavenumbers in the various scattering channels are not too small. The smaller these differences, the larger the radius r has to be chosen at which the series gives a desired accuracy. This difficulty was circumvented by Mercer and Ravenhall[3] by setting all nuclear excitation energies equal to zero. Their method represents a significant improvement over previous methods in that, once the excitation energies are set equal to zero, it includes the effect of coupling at large distances in an essentially exact fashion. In the present formulation the nuclear excitation energies need not be zero, yet the above mentioned difficulty is avoided. It is hoped that the method will provide a basis for computation of dispersion corrections for inelastic electron—nucleus scattering where a large degree of accuracy is required.

2. GENERAL FORMALISM

The projectile is assumed to have no spin. The position of the projectile relative to the center of mass of the nucleus is denoted by r, the spin of the nucleus in state i is $\hbar\mathbf{I}_i$, and the orbital angular momentum of the motion relative to the center of mass is $\hbar\mathbf{L}$. A given total angular momentum $\hbar\mathbf{J}$ can be achieved by coupling

various $\hbar\mathbf{L}$'s to various $\hbar\mathbf{I}_i$'s, and the corresponding radial projectile-target wavefunctions $\psi_{(LI)JM_J}(r)$ are coupled to each other through the equation

$$\left[\frac{d^2}{dr^2} - D_{\alpha\alpha} + K^2_{\alpha\alpha}\right]\psi_\alpha(r) = \sum_{\alpha'} V_{\alpha\alpha'}(r)\psi_{\alpha'}(r). \tag{2.1}$$

Here the index α denotes collectively the quantum numbers L, I, and J. The magnetic quantum number M_J is suppressed because it is the same for all α's. D is a diagonal matrix which contains the centripetal and monopole Coulomb potentials and K is the diagonal matrix of the various wavenumbers $k_{I_i} = [2\mu(E-\epsilon_i)/\hbar^2]^{1/2}$. Here μ is the projectile-nucleus reduced mass, ϵ_i are the nuclear excitation energies, and E is the total energy.

For the case of a point nucleus,

$$D_{\alpha\alpha}(r) = L(L+1)/r^2 + 2\eta_I k_I/r,$$

where η_I are the Coulomb parameters $ze^2/\hbar v_i = \mu z e^2/\hbar k_i$. The coupling potentials $V_{\alpha\alpha'}$ have the same meaning as $V^J_{Il,I'l'}$ defined in Eq. (10) of Ref. 1. They are also equal to $(2\mu/\hbar^2)\langle Y_\alpha | V | Y_{\alpha'}\rangle$, where the V is the sum over the nucleons in the nucleus of the projectile–nucleon potentials and $Y_{(LI)JM_J}(\hat{\mathbf{r}}, \xi_1 \cdots \xi_A)$ is an eigenstate of the angular momentum operators J_z, J^2, L^2, and I^2 and is also an eigenstate of the nuclear Hamiltonian $H_A(\xi_1 \cdots \xi_A)$. The potential $V_{\alpha\alpha'}$ can be expressed in sums over multipole terms λ, which at distances beyond the nuclear surface decrease as $r^{-\lambda-1}$, with $\lambda \geq 1$. The monopole term, $\lambda = 0$, has been removed from $V_{\alpha\alpha'}$ and included in $D_{\alpha\alpha}$. The value of α runs from 1 to n.

There are n linearly independent solutions $\psi^{(s)}_\alpha$, $s = 1, 2, \ldots, n$ of Eq. (2.1) which are regular at the origin, but which may not yet obey the proper asymptotic boundary conditions. If all the components $\psi^{(s)}_\alpha$, $\alpha = 1, \ldots, n$ for a given s, are placed in a column, and if all the columns $s = 1, \ldots, n$ are placed next to each other, one obtains the matrix $\psi(r)$ of the regular solutions,

$$(\psi)_{\alpha s} = \psi^{(s)}_\alpha(r).$$

In matrix form Eq. (2.1) then reads

$$\left[I\frac{d^2}{dr^2} - D(r) + K^2\right]\psi(r) = V(r)\psi(r). \tag{2.2}$$

The matrices $D(r)$, K^2, and $V(r)$ are real and symmetric. At large distances, say $r \geq R_M$, the diagonal potential reduces to the usual point Coulomb form

$$D_{\alpha\alpha}(r) = L_\alpha(L_\alpha+1)/r^2 + 2\eta_\alpha k_\alpha/r, \quad r \geq R_M,$$

where η_α are the Coulomb parameters $ze^2/\hbar v_\alpha = \mu z e^2/\hbar k_\alpha$, and the coupling potential matrix can be expanded in powers of $1/r$,

$$V(r) = \sum_{\lambda \geq 1} V_\lambda r^{-\lambda-1}, \quad r \geq R_M, \tag{2.3}$$

with $\int_{R_M}^\infty |V_{\alpha\beta}(r)|\, dr < \infty$ for each entry $V_{\alpha\beta}(r)$ of $V(r)$. The point Coulomb function[8] which asymptotically contains outgoing waves is denoted by $h_\alpha(r)$,

$$h_\alpha(r) = h(l_\alpha, \eta_\alpha, k_\alpha r) = \phi_\alpha(r)\exp(i\theta_\alpha). \tag{2.4}$$

In the above ϕ_α is the amplitude which at large distances goes to unity and the phase is

$$\theta_\alpha(r) = k_\alpha r - \eta_\alpha \ln(2k_\alpha r) - L_\alpha\pi/2 + \sigma_{L_\alpha,\eta_\alpha}. \tag{2.5}$$

A diagonal matrix $H(r)$ can now be defined by

$$H_{\alpha\alpha}(r) = h_\alpha(r)k_\alpha^{-1/2}, \quad \alpha = 1, \ldots, n, \tag{2.6}$$

and the corresponding matrix of the ingoing point Coulomb waves, h^*_α, is given by $H^\dagger_{\alpha\alpha} \equiv H^*_{\alpha\alpha}$. (In what follows, Hermitian conjugation of a matrix M is denoted by $M^\dagger$, complex conjugation by M^*, transpose of a matrix M by M^T, and M' denotes dM/dr.) Furthermore,

$$H'(r)H^*(r) - H(r)H^{*\prime}(r) = 2iI. \tag{2.7}$$

Employing the method of variation of constants, matrix functions $\mathcal{A}(r)$ and $\mathcal{B}(r)$ are defined by

$$\psi(r) = H(r)\mathcal{A}(r) + H^*(r)\mathcal{B}(r), \tag{2.8a}$$

$$\psi'(r) = H'(r)\mathcal{A}(r) + H^{*\prime}(r)\mathcal{B}(r), \tag{2.8b}$$

which, together with Eq. (2.2) and (2.7) yield for $r \geq R_M$,

$$\mathcal{A}'(r) = U\mathcal{A} + W\mathcal{B}, \tag{2.9a}$$

$$\mathcal{B}'(r) = U^*\mathcal{B} + W^*\mathcal{A}. \tag{2.9b}$$

Here

$$U(r) = \frac{1}{2i}H^*(r)V(r)H(r) \tag{2.10}$$

is skew Hermitian, i.e., $U^\dagger = -U$, and its elements are slowly oscillatory functions of r. The matrix

$$W(r) = \frac{1}{2i}H^*(r)V(r)H^*(r) \tag{2.11}$$

is symmetric, and its elements are rapidly varying functions of r. By means of Eq. (2.7), the "matching" equations for $\mathcal{A}$ and $\mathcal{B}$ in terms of ψ and ψ' are given as

$$\mathcal{A}(r) = -[H^{*\prime}(r)\psi(r) - H^*(r)\psi'(r)]/(2i), \tag{2.12a}$$

$$\mathcal{B}(r) = [H'(r)\psi(r) - H(r)\psi'(r)]/(2i). \tag{2.12b}$$

If $\psi(r)$ is real, which can always be chosen to be the case since the potentials are real (no complex optical potentials are present), then $\mathcal{B}(r) = \mathcal{A}(r)^*$.

Since the coupling potential $V(r)$ has integrable entries for $R_M \leq r < \infty$ and the point Coulomb wavefunctions are uniformly bounded there, the elements of U and W are also integrable for $R_M \leq r < \infty$, in which case the solution matrices $\mathcal{A}$ and $\mathcal{B}$ of (2.9) tend to finite limits at ∞. Furthermore, since ψ is regular at the origin, and its columns are linearly independent solution vectors of Eq. (2.1), then $\mathcal{A}(r)$ and $\mathcal{B}(r)$ are nonsingular for $R_M \leq r < \infty$, even as r tends to ∞, and we can define the matrix function $S(r)$ by

$$S(r) = -\mathcal{A}(r)\mathcal{B}^{-1}(r), \quad R_M \leq r \tag{2.13}$$

and its limit

$$\mathcal{S} = \lim_{r\to\infty} S(r),$$

both of which are unitary and symmetric since ψ is regular and the potentials are real and symmetric. $\mathcal{S}$ represents the scattering matrix associated with Eq. 1. To

see that S and $\mathcal{S}$ are well defined and unitary, note that since ψ is regular and the potentials are Hermitian,

$$\psi^\dagger\psi' - \psi^{\dagger\prime}\psi = 0 \quad \text{for all } r,$$

but

$$\psi^\dagger\psi' = (\mathcal{A}^\dagger H^* + \mathcal{B}^\dagger H)(H'\mathcal{A} + H^{*\prime}\mathcal{B})$$
$$= \mathcal{A}^\dagger H^* H'\mathcal{A} + \mathcal{B}^\dagger H H^{*\prime}\mathcal{B} + \mathcal{A}^\dagger H^* H^{*\prime}\mathcal{B} + \mathcal{B}^\dagger H H'\mathcal{A}$$

so

$$0 = \mathcal{A}^\dagger(H^*H' - H^{*\prime}H)\mathcal{A} + \mathcal{B}^\dagger(HH^{*\prime} - H'H^*)\mathcal{B}$$
$$= 2i[\mathcal{A}^\dagger\mathcal{A} - \mathcal{B}^\dagger\mathcal{B}],$$

where we have used (2.7) and the fact that H is diagonal and so commutes with H'.

Therefore, $\mathcal{A}^\dagger(r)\mathcal{A}(r) = \mathcal{B}^\dagger(r)\mathcal{B}(r)$ for all r. If $\mathcal{B}$ is singular at some point $r_0 \geq R_M$, then for some nonzero constant column vector ξ, $\mathcal{B}(r_0)\xi = 0$. But then $0 = \xi^\dagger\mathcal{B}^\dagger(r_0)\times\mathcal{B}(r_0)\xi = \xi^\dagger\mathcal{A}^\dagger(r_0)\mathcal{A}(r_0)\xi$, so that $\mathcal{A}(r_0)\xi = 0$. But then from (2.8) we see that $\psi(r_0)\xi = 0$ and $\psi'(r_0)\xi = 0$, in which case the vector solution $\phi(r) = \psi(r)\xi$ of (2.2) is identically zero for all r, contradicting the linear independence of the columns of ψ. Therefore, $\mathcal{B}(r)$ is nonsingular for all $r \geq R_M$. Similarly $\mathcal{A}(r)$ is also nonsingular for $r \geq R_M$. Thus $S(r)$ is well defined, and since $\mathcal{A}^\dagger(r)\mathcal{A}(r) = \mathcal{B}^\dagger(r)\mathcal{B}(r)$, $I = (\mathcal{B}^\dagger)^{-1}\mathcal{A}^\dagger\mathcal{A}\mathcal{B}^{-1} = S^\dagger(r)S(r)$, whence $S(r)$ is unitary. Since $\mathcal{A}(r)$ and $\mathcal{B}(r)$ tend to definite limits at ∞, and $S(r)$ is unitary, $\mathcal{S}$ is well defined and unitary.

Since in addition, the potentials are real, then

$$\psi^T\psi' - \psi^{T\prime}\psi = 0 \quad \text{for all } r,$$

in which case one obtains in a manner similar to the above that

$$0 = \psi^T\psi' = \psi^{T\prime}\psi = 2i[\mathcal{B}^T\mathcal{A} - \mathcal{A}^T\mathcal{B}] \quad \text{for all } r \geq R_M.$$

Thus $S^T = -(\mathcal{B}^T)^{-1}\mathcal{A}^T = -\mathcal{A}\mathcal{B}^{-1} = S$, and S is symmetric. Making use of Eq. (2.9) one obtains for S a Riccati equation

$$S'(r) = U(r)S(r) + S(r)U^T(r) - W(r) + S(r)W^*(r)S(r). \tag{2.14}$$

In view of Eqs. (2.10) and (2.11) the above equation can also be written as

$$S'(r) = -(SH - H^*)V(HS - H^*)/2i. \tag{2.15}$$

The integral form of Eq. (2.14) is

$$\mathcal{S} = S(R_M) + \int_{R_M}^{\infty}[U(\xi)S(\xi) + S(\xi)U^T(\xi)]\,d\xi$$
$$- \int_{R_M}^{\infty} W(\xi)\,d\xi + \int_{R_M}^{\infty} S(\xi)W^*(\xi)S(\xi)\,d\xi, \tag{2.16}$$

which can be iterated to yield successively better approximations for $\mathcal{S}$. It will be seen later on, that in order to obtain approximations to $\mathcal{S}$, it is advantageous to first solve the system

$$A'(r) = U(r)A(r), \quad r \geq R_M, \tag{2.17a}$$

$$A(R_M) = I. \tag{2.17b}$$

The solution matrix $A(r)$ of Eq. (2.17) is unitary. This follows from the fact that $A^\dagger(R_M)A(R_M) = I$ and $(A^\dagger(r)\times A(r))' = A^\dagger U^\dagger A + A^\dagger UA = A^\dagger(U^\dagger + U)A$. The latter expression vanishes since U is skew-Hermitian and $A^\dagger A = I$ for all r.

3. COMPARISON WITH THE METHOD OF ALDER AND PAULI

The authors of Refs. 1 and 2 solve the system of equations (2.9) approximately by dropping the rapidly oscillatory terms in W and W^*,

$$A_a'(r) = U(r)A_a(r), \tag{3.1a}$$

$$B_a'(r) = U^*(r)B_a(r), \tag{3.1b}$$

$$\psi_a(r) = HA_a + H^*B_a, \quad r \geq R_M. \tag{3.2}$$

In order to maintain the continuity of ψ and ψ' at the matching point R_M, the initial conditions for A_a and B_a are determined by equating ψ_a and ψ_a', obtained from Eq. (3.3) in terms of A_a, B_a, A_a', and B_a' to ψ and ψ' at $r = R_M$. Making use of Eqs. (3.1) and (3.2), the resulting equations for A_a and B_a at $r = R_M$ are

$$(I + iH^{*2}W^*)A_a = \mathcal{A} - H^{*2}HV\psi/4, \tag{3.3}$$
$$(I - iH^2W)B_a = \mathcal{B} - H^2H^*V\psi/4,$$

where all quantities are to be evaluated at $r = R_M$, and where $\mathcal{A}$ and $\mathcal{B}$ are obtained from Eq. (2.12). Since W decreases with V on the order of $r^{-\lambda_0-1}$, $\lambda_0 \geq 1$, we see that for R_M large enough these equations can be solved explicitly for A_a and B_a, which will be nonsingular. Again, if ψ is real then $B_a(R_M) = A_a^*(R_M)$.

The solutions of Eq. (3.1) can be expressed in terms of $A(r)$, defined by Eq. (2.17), according to

$$A_a(r) = A(r)A_a(R_M), \tag{3.4a}$$

$$B_a(r) = A^*(r)B_a(R_M), \tag{3.4b}$$

and the scattering function analogous to (2.13) is

$$S_a(r) = -A_a(r)B_a^{-1}(r)$$
$$= A(r)[-A_a(R_M)B_a(R_M)]A^T(r), \tag{3.4c}$$

$$S_a(r) = A(r)S_a(R_M)A^T(r). \tag{3.4d}$$

The scattering matrix is again obtained in the limit of large distances

$$S_a(\infty) = \lim_{r\to\infty} S_a(r). \tag{3.5}$$

Code AROSA[1] computes $A(r)$, the solution of Eq. (2.17), for increasing r until, for $r \geq R_A$, the values of $S_a(r)$ change less than a predetermined amount. Thus $S_a(\infty) \approx S_a(R_A)$.

Direct differentiation of the scattering function for this approximation yields

$$S_a'(r) = U(r)S_a(r) + S_a(r)U^T(r). \tag{3.6a}$$

Note that this can also be obtained from Eq. (2.14) by setting the oscillatory term W equal to zero. Further,

$$S_a(R_M) = -(A_aB_a^{-1})_{r=R_M}. \tag{3.6b}$$

This initial value is not exact, since at the matching point R_M we have

$$\psi = HA_a + H^*B_a = (H^* - HS_a)B_a,$$

and

$$\psi = H\mathcal{A} + H^*\mathcal{B} = (H^* - HS)\mathcal{B},$$

so that $\psi B_a^{-1} = H^* - HS_a$ and $\psi\mathcal{B}^{-1} = H^* - HS$, yielding, for $r = R_M$,

$$S - S_a = H^{-1}\psi(B_a^{-1} - \mathcal{B}^{-1}).$$

This is not zero, since $B_a \neq \mathcal{B}$. The error in AROSA's scattering matrix function is

$$\Delta(r) = S(r) - S_a(r). \tag{3.7}$$

In view of Eqs. (2.14) and (3.6a) it satisfies

$$\Delta'(r) = U(r)\Delta(r) + \Delta(r)U^T(r) - W(r) + S(r)W^*(r)S(r), \tag{3.8a}$$

$$\Delta(R_M) = S(R_M) - S_a(R_M). \tag{3.8b}$$

Introducing the matrix $\delta(r)$ by means of the transformation

$$\Delta(r) = A(r)\delta(r)A^T(r), \tag{3.9}$$

one finds that the first two terms on the right-hand side of Eq. (3.8a) cancel in view of Eq. (2.17), and one has

$$\delta'(r) = -A^\dagger WA^* + A^\dagger SW^*SA^*, \tag{3.10a}$$

$$\delta(R_M) = \Delta(R_M). \tag{3.10b}$$

The integral of Eq. (3.10) is

$$\delta(\infty) = \Delta(R_M) - \int_{R_M}^{\infty} A^\dagger WA^*\,d\xi + \int_{R_M}^{\infty} A^\dagger SW^*SA^*\,d\xi. \tag{3.11}$$

An integration by parts of the first integral yields

$$\int_{R_M}^{\infty} A^\dagger WA^*\,d\xi = -Y(R_M) + \int_{R_M}^{\infty} A^\dagger[UY + YU^T]A^*\,d\xi, \tag{3.12}$$

where

$$Y(r) = -\int_r^{\infty} W(\xi)\,d\xi. \tag{3.13}$$

Similarly, for the second integral in Eq. (3.11), one finds

$$\int_{R_M}^{\infty} A^\dagger SW^*SA^*\,d\xi = -S(R_M)Y^*(R_M)S(R_M) - \int_{R_M}^{\infty} A^\dagger[S'Y^*S + SY^*S' + USY^*S + SY^*SU^T]A^*\,d\xi. \tag{3.14}$$

The Euclidean vector norm

$$\|\xi\| = (|\xi_1|^2 + |\xi_2|^2 + \cdots + |\xi_n|^2)^{1/2}$$

for a vector

$$\xi = \begin{bmatrix} \xi_1 \\ \vdots \\ \xi_n \end{bmatrix}$$

induces a natural matrix norm

$$\|M\| = \max\{\|M\xi\|;\ \xi \text{ a unit vector}\}. \tag{3.15a}$$

Since $\|M\xi\|^2 = \xi^\dagger M^\dagger M\xi$ is a positive semidefinite, Hermitian, quadratic form,[9] we see that

$$\|M\|^2 = \text{maximum eigenvalue of } M^\dagger M. \tag{3.15b}$$

We shall find this a convenient norm, since for U a unitary matrix, $\|U\| = 1$, and

$$\|MU\| = \|M\| = \|UM\| \tag{3.16}$$

for any matrix M.

Recalling the expansion for V given in Eq. (2.3), we define

$$v_{\lambda_0} = \max_{r \geq R_M}\{k\|H(r)\|^2\|V(r)\|r^{\lambda_0+1}\}, \tag{3.17}$$

where λ_0 is the index of the leading term $V_{\lambda_0}r^{-\lambda_0-1}$ in that expansion, and k is the smallest of the wavenumbers k_α, $1 \leq \alpha \leq n$, in Eq. (2.1).

Since R_M is greater than the turning points of each of the point Coulomb functions $h_\alpha(r)$, $1 \leq \alpha \leq n$, then $|h_\alpha(r)|$ is nearly 1 for $r \geq R_M$, so that v_{λ_0} is very nearly equal to $\|V_{\lambda_0}\|$.

We, therefore, have for U and W defined in (2.10) and (2.11) that

$$\|U(r)\| = O(v_{\lambda_0}k^{-1}r^{-\lambda_0-1}), \tag{3.18a}$$

$$\|W(r)\| = O(v_{\lambda_0}k^{-1}r^{-\lambda_0-1}). \tag{3.18b}$$

In Appendix A it is established that

$$\|Y(r)\| = O(v_{\lambda_0}k^{-2}r^{-\lambda_0-1}). \tag{3.18c}$$

This result is due to the fact that the integrand W in Eq. (3.13) is a rapidly oscillating function of r, which is also responsible for the additional factor of k^{-1} in (3.18c). Since S is a unitary matrix, it follows from Eq. (2.14) that the norm of $S'(r)$ is of the same order of magnitude as that of the norms of U and W. Hence the norms of the integrands of all three integrals on the right-hand sides of Eqs. (3.12) and (3.14) are of the same order of magnitude, $O(v_{\lambda_0}^2k^{-3}r^{-2\lambda_0-2})$. Thus all three integrals have norms of order $O(v_{\lambda_0}^2k^{-3}r^{-2\lambda_0-1})$. The above argument shows that

$$\|\delta(\infty) - \Delta(R_M)\| = O(\|Y\|) = O(v_{\lambda_0}k^{-2}R_M^{-\lambda_0-1}). \tag{3.19}$$

Further, from Eq. (3.3), it can be seen that

$$\begin{aligned}\|\Delta(R_M)\| &= \|S(R_M) - S_A(R_M)\| \\ &= \|\mathcal{A}(R_M)\mathcal{B}^{-1}(R_M) - A_a(R_M)B_a^{-1}(R_M)\| \\ &= O(v_{\lambda_0}k^{-2}R_M^{-\lambda_0-1}).\end{aligned} \tag{3.20}$$

It is not too surprising that (3.19) and (3.20) are of the same order of magnitude since the error in both was due to the neglect of W. Combining (3.9), (3.19), and (3.20) it is found that the error in the scattering matrix computed by AROSA is

$$\begin{aligned}\|\Delta(\infty)\| &= \|S - S_a(\infty)\| \\ &= O(v_{\lambda_0}k^{-2}R_M^{-\lambda_0-1}) + O(v_{\lambda_0}^2k^{-3}R_M^{-2\lambda_0-1}).\end{aligned}$$

Since in practice $k \approx 1$, $kR_M \gg 1$, this is just

$$\|S - S_a(\infty)\| = O(v_{\lambda_0}k^{-2}R_M^{-\lambda_0-1}). \tag{3.21}$$

From Eq. (3.4d) we see that since A is unitary, the unitarity and symmetry of $S_a(r)$ depends on the unitarity and symmetry of $S_a(R_M) = -A_aB_a^{-1}$. However, since the matching equations (3.3) show that in general A_a differs from $\mathcal{A}$ at R_M, $S_a(R_M)$ is not in general unitary. Since V

and ψ are real, we see that $B_a = A_a^*$, and $S_a(R_M) = -(A_a A_a^{*-1})$ so that $S_a(R_M)S_a^*(R_M) = I$, in which case $S_a(R_M)$ is unitary exactly when it is symmetric.

The matching equations at R_M yield

$$\psi = HA_a + H^*B_a = (H^* - HS_a)B_a \tag{3.22}$$

and

$$\begin{aligned}\psi' &= H'A_a + H^{*\prime}B_a + HA_a' + H^*B_a' \\ &= [(H^{*\prime} + H^*U^*) - (H' + HU)S_a]B_a.\end{aligned} \tag{3.23}$$

Using the regularity condition $\psi^\dagger\psi' - \psi^{\dagger\prime}\psi = 0$, and the invertibility of B_a, we obtain

$$[\psi^\dagger(H' + HU) - \psi^{\dagger\prime}H]S_a = [\psi^\dagger(H^{*\prime} + H^*U^*) - \psi^{\dagger\prime}H^*]. \tag{3.24}$$

If S_a were unitary, then using the fact that U is skew-Hermitian, V is real and symmetric, and ψ is real, we would obtain from (3.24) the equation

$$\begin{aligned}&\psi^T(H^*H)V(H^*H)\psi' - \psi^{T\prime}(H^*H)V(H^*H)\psi \\ &+ \tfrac{1}{2}\psi^T[(H^*H)'V(H^*H) - (H^*H)V(H^*H)']\psi = 0.\end{aligned} \tag{3.25}$$

Certainly, if we were dealing with a single scalar equation, or if by some other means V commuted with H, then this equation would be correct. But there is no reason to expect that it is an identity for all cases. Hence, in general $S_a(r)$ is neither unitary nor symmetric

4. SUCCESSIVE APPROXIMATIONS FOR THE SCATTERING MATRIX

The starting point of the present considerations is the observation that the equation for the exact $S(r)$ matrix function, Eq. (2.14), and that for $S_a(r)$, Eq. (3.6a), differ only by the terms in the rapidly oscillatory matrix W.

In the spirit of Alder and Pauli, it is possible to separate the effect of the slowly oscillating coupling matrix U from that of the rapidly oscillating matrix W by defining a new matrix $R(r)$ by the unitary transformation

$$S(r) = A(r)R(r)A^T(r), \tag{4.1a}$$

$$R(r) = A^\dagger(r)S(r)A^*(r). \tag{4.1b}$$

Note that

$$S(R_M) = R(R_M) \tag{4.2}$$

since $A(R_M) = I$. Eq. (2.14), with the help of Eq. (2.17), reduces to

$$R'(r) = -A^\dagger WA^* + R(A^TW^*A)R. \tag{4.3}$$

According to Eq. (4.1b), R is symmetric and unitary since S is symmetric and unitary, and A is unitary. By integration of Eq. (4.3) one obtains

$$R(r) = S(R_M) - \int_{R_M}^r A^\dagger WA^*\,d\xi + \int_{R_M}^r R(A^TW^*A)R\,d\xi. \tag{4.4}$$

Integration by parts, recalling Eq. (3.13), yields

$$\int_{R_M}^\infty A^\dagger WA^*\,d\xi = -Y(R_M) + \int_{R_M}^\infty A^\dagger[UY + YU^T]A^*\,d\xi \tag{4.5}$$

and

$$\begin{aligned}\int_{R_M}^\infty R[A^TW^*A]R\,d\xi &= -S(R_M)Y^*(R_M)S(R_M) \\ &\quad - \int_{R_M}^\infty [R'A^TY^*AR + RA^TY^*AR']\,d\xi \\ &\quad - \int_{R_M}^\infty RA^T[U^TY^* + Y^*U]AR\,d\xi.\end{aligned} \tag{4.6}$$

Collecting terms, using Eq. (4.1a), and defining

$$\mathcal{S}_1 = A(\infty)[S(R_M) + Y(R_M) - S(R_M)Y^*(R_M)S(R_M)]A^T(\infty), \tag{4.7}$$

one obtains

$$\begin{aligned}\mathcal{S} = \mathcal{S}_1 - A(\infty)\{&\int_{R_M}^\infty A^\dagger[UY + YU^T]A^*\,d\xi \\ &+ \int_{R_M}^\infty [R'A^TY^*AR + RA^TY^*AR']\,d\xi \\ &+ \int_{R_M}^\infty RA^T[U^TY^* + Y^*U]AR\,d\xi\}A^T(\infty).\end{aligned} \tag{4.8}$$

The above equation is exact. However, each of the integrals above is of order $O(v_{\lambda_0}^2 k^{-3} R_M^{-2\lambda_0 - 1})$ as can be seen by arguments similar to those made in connection with Eqs. (3.18)–(3.21), and hence

$$\|\mathcal{S} - \mathcal{S}_1\| = O(v_{\lambda_0}^2 k^{-3} R_M^{-2\lambda_0 - 1}). \tag{4.9}$$

The above represents a considerable improvement over the error in the approximation of Alder *et al*, given by Eq. (3.21). It requires the computation of $Y(R_M) = -\int_{R_M}^\infty W(\xi)\,d\xi$ as the only addition to what is calculated in AROSA. Numerical methods for the calculation of Y in terms of a continued fraction method are described in Ref. 7. Since $S(R_M)$ and $Y(R_M)$ are symmetric matrices, it is clear from Eq. (4.7) that $\mathcal{S}_1$ also is symmetric. Furthermore, direct calculation shows that

$$\mathcal{S}_1^\dagger\mathcal{S}_1 = I + A^*[Y^*Y + SYY^*S - (SY)^2 - (Y^*S)^2]A^T$$

and, therefore,

$$\|\mathcal{S}_1^\dagger\mathcal{S}_1 - I\| = O(v_{\lambda_0}^2 k^{-4} R_M^{-2\lambda_0 - 2}) \tag{4.10}$$

which shows that $\mathcal{S}_1$ is unitary to a higher order than its error given in (4.9). If a better approximation to $\mathcal{S}$ than $\mathcal{S}_1$ is desired, a further examination of the integrals in (4.5) and (4.6) is required. By means of an additional integration by parts, and by defining the matrix

$$X(r) = \int_r^\infty WY^*\,d\xi, \tag{4.11}$$

one obtains as a higher order approximation of $\mathcal{S}$ the result

$$\mathcal{S}_2 = \mathcal{S}_1 + A(\infty)[X(R_M)S(R_M) + S(R_M)X^T(R_M)]A^T(\infty). \tag{4.12}$$

$\mathcal{S}_2$ is obviously symmetric.

Since $X(R_M) + X^\dagger(R_M) = \int_{R_M}^\infty [WY^* + Y^*W]\,d\xi = \int_{R_M}^\infty (YY^*)'\,d\xi = -Y(R_M)Y^*(R_M)$, one sees that

$$X^\dagger(R_M) = -X(R_M) + O(v_{\lambda_0}^2 k^{-4} R_M^{-2\lambda_0 - 2}). \tag{4.13}$$

Using this result together with the assumption that $kR_M \gg 1$, one can show that

$$\|S_2^\dagger S_2 - I\| = O(v_{\lambda_0}^2 k^{-4} R_M^{-2\lambda_0-2}). \tag{4.14}$$

Note that the unitarity of S_2 is not better than that of S_1. However, the accuracy of S_2 is better than that of S_1. Examining the integrals in (4.5) and (4.6) and an additional integration by parts performed by means of the introduction of $X(r)$, one finds

$$\|S - S_2\| = O(v_{\lambda_0}^2 k^{-4} R_M^{-2\lambda_0-2}). \tag{4.15}$$

This is better than (4.9) by another power of $(kR_M)^{-1}$.

The calculation of $X(r)$ is not easy, because first Y has to be computed at all integration points in (4.11). However, since, as shown in the Appendix, $Y^*(r) \approx -(i/2k)W^*(r)$, the integrand $WY^* \approx (-i/2k)WW^*$ is a slowly varying function of r, and the integration steps could be widely spaced. The merit of computing Y and X as compared to only computing Y and using a larger value of R_M should be decided in each particular case. In the numerical example given in Ref. 7, it is shown that the inclusion of only Y, i.e., the use of S_1 rather than $S_a(\infty)$ of Ref. 1 enables one to reduce R_M from 70 to 30 fm. The inclusion of both Y and X would reduce R_M further to 20 fm. However, this may not seem worth the extra effort of calculating X.

5. CALCULATION OF $A(r)$

It is tempting to solve Eq. (2.17) for A by using an asymptotic expansion in powers of $(1/r)$ of the elements of A. This method is essentially the one used by Burke *et al.*,[6] and by Mercer and Ravenhall,[3] in the solution of Eq. (2.1). However, the matrix elements U_{ij} contain phase factors $\theta_i - \theta_j$ which make them slowly varying functions of r. This in turn leads to the requirement that

$$\Delta k r \gg 1, \tag{5.1a}$$

where

$$\Delta k = \min_{k_i \neq k_j} |k_i - k_j|, \tag{5.1b}$$

in order that the error in the asymptotic series is to be kept small.[10] If all the wavenumbers are equal, this difficulty does not arise and the asymptotic expansion can be used. On the other hand, numerical evaluation of $Y(R_M)$ by the methods discussed in Ref. 7 only requires that $R_M \geq (k_i + k_j)^{-1}$. This is much less stringent than (5.1).

Defining $\bar{A}(r)$ by

$$\bar{A}'(r) = U\bar{A}, \tag{5.2a}$$

$$\bar{A}(\infty) = I, \tag{5.2b}$$

we have, by successive iteration, that

$$\bar{A}(r) = I - \int_r^\infty U(s)\bar{A}(s)\,ds = \sum_{\nu=0}^{\infty} \bar{A}_\nu(r), \tag{5.3a}$$

where

$$\bar{A}_0(r) = I \tag{5.3b}$$

and

$$\bar{A}_{\nu+1}(r) = -\int_r^\infty U(s)\bar{A}_\nu(s)\,ds, \quad \nu \geq 0. \tag{5.3c}$$

Since $\|U(r)\| \leq v_{\lambda_0}(2kr^{\lambda_0+1})^{-1}$ for $r \geq R_M$, one readily obtains

$$\|\bar{A}_\nu(r)\| \leq \frac{1}{\nu!}[v_{\lambda_0}/(2k\lambda_0 r^{\lambda_0})]^\nu \quad r \geq R_M, \tag{5.4}$$

which implies absolute and uniform convergence of the series (5.3) for $r \geq R_M$.

For any point $R_A \geq R_M$, we have that the solution $A(r)$ of (2.17) is given by

$$A(r) = \bar{A}(r)\bar{A}^{-1}(R_A)A(R_A). \tag{5.5}$$

In particular

$$A(\infty) = \bar{A}(\infty)\bar{A}^{-1}(R_A)A(R_A) = \bar{A}^{-1}(R_A)A(R_A).$$

Since $\bar{A}(r)$ is also unitary, we see that

$$\bar{A}^{-1}(r) = \bar{A}^\dagger(r) = \sum_{\nu=0}^{\infty} \bar{A}_\nu^\dagger(r).$$

Furthermore, since U is skew-Hermitian, we see that

$$\bar{A}_1^\dagger(r) = -\int_r^\infty U^\dagger(s)\,ds = \int_r^\infty U(s)\,ds.$$

Therefore, defining the matrix $A_1(\infty)$, by

$$A_1(\infty) = [I + \bar{A}_1^\dagger(R_A)]A(R_A) = [I + \int_{R_A}^\infty U(s)\,ds]A(R_A), \tag{5.6}$$

we see that

$$A(\infty) - A_1(\infty) = \left[\sum_{\nu=2}^{\infty} \bar{A}_\nu^\dagger(R_A)\right] A(R_A) = O([v_{\lambda_0}/(2k\lambda_0 R_A^{\lambda_0})]^2). \tag{5.7}$$

If we approximate $A(\infty)$ in (4.7), by $A_1(\infty)$ and define S_F by

$$S_F = A_1(\infty)[S(R_M) + Y(R_M) - S(R_M)Y^*(R_M)S(R_M)]A_1^T(\infty), \tag{5.8}$$

then we have

$$\|S - S_F\| = O(v_{\lambda_0}^2/(k^3 R_M^{2\lambda_0+1})) + O(v_{\lambda_0}^2/(k^2\lambda_0^2 R_A^{2\lambda_0})). \tag{5.9}$$

The first term on the right in (5.9) reflects the error in S due to truncation in W, and the second reflects that due to truncation in U. It is clear that S_F is symmetric, and is unitary up to the order given by (5.9).

If we represent the actual errors in S due to the two truncations as $E_W(R_M)$ and $E_U(R_A)$, so that

$$S = S_F + E_W(R_M) + E_U(R_A), \tag{5.10}$$

then the numerical prescription for computing the approximate scattering matrix S_F to a given accuracy ϵ is as follows.

One solves (2.1) numerically on some interval $[0, R_M]$ for a regular solution matrix ψ and its derivative ψ' at R_M. One then solves (2.17) numerically on some interval $[R_M, R_A]$. In addition one computes the integrals $\int_{R_M}^\infty W(s)\,ds$ and $\int_{R_A}^\infty U(s)\,ds$. These quantities suffice to

construct S_F, given by (5.8). Methods for computing the integrals are discussed in Ref. 7. The points R_M and R_A are determined by requiring that the numerical roundoff error, plus the truncation error $\|E_W(R_M)\| + \|E_U(R_A)\|$ be bounded by ϵ.

Since the error in S_1 due to truncation in W is given to leading order by (4.12), we see that

$$\|S - S_1\| = \|E_W(R_M)\| \lesssim 2\|X(R_M)\| \lesssim v_{\lambda_0}^2/(4k^3(2\lambda_0+1)R_M^{2\lambda_0+1}). \tag{5.11}$$

Furthermore,

$$\|S_1 - S_F\| = \|E_U(R_A)\| \approx 2\|A(\infty) - A_1(\infty) \approx 2\|\bar{A}_2(R_A)\| \leq v_{\lambda_0}^2/(4k^2\lambda_0^2 R_A^{2\lambda_0}). \tag{5.12}$$

Therefore, assuming that R_M is beyond the turning points of the point Coulomb functions $h_\alpha(r)$, and that roundoff error is negligible compared to the truncation error, one can estimate R_M and R_A by requiring that

$$\|E_W(R_M)\| + \|E_U(R_A)\| \leq \epsilon. \tag{5.13}$$

If one wishes simply to balance the two truncation errors, then one obtains an approximate relation between R_A and R_M given by

$$R_A \gtrsim R_M[kR_M(2\lambda_0+1)/\lambda_0^2]^{\lambda_0/2}. \tag{5.14}$$

In the numerical example discussed in Ref. 7, R_A has the approximate value 60 fm, as compared with 400 fm in Ref. 1, in order to obtain an accuracy of 10^{-4}. In Ref. 7 a series method for evaluating $\bar{A}_1(R_A)$ is presented, which does require that R_A be larger than the turning points of the point Coulomb wavefunctions in all the channels involved, but does not require that $\Delta k R_A > 1$.

6. SUMMARY AND CONCLUSIONS

An r dependent scattering matrix, $S(r)$, is defined which approaches the scattering matrix S as $r \to \infty$. The way in which $S(r)$ approaches S is studied by means of a first order nonlinear (Riccati) equation satisfied by S. This equation contains the matrices $U(r)$ and $W(r)$, both of which are obtained from the coupling potential $V(r)$ by multiplication by the matrix of point Coulomb functions. The former, $U(r)$, is a slowly varying function of r, the latter a rapidly varying function, since they contain factors $\exp[i(k_i - k_j)r]$ and $\exp[i(k_i + k_j)r]$, respectively (k_i, $i = 1, \ldots, n$, being the wavenumbers). Both U and W decrease in magnitude as r increases, the leading term being of the order $r^{-\lambda_0-1}$, $\lambda_0 \geq 1$. As a consequence of the difference in the oscillatory behavior of U and W, $S(r)$ can be advantageously transformed so as to separate the effects of the two. Transforming $S(r)$ by $S(r) = A(r)R(r)A^T(r)$, where

$$A' = UA, \quad A(R_M) = I, \tag{6.1}$$

A is unitary, and where

$$R' = -A^\dagger W A^* + R(A^T W^* A)R, \quad R(R_M) = S(R_M), \tag{6.2}$$

R is symmetric and unitary, the effect of the matrices U and W on S can be separately analyzed. An assessment is made of the error introduced in S due to the approximation made in Ref. 1 of neglecting W beyond R_M. More refined approximations for S are obtained by including W to various orders beyond R_M and by approximating A beyond a point $R_A > R_M$ by iterating (6.1). The errors due to the truncation in W and in U are assessed and compared. The final result S_F given in (5.8) requires the calculation of $S(R_M)$, $Y(R_M)$, and $A_1(\infty)$ given in (2.12), (2.13), (3.13), and (5.6). Methods of computing the latter two are described in Ref. 7. The numerical example discussed there shows that S_F provides an approximation for the scattering matrix for the problem discussed in Ref. 1, accurate to 10^{-4}, while reducing R_M from 70 fm to about 30 fm, and reducing R_A from 500 fm to about 60 fm.

APPENDIX A

Set $f(\alpha, \beta; r) = \exp(-i\alpha r)r^{-\beta}$, $\alpha \neq 0$ real, $\mathrm{Re}\beta > 0$. Set $I(\alpha, \beta; r) = \int_r^\infty f(\alpha, \beta; s)\,ds$. Then (A1)

$$I(\alpha,\beta;r) = \int_r^\infty \exp(-i\alpha s)s^{-\beta}\,ds = -\frac{\exp(-i\alpha s)}{i\alpha}s^{-\beta}\Big|_r^\infty - \frac{\beta}{i\alpha}\int_r^\infty \exp(-i\alpha s) \times s^{-\beta-1}\,ds,$$

$$I(\alpha,\beta;r) = \frac{1}{i\alpha}f(\alpha,\beta;r) - \frac{\beta}{i\alpha}I(\alpha,\beta+1;r), \tag{A2}$$

$$I(\alpha,\beta;r) = \frac{1}{i\alpha}f(\alpha,\beta;r) - \frac{\beta}{(i\alpha)^2}f(\alpha,\beta+1;r) + \frac{\beta(\beta+1)}{(i\alpha)^2}I(\alpha,\beta+2;r). \tag{A3}$$

From these we have

$$\left|\frac{i\alpha I(\alpha,\beta;r)}{f(\alpha,\beta;r)} - 1\right| = \left|\frac{\beta}{i\alpha}\frac{f(\alpha,\beta+1;r)}{f(\alpha,\beta;r)} + \frac{\beta(\beta+1)}{i\alpha}\frac{I(\alpha,\beta+2;r)}{f(\alpha,\beta;r)}\right| \leq \left|\frac{\beta}{i\alpha r}\right| + \left|\frac{\beta(\beta+1)r^\beta}{i\alpha}\right| \int_r^\infty s^{-\mathrm{Re}\beta-2}\,ds = \left|\frac{\beta}{\alpha r}\right| + \left|\frac{\beta(\beta+1)}{\alpha}\right| r^{\mathrm{Re}\beta}\frac{r^{-\mathrm{Re}\beta-1}}{(\mathrm{Re}\beta+1)} = \left(1 + \left|\frac{\beta+1}{\mathrm{Re}\beta+1}\right|\right)\left|\frac{\beta}{\alpha r}\right| = O\left(\left|\frac{1}{\alpha r}\right|\right).$$

Thus

$$I(\alpha,\beta;r) = \frac{1}{i\alpha}f(\alpha,\beta;r)\left[1 + O\left(\left|\frac{1}{\alpha r}\right|\right)\right]. \tag{A4}$$

The matrix $W(r) = (1/2i)H^*(r)V(r)H^*(r)$ has entries of the form

$$(W(r))_{\alpha\beta} = \frac{\phi_\alpha^*(r)\phi_\beta^*(r)}{2i(k_\alpha k_\beta)^{1/2}}v_{\alpha\beta}(r)\exp\{-i[\theta_\alpha(r) + \theta_\beta(r)]\}, \tag{A5}$$

where $v_{\alpha\beta}(r)$ is the α,β entry in $V(r)$, and where

$$\theta_\alpha(r) = k_\alpha r - \eta_\alpha \ln r + \mathrm{const}_\alpha,$$

$$\phi_\alpha(r) = 1 + \phi_\alpha^{(1)}r^{-1} + O(r^{-2}),$$

$$v_{\alpha\beta}(r) = v_{\alpha\beta}^{(0)}r^{-\lambda_0-1} + v_{\alpha\beta}^{(1)}r^{-\lambda_0-2} + \cdots = \sum_{\mu=0}^{\infty} v_{\alpha\beta}^{(\mu)}r^{-\lambda_0-1-\mu}.$$

ϕ_α has an asymptotic series representation[8] in r, and $v_{\alpha\beta}(r)$ is analytic in $1/r$, with $\lambda_0 \geq 1$. Consequently

$$\begin{aligned}2i\sqrt{k_\alpha k_\beta}\,(W(r))_{\alpha\beta} = {} & f(k_\alpha + k_\beta, \lambda_0 + 1 - i(\eta_\alpha + \eta_\beta); r) w^{(0)}_{\alpha\beta} \\ & + f(k_\alpha + k_\beta, \lambda_0 + 2 - i(\eta_\alpha + \eta_\beta); r) w^{(1)}_{\alpha\beta} \\ & + f(k_\alpha + k_\beta, \lambda_0 + 3 - i(\eta_\alpha + \eta_\beta); r) w^{(2)}_{\alpha\beta} \\ & + f(k_\alpha + k_\beta, \lambda_0 + 4 - i(\eta_\alpha + \eta_\beta); r) O_{\alpha\beta}(1),\end{aligned} \tag{A6}$$

where $w^{(0)}_{\alpha\beta}$, $w^{(1)}_{\alpha\beta}$, $w^{(2)}_{\alpha\beta}$ are constants, $|w^{(0)}_{\alpha\beta}| = |v^{(0)}_{\alpha\beta}|$, and $O_{\alpha\beta}(1)$ is uniformly bounded as $r \to \infty$. Here f is the function defined in (A1).

From (A6) it is clear that

$$\left|2i\sqrt{k_\alpha k_\beta}(W(r))_{\alpha\beta}\right| = \frac{|v^{(0)}_{\alpha\beta}|}{r^{\lambda_0+1}}(1 + O(1/r)) \quad \text{for } r \geq R_M, \tag{A7}$$

in which case

$$\|W(r)\| = O(v_{\lambda_0}/kr^{\lambda_0+1}), \tag{A8}$$

where

$$k = \min_{1\leq\alpha\leq n} k_\alpha \quad \text{and} \quad v_{\lambda_0} = \max_{r\geq R_M}[k\|H(r)\|^2\|V(r)\|r^{\lambda_0+1}].$$

Integrating (A6) throughout and using (A4) we have

$$\begin{aligned}2i\sqrt{k_\alpha k_\beta} & \int_r^\infty (W(s))_{\alpha\beta}\,ds \\ = {} & f(k_\alpha + k_\beta, \lambda_0 + 1 - i(\eta_\alpha + \eta_\beta); r)\frac{w^{(0)}_{\alpha\beta}}{i(k_\alpha + k_\beta)} \\ & \times\left[1 + O\left(\left|\frac{1}{(k_\alpha + k_\beta)r}\right|\right)\right] + f(k_\alpha + k_\beta, \lambda_0 + 2 - i(\eta_\alpha + \eta_\beta); r) \\ & \times\frac{w^{(1)}_{\alpha\beta}}{i(k_\alpha + k_\beta)r}\left[1 + O\left(\left|\frac{1}{(k_\alpha + k_\beta)r}\right|\right)\right] \\ & + f(k_\alpha + k_\beta, \lambda_0 + 3 - i(\eta_\alpha + \eta_\beta); r)\frac{w^{(2)}_{\alpha\beta}}{i(k_\alpha + k_\beta)} \\ & \times\left[1 + O\left(\left|\frac{1}{(k_\alpha + k_\beta)r}\right|\right)\right] + \int_r^\infty f(k_\alpha + k_\beta, \lambda_0 + 4 \\ & - i(\eta_\alpha + \eta_\beta); s) O_{\alpha\beta}(1)\,ds \\ = {} & \frac{f(k_\alpha + k_\beta, \lambda_0 + 1 - (\eta_\alpha + \eta_\beta); r)}{i(k_\alpha + k_\beta)}\Bigg\{w^{(0)}_{\alpha\beta}\left[1 + O\left(\left|\frac{1}{(k_\alpha + k_\beta)r}\right|\right)\right] \\ & + \frac{w^{(1)}_{\alpha\beta}}{r}\left[1 + O\left(\left|\frac{1}{(k_\alpha + k_\beta)r}\right|\right)\right] + \frac{w^{(2)}_{\alpha\beta}}{r^2}\left[1 + O\left|\frac{1}{(k_\alpha + k_\beta)r}\right|\right] \\ & + \frac{i(k_\alpha + k_\beta)\int_r^\infty f(k_\alpha + k_\beta, \lambda_0 + 4 - i(\eta_\alpha + \eta_\beta); s)}{f(k_\alpha + k_\beta, \lambda_0 + 1 - i(\eta_\alpha + \eta_\beta); r))} O_{\alpha\beta}(1)\,ds\Bigg\}.\end{aligned}$$

Since the last term in braces is $O(|k_\alpha + k_\beta|/r^2)$, which is negligible compared with $O(|1/(k_\alpha + k_\beta)r|)$, we have that

$$\begin{aligned}\int_r^\infty (W(s))_{\alpha\beta}\,ds = {} & \frac{f(k_\alpha + k_\beta, \lambda_0 + 1 - i(\eta_\alpha + \eta_\beta); r)}{-2\sqrt{k_\alpha k_\beta}(k_\alpha + k_\beta)} \\ & \times w^{(0)}_{\alpha\beta}\left[1 + O\left(\left|\frac{1}{\sqrt{k_\alpha k_\beta}(k_\alpha + k_\beta)r}\right|\right)\right].\end{aligned} \tag{A9}$$

This yields the result that

$$\left\| \int_r^\infty W(s)\,ds\right\| = O\left(\frac{v_{\lambda_0}}{k^2 r^{\lambda_0+1}}\right), \quad r \geq R_M. \tag{A10}$$

A similar argument applies to $U(r) = (1/2i)H^*VH$, yielding

$$\left|2i\sqrt{k_\alpha k_\beta}\,(U(r))_{\alpha\beta}\right| = \frac{|v^{(0)}_{\alpha\beta}|}{r^{\lambda_0+1}}\left[1 + O\left(\frac{1}{r}\right)\right], \tag{A11}$$

$$\|U(r)\| = O\left(\frac{v_{\lambda_0}}{kr^{\lambda_0+1}}\right), \quad r \geq R_M, \tag{A12}$$

$$\begin{aligned}\int_r^\infty (U(s))_{\alpha\beta}\,ds = {} & \frac{f(k_\alpha - k_\beta, \lambda_0 + 1 - i(\eta_\alpha - \eta_\beta); r)}{-2\sqrt{k_\alpha k_\beta}(k_\alpha - k_\beta)} \\ & \times u^{(0)}_{\alpha\beta}\left[1 + O\left(\left|\frac{1}{\sqrt{k_\alpha k_\beta}(k_\alpha - k_\beta)r}\right|\right)\right],\end{aligned} \tag{A13}$$

so

$$\left\| \int_r^\infty U(s)\,ds\right\| = O\left(\frac{v_{\lambda_0}}{k(\Delta k)r^{\lambda_0+1}}\right), \tag{A14}$$

where $\Delta k = \min_{\alpha\neq\beta}|k_\alpha - k_\beta|$. The Eqs. (A13) and (A14) are valid for $\Delta k r \gg 1$.

Set $Y(r) = -\int_r^\infty W(s)\,ds$. Then, since in analogy with (A6)

$$(U(r))_{\alpha\beta} = \frac{f(k_\alpha - k_\beta, \lambda_0 + 1 - i(\eta_\alpha - \eta_\beta); r)}{+2i\sqrt{k_\alpha k_\beta}} u^{(0)}_{\alpha\beta}\left[1 + O\left(\frac{1}{r}\right)\right], \tag{A15}$$

we have

$$[UY]_{\alpha\beta} = \left(\sum_{\mu=1}^n \frac{f(k_\alpha - k_\mu, \lambda_0 + 1 - i(\eta_\alpha - \eta_\mu); r) f(k_\mu + k_\beta, \lambda_0 + 1 - i(\eta_\mu + \eta_\beta); r)}{4i(k_\alpha k_\mu^2 k_\beta)^{1/2}(k_\mu + k_\beta)} u^{(0)}_{\alpha\mu} w^{(0)}_{\mu\beta}\right)\left[1 + O\left(\left|\frac{1}{k^2 r}\right|\right)\right]$$

$$= \sum_{\mu=1}^n f(k_\alpha + k_\beta, 2\lambda_0 + 2 - i(\eta_\alpha + \eta_\beta); r)\frac{u^{(0)}_{\alpha\mu} w^{(0)}_{\mu\beta}}{4i(k_\alpha k_\mu^2 k_\beta)^{1/2}(k_\mu + k_\beta)}\left[1 + O\left(\left|\frac{1}{k^2 r}\right|\right)\right]$$

$$[UY]_{\alpha\beta} = \frac{f(k_\alpha + k_\beta, 2\lambda_0 + 2 - i(\eta_\alpha + \eta_\beta); r)}{4i(k_\alpha k_\beta)^{1/2}}\left[1 + O\left(\frac{1}{|k^2 r|}\right)\right]\sum_{\mu=1}^n \left(\frac{u^{(0)}_{\alpha\mu} w^{(0)}_{\mu\beta}}{k_\mu(k_\mu + k_\beta)}\right). \tag{A16}$$

We see by arguments similar to those for $\int_r^\infty (W(s))_{\alpha\beta}\,ds$, that

$$\int_r^\infty [UY]_{\alpha\beta}\,ds = \frac{f(k_\alpha + k_\beta, 2\lambda_0 + 2 - i(\eta_\alpha + \eta_\beta); r)}{-4(k_\alpha k_\beta)^{1/2}(k_\alpha + k_\beta)}\sum_{\mu=1}^n \left(\frac{u^{(0)}_{\alpha\mu} w^{(0)}_{\mu\beta}}{k_\mu(k_\mu + k_\beta)}\right)\left[1 + O\left(\left|\frac{1}{k^3 r}\right|\right)\right]. \tag{A17}$$

Consequently,

$$\| \int_r^\infty U(s)Y(s)\,ds\| = O\left[\left(\frac{v_{\lambda_0}}{k^2 r^{\lambda_0+1}}\right)^2\right] \quad r \geq R_M, \tag{A18}$$

$$(WY^*)_{\alpha\beta} = \left(\sum_{\mu=1}^{n} \frac{f(k_\alpha + k_\mu, \lambda_0 + 1 - i(\eta_\alpha + \eta_\mu); r)}{2i\,(k_\alpha k_\mu)^{1/2}} \frac{f(-k_\mu - k_\beta, \lambda_0 + 1 - i(-\eta_\mu - \eta_\beta); r)}{2(k_\mu k_\beta)^{1/2}(k_\mu + k_\beta)} w_{\alpha\mu}^{(0)} w_{\mu\beta}^{*(0)}\right)\left[1 + O\left(\frac{1}{k^2 r}\right)\right]$$

$$= \left(\sum_{\mu=1}^{n} f(k_\alpha - k_\beta, 2\lambda_0 + 2 - i(\eta_\alpha - \eta_\beta); r) \frac{w_{\alpha\mu}^{(0)} w_{\mu\beta}^{*(0)}}{4i(k_\alpha k_\mu^2)^{1/2}(k_\mu + k_\beta)}\right)\left[1 + O\,\frac{1}{k^2 r}\right]$$

$$= \frac{f(k_\alpha - k_\beta, 2\lambda_0 + 2 - i(\eta_\alpha - \eta_\beta); r)}{4i(k_\alpha k_\beta)^{1/2}}\left[1 + O\left(\frac{1}{k^2 r}\right)\right]\left(\sum_{\mu=1}^{n} \frac{w_{\alpha\mu}^{(0)} w_{\mu\beta}^{*(0)}}{k_\mu(k_\mu + k_\beta)}\right). \tag{A19}$$

Equation (A19) shows that $(WY^*)_{\alpha\beta}$ is a slowly varying function of r compared with say $(UY)_{\alpha\beta}$, since $|k_\alpha - k_\beta|$ is small compared with $k_\alpha + k_\beta$.

[1]F. Roesel, J.X. Saladin, and K. Alder, Comput. Phys. Commun. **8**, 35 (1974); K. Alder, F. Roesel, and J.X. Saladin, J. Phys. Soc. Jpn. Suppl. **34**, 146 (1973).

[2]K. Alder and H.K.A. Pauli, Nucl. Phys. **128**, 193 (1969).

[3]R.L. Mercer and D.G. Ravenhall, Phys. Rev. C **10**, 2022 (1974).

[4]L.D. Tolsma, J. Comput. Phys. **17**, 384 (1975); L.D. Tolsma, in *Proceedings of the International Conference on Nuclear Physics, Munich*, 1973, edited by J. deBoer and H.T. Mang (North-Holland, Amsterdam/American Elsevier, New York, 1973), Vol. 1, p. 385.

[5]D.L. Knirk, J. Chem. Phys. **57**, 4782 (1972); D.L. Knirk, E.F. Hayes, and D.J. Kouri, *ibid.* **57**, 4770 (1972).

[6]P.G. Burke and K. Smith, Rev. Mod. Phys. **34**, 458 (1962); P.G. Burke and H.M. Schey, Phys. Rev. **126**, 147 (1962); P.G. Burke and M.J. Seaton, *Methods in Computational Physics*, edited by B. Alder, S. Fernbach, and M. Rotenberg (Academic, New York, 1971), Vol. 10, p. 1; D.F. Korff, S. Chung, and C.C. Lin, Phys. Rev. A **7**, 545 (1973).

[7]G.H. Rawitscher and C.H. Rasmussen, article submitted to Comp. Phys. Commun.

[8]*Handbook of Mathematical Functions*, edited by M. Abramowitz and I. Stegun (Dover, New York, 1965).

[9]R. Courant and D. Hilbert, *Methods of Mathematical Physics* (Wiley, New York, 1966), Vol. I.

[10]Some very recent work by the authors of Ref. 1, Bull. Am. Phys. Soc. **21**, 12 (1976), suggests that significant improvements may be made concerning restriction (5.1b).

Ordering of the exponential of a quadratic in boson operators. I. Single mode case

C. L. Mehta

Physics Department, Indian Institute of Technology, New Delhi-110029, India and Service de Physique Atomique, Centre d'Etudes Nucléaires de Saclay, BP N° 2, 91190 Gif-Sur-Yvette, France

(Received 12 July 1976)

The Weyl ordered form of the operator $\exp[\alpha \hat{a}^2+\beta \hat{a}^{\dagger 2}+\gamma(\hat{a}^\dagger \hat{a}+\hat{a}\hat{a}^\dagger)]$ is derived in a very simple way. Using this result, we also obtain the normal and antinormal ordered forms as well as the diagonal coherent states representation of this operator.

1. INTRODUCTION

Very often one is interested in writing a given operator in a well-ordered form. Consider an arbitrary operator function $G(\hat{a}, \hat{a}^\dagger)$ of single mode Boson annihilation and creation operators satisfying the commutation relation

$$[\hat{a}, \hat{a}^\dagger]=1. \tag{1.1}$$

One may use this commutation relation to rearrange the operators $\hat{a}$ and $\hat{a}^\dagger$ occuring in G and write it in a desired form. When all powers of creation operator occur to the left of all powers of the annihilation operator, we say that the given expression is in the normal ordered form. Thus for example the normal ordered form of $\hat{a}\hat{a}^\dagger$ is $\hat{a}^\dagger\hat{a}+1$. The expression is in the antinormal ordered form, if all powers of the annihilation operator occur to the left of all powers of the creation operator. If on the other hand the form is completely symmetric in the ordering of $\hat{a}$ and $\hat{a}^\dagger$, such as $(\hat{a}\hat{a}^\dagger+\hat{a}^\dagger\hat{a})$ or $(a^{\dagger 2}a+a^\dagger a a^\dagger+a a^{\dagger 2})$ etc., we say that the operator is in the Weyl ordered form. Several other orderings have also been discussed in the literature.[1-3] The operator ordering plays an important role in phase space description of quantum mechanics,[4] quantum C-number correspondence,[1] etc.

We shall denote by G_N the normal ordered form of $\hat{G}$. On the other hand $:\hat{G}:$ will denote an operator obtained from $\hat{G}$ by arranging all powers of $\hat{a}^\dagger$ to the left of all powers of $\hat{a}$ *without* making use of the commutation relation (1.1). Thus, for example, if $\hat{G}=\hat{a}\hat{a}^\dagger$, then $\hat{G}_N=\hat{a}^\dagger\hat{a}+1$ and $:\hat{G}:=\hat{a}^\dagger\hat{a}$. By definition

$$G_N(\hat{a}, \hat{a}^\dagger)=:G_N(\hat{a}, \hat{a}^\dagger):. \tag{1.2}$$

In a similar way we shall denote by $\hat{G}_A$ the antinormal ordered form of $\hat{G}$ and by $\hat{G}_W$ the Weyl ordered form of $\hat{G}$. We also use the notation "$\hat{G}$" and $\{\hat{G}\}_W$ to denote operators obtained from $\hat{G}$ by simply putting it in the antinormal or Weyl ordered form, respectively, *without* making use of the commutation relation (1.1). It is to be noted that, in general, $\hat{G}\neq$"$\hat{G}$" etc. However, we always have $\hat{G}=\hat{G}_N=\hat{G}_A=\hat{G}_W$, and hence

$$G(\hat{a}, \hat{a}^\dagger)=:G_N(\hat{a}, \hat{a}^\dagger):="G_A(\hat{a}, \hat{a}^\dagger)"=\{G_W(\hat{a}, \hat{a}^\dagger)\}_W. \tag{1.3}$$

We also note that $G_N(v, v^*)$, $G_A(v, v^*)$ and $G_W(v, v^*)$ are the classical functions corresponding to the operator $\hat{G}$ in the normal ordering, antinormal ordering and Weyl's rules of association respectively[5] (cf. also Ref. 1).

Several methods are available to obtain a given ordered form of the operator. These include parametric differentiation,[6] Fourier transform algebra, use of coherent states,[1] etc. In particular, if $|\alpha\rangle$ denotes a coherent state, i.e., a normalized eigenstate of $\hat{a}$ with eigenvalue α,

$$\hat{a}|\alpha\rangle=\alpha|\alpha\rangle, \quad \langle\alpha|\hat{a}^\dagger=\alpha^*\langle\alpha| \tag{1.4}$$

(α being in general a complex number), and if $G_N(\hat{a}, \hat{a}^\dagger)$ is the normal ordered form of $G(\hat{a}, \hat{a}^\dagger)$, we readily find that

$$G_N(\alpha, \alpha^*)=\langle\alpha|G(\hat{a}, \hat{a}^\dagger)|\alpha\rangle. \tag{1.5}$$

The antinormal ordered form of $\hat{G}$ is closely related to its diagonal coherent state representation. We have the relation

$$G(\hat{a}, \hat{a}^\dagger)=(1/\pi)\int G_A(\alpha, \alpha^*)|\alpha\rangle\langle\alpha| d^2\alpha, \tag{1.6}$$

from which we find that[7]

$$G_A(v, v^*)=(1/\pi)\exp(|v|^2)\int\langle-\alpha|\hat{G}|\alpha\rangle \times\exp(|\alpha|^2)\exp(\alpha^* v-\alpha v^*)\, d^2\alpha. \tag{1.7}$$

An expression similar to (1.7) is also known[8,2] for the Weyl ordered form of $\hat{G}$:

$$G_W(v, v^*)=(2/\pi)\exp(2|v|^2)\int\langle-\alpha|G|\alpha\rangle \times\exp[2(\alpha^* v-\alpha v^*)]\, d^2\alpha. \tag{1.8}$$

The Fourier transforms of G_N, G_A, and G_W are related in the following manner. If $\Gamma(\alpha, \alpha^*)$ denotes the Fourier transform of $G(v, v^*)$,

$$\Gamma(\alpha, \alpha^*)=\int G(v, v^*)\exp(\alpha v^*-\alpha^* v)\, d^2v, \tag{1.9}$$

then

$$\Gamma_W(\alpha, \alpha^*)=\exp(|\alpha|^2/2)\Gamma_N(\alpha, \alpha^*) =\exp(-|\alpha|^2/2)\Gamma_A(\alpha, \alpha^*). \tag{1.10}$$

Since the inverse Fourier transform of $\exp(-|\alpha|^2/2)$ is $\exp(-2|v|^2)$, we find from the convolution theorem that

$$G_N(v, v^*)=(2/\pi)\int G_W(v', v'^*)\exp(-2|v-v'|^2)\, d^2v'. \tag{1.11}$$

We also have the relations

$$G_W(v, v^*)=(2/\pi)\int G_A(v', v'^*)\exp(-2|v-v'|^2 d^2v', \tag{1.12}$$

$$G_N(v, v^*) = (1/\pi) \int G_A(v', v'^*) \exp(-|v - v'|^2)\, d^2v'. \tag{1.13}$$

The various ordered forms of the operator $\exp(-\lambda \hat{a}^\dagger \hat{a})$ are well known[1,8,9] and are easily derived using coherent states. We have

$$\exp(-\lambda \hat{a}^\dagger \hat{a}) = :\exp\{(e^{-\lambda} - 1)\hat{a}^\dagger \hat{a}\}: \tag{1.14}$$

$$= \text{“}\exp\{\lambda - (e^{\lambda} - 1)\hat{a}^\dagger \hat{a}\}\text{”} \tag{1.15}$$

$$= \left\{ \frac{2e^{\lambda}}{e^{\lambda}+1} \exp\left(2 \frac{1 - e^{\lambda}}{1 + e^{\lambda}} \hat{a}^\dagger a \right) \right\}_W. \tag{1.16}$$

In the following we derive the various ordered forms of the exponential of a general quadratic:

$$\exp[\alpha \hat{a}^2 + \beta \hat{a}^{\dagger 2} + \gamma(\hat{a}^\dagger \hat{a} + \hat{a}\hat{a}^\dagger)]. \tag{1.17}$$

2. WEYL ORDERED FORM

In order to obtain the Weyl ordered form of the operator $\exp[\alpha \hat{a}^2 + \beta \hat{a}^{\dagger 2} + \gamma(\hat{a}^\dagger \hat{a} + \hat{a}\hat{a}^\dagger)]$, we first consider the special case when $\alpha = \beta = 0$. The Weyl ordered form of the operator

$$\exp[\gamma(\hat{a}^\dagger \hat{a} + \hat{a}\hat{a}^\dagger)] = e^{\gamma} \exp(2\gamma \hat{a}^\dagger \hat{a}) \tag{2.1}$$

may readily be written down using (1.16):

$$\exp[\gamma(\hat{a}^\dagger a + \hat{a}\hat{a}^\dagger)] = \{\operatorname{sech}\gamma \exp(2\hat{a}^\dagger \hat{a} \tanh\gamma)\}_W. \tag{2.2}$$

It may be observed that, even through the fact $\hat{a}^\dagger$ is Hermitian adjoint of $\hat{a}$ has been used in deriving (2.2), the result, being a consequence of rearranging powers of $\hat{a}$ and $\hat{a}^\dagger$, *only* depends on the commutation relation (1.1). Thus, for any two operators $\hat{d}$ and $\hat{c}$ for which $[\hat{d}, \hat{c}] = 1$, we find that

$$\exp[\gamma(\hat{c}\hat{d} + \hat{d}c)] = \{\operatorname{sech}\gamma \exp(2\hat{c}\hat{d} \tanh\gamma\}_W, \tag{2.3}$$

where the right-hand side is a completely symmetric ordered (Weyl ordered) form in $\hat{c}$ and $\hat{d}$.

We now consider the general case

$$\hat{G} = \exp[\alpha \hat{a}^2 + \beta \hat{a}^{\dagger 2} + \gamma(\hat{a}^\dagger \hat{a} + \hat{a}\hat{a}^\dagger)]. \tag{2.4}$$

We introduce two operators $\hat{d}$ and $\hat{c}$ which are linear combinations of $\hat{a}$ and $\hat{a}^\dagger$,

$$\hat{d} = x\hat{a} + y\hat{a}^\dagger, \tag{2.5a}$$

$$\hat{c} = z\hat{a} + t\hat{a}^\dagger, \tag{2.5b}$$

such that x, y, z, t are c-number quantities, such that the commutator

$$[\hat{d}, \hat{c}] = 1, \tag{2.6}$$

and such that

$$\alpha \hat{a}^2 + \beta \hat{a}^{\dagger 2} + \gamma(\hat{a}\hat{a}^\dagger + \hat{a}^\dagger \hat{a}) = \lambda(\hat{c}\hat{d} + \hat{d}\hat{c}). \tag{2.7}$$

It may readily be seen that λ is proportional to the discriminant of the quadratic,

$$\lambda = (\gamma^2 - \alpha\beta)^{1/2}. \tag{2.8}$$

When $\lambda \neq 0$, it is always possible to find x, y, z, and t which satisfy the above conditions. They are, however, not unique. A possible choice is given by

$$x = (1/2\lambda)(\gamma + \lambda), \quad y = \beta/2\lambda, \quad z = \beta(\gamma + \lambda)^{-1}, \quad t = 1. \tag{2.9}$$

We now use (2.3) to obtain a completely symmetric ordered form of $\exp[\lambda(\hat{c}\hat{d} + \hat{d}\hat{c})]$ in $\hat{c}$ and $\hat{d}$. We also observe that since $\hat{c}$ and $\hat{d}$ are both linear combinations of $\hat{a}$ and $\hat{a}^\dagger$, the resulting expression is also completely symmetric in the ordering of $\hat{a}$ and $\hat{a}^\dagger$. Hence we find that

$$\exp[\alpha \hat{a}^2 + \beta \hat{a}^{\dagger 2} + \gamma(\hat{a}^\dagger \hat{a} + \hat{a}\hat{a}^\dagger)] = \exp[\lambda(\hat{c}\hat{d} + \hat{d}\hat{c})]$$

$$= \operatorname{sech}\lambda \{\exp(2\hat{c}\hat{d} \tanh\lambda\}_W$$

$$\equiv \operatorname{sech}\lambda \left\{ \exp\left[\frac{\tanh\lambda}{\lambda} (\alpha \hat{a}^2 + \beta \hat{a}^{\dagger 2} + 2\gamma \hat{a}^\dagger \hat{a}) \right] \right\}_W. \tag{2.10}$$

The case $\lambda = 0$ does not present any difficulty, since in this case we may write

$$\alpha \hat{a}^2 + \beta \hat{a}^{\dagger 2} + \gamma(\hat{a}^\dagger \hat{a} + \hat{a}\hat{a}^\dagger) = (\sqrt{\alpha}\hat{a} + \sqrt{\beta}\hat{a}^\dagger)^2. \tag{2.11}$$

The operator $\exp[\alpha \hat{a}^2 + \beta \hat{a}^{\dagger 2} + \gamma(\hat{a}^\dagger \hat{a} + \hat{a}\hat{a}^\dagger)]$ when expanded is then given by

$$\sum_{n=0}^{\infty} \frac{(\sqrt{\alpha}\hat{a} + \sqrt{\beta}\hat{a}^\dagger)^{2n}}{n!}. \tag{2.12}$$

Each term of this expansion is already in the Weyl ordered form. Hence we find, when $\gamma^2 = \alpha\beta$, that

$$\exp[\alpha \hat{a}^2 + \beta \hat{a}^{\dagger 2} + \gamma(\hat{a}^\dagger \hat{a} + \hat{a}\hat{a}^\dagger)]$$

$$\equiv \{\exp[\alpha \hat{a}^2 + \beta \hat{a}^2 + \gamma(\hat{a}^\dagger \hat{a} + \hat{a}\hat{a}^\dagger)]\}_W. \tag{2.13}$$

Equation (2.13) is in agreement with (2.10) in the limit $\lambda = 0$. Hence we find, in general, that

$$\exp[\alpha \hat{a}^2 + \beta \hat{a}^{\dagger 2} + \gamma(\hat{a}^\dagger \hat{a} + \hat{a}\hat{a}^\dagger)]$$

$$= \operatorname{sech}\lambda \left\{ \exp\left[\frac{\tanh\lambda}{\lambda} (\alpha \hat{a}^2 + \beta a^{\dagger 2} + 2\gamma \hat{a}^\dagger \hat{a}) \right] \right\}_W, \tag{2.14}$$

where λ is given by (2.8).

3. NORMAL ORDERED FORM

The normal ordered form of the exponential of a general quadratic has been derived earlier, using parametric differentiation[6] (cf. also Ref. 10). We may also derive this form by making use of Eqs. (2.14) and (1.11).

From (2.14) and (1.11), we find that

$$G_N(v, v^*) = \frac{2\operatorname{sech}\lambda}{\pi} \int \exp(-2|v - v'|^2) \times \exp\left\{ \frac{\tanh\lambda}{\lambda} (\alpha v^2 + \beta v^{*2} + 2\gamma |v|^2) \right\} d^2v. \tag{3.1}$$

We assume that α, β, γ are such that the integral on the right hand side of (3.1) is well defined. The evaluation of the integral is long, but straightforward, and we obtain on simplification the following expression for G_N:

$$G_N(v, v^*) = \left(\cosh 2\lambda - \frac{\gamma}{\lambda} \sinh 2\lambda \right)^{-1/2} \times \exp\left[\frac{\alpha v^2 + \beta v^{*2} + 2(\gamma - \lambda \tanh\lambda)|v|^2}{2(\lambda \coth 2\lambda - \gamma)} \right]. \tag{3.2}$$

Hence we write

$$\exp[\alpha \hat{a}^2 + \beta \hat{a}^{\dagger 2} + \gamma(\hat{a}^\dagger \hat{a} + \hat{a}\hat{a}^\dagger)]$$

$$= \left(\cosh 2\lambda - \frac{\gamma}{\lambda} \sinh 2\lambda \right)^{-1/2} \times :\exp\left[\frac{\alpha \hat{a}^2 + \beta \hat{a}^{\dagger 2} + 2(\gamma - \lambda \tanh\lambda)\hat{a}^\dagger \hat{a}}{2(\lambda \coth 2\lambda - \gamma)} \right]:. \tag{3.3}$$

It may be remarked that, even though (3.2) was obtained from (3.1), under certain restrictions, the final result, being analytic in α, β, γ, is valid for all values of α, β, γ, even in the limit as λ approaches zero except for the singular case when $\lambda\cosh 2\lambda = \gamma \sinh 2\lambda$.

4. ANTINORMAL ORDERED FORM

The antinormal ordered form of (1.17) may be obtained from Eqs. (1.10) and (2.14) or from Eqs. (1.7) and (3.3). Alternatively, one may obtain this expression also directly from (3.3). We rewrite (3.3) in the following form. Let $\hat{d}$ and $\hat{c}$ be any two operators with $[\hat{d}, \hat{c}] = 1$, then from (3.3) we find that

$$\exp[\alpha\hat{d}^2 + \beta\hat{c}^2 + \gamma(\hat{c}\hat{d} + \hat{d}\hat{c})] = \left(\cosh 2\lambda - \frac{\gamma}{\lambda}\sinh 2\lambda\right)^{-1/2} \times \left\{\exp \frac{\alpha\hat{d}^2 + \beta\hat{c}^2 + 2(\gamma - \lambda\tanh\lambda)\hat{c}\hat{d}}{2(\lambda\coth 2\lambda - \gamma)}\right\}_{cd}, \tag{4.1}$$

where $\{\cdots\}_{cd}$ denotes the ordering such that c always occurs to the left of $\hat{d}$. We now write

$$\hat{d} = \hat{a}^\dagger, \quad \hat{c} = -\hat{a} \tag{4.2}$$

so that $[\hat{d}, \hat{c}] = 1$ is still satisfied. Hence from (4.1) we obtain

$$\exp[\alpha\hat{a}^{\dagger 2} + \beta\hat{a}^2 - \gamma(\hat{a}^\dagger\hat{a} + \hat{a}\hat{a}^\dagger)] = \left(\cosh 2\lambda - \frac{\gamma}{\lambda}\sin 2\lambda\right)^{-1/2} \times {:}\exp\left[\frac{\alpha\hat{a}^{\dagger 2} + \beta\hat{a}^2 - 2(\gamma - \lambda\tanh\lambda)\hat{a}^\dagger\hat{a}}{2(\lambda\coth 2\lambda - \gamma)}\right]{:}. \tag{4.3}$$

Redefining the constants α, β, and λ, we find that

$$\exp[\alpha\hat{a}^2 + \beta\hat{a}^{\dagger 2} + \gamma(\hat{a}^\dagger\hat{a} + \hat{a}\hat{a}^\dagger)] = \left(\cosh 2\lambda + \frac{\gamma}{\lambda}\sinh 2\lambda\right)^{-1/2} \times {:}\exp\left[\frac{\alpha\hat{a}^2 + \beta\hat{a}^{\dagger 2} + 2(\gamma + \lambda\tanh\lambda)\hat{a}^\dagger\hat{a}}{2(\lambda\coth 2\lambda + \gamma)}\right]{:}. \tag{4.4}$$

As before, the result is valid for all values of α, β, and γ except for the singular case when $\lambda\cosh 2\lambda + \gamma\sinh 2\lambda = 0$.

Whenever the operator $\exp\{\alpha\hat{a}^2 + \beta\hat{a}^{\dagger 2} + \gamma(\hat{a}^\dagger\hat{a} + \hat{a}\hat{a}^\dagger)\}$ is a bounded operator, we may write its diagonal coherent state representation using Eqs. (4.4) and (1.6):

$$\exp\{\alpha\hat{a}^2 + \beta\hat{a}^{\dagger 2} + \gamma(\hat{a}^\dagger\hat{a} + \hat{a}\hat{a}^\dagger)\} = \frac{1}{\pi}\left(\cosh 2\lambda + \frac{\gamma}{\lambda}\sinh 2\lambda\right)^{-1/2} \times \int \exp\left(\frac{\alpha v^2 + \beta v^{*2} + 2(\gamma + \lambda\tanh\lambda)|v|^2}{2(\lambda\coth 2\lambda + \gamma)}\right) \times |v\rangle\langle v| d^2v. \tag{4.5}$$

5. SOME GENERALIZATIONS

In this section we derive some generalizations of Eqs. (2.14), (3.3), and (4.4).

We rewrite Eq. (2.14) in the form

$$\exp[\alpha\hat{d}^2 + \beta\hat{c}^2 + \gamma(\hat{c}\hat{d} + \hat{d}\hat{c})] = \left\{\operatorname{sech}\lambda \exp\left[\frac{\tanh\lambda}{\lambda}(\alpha\hat{d}^2 + \beta\hat{c}^2 + 2\gamma\hat{c}\hat{d})\right]\right\}_W, \tag{5.1}$$

where $\hat{d}$ and $\hat{c}$ are any two operators for which $[\hat{d}, \hat{c}] = 1$ and $\lambda = (\gamma^2 - \alpha\beta)^{1/2}$. We now consider a general polynomial

$$\hat{P} = \alpha\hat{a}^2 + \beta\hat{a}^{\dagger 2} + \gamma(\hat{a}^\dagger\hat{a} + \hat{a}\hat{a}^\dagger) + 2\delta\hat{a} + 2\epsilon\hat{a}^\dagger \tag{5.2}$$

and write it in the form

$$\hat{P} = \alpha(\hat{a} + x)^2 + \beta(\hat{a}^\dagger + y)^2 + \gamma[(\hat{a}^\dagger + y)(\hat{a} + x) + (\hat{a} + x)(\hat{a}^\dagger + y)] + \theta. \tag{5.3}$$

One may readily verify, when $\lambda \neq 0$, that

$$x = \lambda^{-2}(\epsilon\gamma - \beta\delta), \tag{5.4}$$

$$y = \lambda^{-2}(\gamma\delta - \alpha\epsilon), \tag{5.5}$$

$$\theta = \lambda^{-2}(\alpha\epsilon^2 + \beta\delta^2 - 2\gamma\epsilon\delta). \tag{5.6}$$

By identifying $\hat{d} = \hat{a} + x$, $\hat{c} = \hat{a}^\dagger + y$, we then obtain, from (5.1), the relation

$$e^{\hat{P}} = e^\theta \operatorname{sech}\lambda \left\{\exp\left[\frac{\tanh\lambda}{\lambda}(\hat{P} - \theta)\right]\right\}_W. \tag{5.7}$$

Proceeding in a strictly analogous manner, we obtain the following generalizations of (3.3) and (4.4) (cf. Ref. 6):

$$e^{\hat{P}} = \frac{e^\theta}{[\cosh 2\lambda - (\gamma/\lambda)\sinh 2\lambda]^{1/2}} \times {:}\exp\left[\frac{\hat{P} - 2\lambda(\hat{a}^\dagger + y)(\hat{a} + x)\tanh\lambda - \theta}{2(\lambda\coth 2\lambda - \gamma)}\right]{:} \tag{5.8}$$

$$= \frac{e^\theta}{(\cosh 2\lambda + \gamma/\lambda \sinh 2\lambda)^{1/2}} \times {:}\exp\left[\frac{P + 2\lambda(\hat{a}^\dagger + y)(\hat{a} + x)\tanh\lambda - \theta}{2(\lambda\coth 2\lambda + \gamma)}\right]{:}, \tag{5.9}$$

where x, y, θ, and λ are given by Eqs. (5.4)—(5.6) and (2.8). Equations (5.7)—(5.9) are valid even in the limit as λ approaches zero.

We introduce a parameter σ and rewrite (5.7) in the form

$$\exp(\sigma\hat{P}) = \exp(\sigma\theta)\operatorname{sech}\sigma\lambda \left\{\exp\left[\frac{\tanh\lambda\sigma}{\lambda}(\hat{P} - \theta)\right]\right\}_W. \tag{5.10}$$

On differentiating (5.10) with respect to σ a number of times and then setting $\sigma = 0$, we may obtain the Weyl ordered form of any power of $\hat{P}$. In fact we may formally write for any arbitrary function of $\hat{P}$,

$$f(\hat{P}) = \left\{\left[f\left(\frac{d}{d\sigma}\right)\exp(\sigma\theta)\operatorname{sech}\sigma\lambda \times \exp\left(\frac{\tanh\lambda\sigma}{\lambda}(\hat{P} - \theta)\right)\right]_{\sigma=0}\right\}_W. \tag{5.11}$$

Similar expressions may be written for the normal and antinormal ordered forms of $f(\hat{P})$.

[1]C. L. Mehta, J. Phys. A (Proc. Phys. Soc. London) 1, 385 (1968).
[2]G. S. Agarawal and E. Wolf, Phys. Rev. D 2, 2161 (1970).
[3]L. Cohen, J. Math. Phys. 7, 781 (1965).
[4]C. L. Mehta, J. Math. Phys. 5, 677 (1964).
[5]J. R. Shewell, Am. J. Phys. 27, 16 (1959).
[6]R. M. Wilcox, J. Math. Phys. 8, 962 (1967).
[7]C. L. Mehta, Phys. Rev. Lett. 18, 752 (1967).
[8]C. L. Mehta, in *Lectures in Theoretical Physics, Vol. 12A*, Proc. of Summer Institute 1969 (Univ. of Colorado Press, Boulder, Colorado), p. 95.
[9]W. H. Louisell, *Radiation and Noise in Quantum Electronics* (McGraw-Hill, New York, 1964).
[10]N. H. McCoy, Proc. Edinburgh Math. Soc. 3, 118 (1932).

Ordering of the exponential of a quadratic in boson operators. II. Multimode case

G. P. Agrawal*

Laboratoire d'Optique Quantique,† Ecole Polytechnique, 91128-Palaiseau, France

C. L. Mehta

Physics Department, Indian Institute of Technology, New Delhi 110029, India
and Service de Physique Atomique, C.E.N. de Saclay 91190-Gif-sur-Yvette, France
(Received 7 September 1976)

We derive the Weyl, the normal, and the antinormal ordered forms of the exponential of a multimode quadratic expression in boson operators. The trace of this exponential operator is also evaluated.

1. INTRODUCTION

In Paper I,[1] one of us obtained the various ordered forms of the exponential of a quadratic in single mode boson operators. It is of interest to consider the more general multimode case, and derive the various ordered forms of the exponential of the general quadratic.

Let $\{\hat{a}_i, \hat{a}_i^\dagger\}$ be a set of boson annihilation and creation operators, satisfying the usual commutation relations

$$[\hat{a}_i, \hat{a}_j^\dagger] = \delta_{ij}, \tag{1.1}$$

$$[\hat{a}_i, \hat{a}_j] = [\hat{a}_i^\dagger, \hat{a}_j^\dagger] = 0 \quad (i,j = 1,2,\ldots,n). \tag{1.2}$$

It is convenient to express the set as a $2n$-dimensional column vector $\hat{A}$,

$$\hat{\mathbf{A}} = \{\hat{a}_1, \ldots, \hat{a}_n, \hat{a}_1^\dagger, \ldots, \hat{a}_n^\dagger\}. \tag{1.3}$$

The commutation relations (1.1) and (1.2) may now be expressed in the form

$$[\hat{A}_i, \hat{A}_j] = z_{ij}, \tag{1.4a}$$

or

$$\hat{\mathbf{A}}\tilde{\hat{\mathbf{A}}} - \tilde{\hat{\mathbf{A}}}\hat{\mathbf{A}} = \mathbf{z}, \tag{1.4b}$$

where $\mathbf{z}$ is the $(2n \times 2n)$ matrix

$$\mathbf{z} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}; \tag{1.5}$$

$\tilde{\mathbf{R}}$ denotes the transpose of the matrix $\mathbf{R}$, and 0 and 1 are $(n \times n)$ null and unit matrices, respectively. The general second order monomial in boson operators (ignoring the constant term)

$$\hat{P} = \sum_{i,j=1}^{n} [\alpha_{ij}\hat{a}_i\hat{a}_j + \beta_{ij}\hat{a}_i^\dagger\hat{a}_j^\dagger + \gamma_{ij}(\hat{a}_i\hat{a}_j^\dagger + \hat{a}_j^\dagger\hat{a}_i)] + 2\sum_{i=1}^{n}(\delta_i\hat{a}_i + \epsilon_i\hat{a}_i^\dagger) \tag{1.6}$$

may be written in the form

$$\hat{P} = \tilde{\hat{\mathbf{A}}}\xi\hat{\mathbf{A}} + 2\tilde{\eta}\hat{\mathbf{A}}, \tag{1.7}$$

where ξ is the $(2n \times 2n)$ symmetric matrix

$$\xi = \begin{pmatrix} \alpha & \gamma \\ \tilde{\gamma} & \beta \end{pmatrix} \tag{1.8}$$

and η is the column vector

$$\eta = \{\delta_1, \ldots, \delta_n, \epsilon_1, \ldots, \epsilon_n\}. \tag{1.9}$$

Matrices α, β, γ are each $(n \times n)$, of which α and β are symmetric.

In this paper we derive the Weyl, the normal, and the antinormal ordered forms of the operator $e^{\hat{P}}$. We use the notation $\hat{G}_W$ for the Weyl ordered form of $\hat{G}$. Similarly $\hat{G}_N$ and $\hat{G}_A$ denote the normal and the antinormal ordered forms, respectively, of $\hat{G}$. On the other hand, $\{\hat{G}\}_W$ will denote the operator obtained from $\hat{G}$ by putting it in the Weyl ordered form without making use of communtation relations (1.4). Similarly $\{\hat{G}\}_N = :\hat{G}:$ and $\{\hat{G}\}_A \equiv$ "$\hat{G}$" denote the normal and the antinormal ordering operations on $\hat{G}$. Thus we have

$$\hat{G}(\hat{\mathbf{A}}) = \{\hat{G}_\mu(\hat{\mathbf{A}})\}_\mu, \quad \mu = W, N, \text{ or } A. \tag{1.10}$$

Our results derived in Secs. 2—4 may be summarized as follows:

$$e^{\hat{P}} = K_\mu\{\exp(\tilde{\hat{\mathbf{A}}}\xi\mathbf{T}_\mu\hat{\mathbf{A}} + 2\tilde{\eta}\mathbf{T}_\mu\hat{\mathbf{A}})\}_\mu, \tag{1.11}$$

where $\mu = W, N$, or A, K_μ is a constant,

$$K_\mu = \left|\frac{\sinh\mathbf{z}\xi}{\mathbf{z}\xi}\right|^{-1/2} |\mathbf{T}_\mu|^{1/2} \exp[\tilde{\eta}(\mathbf{T}_\mu - 1)\xi^{-1}\eta], \tag{1.12}$$

($|\mathbf{R}|$ denotes the determinant of $\mathbf{R}$) and $\mathbf{T}_\mu$ is the matrix

$$T_\mu = [(\sinh\mathbf{z}\xi)/\mathbf{z}\xi][\cosh\mathbf{z}\xi + k_\mu\,\mathbf{y}\mathbf{z}\sinh\mathbf{z}\xi]^{-1}. \tag{1.13}$$

Here $\mathbf{z}$ is the antisymmetric matrix defined in (1.5), $\mathbf{y}$ is the symmetric matrix

$$\mathbf{y} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \tag{1.14}$$

and $k_W = 0$, $k_N = 1$, $k_A = -1$.

It may readily be verified that Eq. (1.11) agrees with previously known special cases. For the single mode case ($n = 1$), Eq. (1.11) reduces to the results derived in Paper I (cf. also Refs. 2 and 3). For the case when $\alpha = \beta = 0$, the normal and the antinormal ordered forms agree with those obtained in Ref. 4. Using methods involving group theory and parametric differentiation, Berezin[5] has also obtained the normal ordered form of $\exp(\hat{P})$ when $\hat{P}$ is Hermitian (i.e., when $\alpha = \beta^*$, $\delta = \epsilon^*$ and $\gamma = \gamma^*$). Equation (1.11) in this special case is in agreement with his result.

We also obtain in Sec. 4, the trace of $\exp(\hat{P})$, whenever it is a trace class operator. We find that

$$\mathrm{Tr}(e^{\hat{P}}) = (1/2)^n\,|\mathbf{y}\sinh\mathbf{z}\xi|^{-1/2}\exp(-\eta\xi^{-1}\eta). \tag{1.15}$$

2. WEYL ORDERED FORM

We first consider the homogeneous quadratic

$$\hat{Q} = \hat{\tilde{\mathbf{A}}} \xi \hat{\mathbf{A}}. \tag{2.1}$$

Let us make a linear symplectic transformation

$$\hat{\mathbf{B}} = \{\hat{d}_1, \ldots, \hat{d}_n, \hat{c}_1, \ldots, \hat{c}_n\} = \mathbf{S}\hat{\mathbf{A}}, \tag{2.2}$$

which reduces the quadratic $\hat{Q}$ to a simpler form

$$\hat{Q} = \sum_{i=1} \lambda_i (\hat{c}_i \hat{d}_i + \hat{d}_i \hat{c}_i) = \hat{\tilde{\mathbf{B}}} \Lambda \hat{\mathbf{B}}. \tag{2.3}$$

The symplectic transformation $\mathbf{S}$ satisfies the identity

$$\mathbf{S}\mathbf{z}\tilde{\mathbf{S}} = \mathbf{z}, \tag{2.4}$$

so that the components of $\hat{\mathbf{B}}$ satisfy the same commutation relations as those of $\hat{\mathbf{A}}$:

$$[\hat{B}_i, \hat{B}_j] = z_{ij}. \tag{2.5}$$

The matrices ξ and Λ are related according as

$$\xi = \tilde{\mathbf{S}} \Lambda \mathbf{S} \tag{2.6}$$

or

$$\Lambda = \tilde{\mathbf{S}}^{-1} \xi \mathbf{S}^{-1}. \tag{2.7}$$

From (2.3) we find that Λ is of the form

$$\Lambda = \begin{pmatrix} 0 & \lambda \\ \lambda & 0 \end{pmatrix}, \tag{2.8}$$

where the $n \times n$ matrix λ is diagonal

$$\lambda_{ij} = \lambda_i \delta_{ij}. \tag{2.9}$$

Thus the required symplectic transformation ($\mathbf{S}^{-1}$) is the one which reduces ξ to the form (2.8). The existence of such a generalized Bogoliubov transformation can be established, since ξ is symmetric. We shall, however, not require an explicit expression for $\mathbf{S}$.

We now proceed to obtain the Weyl ordered form of the exponential

$$\hat{G} = e^{\hat{Q}}. \tag{2.10}$$

From (2.3) we can write

$$\hat{G} = \prod_i \exp[\lambda_i (\hat{c}_i \hat{d}_i + \hat{d}_i \hat{c}_i)]. \tag{2.11}$$

The operators $\hat{c}_i$ and $\hat{d}_i$ satisfy

$$[\hat{d}_i, \hat{c}_i] = 1, \quad i = 1, 2, \ldots, n; \tag{2.12}$$

whereas those with different subscripts commute with each other. From Eq. I(2.3) (of Paper I) we may then write

$$\hat{G} = \prod_i \{\mathrm{sech}\lambda_i \exp(2 \tanh\lambda_i \hat{d}_i \hat{c}_i)\}_W, \tag{2.13}$$

where the subscript W stands for the Weyl ordering operation in $\hat{c}$ and $\hat{d}$. However, since $\hat{c}$ and $\hat{d}$ are linear combinations of $\hat{a}$ and $\hat{a}^*$, the Weyl ordered form in $\hat{c}$ and $\hat{d}$ is also the Weyl ordered form in $\hat{a}$ and $\hat{a}^*$. We rewrite (2.13) in the matrix notation

$$\hat{G} = \{|\Lambda_1|^{1/2} \exp(\hat{\tilde{\mathbf{B}}} \Lambda_2 \hat{\mathbf{B}})\}_W, \tag{2.14}$$

where Λ_1 and Λ_2 are given by

$$\Lambda_1 = \begin{pmatrix} \mathrm{sech}\lambda & 0 \\ 0 & \mathrm{sech}\lambda \end{pmatrix}, \tag{2.15}$$

$$\Lambda_2 = \begin{pmatrix} 0 & \tanh\lambda \\ \tanh\lambda & 0 \end{pmatrix}, \tag{2.16}$$

and $|\Lambda_1|$ stands for the determinant of Λ_1. Using Eqs. (2.8) and (1.5) and observing the fact that $\mathrm{sech}\lambda$ is an even function of λ, we can rewrite (2.15) in the form

$$\Lambda_1 = \mathrm{sech}\begin{pmatrix} \lambda & 0 \\ 0 & -\lambda \end{pmatrix} = \mathrm{sech}(\mathbf{z}\Lambda). \tag{2.17}$$

Further from Eqs. (2.8) and (2.16) we also find that

$$\mathbf{z}\Lambda_2 = \tanh(\mathbf{z}\Lambda). \tag{2.18}$$

From Eqs. (2.14), (2.17), and (2.18) we obtain

$$\hat{G} = |\,\mathrm{sech}\mathbf{z}\Lambda\,|^{1/2} \{\exp[\hat{\tilde{\mathbf{B}}} \tilde{\mathbf{z}} \tanh(\mathbf{z}\Lambda) \hat{\mathbf{B}}]\}_W. \tag{2.19}$$

Using the properties of the symplectic matrices discussed in the Appendix [cf. Eqs. (A7)], we finally rewrite (2.19) in a form which does not contain $\mathbf{S}$ explicitly:

$$\hat{\mathbf{G}} = |\,\mathrm{sech}\mathbf{z}\xi\,|^{1/2} \{\exp(\hat{\tilde{\mathbf{A}}} \xi \mathbf{T}_W \hat{\mathbf{A}})\}_W, \tag{2.20}$$

where

$$\mathbf{T}_W = \mathbf{T}_W(\xi) = (\tanh\mathbf{z}\xi)/\mathbf{z}\xi. \tag{2.21}$$

It is of interest to observe that we may also write

$$\hat{G} = |\,\mathrm{sech}\mathbf{z}\xi\,|^{1/2} \{\exp(\hat{\tilde{\mathbf{A}}} \tilde{\mathbf{T}}_W \xi \hat{\mathbf{A}})\}_W, \tag{2.22}$$

or more symmetrically as

$$\hat{G} = |\,\mathrm{sech}\mathbf{z}\xi\,|^{1/2} \{\exp(\hat{\tilde{\mathbf{A}}} \tilde{\mathbf{T}}_W^{1/2} \xi \mathbf{T}_W^{1/2} \hat{\mathbf{A}})\}_W. \tag{2.23}$$

Hence, if we define

$$\mathbf{A}' = \mathbf{T}_W^{1/2} \mathbf{A}, \tag{2.24}$$

we find that

$$\exp(\hat{\tilde{\mathbf{A}}} \xi \hat{\mathbf{A}}) = |\,\mathrm{sech}\mathbf{z}\xi\,|^{1/2} \{\exp(\hat{\tilde{\mathbf{A}}}' \xi \hat{\mathbf{A}}')\}_W, \tag{2.25}$$

i.e., the Weyl ordered form of the exponential of a homogeneous quadratic is apart from a multiplicative constant, the exponential of the same quadratic in transformed operators.

The transformation matrix $T_W^{1/2}$ is an even function of $\mathbf{z}\xi$. Also from Eqs. (1.5) and (1.8) we find that

$$(\mathbf{z}\xi)^2 = \begin{pmatrix} \tilde{\gamma}^2 - \beta\alpha & \tilde{\gamma}\beta - \beta\gamma \\ \gamma\alpha - \alpha\tilde{\gamma} & \gamma^2 - \alpha\beta \end{pmatrix}. \tag{2.26}$$

Hence, when $\beta\gamma$ and $\gamma\alpha$ are both symmetric, $(\mathbf{z}\xi)^2$ is diagonal and in this case the transformation (2.24) is equivalent to multiplying different components of A by constants (no mixing of the components). Of course, further simplification occurs when $(\mathbf{z}\xi)^2$ is a multiple of unit matrix.

In the special case when $\gamma^2 = \alpha\beta$ and both $\beta\gamma$ and $\gamma\alpha$ symmetric, the matrix $(\mathbf{z}\xi)^2 = 0$. In this singular case

$$(\tanh\mathbf{z}\xi)/\mathbf{z}\xi = \mathrm{sech}\mathbf{z}\xi = 1, \tag{2.27}$$

and the exponential is already in the Weyl ordered form.

We now include the linear terms also in the quadratic. Let $\hat{P}$ be the general second order monomial [Eq. (1.7)]

$$\hat{P} = \hat{\tilde{\mathbf{A}}} \xi \hat{\mathbf{A}} + 2\tilde{\eta}\hat{\mathbf{A}},$$

which may be written as

$$\hat{P} = (\hat{\mathbf{A}} + \tilde{\eta}\xi^{-1})\xi(\hat{\mathbf{A}} + \xi^{-1}\eta) - \tilde{\eta}\xi^{-1}\eta. \tag{2.28}$$

Since $\xi^{-1}\eta$ is a c-number matrix, the components of $\hat{\mathbf{A}} + \xi^{-1}\eta$ satisfy the same commutation relations as those of $\hat{\mathbf{A}}$, and hence, using (2.20), we obtain the following expression for the Weyl ordered form of $e^{\hat{P}}$:

$$e^{\hat{P}} = |\operatorname{sechz}\xi|^{1/2} \exp(\tilde{\eta}(\mathbf{T}_W - 1)\xi^{-1}\eta) \times \{\exp(\hat{\tilde{\mathbf{A}}}\xi\mathbf{T}_W\hat{\mathbf{A}} + 2\tilde{\eta}\mathbf{T}_W\hat{\mathbf{A}})\}_W, \tag{2.29}$$

where the matrix $\mathbf{T}_W$ depends only on ξ and is given by Eq. (2.21).

3. NORMAL ORDERED FORM

The method employed for obtaining the normal ordered form of the exponential in the single-mode case may readily be generalized for the multimode case. The relation [Eq. I(1.11)]

$$G_N(v, v^*) = (2/\pi)\int G_W(v, v^{*\prime}) \exp(-2|v - v'|^2 d^2v' \tag{3.1}$$

valid for the single mode case now generalizes to

$$G_N(\mathbf{V}) = (2/\pi)^n \int G_W(\mathbf{V}') \exp[-(\mathbf{V}^+ - \mathbf{V}'^+)(\mathbf{V} - \mathbf{V}')] d^{2n}\mathbf{V}', \tag{3.2}$$

where $G_N(\hat{\mathbf{A}})$ and $G_W(\hat{\mathbf{A}})$ are the normal and the Weyl ordered forms respectively of $\hat{G}$. $\mathbf{V}$ is the column vector

$$\mathbf{V} = \{v_1, \ldots, v_n, v_1^*, \ldots, v_n^*\}, \tag{3.3a}$$

$\mathbf{V}^+$ its Hermitian adjoint (row vector),

$$\mathbf{V}^+ = (v_1^*, \ldots, v_n^*, v_1, \ldots, v_n), \tag{3.3b}$$

and $d^{2n}V$ stands for

$$d^{2n}V = \prod_{i=1}^{n} d^2v_i = \prod_{i=1}^{n} d(\mathrm{Re}v_i)\ d(\mathrm{Im}v_i). \tag{3.4}$$

If we take $\hat{G}$ to be that given by Eq. (2.10) and use also Eqs. (2.20) and (2.21), we find from (3.2) that

$$G_N(\mathbf{V}) = (2/\pi)^n |\operatorname{sechz}\xi|^{1/2} \exp(-\mathbf{V}^+\mathbf{V}) \times \int \exp\{\tilde{\mathbf{V}}'\xi[(\tanh\mathbf{z}\xi)/\mathbf{z}\xi]\mathbf{V}' - \mathbf{V}'^+\mathbf{V}' + \mathbf{V}'^+\mathbf{V} + \mathbf{V}^+\mathbf{V}'\} d^{2n}V'. \tag{3.5}$$

Observing that

$$\tilde{\mathbf{V}}' = \mathbf{V}'^+\mathbf{y}, \tag{3.6}$$

where $\mathbf{y}$ is the matrix defined in Eq. (1.14), and using the Fourier transform result[6]

$$\int \exp(-\mathbf{V}'^+\mathbf{X}\mathbf{V}' + \mathbf{V}'^+\mathbf{V} + \mathbf{V}^+\mathbf{V}')\, d^{2n}V' = (\pi/2)^n |X|^{-1/2} \exp(\mathbf{V}^+\mathbf{X}^{-1}\mathbf{V}), \tag{3.7}$$

we obtain from (3.5), after simplification, the following expression for $G_N(\mathbf{V})$:

$$G_N(\mathbf{V}) = |\cosh\mathbf{z}\xi - \mathbf{z}\mathbf{y}\sinh\mathbf{z}\xi|^{-1/2} \times \exp[\tilde{\mathbf{V}}\xi\mathbf{T}_N\mathbf{V}], \tag{3.8}$$

where

$$\mathbf{T}_N = T_N(\xi) = (\tanh\mathbf{z}\xi)\,(\mathbf{z}\xi - \mathbf{y}\xi\tanh\mathbf{z}\xi)^{-1}. \tag{3.9}$$

We rewrite (3.8) in terms of the annihilation and creation operators to obtain the required normal ordered form[7]

$$\exp(\hat{\tilde{\mathbf{A}}}\xi\hat{\mathbf{A}}) = |\cosh\mathbf{z}\xi - \mathbf{z}\mathbf{y}\sinh\mathbf{z}\xi|^{-1/2} \times :\exp(\hat{\tilde{\mathbf{A}}}\xi\mathbf{T}_N\hat{\mathbf{A}}):. \tag{3.10}$$

In analogy with the case of Weyl ordering, we may also include the linear terms in the quadratic. Thus we obtain [cf. Eq. (2.28)]

$$\exp(\hat{\tilde{\mathbf{A}}}\xi\hat{\mathbf{A}} + 2\tilde{\eta}\hat{\mathbf{A}}) = |\cosh\mathbf{z}\xi - \mathbf{z}\mathbf{y}\sinh\mathbf{z}\xi|^{-1/2} \exp[\tilde{\eta}(\mathbf{T}_N - 1)\xi^{-1}\eta] \times :\exp(\hat{\tilde{\mathbf{A}}}\xi\mathbf{T}_N\mathbf{A} + 2\eta\mathbf{T}_N\hat{\mathbf{A}}):. \tag{3.11}$$

It is to be observed that, even though Eq. (3.10) and (3.11) have been obtained under certain restrictions,[6] these, being analytic expressions, are valid for all ξ and η except in the singular case when

$$|\cosh\mathbf{z}\xi - \mathbf{z}\mathbf{y}\sinh\mathbf{z}\mathbf{y}| = 0. \tag{3.12}$$

In this case, the normal ordered form does not exist. It may, however, be noted that the Weyl ordered form [Eq. (2.29)] always exists.

4. ANTINORMAL ORDERED FORM

As in the single mode case, we may directly use Eq. (3.10) for obtaining the antinormal ordered form. Since only the commutation relations are of significance in obtaining a particularly ordered form of an operator, Eq. (3.10), viz.,

$$\exp(\hat{\tilde{\mathbf{B}}}\boldsymbol{\zeta}\hat{\mathbf{B}}) = |\cosh\mathbf{z}\boldsymbol{\zeta} - \mathbf{z}\mathbf{y}\sinh\mathbf{z}\boldsymbol{\zeta}|^{-1/2} \times \{\exp[\hat{\tilde{\mathbf{B}}}\boldsymbol{\zeta}\tanh\mathbf{z}\boldsymbol{\zeta}(\mathbf{z}\boldsymbol{\zeta} - \mathbf{y}\boldsymbol{\zeta}\tanh\mathbf{z}\boldsymbol{\zeta})^{-1}\hat{\mathbf{B}}]\}_{cd} \tag{4.1}$$

is valid for any operator column vector

$$\hat{\mathbf{B}} = \{\hat{d}_1, \ldots, \hat{d}_n, \hat{c}_1, \ldots, \hat{c}_n\}, \tag{4.2}$$

where $[\hat{B}_i, \hat{B}_j] = z_{ij}$ and $\{\ \}_{cd}$ denotes the ordering such that all c-operators appear to the left of all $\hat{d}$ operators. We now take

$$\hat{\mathbf{B}} = [\hat{a}_1^+, \ldots, \hat{a}_n^+, -\hat{a}_1, \ldots, -\hat{a}_n] = \mathbf{z}\hat{\mathbf{A}}, \tag{4.3}$$

and

$$\boldsymbol{\zeta} = \mathbf{z}\xi\mathbf{z}^{-1}. \tag{4.4}$$

From (4.1) we then obtain the antinormal ordered form of $\exp(\hat{\tilde{\mathbf{A}}}\xi\hat{\mathbf{A}})$:

$$\exp(\hat{\tilde{\mathbf{A}}}\xi\hat{\mathbf{A}}) = |\cosh\mathbf{z}\xi + \mathbf{z}\mathbf{y}\sinh\mathbf{z}\xi|^{-1/2}\text{“}\exp(\hat{\tilde{\mathbf{A}}}\xi\mathbf{T}_A\hat{\mathbf{A}})\text{”}, \tag{4.5}$$

where

$$\mathbf{T}_A = \mathbf{T}_A(\xi) = (\tanh\mathbf{z}\xi)\ (\mathbf{z}\xi + \mathbf{y}\xi\ \tanh\mathbf{z}\xi). \tag{4.6}$$

As before, if we also include the linear terms, we obtain

$$\exp(\hat{\tilde{\mathbf{A}}}\xi\hat{\mathbf{A}} + 2\tilde{\eta}\hat{\mathbf{A}}) = |\cosh\mathbf{z}\xi + \mathbf{z}\mathbf{y}\sinh\mathbf{z}\xi|^{-1/2} \exp(\tilde{\eta}(\mathbf{T}_A^{-1})\xi^{-1}\eta) \times\text{“}\exp(\hat{\tilde{\mathbf{A}}}\xi\mathbf{T}_A\hat{\mathbf{A}} + 2\tilde{\eta}\mathbf{T}_A\hat{\mathbf{A}})\text{”}. \tag{4.7}$$

Expressions (4.5) and (4.7) are valid for all cases ex-

cept for the singular case

$$|\cosh z\xi + zy \sinh z\xi| = 0. \tag{4.8}$$

If we introduce a parameter k_μ, such that

$$k_W = 0, \quad k_N = 1, \quad \text{and} \quad k_A = -1, \tag{4.9}$$

we may express the Weyl, the normal, and the anti-normal ordered forms of $\exp(\hat{\mathbf{A}}\xi\hat{\mathbf{A}} + 2\tilde{\eta}\hat{\mathbf{A}})$ [Eqs. (2.29), (3.11), and (4.7)] in a single equation, (1.11), given before.

We may use any of the ordered forms of an operator to obtain its trace. For example, if we make use of the identity operator

$$1 = (1/\pi^n) \int |\mathbf{V}\rangle \langle \mathbf{V}| \, d^{2n}V, \tag{4.10}$$

where $|\mathbf{V}\rangle$ is the (n-mode) coherent state,[8] we find that

$$\mathrm{Tr}\hat{G} = (1/\pi^n) \int \langle \mathbf{V}|\hat{G}|\mathbf{V}\rangle d^{2n}V. \tag{4.11}$$

Hence, if $\hat{G}_N$ is the normal ordered form of $\hat{G}$, we obtain

$$\mathrm{Tr}\hat{G} = (1/\pi^n) \int G_N(\mathbf{V}) d^{2n}V. \tag{4.12}$$

Similar results also hold for the Weyl and the anti-normal ordered forms as well. In fact, one may readily verify that[9]

$$(1/\pi^n) \int G_\mu(\mathbf{V}) d^{2n}V$$

does not depend on whether $\mu = W, N$, or A.

Thus, whenever $\exp(\tilde{\mathbf{A}}\xi\hat{\mathbf{A}} + 2\tilde{\eta}\hat{\mathbf{A}})$ is a trace class operator, we find from Eqs. (1.11) that

$$\mathrm{Tr}\exp(\hat{\tilde{\mathbf{A}}}\xi\mathbf{A} + 2\tilde{\eta}\hat{\mathbf{A}})$$
$$= K_\mu \int \exp(\tilde{\mathbf{V}}\xi\mathbf{T}_\mu\mathbf{V} + 2\tilde{\eta}\mathbf{T}_\mu\mathbf{V}) d^{2n}V. \tag{4.13}$$

One may use a result analogous to Eq. (3.7) to carry out the integration on the right-hand side. We find on simplification that

$$\mathrm{Tr}\exp(\hat{\tilde{\mathbf{A}}}\xi\hat{\mathbf{A}} + 2\tilde{\eta}\hat{\mathbf{A}})$$
$$= (1/2^n)|y \sinh z\xi|^{-1/2} \exp(-\tilde{\eta}\xi^{-1}\eta). \tag{4.14}$$

In particular, for a single mode case we obtain

$$\mathrm{Tr}\exp[\alpha\hat{a}^2 + \beta\hat{a}^{+2} + \gamma(\hat{a}^+\hat{a} + \hat{a}\hat{a}^+) + 2(\delta\hat{a} + \epsilon\hat{a}^+)]$$
$$= \tfrac{1}{2}(\sinh\lambda)^{-1} \exp[\lambda^{-2}(\alpha\epsilon^2 + \beta\delta^2 - 2\epsilon\delta\gamma)], \tag{4.15}$$

where $\lambda^2 = \gamma^2 - \alpha\beta$. It is being assumed that the exponential operator on the left-hand side is of trace class, which will certainly be so if α, β, γ are real, $\gamma < 0$, and $\gamma^2 > \alpha\beta$.

APPENDIX: SOME PROPERTIES OF SYMPLECTIC MATRICES

In this appendix, we consider some properties of symplectic matrices. A $(2n \times 2n)$ matrix $\mathbf{S}$ is said to be symplectic if it satisfies the relation

$$\mathbf{S}z\tilde{\mathbf{S}} = z, \tag{A1}$$

where $\tilde{\mathbf{S}}$ denotes the transpose of $\mathbf{S}$ and $\mathbf{z}$ is defined by Eq. (1.5), i.e,

$$\mathbf{z} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}. \tag{A2}$$

From (A1) we find that $\mathbf{S}$ is nonsingular ($|\mathbf{S}| = \pm 1$) and that

$$\mathbf{z} = \mathbf{S}^{-1}\mathbf{z}\tilde{\mathbf{S}}^{-1}, \tag{A3}$$

so that $\mathbf{S}^{-1}$ is also symplectic. In fact, one may readily show from (A1) that $\tilde{\mathbf{S}}$ and $\mathbf{S}^+$ (the Hermitian adjoint) are also symplectic.

If ζ is any $(2n \times 2n)$ symmetric matrix, then $\exp(\mathbf{z}\zeta)$ [or $\exp(\zeta\mathbf{z})$] is symplectic.

If ξ and Λ are any two matrices related by the symplectic transformations

$$\xi = \tilde{\mathbf{S}}\Lambda\mathbf{S}, \tag{A4}$$

then

$$\mathbf{z}\xi = \mathbf{S}^{-1}\mathbf{z}\Lambda\mathbf{S}. \tag{A5}$$

Hence any function of $\mathbf{z}\xi$ is related to the same function of $\mathbf{z}\Lambda$ as

$$f(\mathbf{z}\xi) = \mathbf{S}^{-1}f(\mathbf{z}\Lambda)\mathbf{S}. \tag{A6}$$

In particular we have

$$\tanh(\mathbf{z}\Lambda) = \mathbf{S}\tanh(\mathbf{z}\xi)\mathbf{S}^{-1} \tag{A7a}$$

and

$$\mathrm{sech}(\mathbf{z}\Lambda) = \mathbf{S}\,\mathrm{sech}(\mathbf{z}\xi)\mathbf{S}^{-1}. \tag{A7b}$$

Equations (A7) have been used in deriving Eq. (2.20).

It has been noted in the text that the transformation

$$\hat{\mathbf{B}} = \mathbf{S}\hat{\mathbf{A}} \tag{A8}$$

leaves the commutation relation

$$\hat{\mathbf{A}}\hat{\tilde{\mathbf{A}}} - \hat{\mathbf{A}}\hat{\tilde{\mathbf{A}}} = \mathbf{z}, \tag{A9}$$

invariant if and only if $\mathbf{S}$ is symplectic.

In classical dynamics, if we denote the set of position and momentum variables $q_1, \ldots, q_n, p_1, \ldots, p_n$ by a column vector

$$A = \{q_1, \ldots, q_n, p_1, \ldots, p_n\} \tag{A10}$$

and if B_i $(i = 1, 2, \ldots, 2n)$ are some functions of q's and p's, then one may readily verify that the transformation

$$B_i = B_i(\mathbf{A}) \tag{A11}$$

is canonical if and only if the matrix S, where

$$S_{ij} = \partial B_i/\partial A_j, \tag{A12}$$

is symplectic.

Analogous to symplectic matrices, one may also consider matrices which satisfy the relation

$$\mathbf{R}y\tilde{\mathbf{R}} = y, \tag{A13}$$

where $\mathbf{y}$ is the symmetric matrix

$$y = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}. \tag{A14}$$

These matrices are useful while considering fermion operators, since the transformation

$$\hat{\mathbf{B}} = \mathbf{R}\hat{\mathbf{A}} \tag{A15}$$

leaves the anticommutation relation $\hat{\mathbf{A}}\hat{\tilde{\mathbf{A}}} + \hat{\mathbf{A}}\hat{\tilde{\mathbf{A}}} = \mathbf{y}$ invariant.

*Present address: Department of Physics, The City College of the City University of New York, N.Y. 10031.
†Laboratoire propre du Centre National de la Récherche Scientifique.

[1]C.L. Mehta, J. Math. Phys. 18, xxx (1977).
[2]N.H. McCoy, Proc. Edinburgh Math. Soc. 3, 118 (1932).
[3]R.M. Wilcox, J. Math. Phys. 8, 962 (1967).
[4]C.L. Mehta, J. Phys. A (Proc. Phys. Soc. London) 1, 385 (1968).
[5]F.A. Berezin, *The Method of Second Quantization* (Academic, New York, 1966), p. 143. See also Ref. 7.
[6]C.L. Mehta, in *Progress in Optics, Vol. VIII*, edited by E. Wolf (North-Holland, Amsterdam, 1970), p. 373, Eq. (A.7). Our Eq. (3.7) is valid only if X is a positive definite matrix (the integral on the left-hand side would otherwise diverge). However, as in true in the single mode case, one may use the arguments of analytic continuation to justify the final result [Eq. (3.10 or (3.11)], which is valid even when this restriction is relaxed.
[7]Normal ordered form (3.10) or (3.11) may also be obtained using a result of R. Balian and E. Brezin [Nuovo Cimento B 64, 37 (1969)]. They have shown by group theoretic arguments that $\exp(\hat{P})$ can be expressed as the product $\hat{J}_1\hat{J}_2\hat{J}_3$, where $\hat{J}_1$ contains only the creation operators, $\hat{J}_2$ contains only the annihilation operators, and $\hat{J}_3$ is of the form $\exp(\lambda_{ij}\hat{a}_i+\hat{a}_j)$. The normal ordered form of $\hat{J}_3$ has been obtained in Ref. 4.
[8]R.J. Glauber, Phys. Rev. 131, 2766 (1963).
[9]See, for example, G.S. Agarwal and E. Wolf, Phys. Rev. D 2, 2161 (1970) or Ref. 4.

Unitary analytic representations of SL(3,R)*

Y. Güler

Physics Department, Middle East Technical University, Ankara, Turkey
(Received 18 October 1976)

Unitary, analytic representations of SL(3,R) are studied by operator formalism. It is found that SL(3,R) has two different principal series of representations. Analytic representations are labeled by an integer n and a real number a. The Hilbert space of analytic functions $f(z,x)$ is constructed, and an invariant scalar product is formed.

I. INTRODUCTION

In this paper we determined two different series of principal representations of SL(3,R) by the operator formalism method which was applied to several groups before.[1]

The unitary representations of the complex group SL(3,C) was studied by Bars[2] by the same method which we use. It is a known fact that unitary representations of real unimodular group exhibit some differences in comparison with complex unimodular groups. Real unimodular group admit several principal series of representations. In their paper Gel'fand and Graev[3] showed that SL(n,R) has $(n+1)/2$ principal series of representations if n is odd, and $(n/2)+2$ principal series of representations if n is even. Our aim is to label the analytic, unitary representations of SL(3,R) and to form the representation space.

The unitary representations of SL(3,R) were used by various authors: Dothan, Gell-Mann, and Ne'eman[4,5] used the ladder representations of SL(3,R) which are labeled by $L=0,2,4,\cdots$ to obtain the higher spinned meson and baryon states. They adjoined to three angular momentum operator $\mathbf{L}$ five components of a noncompact operator Q, such that $\mathbf{L}$ and Q generate an SL(3,R) algebra. Higher spins are excited using $\mathbf{L}\equiv\mathbf{J}-\mathbf{S}$ (an internal orbital angular momentum). Change in the L values will cause a change in J. They suggested that these representations could be used as an algebraic model of Regge trajectories. Cusson[6] used SL(3,R) symmetry in nuclear physics. Weaver and Biedenharn[7] studied the nuclear rotational motion assuming transition E2 operators generate SL(3,R) symmetry. Besides Dj. Šijački[8,9] determined the unitary representations of the covering group $\overline{\mathrm{SL}}(3,R)$ of SL(3,R) and presented the group SU(6)$\otimes$SL(3,R) as a model unifying SU(6) quark model and the Regge classification.

This paper is arranged as follows:

In the first chapter the Lie algebra of SL(3,R) is reviewed and the matrix Ω is constructed. In the second and third chapters the commuting operators Z_1 and Z_2 are determined in terms of the generators and the generators are expressed as functions of canonically conjugate operators Z_k and Π_k. In the fourth and fifth chapters representations are labeled and the Hilbert space of analytic functions $f(z,x)$ is formed.

II. THE LIE ALGEBRA OF SL (3, R) AND THE MATRIX Ω

SL(3,R) is a simple group of rank two. It has eight generators which are given as the following in three dimensions:

$$\begin{aligned}&\gamma_1=\epsilon_{11}-\epsilon_{22}, && \gamma_4=\epsilon_{13}, && \gamma_7=\epsilon_{31},\\ &\gamma_2=\epsilon_{12}, && \gamma_5=\epsilon_{33}-\epsilon_{22}, && \gamma_8=\epsilon_{32},\\ &\gamma_3=\epsilon_{21}, && \gamma_6=\epsilon_{23},\end{aligned} \tag{1}$$

where ϵ_{hj} $(h,j=1,2,3)$ is a 3×3 matrix with a one in row h and column j. $\gamma_1,\gamma_2,\gamma_3$ are the generators of SL(2,R) subgroup. In general ϵ_{hj} satisfy the following commutation relation:

$$[\epsilon_{hj},\epsilon_{kl}]=\delta_{jk}\epsilon_{hl}-\delta_{hl}\epsilon_{kj}. \tag{2}$$

Commutation relations of eight generators can easily be found using Eq. (2). The metric matrix $F_{hj}=C_{hkl}C_{jlk}$ for SL(3,R) is an 8×8 nonsingular symmetric matrix with nonzero elements

$$\begin{aligned}&F_{11}=12, && F_{23}=6,\\ &F_{15}=6, && F_{47}=6,\\ &F_{55}=12, && F_{68}=6.\end{aligned} \tag{3}$$

C_{hjk} are the structure constants defined as

$$[\gamma_h,\gamma_j]=C_{hjk}\gamma_k\,. \tag{4}$$

A matrix Ω satisfying the equation

$$U\Omega U^{-1}=\Lambda\Omega\Lambda^{-1} \tag{5}$$

is essential for the operator formalism of SL(3,R). Here U is a representation of SL(3,R) and Λ is its 3×3 representation. Let us define Ω as

$$\Omega=(F^{-1})_{hl}\gamma_h T_l. \tag{6}$$

It is a 3×3 matrix with operator entries. Its explicit form is

$$\Omega=\begin{pmatrix}2T_1-T_5 & 3T_3 & 3T_7\\ 3T_2 & -T_1-T_5 & 3T_8\\ 3T_4 & 3T_6 & 2T_5-T_1\end{pmatrix}. \tag{7}$$

Let ψ be the eigenvector of Ω with eigenvalue 2λ. That is,

$$\Omega\psi=2\lambda\psi. \tag{8}$$

By defining two Z operators Z_1 and Z_2 as

$$Z_1=\psi_1\psi_3^{-1},\quad Z_2=\psi_2\psi_3^{-1} \tag{9}$$

and using Eq. (5) the transformation law for the eigenvector, ψ is obtained as

$$U\psi U^{-1}=\psi'=\Lambda^{-1}\psi C(\Lambda), \tag{10}$$

where ψ_1,ψ_2,ψ_3 are operator components of ψ and $C(\Lambda)$ is a diagonal matrix. Hence

$$UZ_1U^{-1}=Z_1'=(\Lambda_{11}^{-1}Z_1+\Lambda_{12}^{-1}Z_2+\Lambda_{13}^{-1})(\Lambda_{31}^{-1}Z_1+\Lambda_{32}^{-1}Z_2+\Lambda_{33}^{-1})^{-1},$$
$$UZ_2U^{-1}=Z_2'=(\Lambda_{21}^{-1}Z_1+\Lambda_{22}^{-1}Z_2+\Lambda_{23}^{-1})(\Lambda_{31}^{-1}Z_1+\Lambda_{32}^{-1}Z_2+\Lambda_{33}^{-1})^{-1}. \tag{11}$$

III. DETERMINATION OF THE OPERATORS Z_1, Z_2

Operators Z_1 and Z_2 are functions of the generators T_j. The following three homogenous equations which are obtained from the eigenvalue equation (8) are used to determine Z_1 and Z_2. The procedure is the same as in the Ref. 2:

$$(\Omega_{11}-2\lambda)Z_1+\Omega_{12}Z_2+\Omega_{13}=0, \tag{12a}$$

$$\Omega_{21}Z_1+(\Omega_{22}-2\lambda)Z_2+\Omega_{23}=0, \tag{12b}$$

$$\Omega_{31}Z_1+\Omega_{32}Z_2+(\Omega_{33}-2\lambda)=0. \tag{12c}$$

Notice that the coefficients Ω_{hj} are operators. Therefore, one should be careful to determine Z_1 and Z_2. The elements Ω_{hj} satisfy the commutation relation

$$[\Omega_{hj},\Omega_{kl}]=3(\delta_{hl}\Omega_{kj}-\delta_{kj}\Omega_{hl}). \tag{13}$$

Using the above commutation relation and any pair of equations (12), one can eliminate one of the unknowns. As an example we will calculate Z_2 using Eq. (12a) and (12b). Multiplying Eq. (12a) by Ω_{21} and Eq. (12b) by $\Omega_{11}-2\lambda-3$ from the left and subtracting (12b) from (12a) and using the commutation relation $[\Omega_{21},\Omega_{11}]=-3Z_{21}$, we obtain $Z_2^{(a,b)}$ as follows:

$$Z_2^{(a,c)}=[\Omega_{31}\Omega_{12}-(\Omega_{11}-2\lambda-3)\Omega_{32}]^{-1}\times[(\Omega_{11}-2\lambda-3)(\Omega_{33}-2\lambda)-\Omega_{31}\Omega_{13}], \tag{14}$$

$$Z_2^{(b,c)}=[\Omega_{31}(\Omega_{22}-2\lambda)-\Omega_{21}\Omega_{32}]^{-1}[\Omega_{21}(\Omega_{33}-2\lambda)-\Omega_{31}\Omega_{23})], \tag{15}$$

$$Z_1^{(a,b)}=[(\Omega_{22}-2\lambda-3)(\Omega_{11}-2\lambda)-\Omega_{12}\Omega_{21}]^{-1}\times[\Omega_{12}\Omega_{23}-(\Omega_{22}-2\lambda-3)\Omega_{13}], \tag{16}$$

$$Z_1^{(a,c)}=[\Omega_{32}(\Omega_{11}-2\lambda)-\Omega_{12}\Omega_{31}]^{-1}[\Omega_{12}(\Omega_{33}-2\lambda)-\Omega_{32}\Omega_{13}], \tag{17}$$

$$Z_1^{(b,c)}=[\Omega_{32}\Omega_{21}-(\Omega_{22}-2\lambda-3)\Omega_{31}]^{-1}\times[(\Omega_{22}-2\gamma-3)(\Omega_{33}-2\lambda)-\Omega_{32}\Omega_{23}]. \tag{18}$$

To show that all Z_1's and Z_2's are compatible respectively, let us write Z_j in a compact form as in the Ref. 2.

Define

$$K_{ii}(2\lambda)=(\Omega_{jj}-2\lambda-3)(\Omega_{kk}-2\lambda)-\Omega_{kj}\Omega_{jk}, \tag{19}$$

or

$$K_{ij}(2\lambda)=-(\Omega_{kk}-2\lambda-3)\Omega_{ij}+\Omega_{ik}\Omega_{kj}\ ,$$
$$K_{ij}(2\lambda)=-\Omega_{ij}(\Omega_{kk}-2\lambda)+\Omega_{kj}\Omega_{ik},\quad i\neq j. \tag{20}$$

Here i,j,k are in cylic or in anticylic order. In this notation Z_j are obtained as

$$Z_1=K_{3n}^{-1}(2\lambda)K_{1n}(2\lambda), \tag{21}$$

$$Z_2=K_{3n}^{-1}(2\lambda)K_{2n}(2\lambda),\quad n=1,2,3. \tag{22}$$

K_{ij} has the following property:

$$K_{ij}(2\lambda)K_{lm}(2\lambda+3)=K_{lj}(2\lambda)K_{im}(2\lambda+3). \tag{23}$$

Hence

$$Z_1=K_{3n}^{-1}(2\lambda)K_{1n}(2\lambda)=K_{1m}(2\lambda+3)K_{3m}^{-1}(2\lambda+3), \tag{24}$$

$$Z_2=K_{3n}^{-1}(2\lambda)K_{2n}(2\lambda)=K_{2m}(2\lambda+3)K_{3m}^{-1}(2\lambda+3). \tag{25}$$

Equations (24) and (25) show that three Z_j obtained using any two pairs of Eqs. (12a), (12b), (12c) are compatible. Besides using the property of $K_{ij}(2\lambda)$ one can easily show that Z_1 and Z_2 commute. In fact,

$$Z_1Z_2=K_{3n}^{-1}(2\lambda)K_{1n}(2\lambda)K_{2m}(2\lambda+3)K_{3m}^{-1}(2\lambda+3),$$
$$Z_2Z_1=K_{3n}^{-1}(2\lambda)K_{2n}(2\lambda)K_{1m}(2\lambda+3)K_{3m}^{-1}(2\lambda+3), \tag{26}$$
$$[Z_1,Z_2]=K_{3n}^{-1}(2\lambda)[K_{1n}(2\lambda)K_{2m}(2\lambda+3)-K_{2n}(2\lambda)K_{1m}(2\lambda+3)],$$
$$K_{3m}^{-1}(2\lambda+3)=0.$$

IV. DETERMINATION OF GENERATORS IN TERMS OF THE CANONICALLY CONJUGATE OPERATORS Z_k AND Π_k

Ω_{ij} and K_{lm} satisfy the following commutation relation:

$$[\Omega_{ij},K_{lm}]=3(\delta_{im}K_{lj}-\delta_{lj}K_{im}). \tag{27}$$

Hence the commutation relations of Z_k $(k=1,2)$ with Ω_{lm} are obtained as in the Ref. 2:

$$\begin{aligned}
&[\Omega_{11},Z_1]=[Z_1,\Omega_{33}]=[\Omega_{12},Z_2]=-3Z_1,\\
&[\Omega_{11},Z_2]=[\Omega_{12},Z_1]=[\Omega_{21},Z_2]=[\Omega_{22},Z_1]=0,\\
&[\Omega_{31},Z_2]=[\Omega_{32},Z_1]=0,\\
&[\Omega_{21},Z_1]=[Z_2,\Omega_{33}]=[\Omega_{22},Z_2]=-3Z_2,\\
&[\Omega_{13},Z_2]=[\Omega_{23},Z_1]=3Z_1Z_2,\\
&[\Omega_{13},Z_1]=3Z_1^2,\quad [\Omega_{23},Z_2]=3Z_2^2,\\
&[\Omega_{31},Z_1]=-3,\quad [\Omega_{32},Z_2]=-3.
\end{aligned} \tag{28}$$

Now, let

$$\Omega_{11}=\Pi_1Z_1-\lambda,\quad \Omega_{12}=\Pi_2Z_1, \tag{29}$$

$$\Omega_{21}=\Pi_1Z_2,\quad \Omega_{22}=\Pi_2Z_2-\lambda, \tag{30}$$

$$\Pi_1=\Omega_{31},\quad \Pi_2=\Omega_{32}. \tag{31}$$

$\mathrm{Tr}\Omega=0$ gives Ω_{33} as

$$\Omega_{33}=-\Pi_1Z_1-\Pi_2Z_2+2\lambda. \tag{32}$$

The elements Ω_{13} and Ω_{23} are obtained using homogenous equations (12a) and (12b):

$$\Omega_{13}=-\Pi_1Z_1^2+3\lambda Z_1-\Pi_2Z_1Z_2, \tag{33}$$

$$\Omega_{23}=-\Pi_2Z_2^2+3\lambda Z_2-\Pi_1Z_1Z_2. \tag{34}$$

Hence we determined the elements Ω_{kj} in terms of the

canonically conjugate operators, Π_k and Z_k. Using the commutation relations (28), one can easily check that the commutation relation

$$[\Omega_{kj}, \Omega_{kl}] = 3(\delta_{kl}\Omega_{kj} - \delta_{kj}\Omega_{kl}) \tag{35}$$

holds. The generators T_k in terms of Π_k and Z_k are as follows:

$$3T_1 = \Omega_{11} - \Omega_{22} = \Pi_1 Z_1 - \Pi_2 Z_2, \tag{36a}$$

$$3T_2 = \Omega_{21} = \Pi_1 Z_2, \tag{36b}$$

$$3T_3 = \Omega_{12} = \Pi_2 Z_1, \tag{36c}$$

$$3T_4 = \Omega_{31} = \Pi_1, \tag{36d}$$

$$3T_5 = \Omega_{33} - \Omega_{22} = -\Pi_1 Z_1 - 2\Pi_2 Z_2 + 3\lambda, \tag{36e}$$

$$3T_6 = \Omega_{32} = \Pi_2, \tag{36f}$$

$$3T_7 = \Omega_{13} = -\Pi_1 Z_1^2 + 3\lambda Z_1 - \Pi_2 Z_1 Z_2, \tag{36k}$$

$$3T_8 = \Omega_{23} = -\Pi_2 Z_2^2 + 3\lambda Z_2 - \Pi_1 Z_1 Z_2. \tag{36l}$$

V. LABELING OF REPRESENTATIONS

We will label the unitary representations of SL(3, R) by the eigenvalues of two Casimir operators C_1 and C_2. C_1 and C_2 are real multiples of the identity for unitary representations. Defining the second and third order Casimir operators C_1 and C_2 as

$$C_1 = \mathrm{Tr}\Omega^2, \tag{37}$$

$$C_2 = \mathrm{Tr}\Omega^3, \tag{38}$$

one can label the unitary irreducible representations of SL(3, R) by the eigenvalues of the matrix Ω. By letting the eigenvalues λ_1, λ_2 be

$$\lambda_1 = \alpha_1 + i\beta_1, \quad \lambda_2 = \alpha_2 + i\beta_2 \tag{39}$$

and using the condition $\mathrm{Tr}\Omega = 0$, the unitary condition gives

$$\mathrm{Im}(\mathrm{Tr}\Omega^2) = 2\alpha_1\beta_1 + 2\alpha_2\beta_2 + \alpha_2\beta_1 + \alpha_1\beta_2 = 0, \tag{40}$$

$$\mathrm{Im}(\mathrm{Tr}\Omega^3) = 2\alpha_1\alpha_2\beta_1 + 2\alpha_1\alpha_2\beta_2 + \beta_1(\alpha_2^2 - \beta_2^2) + \beta_2(\alpha_1^2 - \beta_1^2) = 0. \tag{41}$$

Equation (40) gives

$$\beta_1/\beta_2 = -(2\alpha_2 + \alpha_1)/(2\alpha_1 + \alpha_2). \tag{42}$$

Inserting β_2 in Eq. (41) and doing simple algebra, we obtain a final condition,

$$(2\alpha_1 + \alpha_2)\{\beta_1[\alpha_2 - \alpha_1][(2\alpha_2 + \alpha_1)^2 + \beta_1^2]\} = 0. \tag{43}$$

This condition gives mainly two classes of representations:

$$\begin{aligned} &\text{(a) } \lambda_1 = \alpha_1 + i\beta_1 \text{ (complex)}, && \text{(b) } \lambda_1 = \alpha_1 \text{ (real)}, \\ &\quad \lambda_2 = -2\alpha_1 \text{ (real)}, && \quad \lambda_2 = \alpha_2 \text{ (real)}, \\ &\quad \lambda_3 = \alpha_1 - i\beta_1 \text{ (complex)}, && \quad \lambda_3 = -\alpha_1 - \alpha_2 \text{ (real)}. \end{aligned} \tag{44}$$

Using the definitions of $Z_1(2\lambda)$, $Z_2(2\lambda)$ and the commutation relations $[\Omega_{kj}, \Omega_{kl}] = 3(\delta_{kl}\Omega_{kj} - \delta_{kj}\Omega_{kl})$, one can easily show that $Z_1(2\lambda)$ and $Z_2(2\lambda)$ are Hermitian operators if λ is real. Besides

$$Z_k(2\lambda) = Z_k^\dagger(2\lambda^*). \tag{45}$$

Let us define a new operator $Z(2\lambda)$ as

$$Z(2\lambda) = (1/\beta)[\Lambda_{31}^{-1}Z_1(2\lambda) + \Lambda_{32}^{-1}Z_2(2\lambda)], \tag{46}$$

where β is a real number. The transformation law of $Z(2\lambda)$ is determined from the transformation law of $Z_1(2\lambda)$ and $Z_2(2\lambda)$. In fact,

$$\begin{aligned} Z'(2\lambda) &= (1/\beta)[\Lambda_{31}^{-1}Z_1'(2\lambda) + \Lambda_{32}^{-1}Z_2'(2\lambda)], \\ Z'(2\lambda) &= (1/\beta)[(\Lambda_{31}^{-1}\Lambda_{11}^{-1} + \Lambda_{32}^{-1}\Lambda_{21}^{-1})Z_1(2\lambda) + (\Lambda_{31}^{-1}\Lambda_{12}^{-1} + \Lambda_{32}^{-1}\Lambda_{22}^{-1}) \\ &\quad \times Z_2(2\lambda) + \Lambda_{31}^{-1}\Lambda_{13}^{-1} + \Lambda_{32}^{-1}\Lambda_{23}^{-1}][\beta Z(2\lambda) + \Lambda_{33}^{-1}]^{-1}. \end{aligned} \tag{47}$$

Letting

$$\begin{aligned} &(1/\beta)[(\Lambda_{31}^{-1}\Lambda_{11}^{-1} + \Lambda_{32}^{-1}\Lambda_{21}^{-1})Z_1(2\lambda) + (\Lambda_{31}^{-1}\Lambda_{12}^{-1} + \Lambda_{32}^{-1}\Lambda_{22}^{-1})Z_2(2\lambda) \\ &\quad = \alpha Z(2\lambda) = (\alpha/\beta)[\Lambda_{31}^{-1}Z_1(2\lambda) + \Lambda_{32}^{-1}Z_2(2\lambda)], \end{aligned} \tag{48}$$

one can define the real parameter α in terms of Λ_{ij}^{-1}, the real and imaginary parts of $Z_1(2\lambda)$ and $Z_2(2\lambda)$. In fact we will later show that real and imaginary parts of $Z(2\lambda)$ are functions of Λ_{ij}^{-1}. Hence α is a rational function of Λ_{ij}^{-1}. Hence the transformation law of $Z(2\lambda)$ is obtained as

$$Z'(2\lambda) = [\alpha Z(2\lambda) + \gamma][\beta Z(2\lambda) + \delta]^{-1}, \tag{49}$$

where

$$\begin{aligned} &\gamma = (1/\beta)(\Lambda_{31}^{-1}\Lambda_{13}^{-1} + \Lambda_{32}^{-1}\Lambda_{23}^{-1}), \quad \delta = \Lambda_{33}^{-1}, \\ &\beta = \text{arbitrary real number}. \end{aligned} \tag{50}$$

VI. CONSTRUCTION OF THE REPRESENTATION SPACE

Let us define the common eigenstates $|z(2\lambda_1), z(2\lambda_3), z(2\lambda_2), \lambda_1, \lambda_2\rangle$ of the commuting operators $Z(2\lambda_1)$, $Z(2\lambda_3)$, $Z(2\lambda_2)$, C_1 and C_2 as the basis of the representation space. Since we know the generators in terms of the canonically conjugate operators Π_k and Z_k, it is an easy job to find the transformation law of the eigenstate. Using the case (a) of (44) and Eq. (45), we will notate $z(2\lambda_1)$ as z and $z(2\lambda_2)$ as η. $z(2\lambda_3)$ is z^*. Hence the transformation law of the eigenstate $|z, z^*, \eta\rangle$ is determined as follows:

$$\begin{aligned} U(\Lambda)|z, z^*, \eta\rangle &= \exp\left(i\sum_{k=1}^{8} b_k T_k(2\lambda_1)\right)\exp\ i\sum_{k=1}^{8} b_k T_k(2\lambda_3) \\ &\quad \times\left(\exp\ i\sum_{k=1}^{8} b_k T_k(2\lambda_2)\right)|z, z^*, \eta\rangle, \end{aligned} \tag{51}$$

where b_k are real parameters and $T_k(2\lambda_k)$ are Hermitian infinitesimal generators of the unitary representation. Noting that $\Pi_k = -\partial/\partial z_k$, we obtain the following infinitesimal transformation:

$$\begin{aligned} U_{\mathrm{inf}}(\Lambda^{-1})|z, z^*, \eta\rangle &= \left(1 - i\sum_{k=1}^{8} b_k T_k(2\lambda_1)\right)\left(1 - i\sum_{k=1}^{8} b_k T_k(2\lambda_3)\right) \\ &\quad \times\left(1 - i\sum_{k=1}^{8} b_k T_k(2\lambda_2)\right)|z, z^*, \eta\rangle, \end{aligned} \tag{52}$$

$$\begin{aligned} &U_{\mathrm{inf}}(\Lambda^{-1})\, z, z^*, \eta\rangle \\ &= \left(1 + i\Delta z_1\frac{\partial}{\partial z_1} + i\Delta z_2\frac{\partial}{\partial z_2}\right)\left(1 + i\Delta z_1^*\frac{\partial}{\partial z_1^*} + i\Delta z_2^*\frac{\partial}{\partial z_2^*}\right) \\ &\quad \times\left(1 + i\Delta\eta_1\frac{\partial}{\partial\eta_1} + i\Delta\eta_2\frac{\partial}{\partial\eta_2}\right)1 + i\lambda_1(b_7 z_1 + b_8 z_2 + b_5) \end{aligned}$$

$$\times[1+i\lambda_3(b_7z_1^*+b_8z_2^*+b_5)][i+i\lambda_2(b_7\eta_1+b_8\eta_2+b_5)]$$
$$\times|z,z^*,\eta\rangle, \tag{53}$$

$$U_{\rm inf}(\Lambda^{-1})|z,z^*,\eta\rangle$$
$$=[1+b_7z_1+b_8z_2+b_5]^{i\lambda_1}[1+b_7z_1+b_8z_2+b_5]^{i\lambda_3}$$
$$\times[1+b_7\eta_1+b_8\eta_2+b_5]^{i\lambda_2}|z',z'^*,\eta'\rangle, \tag{54}$$

where

$$\Delta z_1=b_1z_1+b_2z_2+b_4-b_5z_1-b_7z_1^2-b_8z_1z_2,$$
$$\Delta z_2=-b_1z_2+b_3z_1-2b_5z_2+b_6-b_7z_1z_2-b_8z_2^2. \tag{55}$$

On the other hand the infinitesimal form of 3×3 representation of $SL(3,R)$ is as follows:

$$\Lambda_{\rm inf}^{-1}=\begin{pmatrix}1+ib_1 & ib_2 & ib_4\\ ib_3 & 1-ib_1-ib_5 & ib_6\\ ib_7 & ib_8 & 1+ib_5\end{pmatrix}. \tag{56}$$

Integrating Eq. (54) and using Eq. (56), one obtains transformation law for the eigenstate $|z,z^*,\eta\rangle$:

$$U(\Lambda^{-1})|z,z^*,\eta\rangle$$
$$=[\Lambda_{31}^{-1}z_1+\Lambda_{32}^{-1}z_2+\Lambda_{33}^{-1}]^{i\lambda_1}[\Lambda_{31}^{-1}z_1+\Lambda_{31}^{-1}z_2+\Lambda_{33}^{-1}]^{i\lambda_3}$$
$$\times[\Lambda_{31}^{-1}\eta_1+\Lambda_{32}^{-1}\eta_2+\Lambda_{33}^{-1}]^{i\lambda_2}|z',z'^*,\eta'\rangle, \tag{57}$$

$$U(\Lambda^{-1})|z,z^*,\eta\rangle$$
$$=[\beta z+\delta]^{i\lambda_1}[\beta z^*+\delta]^{i\lambda_1^*}[\beta\eta+\delta]^{i\lambda_2}|z',z'^*,\eta'\rangle, \tag{58}$$

where

$$z'=[\alpha z+\gamma][\beta z+\delta]^{-1},\quad \eta'=[\alpha\eta+\gamma][\beta\eta+\delta]^{-1}.$$

A wavepacket $|K\rangle$ is written in terms of the basis states $|z,z^*,\eta\rangle$ as

$$|K\rangle=\int f(z,z^*,\eta)|z,z^*,\eta\rangle\,d\sigma\,d\eta, \tag{59}$$

where $d\sigma dn=dxdy/y^2d\eta$ is the right invariant measure and $z=x+iy$. The transformation law of the "component" $f(z,z^*,\eta)$ is obtained using the transformation law for the basis $|z,z^*,\eta\rangle$. In fact,

$$U(\Lambda^{-1})|K\rangle=\int f(z,z^*,\eta)U(\Lambda^{-1})|z,z^*,\eta\rangle\,d\sigma\,d\eta, \tag{60}$$

$$U(\Lambda)f(z,z^*,\eta)=[\beta z+\delta]^{i\lambda_1}[\beta z^*+\delta]^{i\lambda_3}[\beta\eta+\delta]^{i\lambda_2}$$
$$\times f[(\alpha z+\gamma)(\beta z+\delta)^{-1},\ (\alpha z^*+\gamma)(\beta z^*+\delta)^{-1},$$
$$(\alpha\eta+\gamma)(\beta\eta+\delta)^{-1}|. \tag{61}$$

By letting $\lambda_1=\frac{1}{2}(a+ib)$ and using the condition (a) of (44), the transformation law (61) becomes

$$U(\Lambda)f(z,z^*,\eta)=\left(\frac{\beta z+\delta}{|\beta z+\delta|}\right)^{-b}\left(\frac{\beta\eta+\delta}{|\beta z+\delta|}\right)^{-ia}f[(\alpha z+\gamma)(\beta z+\delta)^{-1},$$
$$(\alpha z^*+\gamma)(\beta z^*+\delta)^{-1},\ (\alpha\eta+\gamma)(\beta\eta+\delta)^{-1}|. \tag{62}$$

VII. ANALYTIC REPRESENTATIONS

In their paper Gel'fand and Graev[2] determined the principal series of representations of $SL(n,R)$ in the space of the analytic functions $f(z,x)$. The real parameters α,β,γ, and δ appeared in the transformation law of the functions $f(z,x)$. Besides, the real and the imaginary parts of the complex variable z are obtained in terms of these parameters. Indeed z is obtained as

$$z=(\alpha\beta-\gamma\delta/(\beta^2+\delta^2)+i(\alpha\delta-\beta\gamma)/(\beta^2+\delta^2). \tag{63}$$

These real parameters appear in the subgroup $\hat{K}_m$ with elements $\hat{k}$ given as

$$\hat{k}=\begin{pmatrix}\alpha & \beta & e\\ \gamma & \delta & f\\ 0 & 0 & m\end{pmatrix}. \tag{64}$$

On the other hand the canonical decomposition $g=\hat{k}x$ of $GL(n,R)$ enables one to obtain the real parameters α,β,γ, and δ in terms of the elements of g. Here the matrix x is the element of the subgroup X_m. Its explicit form is

$$x=\begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ a & b & 1\end{pmatrix}. \tag{65}$$

Hence, letting $|\beta z+\delta|^2=|\alpha\delta-\beta\gamma|$ and $(\beta\eta+\delta)^2=|\alpha\delta-\beta\gamma|^3$, we can express the variables z and η in terms of Λ_{kj}^{-1} which is the case in Ref. 2. So the transformation laws of functions $f(z,z^*,\eta)$ becomes

$$U(\Lambda)f(z,z^*,\eta)=\left(\frac{\beta z+\delta}{|\alpha\delta-\beta\gamma|^{1/2}}\right)^{-b}|\alpha\delta-\beta\gamma|^{-ia}$$
$$\times f(z',z'^*,\eta'). \tag{66}$$

Analyticity requires that analytic function $f(z,z^*,\eta)$ should remain analytic after the transformation. Assuming analytic continuation is possible, one should remove the branch cuts on the real axis. This removal imposes the condition b is an integer n. The Hilbert space L_2 of square integrable functions $f(z,z^*,\eta)$ forms the representation space. An invariant scalar product in this space is given as

$$(f_1,f_2)=c\int f_1(z,z^*,\eta)f_2^*(z,z^*,\eta)|\,\mathrm{Im}\,z\,|^{n-2}dx\,dy\,d\eta. \tag{67}$$

CONCLUSION

Principal series of representations of $SL(3,R)$ are determined by the operator formalism method. It is found that $SL(3,R)$ has two series of principal representations. Case (b) of (44) corresponds to the series d_0 and (a) corresponds to the series d_1 of Ref. 3. Analytic representations are determined in the Hilbert space of functions $f(z,x)$ analytic in the upper or lower half plane. It is shown that they are labeled by an integer n and a real number a.

ACKNOWLEDGMENT

I thank Dr. Mehmet Koca for his proposal to apply the operator formalism to $SL(3,R)$. I also thank the

Scientific and Technical Research Council of Turkey for partial support of this work.

*Research partially supported by the Scientific and Technical Research Council of Turkey.

[1]Y. Güler, J. Math. Phys. 17, 554 (1976).
[2]I. Bars, J. Math. Phys. 14, 759 (1973).
[3]I.M. Gel'fand and M.I. Graev, Am. Math. Soc. Transl., Ser 2, 2, 147 (1956).
[4]Y. Dothan, M. Gell-Mann, and Y.Ne'eman, Phys. Lett. 17, 148 (1965).
[5]Y. Dothan and Y. Ne'eman, "Band Spectra Generated by Non-Compact Algebra," in *Symmetry Groups in Nuclear and Particle Physics*, edited by F.J. Dyson (Benjamin, New York, 1966).
[6]R.Y. Cusson, Nucl. Phys. A **114**, 289 (1968).
[7]L. Weaver and L.C. Biedenharn, Nucl. Phys. A 185, 1 (1972).
[8]Dj. Šijački, J. Math. Phys. **16**, 298 (1975).
[9]Dj. Šijački, Phys. Lett. B **62**, 323 (1976).

Evolution of isometries in the Bondi formalism

R. Berezdivin* and L. Herrera†

Departamento de Física, Facultad de Ciencias, Universidad Central de Venezuela, Caracas, Venezuela
(Received 2 August 1976)

It is shown that if an additional symmetry, assumed part of the BMS group, is imposed in the Bondi formalism at one retarded time, and if gravitational radiation is absent, then the symmetry will evolve to fill the region of space–time where the Bondi metric is nonsingular. Furthermore, that region will admit a static Weyl metric. There is no necessary evolution if there is radiation present. The evolution of vector field which are nearly isometric is then examined: These evolve as small perturbations off a Weyl metric. A simple and nonlinear but approximate energy formula is written in terms of a quadrupole moment.

I. INTRODUCTION

Recent interest in relativistic stellar structure and black holes has been concentrated in the rotating case where the Kerr metric is, almost certainly, the end point of collapse. Nonrotating spherical stars and collapse had been previously worked on and basically understood. Nevertheless, nonrotating but nonspherical possibilities, except small perturbations off Schwarzschild, have been usually ignored. Essentially one would not expect high nonsphericity to maintain itself without rotation. Although this may be so, some nonlinearities in the perturbations may be more easily gotten by starting with exact solutions. The axisymmetric static exterior solutions would be part of a Weyl space–time; at least some of these should be able to be fitted to interior solution generated by some source that somehow avoids sphericity through anisotropic pressures and stresses. Of course, the rotating cases are, probably, more important in astrophysics.

A recent calculation[1] of the effect of the quadrupole moment of the sun on the proposed gyroscope experiment, using a Weyl metric, verified the previous linearized approaches but showed various nonlinear effects, which in this case are too small to be measured. There is further recent interest[2] in Weyl space–times. Perturbations of these may show radiation and other nonlinear effects, and recent interest[3] in the nonstatic case also exists. Fittings of Weyl space–time to an internal solutions are known,[4] and recently a special case has been examined and fitted.[5]

We[6] have previously examined a radiation formalism in an axially symmetric nonrotating (reflection symmetric) asymptotically flat space–time, due to Bondi, Van der Burg, and Metzner,[7] specifically as to restrictions imposed in the results by a further isometry, and found that the radiation was eliminated. Here the symmetry is weakened to an initial time: We impose an arbitrary isometry at one retarded time and see if it evolves and what it restricts. Previous evolution of nonlightlike isometries off timelike hypersurfaces, assuming analyticity, has been demonstrated,[8] and here those results will be complemented for lightlike hypersurfaces. If there is gravitational radiation, it will be argued that the symmetry will not evolve. Otherwise, and if we assume that the initial isometry is part of the BMS group,[7] and if certain smoothness conditions (see later) are assumed, it will be shown to evolve to fill out the space–time. Furthermore, previous results[6] are strengthened by showing that the region of space–time treated must necesarily be of Weyl type. Of course, results hold for large enough, but not necessarily infinite, distances from the source where the coordinate system used is well defined.[7] The initial cone used is actually the future light cone of a point: This causes no formal problems and is intuitively acceptable.[9]

The evolution of small asymmetries in axially symmetric nonrotating space–times is then studied. These evolve as arbitrary axially symmetric nonrotating perturbations off Weyl. Using the Bondi formalism, we get a radiated energy formula, which include nonlinearities, radiative wavetails, in terms of a quadrupole moment constructed from the field. The results are similar to those found previously[10] using the Newman–Penrose formalism.[11]

We use Bondi's formalism throughout. Reference 7 will be called Paper B, and formulas from it will be preceded by a B and in parenthesis. Similarly for Ref. 6, called BH, where the present notation is taken and for Ref. 15 denoted by S. Section II treats the main problem, with Secs. III and IV as special cases. Section V treats the small asymmetries. Appendix A has equations (BH18). Appendix B is another special case.

II. FIRST PART. INITIAL ISOMETRY IN BONDI: GENERAL CASE

The region of interest M will be far enough from all bodies, and we will call it space–time. M is axially and reflection symmetric, and asymptotically flat.[7] An additional isometry is imposed on M at one "Bondi retarded time," $u=u_0$.[7] Absence of radiation is also imposed throughout M, $c_0=0$. It will be shown that the initial isometry evolves off the light cone to fill out M. Furthermore, M will necessarily admit a static asymptotically flat Weyl metric. The problem of the evolution of symmetries was treated in other circumstances with different methods.[6,12,13] In Refs. 12 and 13 there was an assumed timelike isometry before a retarded time, $u=u_0$ (but in a more general M) and no radiation afterwards: They showed that the isometry evolved. Reference 6 is described in the Introduction.

Here the timelike case is straightforward if the initial-Killing vector field is $\partial/\partial u$, u the "Bondi retarded time." For if it is so, taking, without loss of generality, $u_0=0$, Killing's equations at $u=0$ give $M_0(0)=C_0(0)=N_0(0)$ $[f(0)\equiv f(0,\theta,\phi)]$, and all higher metric

coefficients have zero derivative at $u=0$. With $c_0(u)=0$, (B35) gives $M_0(u)=0$, (B36) gives $N_{00}(u)=0$ so that $N_0(u)=0$, (B34) gives $C_{00}(u)=0$ so $C_0(u)=0$. Similarly for the higher order terms so that $\partial/\partial u$ is Killing at all u. It is also easy to see that if $c_0(u)\neq 0$ (but is analytic in u, which tends to happen[7]) then it would not necessarily evolve. For $c_{00}(0)$ could be unequal to zero and there would be no way to force $M_{00}(0)$, from (B35), to be 0. An isometry at $u=0$ and at $u=du$ will still leave $M_{000}(0)$ free. Thus, as in preceding studies, absence of radiation or shock waves are necessary for the evolution.

A. Asymptotic conditions and ϕ dependence

With $\eta(0)$ as our Killing vector field at $u=0$, A, B, f, g as the same functions they denoted on BH except now defined only at $u=0$ so far, and since δ^α_ϕ, the axial rotation generator, is also Killing, $[\eta^\alpha, \delta^\alpha_\phi]^\beta = \eta^\beta{}_{,\phi}$ must also be Killing; hence its (A, B, f, g) is $(A_\phi, B_\phi, f_\phi, g_\phi)$, at $u=0$. u is taken to be 0 throughout this subsection (IIA). Assuming (A, B, f, g) smooth enough in ϕ in $[0, 2\pi]$, a Fourier series expansion may be performed, e.g., $f=\sum_{k=0}^{\infty} f_k \exp(ik\phi)$ so that $\eta^\alpha = \sum_{k=0}^{\infty} \eta^\alpha_k \exp(ik\phi)$ and then

$$\pounds_\eta g_{\alpha\beta} = \sum_{k=0}^{\infty} \pounds_{\eta_k e^{ik\phi}} g_{\alpha\beta} = \sum_{k=0}^{\infty} \exp(ik\phi)(\pounds_{\eta_k} g_{\alpha\beta} + 2ik\delta^\phi_{(\alpha}\eta_{\underline{k}\beta)} = 0, \quad \text{(II. 1)}$$

where some known properties[14] of ϕ, the Lie derivative, have been used and where $\underline{k}$ will denote that k is not symmetrized. Since the $\exp(ik\phi)$ are linearly independent on $L^2(0, 2\pi, R)$ for each k, we have a Killing vector,

$$\exp(-ik\phi)\,\pounds_{\eta_k e^{ik\phi}} g_{\alpha\beta} = \pounds_{\eta_k} g_{\alpha\beta} + 2ik\delta^\phi{}_{(\alpha}\eta_{\underline{k}\beta)} = 0. \quad \text{(II. 2)}$$

Now, as $r\to\infty$ the Killing vector field will be assumed to be one of the generators of the BMS group. This is necessary if we are to remain in asymptotically flat space—time. Then, since the BMS group has its (A, B, f, g) independent of ϕ except for spatial rotations where $k=1$ terms come in,[7,15] the $k=0$ or $k=1$ terms will appear, and we will take as our vectors the $k=0$ $+k=1$ components only. Since $R(3)$ has $B_\phi=0$ and as will be seen in the next subsection, $B_r=0$, our B must be the same it is at $r=\infty$ so we will take $B_\phi=0$,

$$f = {}_0 f(u,\theta,r) + {}_{(-)}f(u,\theta,r)\sin\phi + {}_{(+)}f(u,\theta,r)\cos\phi$$

and A, g similarly. However, in Appendix C it is shown that A may be taken independent of ϕ, and so we will do so henceforth.

B. Further asymptotic conditions

$\eta(0)$ gives rise to an asymptotic vector field, assumed nonzero, $\eta(u=0,\ r\to\infty)\equiv\eta_\infty(0)$. $\eta_\infty(0)$ is then part of the BMS group at all u. If $\eta(u)$ exists, it will go to $\eta_\infty(u)$ as $r\to\infty$ and so its (A, B, f, g) would have a known u dependence[15] as $r\to\infty$. $\eta(u)$ will have to be defined with that dependence of u as $r\to\infty$. Thus, (SIII. 2) gives $B_r=0$ at *all* u and r; we will then define B off $u=0$ such that $B_r=0$. Similarly for all the other coefficients. Of course, the consistency of all these definitions will have to be shown. (SIII. 2) gives $f^{(-N)} = g^{(-N)} = 0$ for $N>1$ and (BH19e) and (BH19f) at all u and r. Were we to expand everything in (BH19a) in inverse powers of r, the terms to $O(1)$ and below in r must give 0: at all u because of the first of (SIII. 3); at all r because the coefficients in the series are independent of r. Similarly we get $A^{(-M)}=0$ for $N>1$ $\forall\, u, r$ and for all the other equations in (BH19) we will get, for all u,

$$\pounds_\eta g_{00} = 0 \text{ to } O(1), \quad \text{(II. 3a)}$$

$$\pounds_\eta g_{22} = 0 \text{ to } O(r^2), \quad \text{(II. 3c)}$$

$$\pounds_\eta g_{33} = 0 \text{ to } O(r^2), \quad \text{(II. 3d)}$$

$$\pounds_\eta g_{23} = 0 \text{ to } O(r^2), \quad \text{(II. 3g)}$$

$$\pounds_\eta g_{01} = 0 \text{ to } O(r^{-1}), \quad \text{(II. 3h)}$$

$$\pounds_\eta g_{02} = 0 \text{ to } O(1), \quad \text{(II. 3i)}$$

$$\pounds_\eta g_{03} = 0 \text{ to } O(1), \quad \text{(II. 3j)}$$

(II. 3b) is $B_r=0$ $\forall\, u$, (II. 3e) and(II. 3f) are (BH19e) and (BH19f) respectively, $\forall\, u$. Of course, (II. 3) hold at $u=0$ for all orders of r. The form of $\pounds_\eta g_{\alpha\beta}$ in (II. 3) are copied in Appendix A from (BH18).

C. Solution for the vector field: General case

The procedure will continue as follows: (A, B, f, g) will be defined so that (II. 3) hold and so that we get the vector field $\eta(0)$. The higher orders necessary in (II. 3), to show that $\eta(u)$ is Killing, i.e., the rest of (BH19), will be shown to hold. To start we will take $B=1$. From the no-shock condition B will be a regular function of $\cos\theta$ [including $B(u=0)$]: In Appendix C it is shown that B is a constant. Without loss of generality this may be taken to be 0 or 1. If $B=0$ and $B=1$ occur at different regions in the same initial problem, we solve for each initial region separately. Since, as we will see later, in all cases the result is the same, namely evolution and a Weyl metric, this does not affect the result. $B=0$ is treated later. Lind[16] uses a coordinate-tetrad transformation to make $B=1$ (or 0 if not timelike); this, however, could not directly be done here as it could change the metric conditions.

Also notice that (BH24) through (BH39) hold at all u since they are gotten in accordance with (II. 3). Similarly (BH40e), (BH40k), (BH40i) and (BH40j) also hold at all u. The rest of (BH40) hold only at $u=0$.

Now, (II. 3f) gives

$$g_r/g = -(e^\gamma/r)_r/(e^\gamma/r)$$

so that

$$g = re^{-\gamma}e^{-\tilde{A}}$$

with

$$\exp(\tilde{A}) = \exp({}_{(0)}\tilde{A}) + \exp({}_{(-)}\tilde{A}\sin\phi) + \exp({}_{(+)}\tilde{A}\cos\phi)$$

${}^{0}_{(\pm)}\tilde{A}$ arbitrary functions of u, θ. Then (II. 3j) gives since $\tilde{A}_\phi$ is a $\cos\phi$ or $\sin\phi$ term

$${}_{(0)}g_u/{}_{(0)}g + \gamma_u = 0 \text{ to } O(1)$$

so

$${}_{(0)}\tilde{A}_u = 0.$$

Also, (II. 3g) gives, similarly,

$$-\sin\theta e^{-\gamma}\partial_\theta(e^\gamma/\sin\theta) - {}_{(0)}g_\theta/{}_{(0)}g = 0 \text{ to } O(r^2)$$

so that

$$ {}_{(0)}\tilde{A} = -\ln|\sin\theta| + c_1' $$

and

$$ {}_{(0)}g = c_1 r \sin\theta e^{-\gamma}. \tag{II.4} $$

Notice that (II. 3g) holds for all orders of r if there is no ϕ dependence. Similarly, (II. 3e) gives

$$ (e^{-\gamma}/r)f_r + (e^{-\gamma}/r)_r f + U_r = 0 $$

so that

$$ f = -Ure^{\gamma} + H(\theta, \phi, u)\, re^{\gamma}. \tag{II.5} $$

H is arbitrary. But (BH32) gives $f_u^{(-1)} = 0$ and $B = 1$, (BH37) and (BH39) give $f_\theta^{(-1)} = 0$, so that with ${}_{(\pm)}^{0}l = \text{const}$, $H = f^{(-1)} = l(\phi) = {}_{(0)}l + {}_{(-)}l\sin\phi + {}_{(+)}l\cos\phi$. In Sec. IV, ${}_{(+)}l \neq 0$ are treated. ${}_{(+)}l = {}_{(-)}l = 0$, ${}_{(0)}l = 0$ is treated in Appendix B. Here we take $l = 0$ and then (II. 3j) and (II. 3g) give also ${}_{(+)}g = {}_{(-)}g = 0$ as well as

$$ f = -Ure^{\gamma}. \tag{II.6} $$

Next, notice that (SIII2) last gives a relation valid at all u and r between (BH18c), (BH18d), and (BH18g). (BH18g) is already zero for all u and r so that if we had (BH19c) holding at all u and r, (BH19d) would do likewise. Thus we define A for all u and r by (BH19c). It can be solved for A unambiguously. Then all of (BH19) hold for all u and r except possibly (BH19a), (BH19h) and (BH19i). We will see, with our (A, B, f, g) whether they indeed hold, i. e., if they evolve. The definitions of (A, B, f, g) in BH give

$$ \eta^{\alpha} = (B, Ae^{-2\beta} - \tfrac{1}{2}BVr^{-1},\ BU + fe^{-\gamma}/r,\ ge^{\gamma}/r\sin\theta) \tag{II.7} $$

so that using (II. 4), (II. 6), and $B = 1$,

$$ \eta^{\alpha} = (1, Ae^{-2\beta} - \tfrac{1}{2}Vr^{-1}, 0, c_1). \tag{II.8} $$

However, $\eta^{\alpha} = (0, 0, 0, c_1)$ is Killing, and hence setting $c_1 = 0$ will affect nothing. Also, (BH19c) gives, on rearranging, using (II. 6) and $B = 1$,

$$ (Ae^{-2\beta} - \tfrac{1}{2}Vr^{-1})(\gamma_r + 1/r) = -\gamma_u \tag{II.9} $$

while (BH19d) similarly gives

$$ (Ae^{-2\beta} - \tfrac{1}{2}Vr^{-1})(1/r - \gamma_r) = \gamma_u \tag{II.10} $$

so that one must clearly have for the definition of A

$$ Ae^{-2\beta} - \tfrac{1}{2}Vr^{-1} = 0. \tag{II.11} $$

Hence,

$$ \eta^{\alpha} = \delta_u^{\alpha}. \tag{II.12} $$

From arguments in Sec. I this will evolve to fill out the region M of space–time so that the remaining three equations will evolve. The definition of A by (II. 11) will not be inconsistent with any of the three remaining equations: E. g., (BH19h) to $O(1/r^2)$ gives (BH79) which gives $A^{(1)} = -M$. But (II. 11) gives just that; further (BH19h) gives, using (II. 11), $\beta_u = 0$, and is thus consistent. Similarly (BH19a) gives $\partial_u(Vr^{-1}e^{2\beta} - U^2r^2e^{2\gamma}) = 0$ and (BH19i) has $\gamma_u = 0$, both consistent. The definition will, of course, be consistent with the initial data.

In the cases considered so far it has emerged that the isometry evolves and M must be Weyl.

III. SPECIAL CASE: $B = 0$, ${}_{(\pm)}l = 0$

Checking back on the discussion leading to (II. 8), g can be taken to be 0. (II. 3e) gives

$$ f = H(\theta, u)\, re^{\gamma} $$

and then (BH24) gives $H = 0$. (BH10c) gives $A = 0$ at $u = 0$ and (BH19a) gives $A_u = 0$ at $u = 0$. A can be defined to equal zero without inconsistency. This is a trivial subcase.

IV. SPECIAL CASE: ${}_{(-)}l, {}_{(+)}l \neq 0$

Since $(\cos\phi, \sin\phi)$ are linearly independent in $L^2(0, 2\pi, R)$, (II. 3j) and (II. 3g) give respectively, at $u = 0$, for $B = 1$ or $B = 0$,

$$ -{}_{(\pm)}\tilde{A}_u re^{-\gamma}\exp(-(\pm)\tilde{A}) \pm {}_{(\mp)}fUe^{2\gamma}/\sin\theta = 0 \tag{IV.1} $$

and

$$ -(\gamma_\theta - \cot\theta)re^{-\gamma}\exp(-(\pm)\tilde{A}) - re^{-\gamma}\exp(-(\pm)\tilde{A})(-\gamma_\theta - {}_{(\pm)}A_\theta) $$
$$ \mp {}_{(\mp)}fe^{2\gamma}/\sin\theta = 0 \tag{IV.2} $$

with the obvious notation

$$ f = {}_{(0)}f + {}_{(-)}f\sin\phi + {}_{(+)}f\cos\phi \tag{IV.3} $$

and so on. (IV. 2) may be written as

$$ \exp(-(\pm)\tilde{A})({}_{(\pm)}\tilde{A}_\theta + \cot\theta) = \mp {}_{(\mp)}le^{4\gamma}/\sin\theta \tag{IV.4} $$

with ${}_{(\pm)}l$ constants. Clearly $a\cos\phi$ in f gives $a\sin\phi$ in g and vice versa. Now, since the left-hand side of (IV. 4) is independent of r, the right-hand side will also be so at $u = 0$, i. e., $\gamma = 0$ at $u = 0$. (IV. 1) gives $U = 0$, so $C = N = 0$; then, (B22) gives $\beta = 0$ and (B24) gives $V = r - 2M$, $M_0 = 0$ from (B35). (IV. 1) also gives ${}_{(\pm)}\tilde{A}_u = 0$ at $u = 0$. If $B = 1$, (BH19c) and (BH19d) give respectively, at $u = 0$

$$ (A - \tfrac{1}{2}Vr^{-1})(1/r) = -\gamma_u, \tag{IV.5} $$

$$ (A - \tfrac{1}{2}Vr^{-1})(1/r) = \gamma_u - {}_{(0)}l\cos\theta/\sin\theta \tag{IV.6} $$

so that ${}_{(0)}l = 0$, $\gamma_u = 0$ and ${}_{(0)}A = Vr^{-1}/2$. Then (BH19a) for ${}_{(\pm)}f$ gives $M_2 = 0$ and then (B36) gives $N_0 = 0$, all at $u = 0$. An examination of Paper B then clearly shows that, since $c_0(u) = 0$, the initial value problem evolves uniquely as spherically symmetric, i. e., $\gamma = \beta = U = 0$, etc., continue at all u. If $B = 0$, (BH19c) gives $A = 0$, and (BH19d) gives ${}_{(0)}l = 0$ both at $u = 0$. (BH19a) for ${}_{(\pm)}f$ gives $M_2 = 0$ and then as before the metric evolves spherically. Birkhoff theorem, of course, gives us a static space–time.

V. SECOND PART. SMALL ANISOMETRIES: SMALL PERTURBATIONS OFF WEYL

Since imposing an isometry in Bondi gives us a Weyl space–time, one becomes interested in the next best thing: imposing an "almost symmetry." If by this is meant taking a vector field η and taking $\mathcal{L}_{(\eta+\delta\eta)}g_{\alpha\beta} = O(\epsilon)$, where ϵ is a small parameter which will denote the magnitude of the symmetry breaking, then this will be a space–time very near to a Weyl one, i. e., $g_{\alpha\beta} = {}_{w}g_{\alpha\beta} + \epsilon h_{\alpha\beta}$ and

$$ \mathcal{L}_{(\eta+\delta\eta)}g_{\alpha\beta} = \mathcal{L}_{\eta w}g_{\alpha\beta} + \epsilon\mathcal{L}_{\eta}h_{\alpha\beta} + \epsilon\mathcal{L}_{(\delta\eta)w}g_{\alpha\beta}. $$

The last term may be removed by a gauge transformation. We see that if we wish to stick to our definition of "almost symmetry," η will have to be a Killing vector field of ${}_{w}g_{\alpha\beta}$; if we do not necessarily wish a timelike

vector or axial symmetry, the base Weyl space–time will have to be thus restricted. The equations for any perturbation $h_{\alpha\beta}$ may be written down[17] and compared[18] with those for $\mathcal{L}_\eta h_{\alpha\beta}$, and one sees that equations for both, in vacuum, are identical. Metric perturbations of Schwarzchild space–times have been treated extensively in the literature,[19,20] spurred by interest in radiation near nonrotating black holes. Metric perturbations in Weyl space–times have also been carried out.[21] The expansions used have been essentially of two types: a double series of multipole moments–multipole moments,[21] and in a $1/r$ series and later in multipole moments. Objections and convergence problems of the latter method are not important if one remains far away (but *not* at ∞) from the source. Here we carry out a $1/r$ expansion and calculate perturbations far away in a way that, very simply, would show some second order perturbations effects had one started with linear theory.[21] Of course, higher order perturbations in this scheme could be carried out, and the convergence is an open question. We start with Weyl and perturb it, keeping the axial symmetry and nonrotation, and thus use Bondi's method linearized about Weyl. In a very simple way nonlinear effects appear. The only drawback seems to be that one never really relates masses and so on to exact sources (e.g., fluid or kinetic densities). This is a standing unsolved problem in all of gravitation.

We start with a Weyl $s-t$ transformed to a coordinate system where $c=0$.[7,22] A perturbation will involve a small change in c to $\epsilon f(u,\theta)$, ϵ a small parameter to be specified later. We will keep only terms of $O(\epsilon)$. (B35) gives

$$M=\tfrac{1}{2}\epsilon F+m_w, \tag{V.1}$$

where

$$F=f_{22}+3f_2\cot\theta-2f, \tag{V.2}$$

and m_w is the constant mass in the Weyl unperturbed case. (B36) then gives

$$N_0=-\tfrac{1}{6}\epsilon F_2. \tag{V.3}$$

(B34) gives

$$4C_{00}=\epsilon[2m_w f_0-\tfrac{1}{6}(F_2\cot\theta-F_{22})]. \tag{V.4}$$

We will now calculate the quadruple moment. There are different nonequivalent definitions and we used Eq. (3) in Newman and Unti,[23]

$$Q(u)=3l\int_0^\pi CP_2^2(\cos\theta)\sin\theta\,d\theta \tag{V.5}$$

(l a numerical factor so that the linear approximation should give the right answer), where P_2^2 is the $\binom{2}{2}$ associated Legendre polynomial

$$P_2^2=3\sin^2\theta; \tag{V.6}$$

then, using (V.4),

$$\ddot{Q}=l9\epsilon\int_0^\pi[2m_w f_0-\tfrac{1}{6}(F_2\cot\theta-F_{22})]\sin^3\theta\,d\theta \tag{V.7}$$

This integral will of course depend on f. So will F. Expanding f in Legendre polynomials in $\cos\theta$ with u dependent coefficients

$$f=\sum_{n=0}^{\infty} d_n(u)P_n(\cos\theta). \tag{V.8}$$

Then, from (V.2), with $\Psi'\equiv d\Psi/d(\cos\theta)$

$$F=\sum_{n=0}^{\infty} d_n\{[1-\cos^2\theta]P_n''-4\cos\theta P_n'-2P_n\}$$

and since P_n satisfies its differential equation

$$F=\sum_{n=0}^{\infty} d_n\{-2\cos\theta P_n'-[2+n(n+1)]P_n\}; \tag{V.9}$$

since $P_n'\cos\theta=\sum_{l=0}^{n}K_{ln}P_l$,

$$F=\sum_{n=0}^{\infty} a_nP_n \tag{V.10}$$

with

$$a_n=-2\sum_{m=n}^{\infty} d_mK_{mn}-d_n[2+n(n+1)], \tag{V.11}$$

so that if f has $d_n=0$ for $n>k$, $a_n=0$ for $n>k$ also. Since

$$F_2\cot\theta-F_{22}=(F')_2\sin\theta,$$

we get, integrating by parts,

$$\int_0^\pi(-F_{22}+F_2\cot\theta)\sin^3\theta\,d\theta=-8\int_{-1}^{1}FP_2(\cos\theta)\,d\cos\theta). \tag{V.12}$$

Also,

$$\int_0^\pi f_0\sin^3\theta\,d\theta=\tfrac{2}{3}\int_{-1}^{1}f_0[P_0-P_2]d(\cos\theta). \tag{V.13}$$

Thus, we see from (V.11), (V.12), and (V.13), and the orthogonality of the Legendre polynomials, that the only contributions to $\ddot{Q}$ will be from the $n=0$ and $n=2$ terms of f and the $n=2$ term of $F(i_e,d_0,d_2,a_2)$. Higher terms of f, i.e., d_4,d_6 and so on, will only contribute in a_2. This is because

$$\int_{-1}^{1}\sum_{n=0}^{\infty}-2d_n\cos\theta P_n'P_2\,d(\cos\theta)$$

$$=2\sum_{n=0}^{\infty}[d_n\{\tfrac{6}{5}\delta_{n2}+2\delta_{n0}\}-2\ (\text{if } n \text{ even})].$$

Hence,

$$\ddot{Q}=4\epsilon l[6m_w(d_0^{\cdot}-\tfrac{1}{5}d_2^{\cdot})+9(-\tfrac{72}{5}d_2)], \tag{V.14}$$

where the superscript dot indicates d/du. This is exact to $O(\epsilon)$ and relates the quadrupole moments to two coefficients of c. The lowest order effect in m_0 is, from (B58),

$$m_0=-\epsilon^2[d_0^{\cdot 2}+\tfrac{1}{3}d_1^{\cdot 2}+\tfrac{1}{5}d_2^{\cdot}+\cdots]. \tag{V.15}$$

It is clear that the $d_1^{\cdot}$, $d_i^{\cdot}$ ($i>2$) are independent of Q, even at higher orders of ϵ. Thus, at least part of the radiated energy is not related to quadrupole moment change. If we, however, take the 0 and 2 terms of f assuming the $i>2$ terms and d_1 to be much less than the others (or simply ignoring their effects on m_0) and if we let time (u) derivatives be much less than $1/m_w$, i_e, for any coefficient d_i, $\ddot{d}_i m_w \ll d_i^{\cdot}$, so that the background should be strong and the change slow, but not totally negligible, and so that ϵ could be roughly taken as the energy radiated/total energy and $\epsilon \ll m_w/T$, where

$1/T=$ rate of energy radiated/total energy

as that the approximation is consistent, then

$$\dddot{Q}^{\cdot}=(4l)^2\epsilon^2[6m_w(d_0^{\cdot\cdot}-\tfrac{1}{5}d_2^{\cdot\cdot})-\tfrac{72}{5}d_2^{\cdot}]^2$$

$$\approx\epsilon^2l^2\{-[(72)^2\cdot 16/25]\dot{d}_2^2+12m_w(-\tfrac{1}{5}d_2^{\cdot\cdot})\ddot{Q}^{\cdot}/l\epsilon\} \tag{V.16}$$

where we have set in the last time $\ddot{Q}_2^{\cdot\cdot} \gg \ddot{Q}_0^{\cdot\cdot}$. Then using (V. 15),

$$-\ddot{Q}^{\cdot 2} \approx [(72)^2/5]16l^2 m_0 + \tfrac{12}{72} m_w \epsilon \dddot{Q}^{\cdot\cdot\cdot} \ddot{Q}^{\cdot\cdot},$$

or

$$m_0 \approx -[5/16l^2(72)^2]\dddot{Q}^{\cdot\cdot\cdot 2} - [60/(72)^3] m_w \dddot{Q}^{\cdot\cdot\cdot} \ddddot{Q}^{\cdot\cdot\cdot\cdot}. \quad \text{(V. 17)}$$

The first term is the usual one that relates radiated energy to change in quadrupole moment (different numbers depending on "normalization" l). The second, since it has a term m_w, is an interaction of the radiation with the background.[19] This approximate expression is similar to those in Refs. 19–21 and represent radiative tails. Further multipole moments may be defined as in Ref. 23 so as to write more exact expressions for m_0 in terms of them from (V. 14). The advantage in the perturbative approach here used is the mathematical simplicity. The disadvantage is the usual not-apparent interpretation, including that of Q and the approximations used. If Q were related to source densities, Eq. (V. 17) could be used to calculate the energy radiated by a massive axially symmetric static object changing slowly. More exact expressions are needed to disentangle further the nonlinearities, but it is helpful that even in this simple approximation to $O(\epsilon)$ some appear. Care should be exercised in that there are different (usually numerical) definitions for Q.[7,19,23] Also if instead of Bondi's m_0 we used Newman and Unti's,[23] there is a relation m_0 (Bondi) $= m_0$ (N. U.) $- 1/2\int_0^\pi (c_0^2 + cc_{00}) \sin\theta\, d\theta$ so that we may again get an approximate different expression for radiated energy. For a discussion of mass and multipole moments in relation to Bondi's formalism see Refs. 23 and 24. For a nice invariant treatment and generalization of Bondi's formalism see Ref. 25.

APPENDIX A

Some errors in (BH) have been taken care of here:

$$\begin{aligned}\pounds_\eta g_{00} = {} & f[e^{-\gamma}\partial_\theta(Vr^{-2}e^{2\beta} - U^2 re^{2\gamma}) + 2Ure^{2\gamma}\partial_u e^{-\gamma}] \\ & + 2f_u Ure^{\gamma} + B\partial_u(Vr^{-1}e^{2\beta} - U^2 r^2 e^{2\gamma}) \\ & + (Ae^{-2\beta} - \tfrac{1}{2}BVr^{-1})\partial_r(Vr^{-1}e^{2\beta} - U^2 r^2 e^{2\gamma}) \\ & + BU\partial_\theta(Vr^{-1}e^{2\beta} - U^2 r^2 e^{2\gamma}) + 2(Vr^{-1}e^{2\beta} - U^2 r^2 e^{2\gamma}) B_u \\ & + 2e^{2\beta}\partial_u(Ae^{-2\beta} - \tfrac{1}{2}BVr^{-1}) + 2Ur^2 e^{2\gamma}\partial_u(BU),\end{aligned} \quad \text{(BH18a)}$$

$$\pounds_\eta g_{11} = e^{2\beta}\partial_r B, \quad \text{(BH18b)}$$

$$\begin{aligned}\pounds_\eta g_{22} = {} & -2f_\theta re^{\gamma} + B\partial_u(-r^2 e^{2\gamma}) - (Ae^{-2\beta} - BVr^{-1}/2)\partial_r(r^2 e^{2\gamma}) \\ & - BU\partial_\theta(r^2 e^{2\gamma}) + 2Ur^2 e^{2\gamma}\partial_\theta B - 2r^2 e^{2\gamma}\partial_\theta(BU),\end{aligned} \quad \text{(BH18c)}$$

$$\begin{aligned}\pounds_\eta g_{33} = {} & (fe^{-\gamma}/r)\partial_\theta(-r^2 e^{-2\gamma}\sin^2\theta) - 2g_\phi re^{-\gamma}\sin\theta \\ & + B\partial_u(-r^2 e^{-2\gamma}\sin^2\theta) + (Ae^{-2\beta} - \tfrac{1}{2}BVr^{-1})\partial_r \\ & (-r^2 e^{-2\gamma}\sin^2\theta) + BU\partial_\theta(-r^2 e^{-2\gamma}\sin^2\theta),\end{aligned} \quad \text{(BH18d)}$$

$$\begin{aligned}\pounds_\eta g_{12} = {} & f[-r^2 e^{2\gamma}\partial_r(e^{-\gamma}/r)] - f_r re^{\gamma} + e^{2\beta}B_\theta - r^2 e^{2\gamma}(BU)_r \\ & + Ur^2 e^{2\gamma}B_r,\end{aligned} \quad \text{(BH18e)}$$

$$\pounds_\eta g_{13} = -g_r re^{-\gamma}\sin\theta - gr^2 e^{-2\gamma}\sin\theta(e^{\gamma}/r)_r + e^{2\beta}B_\phi, \quad \text{(BH18f)}$$

$$\pounds_\eta g_{23} = -gre^{-2\gamma}\sin^2\theta\partial_\theta(e^{\gamma}/\sin\theta) - f_\phi re^{\gamma} - g_\theta re^{-\gamma}\sin\theta, \quad \text{(BH18g)}$$

$$\begin{aligned}\pounds_\eta g_{01} = {} & f[(e^{-\gamma}/r)\partial_\theta(e^{2\beta}) + Ur^2 e^{2\gamma}\partial_r(e^{-\gamma}/r)] + f_r Ure^{\gamma} \\ & + B\partial_u e^{2\beta} + (Ae^{-2\beta} - \tfrac{1}{2}BVr^{-1})\partial_r e^{2\beta} + BU\partial_\theta e^{2\beta} \\ & + e^{2\beta}\partial_r(Ae^{-2\beta} - \tfrac{1}{2}BVr^{-1}) + Ur^2 e^{2\gamma}BU_r + e^{2\beta}B_u \\ & + B_r(Vr^{-1}e^{2\beta}),\end{aligned} \quad \text{(BH18h)}$$

$$\begin{aligned}\pounds_\eta g_{02} = {} & f[\partial_\theta(Ure^{\gamma}) - re^{2\gamma}\partial_u e^{-\gamma}] - f_u re^{\gamma} + f_\theta Ure^{\gamma} \\ & + B\partial_u(r^2 Ue^{2\gamma}) + (Ae^{-2\beta} - \tfrac{1}{2}BVr^{-1})\partial_r(Ur^2 e^{2\gamma}) \\ & + (Vr^{-1}e^{2\beta} - U^2 r^2 e^{2\gamma})B_\theta + e^{2\beta}\partial_\theta(Ae^{-2\beta} - \tfrac{1}{2}BVr^{-1}) \\ & + Ur^2 e^{2\gamma}\partial_\theta(BU) + Ur^2 e^{2\gamma}B_u - r^2 e^{2\gamma}(BU)_u \\ & + BU_\theta(r^2 e^{2\gamma}U),\end{aligned} \quad \text{(BH18i)}$$

and, lastly,

$$\pounds_\eta g_{03} = A_\phi - g(re^{-\gamma}\sin\theta\gamma_u) + f_\phi Ure^{\gamma} - g_u re^{-\gamma}\sin\theta. \quad \text{(BH18j)}$$

APPENDIX B

$_{(\pm)}l = 0, \quad _{(0)}l \neq 0, \quad B = 1.$

$_{(0)}l = K_1$, (BH40e) and equations preceding (II. 6) give

$$f^{(1)} = 2K_1 c + (c_2 + 2c\cot\theta).$$

(BH40i), while using (BH24), gives

$$-K_1(c_2 + 2c\cot\theta)_\theta + A_\theta^{(0)} = 0. \quad \text{(B1)}$$

(BH40c) and (BH40d) actually hold for all u since all the terms, from (BH34) and with $g^{(0)} = 0$, are independent of u. Thus

$$-f_\theta^{(0)} = c_\theta K_1 - f^{(0)}\cot\theta$$

and (II. 5), with $H = K_1$, gives

$$f^{(0)} = K_1 c$$

so that $K_1 = 0$ or $2c_2 + c\cot\theta = 0$. Ignoring the already treated former case, and with $A^{(0)} - 1/2 = -K_1 c_2$, and using (B. 1), $c = 0$. (II. 4) gives $g^{(-1)} = c_1$, $g^{(0)} = g^{(1)} = 0$, and there is also $A^{(-1)} = f^{(0)} = f^{(1)} = 0$, $A^{(0)} = 1/2$. (BH40a) gives $A_u^{(1)} = K_1 M_2$ at $u = 0$, (BH19h) gives $A^{(1)} = M$ at $u = 0$, and defining A by (BH19c) as before, $A^{(1)} = -M$ at all u so that $M_2 = 0$, using (B35). As before, one one may get

$$\eta^\alpha = \delta_u^\alpha + K_1\delta_\theta^\alpha, \quad \text{(B2)}$$

which means $g_{\alpha\beta} = g_{\alpha\beta}(u - \theta/K_1)$; $\alpha = \beta = 0$ give clearly $M_2 = 0$ at $u = 0$ anyway: Again there is no inconsistency nor additional imposed metric conditions. (B36) now gives $N_0 = 0$, (B34) $C_{00} = 0$, (BH19i) gives $U_\theta = 0$, (B2) then gives $U_u = 0$, all at $u = 0$. One may continue using (B2) and equations in B and BH to get zero for all functions. This is a trivial subcase, flat space–time.

APPENDIX C

First, (II. 4) will still hold as well as the equations before it. (IV. 1) will also hold, and so will (IV. 4) except one would have, to start, on the right-hand side of (V. 4) $\pm_{(\pm)} fe^{3\gamma}/r\sin\theta$ and on the left of (IV. 1) $\pm_{(\pm)} Ae^{\gamma}/r\sin\theta$. The argument below it would give

$$_{(\pm)}f = _{(\pm)}G(u,\theta)\, re^{-3\gamma}.$$

(BH19e) gives ${}_{(\pm)}f = {}_{(\pm)}H(u,\theta)re^{\gamma}$ so that, since $[{}_{(\pm)}G/{}_{(\pm)}H]_r = 0$, $\gamma = 0$. Then (IV.1) gives ${}_{(\pm)}\tilde{A}_u = 0$, ${}_{(\pm)}G\overset{(n)}{U} + {}_{(\mp)}A^{(n-2)} = 0$; with $n=1$ we get ${}_{(\pm)}A^{(-1)} = 0$ so that ${}_{(\pm)}f^{(-1)} = \text{const} = G$. If ${}_{(\pm)}G = 0$, then ${}_{(\pm)}g = 0$ and ${}_{(\pm)}A = 0$ and there is no ϕ dependence (treated below). But (BH19h) gives

$$ {}_{(\pm)}A_r + 2e^{2\beta}\beta_\theta = 0 \quad \text{so that} \quad A^{(1)} = 0, $$

and since (B22) gives $\beta = 0$, we must have ${}_{(\pm)}A = 0$ from (BH40c), and thus $U = 0$ at $u = 0$. (BH19c) and (BH19d) give with $B \neq 0$, respectively, (IV.5) and (IV.6) with the former's right-hand side changed to $\gamma_u - {}_{(0)}f\cos\theta/r\sin\theta$ and both left-hand sides changed $V \to BV$, and $A \to {}_{(0)}A$. (BH19e) also implies ${}_{(0)}f = T(u,\theta)r$, and so, comparing (IV.5) and (IV.6), we get $\gamma_u = 0$, $T = 0$. Everything, so far, is at $u = 0$. The argument below (IV.6), from "then (BH19)" on, is then valid. Thus, if there is a ϕ dependence on (f, g) we must have spherical symmetry, and ${}_{(\pm)}A = 0$.

If there is no ϕ dependence anywhere the argument above does not follow. One proceeds as follows in order to set $B = 1$ or 0.

Taking ∂_u on (BH40c) and (BH40d) (we will define $A^{(0)}$ for all u in such a way) gives $K = c_5\cos\theta$ (c_5 const), and previous equations in BH then give $f^{(-1)} = c_5\sin\theta$, $A^{(-1)} = -c_5\cos\theta$. (BH40c) and (BH40d) give (without ∂_u)

$$ -c_5(c_\theta\sin\theta + c\cos\theta) + m/2 + m'' = c_5(c_\theta\sin\theta - c\cos\theta) + m'\cot\theta + m/2 \tag{C1} $$

while (BH40i) gives

$$ c_5\{\sin\theta\overset{(2)}{U}_\theta - c\sin\theta + 2\cos\theta\overset{(2)}{U}\} + n' + m'/2 = 0 \tag{C2} $$

with

$$ n = -c_5(c\sin\theta)_\theta + m/2 + m''. \tag{C3} $$

If one now expands $c/\sin^2\theta$ in $\cos\theta$ near $\theta = 0$ ($c/\sin^2\theta$ must be regular[7] there) and m in a Taylor series at $\theta = 0$ (B must be regular also),

$$ c = d^{(2)}\theta^2 + d^{(4)}\theta^4 + \cdots, $$

$$ m = m^{(0)} + m^{(1)}\theta + \cdots, $$

and uses the above two equations for c and m, one gets $c_5 d^{(2)} = 0$, $m^{(0)}$ arbitrary, $m^{(1)} = m^{(3)} = 0$, $8m^{(4)} + \frac{2}{3}m^{(2)} = 0, \cdots$. Since $d^{(2)}$ may be taken different from 0 (by a supertranslation[7] $K = 1$, $\alpha = \alpha(\bar{\theta})$, one can always do so; also set $c \neq 0$), we must have $c_5 = 0$. Then one may solve exactly above to get $B = \tilde{c}_8\cos\theta + \tilde{c}_9$ and if $\tilde{c}_8 \neq 0$, take $B = c_9 + \cos\theta$; now, (BH19e) gives

$$ fe^{-\gamma}/r = -BU - \int(e^{2\beta-2\gamma}/r^2)\sin\theta\,dr, \tag{C4} $$

and with

$$ L = -re^{\gamma}\int(e^{2\beta-2\gamma}/r^2)\sin\theta\,dr, $$

(BH19c) and (BH19d) give, respectively,

$$ (Ae^{-2\beta} - BVr^{-1}/2)(1/r + \gamma_r) = -U\sin\theta - L_\theta/re^{\gamma} - B\gamma_u \tag{C5} $$

and

$$ (Ae^{-2\beta} - BVr^{-1}/2)(1/r\gamma_r) = B\gamma_u - (Le^{\gamma}/2r\sin^2\theta)\,\partial_\theta(e^{-2\gamma}\sin^2\theta). \tag{C6} $$

Checking both to $O(1/r^2)$, we get, respectively,

$$ -c(A^{(0)} - B/2) + (A^{(1)} - BM) = (c\sin\theta)_\theta, $$

$$ +c(A^0 - B/2) + (A^{(1)} - BM) = (c\sin\theta)_\theta, $$

so that $c(A^{(0)} - B/2) = 0$. With c set not equal to 0 by a supertranslation, $A^{(0)} - B/2 = -\cos\theta = 0$. This is nonsense so that $\tilde{c}_8$ must have been 0. Thus $b = c_9$ constant.

*Present address: Inco Inc., 7916 Westpark Dr., McLean, Va.

†Supported in part by Consejo Nacional de Investigaciones Cientificas y Tecnologicas, Caracas, Venezuela.

[1]D. Yaacobi, J. Math. Phys. **16**, 1145 (1975).
[2]W. Halliday, J. Math. Phys. **16**, 322 (1975).
[3]R. A. d'Inverno, J. Math. Phys. **16**, 670, 674 (1975).
[4]W. C. Hernández, Phys. Rev. **153**, 1359 (1967).
[5]R. Berezdivin, L. Herrera, and L. Witten, in preparation (1976).
[6]R. Berezdivin, L. Herrera, Phys. Rev. D **11**, 2063 (1975).
[7]H. Bondi, M.G.J. Van der Burg, and A.W.K. Metzner, Proc. Roy. Soc. (London) A **269**, 21 (1962). For the BMS group and the "retarded time," see also R. Penrose, in *Group Theory in Non-Linear Problems*, edited by A.O. Barut (Reidel, Boston, 1975).
[8]R. Berezdivin, J. Math. Phys. **15**, 1963 (1974).
[9]W.B. Bonnor, in *Atti del Convegno sulla Relativita Generale* (Barbera, Firenze, 1965).
[10]A. Janis and E. Newman, J. Math. Phys. **6**, 902 (1965); R. Torrence and A. Janis, J. Math. Phys. **8**, 1355 (1967). See also Ref. 20.
[11]E. Newman and R. Penrose, J. Math. Phys. **3**, 566 (1962).
[12]A. Papapetrou, J. Math. Phys. **6**, 1405 (1965).
[13]I. Moret-Bailly, Ann. Inst. H. Poincaré A **9**, 395 (1968).
[14]R. L. Bishop and S.I. Goldberg, *Tensor Analysis on Manifolds* (Macmillan, New York, 1968).
[15]R.K. Sachs, Phys. Rev. **128**, 2851 (1962).
[16]R.W. Lind, J. Math. Phys. **16**, 39 (1975).
[17]D.W. Sciama, P.C. Waylen, and R.C. Gilman, Phys. Rev. **187**, 1762 (1969).
[18]R. Berezdivin, Rev. Méx. Fís. **23**, 365 (1974).
[19]J.M. Bardeen and W.H. Press, J. Math. Phys. **14**, 7 (1973). See also references therein. For a review see, e.g., W.H. Press, Ann. N.Y. Acad. Sci. **224**, 272 (1974).
[20]W.E. Couch, R.J. Torrence, A.I. Janis, and E.T. Newman, J. Math. Phys. **9**, 483 (1968).
[21]W.B. Bonnor and M.A. Rottemberg, Proc. Roy. Soc. (London) A **289**, 247 (1965). Also see W.B. Bonnor, in *Ondes et radiations gravitationnelles* (CNRS, Paris, 1973).
[22]M.G.J. Van der Burg, Proc. Roy. Soc. A **294**, 112 (1966).
[23]E.T. Newman and T.W.J. Unti, J. Math. Phys. **6**, 1806 (1965).
[24]H. Bondi, in *Atti del Convegno sulla Relativita Generale* (Barbera, Firenze, 1965). See mainly the discussion at the end.
[25]R.K. Sachs, Proc. Roy. Soc. A **270**, 103 (1962).

Time-dependent dynamical symmetry mappings and associated constants of motion for classical particle systems. II

Gerald H. Katzin,* Jack Levine,† and Robert N. Sane*

Department of Physics and Department of Mathematics,† North Carolina State University, Raleigh, North Carolina 27607*
(Received 3 September 1976)

This paper is a continuation of a previous paper with a similar title [J. Math. Phys. **17**, 1345 (1976)]. In this paper we develop further properties of time-dependent symmetries of dynamical systems expressible in the form (a) $E^i(\ddot{x},\dot{x},x,t) \equiv E^i(\ddot{x}^1,...,\ddot{x}^n;\ \dot{x}^1,...,\dot{x}^n;\ x^1,...,x^n;t) = 0$. Such dynamical symmetries are based upon infinitesimal transformations of the form (b) $\bar{x}^i = x^i + \delta x^i$, $\delta x^i \equiv \xi^i(x,t)\delta a$, (c) $\bar{t} = t + \delta t$, $\delta t \equiv \xi^0(x,t)\delta a$, which satisfy the condition (d) $\delta E^i = 0$ whenever $E^j = 0$. It is shown that if $(\xi^i_A,\ \xi^0_A)$, $A = 1,...,\rho$, is a complete set of solutions of the symmetry equations as determined by (d), then these solutions generate a ρ-parameter complete group of symmetry mappings, and the group structure implies linear dependency relations between first and second derived time-dependent constants of motion as obtained by a related integral theorem. The complete groups of time-dependent symmetry mappings are obtained for all conservative systems ($n > 1$) with spherically symmetric potentials. These groups are classified into six types according to the associated form of the potential. A similar analysis leads to three types of Noether symmetries. In the case where (a) takes the form (e) $E^i(\ddot{x},\dot{x},x) = 0$, it is shown that if $(\xi^i,\ \xi^0)$ defines a symmetry mapping then in general $(\partial^K\xi^i/\partial t^K,\ \partial^K\xi^0/\partial t^K)$, $K = 1,2,...$, will also define symmetry mappings; similar properties are shown for Noether symmetries. These results when applied to a large class of time-dependent constant of motion defined in terms of $(\xi^i,\ \xi^0)$ lead to further contants of motion.

1. INTRODUCTION

With respect to a classical particle dynamical system a dynamical symmetry is defined as a mapping of the set of system trajectories into itself.

In a previous paper[1] (Paper I of this series) we developed a gauge invariant formulation of time-dependent symmetry mappings and associated constants of motion for Lagrange's equations. In this paper we continue this work by developing further properties of time-dependent dynamical symmetries, and exemplifying several aspects of the theory.

We shall now consider dynamical systems expressible in the form[2]

$$E^i(\ddot{x},\dot{x},x,t) \equiv E^i(\ddot{x}^1,\ldots,\ddot{x}^n;\ \dot{x}^1,\ldots,\dot{x}^n;\ x^1,\ldots,x^n;\ t) = 0. \tag{1.1}$$

In Ref. 1 we defined Type I (infinitesimal) mappings by

$$\bar{x}^i = x^i + \delta x^i, \quad \delta x^i \equiv \xi^i(x,t)\,\delta a, \tag{1.2}$$

$$\bar{t} = t + \delta t, \quad \delta t \equiv \xi^0(x,t)\,\delta a. \tag{1.3}$$

Based upon such mappings the δ variations $\delta\dot{x}^i$ and $\delta\ddot{x}^i$ were defined respectively by[1]

$$\delta\dot{x}^i \equiv \frac{d\bar{x}^i}{d\bar{t}} - \frac{dx^i}{dt} = (\dot{\xi}^i - \dot{x}^i\dot{\xi}^0)\delta a, \tag{1.4}$$

$$\delta\ddot{x}^i \equiv \frac{d^2\bar{x}^i}{d\bar{t}^2} - \frac{d^2x^i}{dt^2} = (\ddot{\xi}^i - \dot{x}^i\ddot{\xi}^0 - 2\ddot{x}^i\dot{\xi}^0)\delta a. \tag{1.5}$$

For any function $G(\ddot{x},\dot{x},x,t)$, δG is defined by

$$\delta G \equiv \frac{\partial G}{\partial \ddot{x}^i}\,\delta\ddot{x}^i + \frac{\partial G}{\partial \dot{x}^i}\,\delta\dot{x}^i + \frac{\partial G}{\partial x^i}\delta x^i + \frac{\partial G}{\partial t}\delta t. \tag{1.6}$$

A Type I mapping will define a symmetry mapping of the dynamical system (1.1) if

$$\delta E^i = 0 \ \text{(whenever } E^i = 0). \tag{1.7}$$

To obtain the explicit conditions on the mapping functions $\xi^\alpha(x,t)$, $(\equiv \xi^0,\ \xi^i)$, $\alpha = 0,1,\ldots,n$, in order that (1.7) be satisfied we proceed as follows:

(a) We expand (1.7) by means of (1.6) in which δx^i, δt, $\delta\dot{x}^i$, and $\delta\ddot{x}^i$ are expressed by (1.2), (1.3), (1.4), and (1.5), respectively.

(b) In the resulting equation we eliminate the $\ddot{x}^i$ terms by means of (1.1) [which we assume to be solvable for $\ddot{x}^i$]. Since $\delta\dot{x}^i$ and $\delta\ddot{x}^i$ are linear and homogeneous in the ξ^α, $\xi^\alpha_{,\beta}$, $\xi^\alpha_{,\beta\gamma}$ the equations so obtained will be of the form

$$G^i_\alpha(\dot{x},x,t)\xi^\alpha + G^{\beta i}_\alpha(\dot{x},x,t)\xi^\alpha_{,\beta} + G^{\beta\gamma i}_\alpha(\dot{x},x,t)\xi^\alpha_{,\beta\gamma} = 0. \tag{1.8}$$

(c) The explicit symmetry equations for the $\xi^\alpha(x,t)$ are obtained by considering (1.8) as identically zero in the $\dot{x}^i$ variables (since otherwise they would impose constraints on the dynamical system). The $\xi^\alpha(x,t)$ which satisfy these symmetry equations will be referred to as symmetry solutions.

By means of (1.2), (1.3) such symmetry solutions determine the above-mentioned Type I symmetry mappings of a dynamical system (1.1).

In Sec. 2 it is shown that if a dynamical system of the form (2.1) admits a symmetry mapping defined by $[\xi^i(x,t),\ \xi^0(x,t)]$, then in general it also admits a symmetry mapping defined by $[\partial^K\xi^i/\partial t^K,\ \partial^K\xi^0/\partial t^K]$, $K = 1,2,\cdots$. A similar property is shown to hold for Type I Noether symmetries. These results when applied to a large class of time-dependent constants of motion based upon the functions $[\xi^i,\ \xi^0]$ lead to further constants of motion.

In Sec. 3 it is shown that if $[\xi^i_A(x,t),\ \xi^0_A(x,t)]$, $A = 1,2,\ldots,\rho$, is a complete set of solutions of the

symmetry equations based upon a dynamical system (1.1), then these solutions generate a ρ-parameter (complete) group of symmetry mappings. It is shown that the group structure implies linear dependency relations between first and second derived time-dependent constants of motion obtained by use of a related integral theorem.[3,4]

In Sec. 4 we give an example of a class of time-dependent Lagrangians which satisfy condition [Ref. 1, (4.12)]. (It was shown in Ref. 1, Sec. 4 that if the dynamical systems defined by such Lagrangians admit Type I symmetries, then a time-dependent constant of motion C_1 exists.)

In Secs. 5—9 explicit time-dependent symmetry equations for conservative systems are derived. Based upon the solutions of these equations the complete groups of time-dependent symmetry mappings are obtained for all such systems ($n>1$) with spherically symmetric potentials. These groups are classified into six types according to the form of the potential.

In Sec. 10 a similar procedure is carried out for Type I Noether symmetries.

2. TIME DERIVATIVES OF SYMMETRY SOLUTIONS

In this section we restrict the dynamical equations (1.1) to be of the form

$$E^i(\ddot{x}, \dot{x}, x)=0. \tag{2.1}$$

We shall first show that if $\xi^\alpha(x,t)$ is a symmetry solution of the dynamical system (2.1) so also will be $\xi^\alpha_{,t}$. Since (2.1) does not contain t explicitly, it follows from the discussion leading to (1.8) that (1.8) now takes the form

$$G^i_\alpha(\dot{x}, x)\xi^\alpha + G^{\beta i}_\alpha(\dot{x}, x)\xi^\alpha_{,\beta} + G^{\beta\gamma i}_\alpha(\dot{x}, x)\xi^\alpha_{,\beta\gamma}=0. \tag{2.2}$$

To find the explicit conditions that the functions ξ^α define a symmetry solution, we require that (2.2) be identically zero in the $\dot{x}^i$ variables. We assume that the resulting equations in the unknown quantities ξ^α so obtained (referred to as the symmetry equations) by this requirement are satisfied by the solution

$$\xi^\alpha = f^\alpha(x,t). \tag{2.3}$$

Hence, if we substitute for ξ^α in (2.2) by use of (2.3), the resulting equations will be identically zero in the $\dot{x}^i$, i.e.,

$$F^i \equiv G^i_\alpha f^\alpha + G^{\beta i}_\alpha f^\alpha_{,\beta} + G^{\beta\gamma i}_\alpha f^\alpha_{,\beta\gamma} \equiv 0 \quad (\text{in } \dot{x}^i). \tag{2.4}$$

From (2.4) it follows that

$$\frac{\partial F^i}{\partial t} \equiv 0 \tag{2.5}$$

in the $\dot{x}^i$. From (2.4) and (2.5) it follows that

$$G^i_\alpha(f^\alpha_{,t}) + G^{\beta i}_\alpha(f^\alpha_{,t})_{,\beta} + G^{\beta\gamma i}_\alpha(f^\alpha_{,t})_{,\beta\gamma} \equiv 0 \quad (\text{in } \dot{x}^i). \tag{2.6}$$

Hence $\xi^\alpha_{,t} = f^\alpha_{,t}$ will also define a symmetry solution.

In a similar manner it follows that in general $f^\alpha_{,tt}$, $f^\alpha_{,ttt}$, $\cdots$ will also be symmetry solutions.

This result can be stated by the following theorem.

Theorem 2.1: If a dynamical system which is characterized by (2.1) admits a symmetry mapping (1.2), (1.3) defined by $[\xi^i(x,t), \xi^0(x,t)]$, then in general it also admits symmetry mappings defined by $[\partial^K\xi^i/\partial t^K, \partial^K\xi^0/\partial t^K]$, $K=1,2,3,\cdots$.

We now assume the case where the dynamical equations (2.1) are expressible in the form of Lagrange's equations

$$\Lambda_i(L)=0, \tag{2.7}$$

based on the Lagrangian $L=L(\dot{x},x)$.

A Type I mapping defined by (1.2), (1.3) is called a Type I Noether mapping if there exist functions $[\xi^\alpha(x,t), \psi(x,t)]$ such that[5]

$$\delta L + L\frac{d}{dt}\delta t + \frac{d\psi}{dt}=0. \tag{2.8}$$

It is clear that the expanded form of the left-hand side of (2.8) will be linear and homogeneous in the quantities ξ^α, $\xi^\alpha_{,\beta}$, and $\psi_{,\alpha}$. Hence by an argument similar to that used in the proof of Theorem 2.1 we obtain the additional theorem

Theorem 2.2: If a dynamical system based on a Lagrangian $L(\dot{x},x)$ admits a Type I Noether symmetry mapping defined by $[\xi^\alpha(x,t), \psi(x,t)]$, then in general it also admits a Noether symmetry mapping defined by $[\partial^K\xi^\alpha/\partial t^K, \partial^K\psi/\partial t^K]$, $K=1,2,3,\cdots$.

We may apply Theorem 2.1 to the case where (2.1) takes the form of Hamilton's equations[6]

$$E^A(\dot{x},x) \equiv \dot{x}^A - \eta^{AB}H_{,B}=0, \quad A,B=1,\ldots,2n, \tag{2.9}$$

where

$$H(x,t)\equiv A(x)+B(t). \tag{2.10}$$

In this case (1.7) is again linear and homogeneous in $[\xi^A(x,t), \xi^0(x,t)]$, and hence if these quantities define a symmetry mapping of the Hamiltonian system so in general will $[\partial^m\xi^A/\partial t^m, \partial^m\xi^0/\partial t^m]$, $m=1,2,3,\cdots$, define a symmetry.[7]

It is well known in Hamiltonian mechanics that if $H(x,t)$ is of the form (2.10), then if $I(x,t)$ is a constant of motion then $\partial I/\partial t$ will also be a constant of motion.

A similar result is easily shown to hold for dynamical systems represented by equations of the form (2.1).[8] To prove this, assume $I(\dot{x},x,t)$ is a constant of motion of such a dynamical system; then

$$\frac{d}{dt}\left(\frac{\partial I}{\partial t}\right) = \frac{\partial^2 I}{\partial x^i \partial t}\dot{x}^i + \frac{\partial^2 I}{\partial \dot{x}^i \partial t}\ddot{x}^i + \frac{\partial^2 I}{\partial t\,\partial t}$$
$$= \frac{\partial}{\partial t}\left(\frac{\partial I}{\partial x^i}\dot{x}^i + \frac{\partial I}{\partial \dot{x}^i}\ddot{x}^i + \frac{\partial I}{\partial t}\right) = \frac{\partial}{\partial t}\left(\frac{dI}{dt}\right)=0. \tag{2.11}$$

We state the following theorem.

Theorem 2.3: If a dynamical system of the form (2.1) admits a constant of motion $I(\dot{x},x,t)$, then $\partial I(\dot{x},x,t)/\partial t$ is a constant of motion.

Corollary 2.3: If a dynamical system (2.1) admits a Type I symmetry mapping (1.2), (1.3) defined by $\xi^\alpha(x,t)$ and an associated constant of motion of the form

$$I \equiv I[\xi^\alpha, \xi^\alpha_{,\beta}, \xi^\alpha_{,\beta\gamma}; \dot{x}, x]$$
$$\equiv I_\alpha(\dot{x},x)\xi^\alpha + I^\beta_\alpha(\dot{x},x)\xi^\alpha_{,\beta} + I^{\beta\gamma}_\alpha(\dot{x},x)\xi^\alpha_{,\beta\gamma}, \tag{2.12}$$

then the constant of motion $\partial I/\partial t$ is of the form

$$\partial I/\partial t = I[(\xi^\alpha_{,t}), (\xi^\alpha_{,t})_{,\beta}, (\xi^\alpha_{,t})_{,\beta\gamma}; \dot{x}, x], \tag{2.13}$$

where by Theorem 2.1 $\xi^\alpha_{,t}$ also defines a Type I symmetry.

As an illustration of a constant of motion of the form (2.12) we mention the function C_1 defined by Ref. 1, (4.6) (with $\gamma_1 = 0$), where the dynamical system (2.1) is based on Lagrangian $L = L(\dot{x}, x)$.

We note that a statement similar to that of Corollary 2.3 holds for the Noether constant of motion C_2 defined by [see Ref. 1, (4.8)]

$$C_2 \equiv C_2[\xi^\alpha(x,t), \psi(x,t); \dot{x}, x] = \frac{\partial L}{\partial \dot{x}^i}\,\xi^i - \left(\frac{\partial L}{\partial \dot{x}^i}\dot{x}^i - L\right)\xi^0 + \psi, \tag{2.14}$$

(where $[\xi^\alpha, \psi]$ define a Type I Noether mapping) in that the constant of motion

$$C_{2,t} = C_2[(\xi^\alpha_{,t}), (\psi_{,t}); \dot{x}, x], \tag{2.15}$$

where again (2.1) is based on the Lagrangian $L(\dot{x}, x)$.

3. GROUP PROPERTIES OF TYPE I SYMMETRIES

In this section we shall first show that if ξ^α_A, ξ^α_B define any two Type I symmetry mappings of the dynamical equation (1.1), then[9]

$$\xi^\alpha_{AB} \equiv \pounds_A \xi^\alpha_B \equiv \xi^\alpha_{B,\beta}\xi^\beta_A - \xi^\alpha_{A,\beta}\xi^\beta_B \equiv \Delta_A\xi^\alpha_B - \Delta_B\xi^\alpha_A \tag{3.1}$$

will also define a symmetry mapping of (1.1), where for any function $M(\ddot{x}, \dot{x}, x, t)$

$$\Delta_A M \equiv \delta_A M/\delta a, \tag{3.2}$$

with $\delta_A M$ based upon ξ^α_A.

To prove the above stated property concerning ξ^α_{AB}, we note the commutator expression defined by

$$[\Delta_A, \Delta_B]M \equiv (\Delta_A\Delta_B - \Delta_B\Delta_A)M \equiv \Delta_A(\Delta_B M) - \Delta_B(\Delta_A M) \tag{3.3}$$

gives on expansion (by means of a lengthy but straightforward calculation)

$$[\Delta_A, \Delta_B]M = [(\ddot{\xi}^j_{AB} - \dot{x}^j\ddot{\xi}^0_{AB} - 2\ddot{x}^j\dot{\xi}^0_{AB})\partial/\partial\ddot{x}^j + (\dot{\xi}^j_{AB} - \dot{x}^j\dot{\xi}^0_{AB})\partial/\partial\dot{x}^j + \xi^\alpha_{AB}\partial/\partial x^\alpha]M. \tag{3.4}$$

From (1.2)—(1.5) Eq. (3.4) can be written in the form

$$[\Delta_A, \Delta_B]M = [(\Delta_{AB}\ddot{x}^j)\partial/\partial\ddot{x}^j + (\Delta_{AB}\dot{x}^j)\partial/\partial\dot{x}^j + (\Delta_{AB}x^\alpha)\partial/\partial x^\alpha]M. \tag{3.5}$$

It is observed that (3.5) may be rewritten in the form

$$[\Delta_A, \Delta_B]M = \Delta_{AB}M, \tag{3.6}$$

where the operator Δ_{AB} is defined as in (3.2) in terms of ξ^α_{AB}.

Since ξ^α_A, ξ^α_B define symmetry mappings, we have from (1.7) and (3.2) that

$$\Delta_A E^i = 0, \quad \Delta_B E^i = 0 \quad [\text{whenever } E^i(\ddot{x}, \dot{x}, x, t) = 0]. \tag{3.7}$$

Hence it follows that

$$\Delta_A(\Delta_B E^i) = 0, \quad \Delta_B(\Delta_A E^i) = 0 \quad (\text{whenever } E^i = 0). \tag{3.8}$$

It therefore follows from (3.3) and (3.6)

$$\Delta_{AB}E^i = 0 \quad (\text{whenever } E^i = 0). \tag{3.9}$$

We summarize the above results in the following theorem.

Theorem 3.1: If ξ^α_A, ξ^α_B define Type I symmetry mappings of the dynamical equations (1.1), then ξ^α_{AB} defined by (3.1) will also define a Type I symmetry mapping.

Assume now that the symmetry equations (refer to Sec. 1) associated with the dynamical equations (1.1) admit a complete set of solutions $\xi^\alpha_A(x,t)$, $A = 1, 2, \ldots, \rho$. By Theorem 3.1, ξ^α_{AB} defined by (3.1) is also a solution of the symmetry equations for any choice of A, B. Hence we must be able to express ξ^α_{AB} in the form

$$\xi^\alpha_{AB} = C^J_{AB}\xi^\alpha_J, \tag{3.10}$$

where C^J_{AB} are constants.

If now we define the operators X_A by

$$X_A \equiv \xi^\alpha_A\partial/\partial x^\alpha, \tag{3.11}$$

then by means of (3.1) we may express (3.10) in the form

$$[X_A, X_B] = C^J_{AB}X_J. \tag{3.12}$$

Hence we can state the following theorem

Theorem 3.2: If the symmetry equations[10] of a dynamical system (1.1) admit a complete set of solutions ξ^α_A, $A = 1, 2, \ldots, \rho$, then these solutions generate a ρ-parameter group G_ρ of symmetry mappings (in the $n+1$ variables x^i, t).

Remark 1: If we assume the conditions of Theorem 3.2 are satisfied, then it follows from (3.10) and the definition of the operator Δ_{AB} that (3.6) can be expressed in the form

$$[\Delta_A, \Delta_B]M = C^J_{AB}\Delta_J M. \tag{3.13}$$

If we interpret $\Delta_A M(\ddot{x}, \dot{x}, x, t)$ as the generators (in the $3n+1$ variables $\ddot{x}^i, \dot{x}^i, x^i, t$) of the second extension of the group G_ρ, then (3.13) is a statement of a well-known group property.[11]

Remark 2: If in (3.13) we regard $M \equiv I(\dot{x}, x, t)$, where I is a constant of motion of (1.1), then (3.13) can be written in the form

$$I_{AB} - I_{BA} = C^J_{AB}I_J, \tag{3.14}$$

where

$$I_A \equiv \Delta_A I, \quad I_{AB} \equiv \Delta_A I_B \tag{3.15}$$

are by the related integral theorem[3,4] the first and second derived constants of motion (based on I) respectively.

4. A LAGRANGIAN ILLUSTRATING COROLLARY 4.1 OF PAPER I

In Paper I[1] it was shown that if a dynamical system defined by Lagrange's equations

$$\Lambda_i(L) = 0, \quad L(\dot{x}, x, t) \tag{4.1}$$

admits a Type I symmetry, then a sufficient condition

that the dynamical system admits the constant of motion C_1 given by Ref. 1, (4.6), is that there exist functions $A_{ij}(\ddot{x}, \dot{x}, x, t)$ such that[12]

$$\frac{\partial \Lambda_j(L)}{\partial t} = \sum_{i=1}^{n} A_{ji} \Lambda_i(L). \tag{4.2}$$

In this section we obtain conditions that the class of Lagrangians defined by

$$L \equiv \sum_{i=1}^{n} \exp(\theta_i t)\, g_i(\dot{x}, x), \quad \theta_i(\dot{x}, x) \neq \theta_j(\dot{x}, x), \quad i, j = 1, 2, \ldots, n, \tag{4.3}$$

will satisfy (4.2) for some A_{ji}.

If we evaluate the left side of (4.1) using (4.3), we obtain

$$\Lambda_j(L) = \sum_{i=1}^{n} \exp(\theta_i t) F_{ji}(\ddot{x}, \dot{x}, x, t), \tag{4.4}$$

where

$$F_{ji} \equiv \sum_{k=1}^{n} \left[\left(t G_{ij} \frac{\partial \theta_i}{\partial \dot{x}^k} + \frac{\partial G_{ij}}{\partial \dot{x}^k} \right) \ddot{x}^k + \left(t G_{ij} \frac{\partial \theta_i}{\partial x^k} + \frac{\partial G_{ij}}{\partial x^k} \right) \dot{x}^k + \frac{\partial G_{ij}}{\partial t} + \theta_i G_{ij} - H_{ij} \right], \tag{4.5}$$

$$G_{ij}(\dot{x}, x, t) \equiv t g_i \frac{\partial \theta_i}{\partial \dot{x}^j} + \frac{\partial g_i}{\partial \dot{x}^j}, \tag{4.6}$$

$$H_{ij}(\dot{x}, x, t) \equiv t g_i \frac{\partial \theta_i}{\partial x^j} + \frac{\partial g_i}{\partial x^j}. \tag{4.7}$$

Substitution of (4.4) in (4.2) gives

$$\sum_{k=1}^{n} \exp(\theta_k t) \left[\theta_k F_{jk} + \frac{\partial F_{jk}}{\partial t} - A_{ji} F_{ik} \right] = 0. \tag{4.8}$$

A sufficient condition that (4.8) be satisfied is obtained by equating to zero each of the bracketed expressions in (4.8). This gives the system of linear equations (in the unknowns A_{jk}).

$$\sum_{i=1}^{n} A_{ji} F_{ik} = \frac{\partial F_{jk}}{\partial t} + \theta_k F_{jk}. \tag{4.9}$$

For a fixed j and with k assuming the values $1, \ldots, n$, we obtain n equations in the n unknowns A_{ji} $(i = 1, \ldots, n)$. This set of n equations will have a solution for these A_{ji} if

$$|F_{ik}| \neq 0 \quad \text{(along a trajectory)}. \tag{4.10}$$

It is clear that if (4.10) is satisfied, all the A_{ji} $(i, j = 1, \ldots, n)$ can be determined.

5. TIME-DEPENDENT SYMMETRIES OF CONSERVATIVE DYNAMICAL SYSTEMS

In this section we derive the explicit time-dependent symmetry equations for the conservative system defined by the Lagrangian

$$L(\dot{x}, x) \equiv \tfrac{1}{2} g_{ij}(x) \dot{x}^i \dot{x}^j - V(x), \tag{5.1}$$

where g_{ij} defines the metric of a Riemannian (configuration) space.

From (5.1) Lagrange's equations take the form

$$\Lambda_i(L) = g_{ij}(\ddot{x}^j + \Gamma^j_{ab} \dot{x}^a \dot{x}^b + g^{ja} V_{,a}) = 0, \tag{5.2}$$

where $\Gamma^j_{ab}(x)$ is the Christoffel symbol based on the g_{ij}. By inspection it is seen that (5.2) may be expressed in the equivalent form [see (1.1)]

$$E^i = \ddot{x}^i + \Gamma^i_{ab} \dot{x}^a \dot{x}^b + g^{ia} V_{,a} = 0. \tag{5.3}$$

To obtain the symmetry equations by use of (1.7), we calculate δE^i from (5.3) and, as outlined in Sec. 1, use (1.2), (1.4), (1.5) to eliminate the δx^i, $\delta \dot{x}^i$, $\delta \ddot{x}^i$ terms, respectively. In the resulting equations, (5.3) is used to eliminate the $\ddot{x}^i$ terms. This procedure leads to

$$\begin{aligned} &(\xi^0_{,j} \Gamma^j_{km} - \xi^0_{,km}) \dot{x}^k \dot{x}^m \dot{x}^i + (\xi^i_{,jk} - \xi^i_{,m} \Gamma^m_{jk} - 2\xi^0_{,tj} \delta^i_k + 2\Gamma^i_{mk} \xi^m_{,j} \\ &\quad + \Gamma^i_{jk,m} \xi^m) \dot{x}^j \dot{x}^k + (\xi^0_{,j} g^{jk} V_{,k} \delta^i_m + 2\xi^0_{,m} g^{ij} V_{,j} - \xi^0_{,tt} \delta^i_m \\ &\quad + 2\xi^i_{,tm} + 2\Gamma^i_{jm} \xi^j_{,t}) \dot{x}^m + \xi^i_{,tt} - \xi^i_{,j} g^{jk} V_{,k} + 2\xi^0_{,t} g^{ij} V_{,j} \\ &\quad + \xi^m g^{ij} V_{,jm} + g^{ij}_{,m} \xi^m V_{,j} = 0. \end{aligned} \tag{5.4}$$

The symmetry equations in ξ^i, ξ^0 are obtained by requiring that (5.4) be identically zero in the $\dot{x}^i$. As a consequence we obtain

$$\xi^i_{,tt} + 2\xi^0_{,t} g^{ij} V_{,j} + g^{ij} V_{;jk} \xi^k - g^{jk} V_{,j} \xi^i_{;k} = 0, \tag{5.5}$$

$$\xi^0_{,j} g^{jk} V_{,k} \delta^i_m + 2\xi^0_{,m} g^{ij} V_{,j} - \xi^0_{,tt} \delta^i_m + 2\xi^i_{,t;m} = 0, \tag{5.6}$$

$$\pounds_\xi \Gamma^i_{jk} - \delta^i_k \xi^0_{,tj} - \delta^i_j \xi^0_{,tk} = 0, \tag{5.7}$$

$$\xi^0_{;ij} = 0, \tag{5.8}$$

where

$$\pounds_\xi \Gamma^i_{jk} \equiv \xi^i_{,jk} + \xi^a_{,k} \Gamma^i_{aj} + \xi^a_{,j} \Gamma^i_{ak} - \xi^i_{,a} \Gamma^a_{jk} + \xi^a \Gamma^i_{jk,a}. \tag{5.9}$$

For our purposes, unless otherwise stated, we now assume in this and the remaining sections that the configuration space be Euclidean referred to rectangular coordinates x^i (which implies that $g_{ij} = g^{ij} = \delta_{ij}$ and $\Gamma^i_{jk} = 0$). As a consequence the symmetry equations (5.5)–(5.8) reduce to the form

$$\xi^i_{,tt} + 2\xi^0_{,t} V_{,i} + V_{,ik} \xi^k - V_{,k} \xi^i_{,k} = 0, \tag{5.10}$$

$$\xi^0_{,j} V_j \delta^i_m + 2\xi^0_{,m} V_{,i} - \xi^0_{,tt} \delta^i_m + 2\xi^i_{,tm} = 0, \tag{5.11}$$

$$\xi^i_{,jk} - \delta^i_k \xi^0_{,tj} - \delta^i_j \xi^0_{,tk} = 0, \tag{5.12}$$

$$\xi^0_{,ij} = 0. \tag{5.13}$$

If we define $\phi \equiv \xi^0_{,t}$, then (5.12) takes the form

$$\xi^i_{,jk} = \phi_{,j} \delta^i_k + \phi_{,k} \delta^i_j. \tag{5.14}$$

This is recognized as defining a (time-dependent) projective collineation (in a flat space) with the known solution[13]

$$\xi^i(x, t) = a_j(t) x^j x^i + B^i_j(t) x^j + C^i(t). \tag{5.15}$$

From (5.13) we obtain

$$\xi^0(x, t) = A_j(t) x^j + B(t). \tag{5.16}$$

Hence by use of (5.15), (5.16) in (5.12) there is obtained[2] $a_i = A'_i$, and (5.15) is expressible as

$$\xi^i = A'_j x^j x^i + B^i_j x^j + C^i. \tag{5.17}$$

There remains to be considered (5.10) and (5.11). [Note that (5.16), (5.17) hold for any potential $V(x)$.]

Use of (5.16), (5.17) in (5.10), (5.11) leads to the respective conditions

$$A_j'''x^jx^i + B_j^{i\prime\prime}x^j + C^{i\prime\prime} - [A_k'x^i + A_j'x^j\delta_k^i + B_k^i]V_{,k} + 2(A_j'x^j + B')V_{,i} + (A_j'x^jx^k + B_j^kx^j + C^k)V_{,ki} = 0, \quad (5.18)$$

$$A_jV_{,j}\delta_m^i + 2A_mV_{,i} - (A_j''x^j + B'')\delta_m^i + 2(A_m''x^i + A_k''x^k\delta_m^i + B_m^{i\prime}) = 0. \quad (5.19)$$

If in (5.19) we put $i = m$ and sum, the result can be solved for $A_jV_{,j}$. If this solution is used in (5.19), we find

$$A_jV_{,i} = [1/(n+2)](B'' + B_k^{k\prime})\delta_j^i - B_j^{i\prime} - A_j''x^i. \quad (5.20)$$

Thus the solution to the symmetry equations (5.10)–(5.13) has been reduced to solving (5.18), (5.20) for the quantities A_i, B, B_j^i, C^i, V. The forms of ξ^0, ξ^i will then be given by (5.16), (5.17) respectively.

6. SYMMETRY SOLUTIONS FOR THE CASE $V(x) = V_0 = \text{CONST}$

Note that in this case the dynamical equations (5.3) reduce to the equations of the geodesics.

When $V = V_0 = \text{const}$, we have from (5.18), (5.20) respectively

$$A_j'''x^jx^i + B_j^{i\prime\prime}x^j + C^{i\prime\prime} = 0, \quad (6.1)$$

$$(A_j''x^j - B'')\delta_j^i + 2A_j''x^i + 2B_j^{i\prime} = 0. \quad (6.2)$$

From (6.2) we obtain

$$A_i'' = 0. \quad (6.3)$$

When this result is used in (6.1), (6.2) we find

$$C^{i\prime\prime} = 0, \quad B''' = 0, \quad B_j^i = \tfrac{1}{2}B'\delta_j^i + \alpha_j^i \quad (\alpha_j^i = \text{const}). \quad (6.4)$$

Hence from (6.3), (6.4) we have

$$A_i = \sigma_i t + \tau_i, \quad C^i = \nu_0^i t + \nu_1^i, \quad B = \mu_0 t^2 + \mu_1 t + \mu_2, \quad B_j^i = (\mu_0 t + \tfrac{1}{2}\mu_1)\delta_j^i + \alpha_j^i, \quad (6.5)$$

where in (6.5) σ_i, τ_i, ν_0^i, ν_1^i, μ_0, μ_1, μ_2, α_j^i are arbitrary constants. The solutions for ξ^i, ξ^0 can now be obtained by use of (6.5) in (5.16), (5.17) (see Type I solution, Sec. 9).

We thus have the theorem:

Theorem 6.1: The most general time-dependent symmetry mappings of a conservative system with a constant potential [with a Euclidean configuration space referred to rectangular coordinates (x^i)] are determined by (5.16), (5.17), (6.5).

7. SOLUTIONS OF SYMMETRY EQUATIONS (5.18), (5.20) WITH AT LEAST ONE $A_i \neq 0$ [14]

In this situation there is no loss of generality in taking, for example, $A_1 \neq 0$.

From (5.20) we find

$$A_jV_{,ik} = 0 \quad (i \neq k). \quad (7.1)$$

Use of $j = 1$ in (7.1) gives $V_{,ik} = 0$ $(i \neq k)$, and hence $V(x)$ must be of the separable form

$$V(x) = \sum_{i=1}^{n} V_i(x^i) \quad (V_i \text{ is a function of } x^i \text{ only}). \quad (7.2)$$

Again, from (5.20) with $j = 1$, we find by differentiation with respect to x^i

$$V_{,ii} = V_i''(x^i) = A_1''(t)/A_1(t) \equiv 2a \quad (i = 1, \ldots, n;\ a = \text{const}), \quad (7.3)$$

so that we may write

$$V = ar^2 + \beta_ix^i + \gamma \quad (\beta_i,\ \gamma = \text{const}), \quad (7.4)$$

where

$$r^2 \equiv \sum_{i=1}^{n} (x^i)^2. \quad (7.5)$$

Equation (7.4) gives the most general form of potential when at least one of the $A_i \neq 0$ in (5.18), (5.20). This implies (when at least one $A_i \neq 0$) that the only spherically symmetric potential, that is, $V = V(r)$, is given by

$$V = ar^2 + b \quad (a,\ b = \text{const}). \quad (7.6)$$

We continue with the solutions of (5.18), (5.20) assuming the potential V has the form (7.6). If then (7.6) is substituted in (5.18), (5.20), the following conditions are obtained:

$$A_i'' + 2aA_i = 0, \quad (7.7)$$

$$B''' + 8aB' = 0, \quad (7.8)$$

$$C^{i\prime\prime} + 2aC^i = 0, \quad (7.9)$$

$$B_j^i = \tfrac{1}{2}\delta_j^iB' + \alpha_j^i, \quad \alpha_j^i = \text{const}. \quad (7.10)$$

[Note that the conditions (6.3), (6.4) (for $V = V_0$) correspond to the case $a = 0$.]

The above results are summarized in the theorem stated below.

Theorem 7.1: If a conservative dynamical system defined by the Lagrangian (5.1) [in a Euclidean configuration space referred to rectangular coordinates (x^i)] admits a time-dependent symmetry defined by functions ξ^i, ξ^0, these functions must have the forms given by (5.16), (5.17) respectively. If there is in (5.16) at least one $A_i \neq 0$, then the potential $V(x)$ must have the form (7.4). In this case the most general spherically symmetric potential is given by (7.6), and for this potential the conditions on the coefficients in (5.16), (5.17) are given by (7.7)–(7.10).

Since in a later paper we plan to discuss symmetry mappings of conservative systems with separable potentials of the general form (7.2), we will postpone further discussion of the potential (7.4) to this later paper.

8. SOLUTIONS OF SYMMETRY EQUATIONS (5.18), (5.20) WITH ALL $A_i = 0$, AND WITH SPHERICALLY SYMMETRIC POTENTIALS

When all $A_i = 0$ in (5.16), (5.17), we obtain

$$\xi^0 = B(t), \quad \xi^i = B_j^i(t)x^j + C^i(t). \quad (8.1)$$

Equations (5.18), (5.20) now reduce to the respective forms

$$B_j^{i\prime\prime}x^j + C^{i\prime\prime} - WB_j^ix^j + 2WB'x^i + (B_j^kx^j + C^k) \times [(W'/r)x^kx^i + W\delta_k^i] = 0, \quad (8.2)$$

$$(B'' + B_j^{j\prime})\delta_k^i = (n+2)B_k^{i\prime}, \quad (8.3)$$

where

$$W \equiv V'/r. \tag{8.4}$$

From (8.3) it follows that

$$B^i_j = \tfrac{1}{2}\delta^i_j B' + \alpha^i_j, \quad \alpha^i_j = \text{consts}. \tag{8.5}$$

From (8.5) and (8.2) there is obtained

$$C^{i\prime\prime} + WC^i + x^i[(W'/r)(C^j x^j) + \tfrac{1}{2}B''' + B'(2W + \tfrac{1}{2}rW') + (W'/r)(\alpha^j_k x^j x^k)] = 0. \tag{8.6}$$

As (8.3) is satisfied identically by means of (8.5), Eq. (8.6) is the only remaining equation to be considered.

If (8.6) be differentiated with respect to x^j $(j \neq i)$, the result implies that[14]

$$W'(x^j C^i - x^i C^j) = 0. \tag{8.7}$$

The choice $W' = 0$ requires V to be of the form (7.6), and it can be shown that (7.7)–(7.10) will still hold (with all $A_i = 0$). If $W' \neq 0$, then (8.7) implies that

$$C^i = 0. \tag{8.8}$$

With (8.8) used in (8.6) we have

$$B''' + (4W + rW')B' + (2W'/r)(\alpha^j_k x^j x^k) = 0. \tag{8.9}$$

Equation (8.9) implies

$$\alpha^i_j + \alpha^j_i = 2\alpha\delta_{ij} \quad (\alpha = \text{const}), \tag{8.10}$$

where

$$\alpha \equiv \alpha^1_1 = \alpha^2_2 = \cdots = \alpha^n_n. \tag{8.11}$$

From (8.4), (8.10), (8.11) used in (8.9) there results

$$B''' + (V'' + 3V'/r)B' + 2\alpha(V'' - V'/r) = 0, \tag{8.12}$$

which is the only remaining condition on $V(r)$ and $B(t)$.

Differentiation of (8.12) with respect to r and t leads to

$$B''(V'' + 3V'/r)' = 0. \tag{8.13}$$

Now if $(V'' + 3V'/r)' = 0$, (8.12) and (8.4) will give the condition $\alpha W' = 0$. Since we are assuming $W' \neq 0$, this implies $\alpha = 0$. The condition $(V'' + 3V'/r)' = 0$, (with $W' \neq 0$) implies that

$$V = k_0 r^2 + k_1/r^2 + k_2, \quad k_1 \neq 0 \quad (k_0, k_1, k_2 = \text{consts}). \tag{8.14}$$

Since now $\alpha = 0$, (8.12) reduces to

$$B''' + 8k_0 B' = 0, \tag{8.15}$$

from which B can be found. Then ξ^i, ξ^0 can be found by means of (8.1), (8.5), (8.8), and (8.10) with $\alpha = 0$. (See Type III of Sec. 9.)

The other possibility of (8.13), $B'' = 0$, gives

$$B = b_0 t + b_1 \quad (b_0, b_1 = \text{consts}), \tag{8.16}$$

and hence (8.12) reduces to

$$(b_0 + 2\alpha)V'' + (3b_0 - 2\alpha)V'/r = 0. \tag{8.17}$$

Solutions of (8.17) are of three types,

$$V = c_0 \ln r + c_1, \quad (c_0 \neq 0, \ b_0 = 2\alpha), \tag{8.18}$$

$$V = c_0 r^p + c_1,$$

$$[c_0 \neq 0, \ p(b_0 + 2\alpha) + 2(b_0 - 2\alpha) = 0, \ p(p^2 - 4) \neq 0], \tag{8.19}$$

$$V(r) = \text{arbitrary} \quad (\alpha = b_0 = 0). \tag{8.20}$$

[The restriction $p(p^2 - 4) \neq 0$ in (8.19) is to exclude duplication of previous solutions.]

The corresponding solutions for ξ^i, ξ^0 may be found in the general summary given in Sec. 9.

9. SUMMARY OF COMPLETE GROUPS OF TIME-DEPENDENT SYMMETRIES ADMITTED BY CONSERVATIVE SYSTEMS WITH SPHERICALLY SYMMETRIC POTENTIALS

Conservative systems $(n > 1)$ are classified into six types according to the form of their associated spherically symmetry potentials. For each type the corresponding (ξ^0, ξ^i) symmetry solution is given along with the complete group generated by the solution.

It is found convenient to represent the symbols of these groups in terms of the following operators:

$$P_\alpha \equiv p_\alpha, \tag{9.1}$$

$$Q_{\alpha\beta} \equiv x^\alpha p_\beta, \tag{9.2}$$

$$S_{ij} \equiv x^i p_j - x^j p_i, \tag{9.3}$$

$$U \equiv x^i p_i, \quad U^* \equiv x^\alpha p_\alpha, \tag{9.4}$$

where

$$p_i \equiv \partial/\partial x^i, \quad p_0 \equiv \partial/\partial t, \quad x^0 \equiv t. \tag{9.5}$$

Note that

$$U = \sum_{i=1}^{n} Q_{ii}, \quad S_{ij} = Q_{ij} - Q_{ji}. \tag{9.6}$$

A group of ρ parameters is represented by a set of symbols indicated by expressions of the form $[X_1, X_2, \ldots, X_\rho]$. It will be noted that each complete symmetry group will contain the rotation, time-translation (sub)group $[S_{ij}, P_0]$. Each of the groups listed below is a complete group.

Type I

$$V = V_0,$$

$$\xi^0 = (\sigma_i t + \tau_i)x^i + \mu_0 t^2 + \mu_1 t + \mu_2,$$

$$\xi^i = (\sigma_j x^j)x^i + (\mu_0 t + \tfrac{1}{2}\mu_1)x^i + \alpha^i_j x^j + \nu^i_0 t + \nu^i_1.$$

Complete group $= [Q_{\alpha\beta}, Q_\alpha, P_\alpha]$, $\rho = (n+1)(n+3)$, where $Q_\alpha \equiv x^\alpha U^*$. This is the general projective group in $n+1$ variables. (The basis of the group given above is obtained by a suitable change of the original basis.)

Type II(a)

$$V = ar^2 + b \quad (a > 0),$$

$$\xi^0 = (a_j \cos\lambda t + b_j \sin\lambda t)x^j + (1/2\lambda)(c_1 \sin 2\lambda t - c_2 \cos 2\lambda t) + c_3,$$

$$\xi^i = (-\lambda a_j \sin\lambda t + \lambda b_j \cos\lambda t)x^j x^i + d_i \cos\lambda t + e_i \sin\lambda t + \tfrac{1}{2}(c_1 \cos 2\lambda t + c_2 \sin 2\lambda t)x^i + \alpha^i_j x^j, \quad (\lambda \equiv \sqrt{2a}).$$

Complete group $= [Q_{ij}, D_i, E_i, F_i, G_i, L_1, L_2, P_0]$,

$$\rho = (n+1)(n+3),$$

$$D_i \equiv x^i[(\cos\lambda t)P_0 - \lambda(\sin\lambda t)U],$$

$$E_i \equiv x^i[(\sin\lambda t)P_0 + \lambda(\cos\lambda t)U],$$
$$F_i \equiv (\cos\lambda t)P_i,\ G_i \equiv (\sin\lambda t)P_i,$$
$$L_1 \equiv \tfrac{1}{2}[(\cos 2\lambda t)U + (1/\lambda)(\sin 2\lambda t)P_0],$$
$$L_2 \equiv \tfrac{1}{2}[(\sin 2\lambda t)U - (1/\lambda)(\cos 2\lambda t)P_0].$$

Type II(b)

$$V = ar^2 + b \quad (a < 0),$$
$$\xi^0 = [a_j \exp(\lambda t) + b_j \exp(-\lambda t)]x^j + (1/2\lambda)(c_1 \exp(2\lambda t) - c_2 \exp(-2\lambda t)) + c_3,$$
$$\xi^i = \lambda[a_j \exp(\lambda t) - b_j \exp(-\lambda t)]x^j x^i + d_i \exp(\lambda t) + e_i \exp(-\lambda t) + \tfrac{1}{2}[c_1 \exp(2\lambda t) + c_2 \exp(-2\lambda t)]x^i + \alpha^i_j x^j \quad (\lambda = \sqrt{-2a}).$$

Complete group $= [Q_{ij}, H_i, I_i, J_i, K_i, M_1, M_2, P_0]$,
$\rho = (n+1)(n+3)$,
$H_i \equiv \exp(\lambda t)x^i(P_0 + \lambda U)$, $I_i \equiv \exp(-\lambda t)x^i(P_0 - \lambda U)$,
$J_i \equiv \exp(\lambda t)P_i$, $K_i \equiv \exp(-\lambda t)P_i$,
$M_1 \equiv (1/2\lambda)\exp(2\lambda t)(P_0 + \lambda U)$,
$M_2 = -(1/2\lambda)\exp(-2\lambda t)(P_0 - \lambda U)$.

Type III(a)

$$V = k_0 r^2 + k_1/r^2 + k_2 \quad (k_1 \neq 0,\ k_0 > 0),$$
$$\xi^0 = (1/\mu)(a_1 \sin\mu t - a_2 \cos\mu t) + a_3 \quad (\mu = \sqrt{8k_0}),$$
$$\xi^i = \tfrac{1}{2}(a_1 \cos\mu t + a_2 \sin\mu t)x^i + \alpha^i_j x^j \quad (\alpha^i_j + \alpha^j_i = 0).$$

Complete group $= [S_{ij}, N_1, N_2, P_0]$, $\rho = [n(n-1)/2] + 3$,
$N_1 \equiv \tfrac{1}{2}(\cos\mu t)U + (1/\mu)(\sin\mu t)P_0$,
$N_2 \equiv \tfrac{1}{2}(\sin\mu t)U - (1/\mu)(\cos\mu t)P_0$.

Type III(b)

$$V = k_0 r^2 + k_1/r^2 + k_2 \quad (k_1 \neq 0,\ k_0 < 0),$$
$$\xi^0 = (1/\mu)[a_1 \exp(\mu t) + a_2 \exp(-\mu t)] + a_3 \quad (\mu = \sqrt{-8k_0}),$$
$$\xi^i = \tfrac{1}{2}[a_1 \exp(\mu t) - a_2 \exp(-\mu t)]x^i + \alpha^i_j x^j \quad (\alpha^i_j + \alpha^j_i = 0).$$

Complete group $= [S_{ij}, R_1, R_2, P_0]$, $\rho = [n(n-1)/2] + 3$,
$R_1 = \tfrac{1}{2}\exp(\mu t)U + (1/\mu)\exp(\mu t)P_0$,
$R_2 = -\tfrac{1}{2}\exp(-\mu t)U + (1/\mu)\exp(-\mu t)P_0$.

Type III(c)

$$V = k_1/r^2 + k_2 \quad (k_1 \neq 0),$$
$$\xi^0 = \mu_0 t^2 + \mu_1 t + \mu_2,$$
$$\xi^i = (\mu_0 t + \tfrac{1}{2}\mu_1)x^i + \alpha^i_j x^j \quad (\alpha^i_j + \alpha^j_i = 0).$$

Complete group $= [S_{ij}, t(tP_0 + U), tP_0 + \tfrac{1}{2}U, P_0]$,
$\rho = [n(n-1)/2] + 3$.

Type IV

$$V = a_0 \ln r + a_1 \quad (a_0 \neq 0),$$
$$\xi^0 = b_0 t + b_1,$$
$$\xi^i = b_0 x^i + \sum_{j \neq i} \alpha^i_j x^j, \quad \alpha^i_j + \alpha^j_i = b_0 \delta_{ij}.$$

Complete group $= [S_{ij}, P_0, U^*]$, $\rho = n(n-1)/2 + 2$.

Type V

$$V = c_0 r^p + c_1 \quad [c_0 \neq 0,\ p(p^2-4) \neq 0],$$
$$\xi^0 = -2\alpha[(p-2)/(p+2)]t + b_1,$$
$$\xi^i = -\alpha[(p-2)/(p+2)]x^i + \alpha^i_j x^j, \quad \alpha^i_j + \alpha^j_i = 2\alpha\delta_{ij}.$$

Complete group $= [S_{ij}, P_0, Y]$, $\rho = [n(n-1)/2] + 2$,
where
$$Y = 2U + (2-p)tP_0.$$

Type VI

$V(r) =$ arbitrary (but not one of the previous types),
$$\xi^0 = a, \quad \xi^i = \alpha^i_j x^j, \quad \alpha^i_j + \alpha^j_i = 0.$$

Complete group $= [S_{ij}, P_0]$, $\rho = [n(n-1)/2] + 1$. As noted above this group is a subgroup of all the previous groups.

10. TYPE I NOETHER SYMMETRIES OF CONSERVATIVE SYSTEMS WITH SPHERICALLY SYMMETRIC POTENTIALS

In Sec. 2 a Type I (time-dependent) Noether symmetry was defined by means of (2.8). In the present section we first obtain the explicit form of the Type I Noether symmetry equations for the conservative system defined by the Lagrangian (5.1).

By use of (1.2), (1.3), (1.4), (1.6), and (5.1) it is found that (2.8) expands to the form[5]

$$-\tfrac{1}{2}\xi^0_{,j} g_{ik}\dot{x}^i\dot{x}^j\dot{x}^k + \tfrac{1}{2}(\pounds g_{jk} - \xi^0_{,t} g_{jk})\dot{x}^j\dot{x}^k + (g_{ij}\xi^i_{,t} - V\xi^0_{,j} + \psi_{,j})\dot{x}^j - (V_{,i}\xi^i + V\xi^0_{,t} - \psi_{,t}) = 0, \quad (10.1)$$

where the Lie derivative

$$\pounds g_{jk} \equiv g_{jk,i}\xi^i + g_{ik}\xi^i_{,j} + g_{ij}\xi^i_{,k}. \quad (10.2)$$

Equation (10.1) must hold identically in the $\dot{x}^i$, and hence this leads to the Noether symmetry equations[15]

$$\xi^0_{,j} g_{ik} + \xi^0_{,k} g_{ji} + \xi^0_{,i} g_{kj} = 0, \quad (10.3)$$
$$\pounds g_{jk} - \xi^0_{,t} g_{jk} = 0, \quad (10.4)$$
$$g_{ji}\xi^i_{,t} - V\xi^0_{,j} + \psi_{,j} = 0, \quad (10.5)$$
$$V_{,i}\xi^i + V\xi^0_{,t} - \psi_{,t} = 0. \quad (10.6)$$

From (10.3) it follows that

$$\xi^0 = \xi^0(t) \equiv B(t). \quad (10.7)$$

By use of (10.7) Eqs. (10.4), (10.5), (10.6) reduce respectively to

$$\pounds g_{ij} = B'(t) g_{ij}, \quad (10.8)$$
$$g_{ij}\xi^i_{,t} + \psi_{,j} = 0, \quad (10.9)$$
$$V_{,i}\xi^i + VB' = \psi_{,t}. \quad (10.10)$$

It is to be noted that (10.8) defines a time-dependent homothetic symmetry.

We now restrict the space to be Euclidean referred to rectangular coordinates (x^i) and assume a spherically symmetry potential $V(r)$. In this situation (10.8) has the known solution[16]

$$\xi^i = \tfrac{1}{2}B'x^i + b^i_j(t)x^j + C^i(t) \quad (b^i_j + b^j_i = 0), \tag{10.11}$$

and (10.9) becomes

$$\xi^i_{,t} = -\psi_{,i}. \tag{10.12}$$

From (10.11), (10.12) there results

$$\tfrac{1}{2}B''x^i + b^{i\prime}_j x^j + C^{i\prime} + \psi_{,i} = 0, \tag{10.13}$$

and (10.13) implies

$$\psi_{,ij} - \psi_{,ji} = b^{i\prime}_j - b^{j\prime}_i = 0, \tag{10.14}$$

which by the skew symmetry of b^i_j [see (10.11)] means

$$b^i_j = \text{const}. \tag{10.15}$$

The solution for ψ is now found from (10.13), (10.15) to be of the form

$$\psi = -\tfrac{1}{4}r^2B'' - C^{i\prime}x^i + D(t). \tag{10.16}$$

There remains (10.10) to be considered. From (10.10), (10.11), (10.16) we obtain

$$(V'/r)(x^iC^i + \tfrac{1}{2}r^2B') + VB' = D' - \tfrac{1}{4}r^2B''' - x^iC^{i\prime\prime}, \tag{10.17}$$

from which we determine $B(t)$, $C^i(t)$, $D(t)$, and $V(r)$. We omit the details since the calculations are similar to those of Secs. 5–8. We give the summary of results below.

It was shown in Ref. 1, Sec. 3, that every Type I Noether symmetry will be a (general) Type I symmetry as defined in Sec. 1. Hence for a given dynamical system this implies that any complete group of Type I Noether symmetries will be a (sub)group of the complete group of general Type I symmetries associated with the system.

The solution to (10.17) result in three cases which we refer to as Types N-1, N-2, N-3 respectively.

Type N-1 is identical to Type III of Sec. 9. In addition, the value of ψ is given by

$$\psi = -\tfrac{1}{4}r^2B'' + k_2B + k_3, \quad k_3 = \text{const}, \tag{10.18}$$

where k_2 is a constant appearing in the potential of Type III and where the function $B(t)$ satisfies (8.15).

For Type N-2, $V(r) = ar^2 + b$, $a, b = \text{const}$,

$$\xi^0 = B(t), \quad \xi^i = \tfrac{1}{2}B'x^i + \alpha^i_j x^j + C^i(t), \tag{10.19}$$
$$\alpha^i_j = \text{consts}, \quad \alpha^i_j + \alpha^j_i = 0,$$

where B satisfies (7.8), C^i satisfies (7.9), and

$$\psi = -\tfrac{1}{4}r^2B'' - x^iC^{i\prime} + bB + c, \quad c = \text{const}. \tag{10.20}$$

There are three subcases corresponding to the value of the constant a:

$$(a > 0) \quad \text{complete group} = [S_{ij}, F_i, G_i, L_1, L_2, P_0], \tag{10.21}$$

$$(a < 0) \quad \text{complete group} = [S_{ij}, J_i, K_i, M_1, M_2, P_0], \tag{10.22}$$

$$(a = 0) \quad \text{complete group} = [S_{ij}, tU^*, tP_0 + \tfrac{1}{2}U, tP_i, P_i, P_0]. \tag{10.23}$$

The group (10.21) is a subgroup of Type II(a); the group (10.22) is a subgroup of Type II(b); the group (10.23) is a subgroup of Type I.

For Type N-3, $V(r)$ is an arbitrary function other than the potentials of Types N-1 and N-2. For Type N-3 the functions ξ^i, ξ^0 and its complete group are identical to those given in Type VI. In addition $\psi = \text{const}$.

[1] G.H. Katzin and J. Levine, J. Math. Phys. 17, 1345 (1976).
[2] Unless otherwise indicated, lower case Latin indices will have the range $1, \ldots, n$, Greek indices will have the range $0, 1, \ldots, n$, and the Einstein summation notation is used. A dot $(\cdot)$ indicates total time derivative d/dt. A comma $(,)$ indicates partial differentiation. A prime $(')$ indicates differentiation with respect to the indicated argument. A semicolon $(;)$ indicates covariant differentiation.
[3] For a discussion of related integral theorems see Theorem 4.2 of Ref. 1 and also G.H. Katzin and J. Levine, J. Math. Phys. (a) **15**, 1460 (1974); (b) **16**, 548 (1975).
[4] For a discussion of linear dependence between derived integrals see Ref. 3(b) and also G.H. Katzin, J. Math. Phys. 14, 1213 (1973).
[5] See Sec. 3 of Ref. 1. Note that ψ of (2.8) is related to $\delta\Omega$ of Ref. 1, (3.11) by $\psi\delta a = \delta\Omega$.
[6] See Ref. 3(b) for notation.
[7] This result is a generalization of Corollary 6.1 of Ref. 3(b).
[8] This may be a known result, but a proof is given for completeness.
[9] The symbol $\pounds_A$ denotes the Lie derivative with respect to the vector ξ^α_A.
[10] See Sec. 1 for the procedure which leads to these equations.
[11] Refer to L.P. Eisenhart, *Continuous Groups of Transformations* (Princeton U.P., Princeton, N.J., 1933), Chap. 2.
[12] In this section the Einstein summation convention is not used.
[13] For the familiar time independent case see L.P. Eisenhart, *Non-Riemannian Geometry* (American Mathematical Society, New York, 1927), Vol. III, p. 127.
[14] The case $n = 1$ is an exception and will be considered in a later paper.
[15] It can be shown that (5.10)–(5.13) are satisfied as a consequence of (10.3)–(10.6). This is in agreement with Theorem 3.2 of Ref. 1.
[16] J. Levine and G.H. Katzin, J. Math. Phys. 14, 1886 (1973).

An extended Levinson's theorem*

T. A. Osborn† and D. Bollé‡

Stanford Linear Accelerator Center, Stanford University, Stanford, California 94305
(Received 18 August 1976; revised manuscript received 15 October 1976)

We investigate the form Levinson's theorem takes when the two-body scattering amplitude is not decomposed into partial waves. It is found that the theorem changes its structure in this case and is not merely the sum over angular momentum of the well-known partial wave results. The energy dependent quantity that replaces the partial wave phase shift turns out to be the trace of the two-body time delay operator. This extended version of the theorem remains valid for scattering by nonspherically symmetric potentials.

I. INTRODUCTION

We study an extension of Levinson's theorem for two-particle scattering. This extension states the theorem as a moment property of the trace of the two-body time delay operator. In the form obtained here the theorem is valid for the entire—nonpartial wave decomposed—amplitude. The resultant form of the theorem found here is *not* what one would surmise on the basis of simply summing the well-known partial wave statements in terms of phase shifts. Our derivation will be rigorously carried out for the class of local potentials that belong to $L^1 \cap L^2$.

To begin with, we list the known features of time delay in two-particle potential scattering which we must employ in this analysis. This outline is too brief to be a balanced introduction to the theory of time delay concepts in scattering. Such a general discussion is found in Ref. 1, which also gives a survey of the recent literature on this topic.

The scattering system studied here is characterized by an interacting Hamiltonian h and an asymptotic Hamiltonian h_0. In these two Hamiltonians the center-of-mass motion has been removed. If $\mathbf{x}$ is the vector separation of the two particles, then h and h_0 act on a Hilbert space $\mathcal{H}$ composed of square integrable functions of $\mathbf{x}$. On $\mathcal{H}$ one defines the Møller wave operator by the strong limit,

$$\Omega^{(\pm)} = \text{s-}\lim_{t \to \mp\infty} \exp(iht)\exp(-ih_0 t), \tag{I. 1}$$

where t is the real parameter denoting time. Each f in $\mathcal{H}$ may correspond to a possible incident wavepacket. The symbol $\phi(t)$ will always represent the time dependent noninteracting wavepacket associated with f. Likewise $\psi(t)$ will be the function that is the fully interacting wavepacket evolving in time according to h. These two functions are given by

$$\phi(t) = \exp(-ih_0 t)f, \tag{I. 2}$$

$$\psi(t) = \exp(-iht)\Omega^{(\pm)}f. \tag{I. 3}$$

The time delay of a scattering process is defined by the following construction. Let us describe a family of projection operators that is specified by the equations

$$\begin{aligned} P(R)g(\mathbf{x}) &= g(\mathbf{x}), \quad x \leq R, \\ &= 0, \qquad x > R, \end{aligned} \tag{I. 4}$$

where g is any function belonging to $\mathcal{H}$. Thus $P(R)$ projects any function onto a sphere of radius R measured from the collision center at $\mathbf{x}=0$. Given an incident wavepacket f and a specific value of R the time delay is determined by the expression,

$$T(R,f) = \int_{-\infty}^{\infty} dt[(\psi(t), P(R)\psi(t)) - (\phi(t), P(R)\phi(t))]. \tag{I. 5}$$

The inner product is that of $\mathcal{H}$. The second member of the integrand gives the probability that at time t the wavepacket $\phi(t)$ is inside the sphere $P(R)$. The integral over t of this real quantity is just the total time $\phi(t)$ spends inside the sphere $P(R)$. The same interpretation applies to the first inner product involving $\psi(t)$. Consequently, $T(R,f)$ is the difference of time the two waves reside in the sphere.

Consider now the description of the scattering problem in momentum space. The relative two-particle momentum will be the vector $\mathbf{p}$. The corresponding kinetic energy of relative motion will be $E = \mathbf{p}^2/2\mu$, where μ is the reduced mass of the two particles. The symbol $\hat{p}$ will denote the unit vector direction of $\mathbf{p}$. We introduce a Hilbert space $\mathcal{H}_-$ of L^2 functions of $\hat{p}$—namely that space determined by the inner product

$$(g,g')_- = \int g(\hat{p})^* g'(\hat{p})\, d\hat{p}. \tag{I. 6}$$

The theory of time delay allows one to construct a family of operators $q(E,R)$ acting on $\mathcal{H}_-$. This family has the property[2,3]

$$T(R,f) = \int_0^{\infty} dE\, \mu p \int\int d\hat{p}\, d\hat{p}'\, f^*(p\hat{p})\langle \hat{p}|q(E,R)|\hat{p}'\rangle f(p\hat{p}'), \tag{I. 7}$$

where $p = \sqrt{2\mu E}$. In expression (I. 7) $\langle \hat{p}|q(E,R)|\hat{p}'\rangle$ is the kernel representation of the operator $q(E,R)$. Furthermore, for well-behaved potentials, the $R \to \infty$ limit of $T(R,f)$ exists and is associated with an operator $q(E)$, viz.,

$$\lim_{R\to\infty} T(R,f) = \int_0^{\infty} dE\, \mu p \int\int d\hat{p}\, d\hat{p}'\, f^*(p\hat{p})\langle \hat{p}|q(E)|\hat{p}'\rangle f(p\hat{p}'). \tag{I. 8}$$

The operator $q(E)$ is known[2] to be simply determined by the S matrix. The full S matrix that acts on $\mathcal{H}$ is defined by the product of wave operators:

$$S = \Omega^{(-)\dagger}\Omega^{(+)}. \tag{I. 9}$$

If one takes the momentum space matrix elements of Eq. (I. 9), one is led to a natural definition of a reduced, energy dependent S matrix, $s(E)$, that acts on $\mathcal{H}_-$. The operator $s(E)$ is specified by its kernel $\langle \hat{p}|s(E)|\hat{p}'\rangle$, which is determined from S by the expression

$$\langle \mathbf{p}|S|\mathbf{p}'\rangle = [\delta(E-E')/\mu p]\langle \hat{p}|s(E)|\hat{p}'\rangle, \qquad \text{(I. 10)}$$

for $E = \mathbf{p}^2/2\mu$. The energy dependent delta function, of course, indicates the physical conservation of energy in the scattering process. In terms of $s(E)$ the operator $q(E)$ may be expressed[2,4] as,

$$q(E) = -is^{\dagger}(E)\frac{d}{dE}s(E). \qquad \text{(I. 11)}$$

It is interesting to note here that the structure of Eq. (I. 11) is such that the unitarity of $s(E)$ implies that $q(E)$ must be Hermitian.

The feature of time delay that is vital for our analysis is known as the spectral property. Let $r(z) = (h-z)^{-1}$ and $r_0(z) = (h_0 - z)^{-1}$ be the resolvents of h and h_0 defined for complex energy z. Then the spectral property is the relation

$$2\,\mathrm{Im}\,\mathrm{Tr}[r(E+i0) - r_0(E+i0)] = \mathrm{tr}[q(E)]. \qquad \text{(I. 12)}$$

In this equation Tr is the trace on $\mathcal{H}$ and tr is the trace on $\mathcal{H}_a$. This relation has a simple physical interpretation. The right-hand side is just the trace of the time delay operator $q(E)$ and is proportional to the total time delay experienced by an incident plane wave of energy E. The left-hand side is the change of state density produced by the interaction v. In fact we shall require a somewhat more general version of Eq. (I. 12), specifically

$$\mathrm{Im}\,\mathrm{Tr}[r(E+i\eta) - r_0(E+i\eta)] = \sum_{i=1}^{N} \frac{\eta}{|E_i + E + i\eta|^2} + \frac{1}{2\pi}\int_0^{\infty} dE' \frac{\eta}{(E-E')^2+\eta^2}\,\mathrm{tr}[q(E')]. \qquad \text{(I. 13)}$$

The E_i appearing here are the negative of the eigenvalues of h. This equation is given explicitly in Ref. 5. The spectral property is readily obtained from Eq. (I. 13) by letting the imaginary parameter η go to zero. The advantage inherent in this version of the spectral property is that it allows one to estimate how rapidly $\mathrm{Im}\,\mathrm{Tr}[r(E+i\eta) - r_0(E+i\eta)]$ approaches its $\eta = 0$ value.

Throughout this study we will consistently assume that the potential belongs to $L^1 \cap L^2$. This means that $v(\mathbf{x})$ is such that

$$\int d\mathbf{x}|v(\mathbf{x})| = B_1 < \infty, \quad \left(\int d\mathbf{x}|v(\mathbf{x})|^2\right)^{1/2} = B_2 < \infty. \qquad \text{(A)}$$

This class of potentials is broad enough to include most cases of physical interest. However both hard core potentials and Coulomb potentials are excluded by (A). We note that the L^1 restriction of (A) dictates that the power behavior of v for $|\mathbf{x}|$ very large must be like $\sim |\mathbf{x}|^{-3-\delta}$, where δ is an arbitrarily small positive number. The L^2 requirement of (A) implies that most severe local singularities can be $\sim |\mathbf{x}|^{-3/2+\delta}$.

The time delay formalism has been rigorously studied under assumption (A). In particular Kato[6] has proved that the wave operators $\Omega^{(\pm)}$ in Eq. (I. 1) exist when (A) holds. Jauch, Sinha, and Misra[3] prove the existence of the limit given in Eq. (I. 8). Equation (I. 13) is found in Ref. 5. This equation, which is central to our discussion, may also be easily inferred from the results of Jauch, Sinha, and Misra.[7] Another rigorous analysis of the time delay formalism above has been recently given by Martin[4] for slightly different assumptions on the potential.

One may question whether or not it is necessary to treat this problem in a rigorous fashion. For example, is not Levinson's theorem valid so long as the potential falls off more rapidly than the Coulomb force? Two observations indicate why a careful and detailed analysis is necessary. First, as indicated at the beginning of this section, a simple sum of known partial wave results does not lead to the correct answer for the entire scattering problem. The form of the answer is sensitive to the order of integration and limiting processes; thus each change of order must be justified. A second observation emphasizes the need to specify precisely the behavior of the potential. Suppose one considers the following central potential:

$$v_1(r) = (\lambda_1/r^2)\int_0^r dr'\, g(r')^2 + \lambda_2/r^2, \quad \lambda_1 > 0, \qquad \text{(I. 14)}$$

where λ_1 and λ_2 are real parameters and $g(r)$ is an arbitrary real function. In this case one can prove[8] that the momentum derivative of all partial wave phase shifts is positive for all k, so that

$$\int_0^{\infty} dk \frac{d}{dk}\delta_l(k) > 0, \quad \text{all } l. \qquad \text{(I. 15)}$$

By way of contrast the Levinson's theorem for partial wave phase shifts states

$$\int_0^{\infty} dk \frac{d}{dk}\delta_l(k) = \delta_l(\infty) - \delta_l(0) = -\pi N_l, \qquad \text{(I. 16)}$$

where N_l is the number of two-body bound states with angular momentum l. Thus potentials of the form v_1 violate the theorem for every partial wave. Sufficient conditions for the existence of the partial wave form of Levinson's theorem are that the moments

$$M_i = \int_0^{\infty} dr\, r^i |v(r)|, \quad i = 1, 2, \qquad \text{(I. 17)}$$

be finite.[9] By this criteria we see that potentials like v_1 fall off too slowly in r to lead to a reasonable phase shift behavior. Also in the extended case, the potential v_1 would be excluded by condition (A).

The proof we shall give of our extended Levinson's theorem is based on two elements. One is the spectral property Eq. (I. 13). The second is the analytic behavior of $\mathrm{Tr}[r(z) - r_0(z)]$ in the complex z plane. Section II of this paper gives a rigorous proof of the various aspects of the analytic behavior we need. Section III combines this analytic behavior with Eq. (I. 13) to complete the proof. In Sec. IV we give a general discussion of these results and also describe a second approach to the problem that is based on the asymptotic completeness of the wave operators. A quick, albeit nonrigorous, understanding of our results may be obtained by just reading Sec. IV.

II. ANALYTIC PROPERTIES OF Tr $[r(z) - r_0(z)]$

This section is devoted to the study of $\mathrm{Tr}[r(z) - r_0(z)]$. We always assume condition (A) is obeyed by the potential v. One very useful consequence of (A) is that it implies that our potential v is in the Rollnik class that Simon[10] has extensively studied, viz.,

$$\iint d\mathbf{x}\,d\mathbf{y}\,\frac{|v(\mathbf{x})|\,|v(\mathbf{y})|}{|\mathbf{x}-\mathbf{y}|^2}=B_v<\infty. \tag{II. 1}$$

Our analysis will make extensive use of the well-known operators $r_0(z)$, $Vr_0(z)$, and $A(z)$. These operators all act on $\mathcal{H}$ and depend parametrically on z. They are conveniently defined by their kernel representations in coordinate space:

$$\langle \mathbf{x}|r_0(z)|\mathbf{y}\rangle=\frac{\mu}{2\pi}\,\frac{\exp(ik|\mathbf{x}-\mathbf{y}|)}{|\mathbf{x}-\mathbf{y}|}, \tag{II. 2}$$

$$\langle \mathbf{x}|Vr_0(z)|\mathbf{y}\rangle=\frac{\mu}{2\pi}v(\mathbf{x})\frac{\exp(ik|\mathbf{x}-\mathbf{y}|)}{|\mathbf{x}-\mathbf{y}|}, \tag{II. 3}$$

$$\langle \mathbf{x}|A(z)|\mathbf{y}\rangle=\frac{\mu}{2\pi}v^{1/2}(\mathbf{x})\frac{\exp(ik|\mathbf{x}-\mathbf{y}|)}{|\mathbf{x}-\mathbf{y}|}\,|v(\mathbf{y})|^{1/2}, \tag{II. 4}$$

where $k=\sqrt{2\mu z}$, $v^{1/2}(\mathbf{x})=|v(\mathbf{x})|^{1/2}\operatorname{sgn}[v(\mathbf{x})]$, and $A(z)=V^{1/2}r_0(z)|V|^{1/2}$. The set of points in the z plane a distance δ or greater from positive real axis we will denote by Π_δ. The symbol Π_0 will denote the cut z-plane obtained by letting $\delta\to 0$. For $z\in\Pi_\delta$ or Π_0 then k clearly belongs to the upper-half complex k plane.

We shall use three different norms to describe operators on $\mathcal{H}$. First, the usual operator norm will be represented by the symbol $\|\cdot\|$. Second, we define the Schmidt norm of an operator A on $\mathcal{H}$ by

$$\|A\|_2=\left(\int\int d\mathbf{x}\,d\mathbf{y}\,|A(\mathbf{x},\mathbf{y})|^2\right)^{1/2}, \tag{II. 5}$$

where $A(\mathbf{x},\mathbf{y})$ is the kernel generated by A. The class of all operators on $\mathcal{H}$ with finite Schmidt norm is called the Schmidt class. The class is denoted by β_2. Our third norm is the trace norm defined by

$$\|A\|_1=\sum_i^\infty(\phi_i,|A|\phi_i), \tag{II. 6}$$

where $A=(A^\dagger A)^{1/2}$. When A has finite $\|A\|_1$, it belongs to the trace class of operators on $\mathcal{H}$. The trace class is labeled β_1. When $A\in\beta_1$, then the operator has a well-defined trace given by the sum

$$\mathrm{Tr}A=\sum_i^\infty(\phi_i,A\phi_i). \tag{II. 7}$$

Of course, this sum is independent of the basis set $\{\phi_i\}$. Our analysis will frequently use the following general properties of the trace and the Schmidt operator.

(i) $A\in\beta_1$ if and only if it can be written as the product $A=BC$, where $B\in\beta_2$ and $C\in\beta_2$. Furthermore, $\|A\|_1\le\|B\|_2\|C\|_2$.

(ii) If $B\in\beta_2$, then $B^\dagger\in\beta_2$.

(iii) If $B\in\beta_2$ and $C\in\beta_2$, then $\mathrm{Tr}BC=\mathrm{Tr}CB$.

(iv) If $B\in\beta_1$ and A has finite norm $\|A\|$, then $BA\in\beta_1$ and $AB\in\beta_1$ and $\|BA\|_1\le\|A\|\|B\|_1$.

(v) If $A\in\beta_1$, then $|\mathrm{Tr}A|\le\sum_i^\infty|(\phi_i,A\phi_i)|\le\|A\|_1$.

(vi) If B and C are Schmidt class, then $\mathrm{Tr}BC$ has the representation

$$\mathrm{Tr}BC=\int\int d\mathbf{x}\,d\mathbf{y}\,B(\mathbf{y},\mathbf{x})C(\mathbf{x},\mathbf{y}),$$

where $B(\mathbf{y},\mathbf{x})$ and $C(\mathbf{x},\mathbf{y})$ are the L^2 kernels generated by the operators B and C.

Shatten[11] gives proofs of all these statements.

For later convenience let us collect here some well-known estimates for norms of the operators occurring in this problem. We shall show that the operators $V^{1/2}r_0(z)$, $|V|^{1/2}r_0(z)$, and $r_0(z)V$ are Schmidt class for all $z\in\Pi_\delta$. Consider the first operator in the list above. If we employ the integral form of the Schmidt norm to compute $\|V^{1/2}r_0(z)\|_2$, we have

$$\|V^{1/2}r_0(z)\|_2=\frac{\mu}{2\pi}\left(\iint d\mathbf{x}\,d\mathbf{y}\,\frac{|v(\mathbf{x})|\exp(-2\,\mathrm{Im}k|\mathbf{x}-\mathbf{y}|)}{|\mathbf{x}-\mathbf{y}|^2}\right)^{1/2}$$
$$=\mu\left(\frac{B_1}{2\pi\,\mathrm{Im}k}\right)^{1/2}. \tag{II. 8}$$

The same expression holds for the norm $\|\,|V|^{1/2}r_0(z)\|_2$. Similar considerations show that

$$\|Vr_0(z)\|_2=\mu B_2(1/2\pi\,\mathrm{Im}k)^{1/2}. \tag{II. 9}$$

Now let us examine the operator $A(z)$ given in Eq. (II. 4). This operator is Schmidt class in the entire z plane Π_0

$$\|A(z)\|_2$$
$$=\frac{\mu}{2\pi}\left(\iint d\mathbf{x}\,d\mathbf{y}\,\frac{|v(\mathbf{x})|\,|(v(\mathbf{y})|}{|\mathbf{x}-\mathbf{y}|^2}\exp(-2\,\mathrm{Im}k|\mathbf{x}-\mathbf{y}|)\right)^{1/2}$$
$$\le\frac{\mu}{2\pi}B_v^{1/2}, \tag{II. 10}$$

where B_v is the constant entering the Rollnik bound on v. Another useful bound pertains to $A(z)^2$. One may show, using the Riemann–Lebesque lemma, that

$$\lim_{|\mathrm{Re}k|\to\infty}\|A(z)^2\|_2=0. \tag{II. 11}$$

The convergence is uniform in $\mathrm{Im}k\ge 0$. Equation (II. 11) means that there exists a finite k_v such that for all $|\mathrm{Re}k|>k_v$ then $\|A(z)^2\|_2<\frac{1}{2}$. The number k_v depends only on v. We will not write out the proof of Eq. (II. 11) and the estimate for $\|A(z)^2\|_2$. Theorem I. 23 of Simon's book is very nearly result (II. 11). The difference is that Simon requires k to be real. It is a simple modification of Simon's proof to extend it to complex k in the upper half plane and to show that the convergence is uniform in $\mathrm{Im}k$.

Lemma 1: Let the potential v satisfy (A). For all positive integers n, the operators $r_0(z)[Vr_0(z)]^n$ are trace class for $z\in\Pi_\delta$. The function $\mathrm{Tr}\{r_0(z)[Vr_0(z)]^n\}$ is an analytic function of z in the Π_δ domain. Furthermore, the order of the trace operation and d/dz may be freely interchanged in Π_δ.

Proof: We first establish $r_0[Vr_0]^n$ is trace class. Consider $n=1$. This operator may be written as the product of $r_0|V|^{1/2}$ and $V^{1/2}r_0$, each of which in Schmidt class in Π_δ. Thus employing property (i) of the trace gives

$$\|r_0Vr_0\|_1\le\|r_0|V|^{1/2}\|_2^2. \tag{II. 12}$$

Estimate (II. 8) tells us the right-hand side is finite. For $n>1$ we may write

$$\|r_0[Vr_0]^n\|_1\le\|r_0Vr_0\|_1\|Vr_0\|_2^{n-1}. \tag{II. 13}$$

In obtaining (II. 13) we have used trace property (iv) together with the general inequality $\|A\|\le\|A\|_2$. Estimates (II. 8) and (II. 9) then imply that the right-hand side of (II. 13) is finite.

Next consider the analyticity of $\mathrm{Tr}\{r_0(z)[Vr_0(z)]^n\}$. Set $n=1$. As noted above the operator is the product of two Schmidt operators $r_0|V|^{1/2}$ and $V^{1/2}r_0$ in the domain Π_δ. Invoking property (vi) of the trace for the operator r_0Vr_0 gives us

$$\mathrm{Tr}(r_0Vr_0)=\left(\frac{\mu}{2\pi}\right)^2\iint d\mathbf{x}\,d\mathbf{y}\,\frac{v(\mathbf{x})}{|\mathbf{y}-\mathbf{x}|^2}\exp(2ik|\mathbf{y}-\mathbf{x}|). \tag{II. 14}$$

For values of k restricted by the condition $\mathrm{Im}k\geq\delta'\geq 0$ the expression

$$|v(\mathbf{x})|\exp(-2\delta'|\mathbf{y}-\mathbf{x}|)/|\mathbf{y}-\mathbf{x}|^2$$

bounds the integrand above uniformly in k. This bound is absolutely integrable with respect to $(\mathbf{x},\mathbf{y})$. Thus the integral in Eq. (II. 14) defines an analytic function of k and thus z. This argument may be extended to show $\mathrm{Tr}\{r_0[Vr_0]^n\}$ is analytic for all n and $z\in\Pi_\delta$.

Finally let us examine the differential properties of $\mathrm{Tr}[Vr_0(z)]^n$. The trace diverges for $n=1$, but is well defined for $n\geq 2$. Consider the case $n=2$. If we examine the integral representation of $\mathrm{Tr}[Vr_0(z)]^2$, the Rollnik condition (II. 1) guarantees that the integral is uniformly convergent in Π_δ. Thus we can differentiate under the integral to obtain

$$\frac{d}{dz}\mathrm{Tr}[Vr_0(z)]^2=\left(\frac{\mu}{2\pi}\right)^2\iint d\mathbf{x}\,d\mathbf{y}\,\frac{2i\mu v(\mathbf{x})v(\mathbf{y})}{k|\mathbf{y}-\mathbf{x}|}\exp(2ik|\mathbf{y}-\mathbf{x}|). \tag{II. 15}$$

One then observes that the Hilbert identity for $r_0(z)$,

$$r_0(z_1)-r_0(z_2)=(z_1-z_2)r_0(z_1)r_0(z_2), \tag{II. 16}$$

implies the operator relation

$$\frac{dr_0(z)}{dz}=r_0(z)^2. \tag{II. 17}$$

The kernel form of this last identity is

$$\frac{i\mu}{k}\exp(ik|\mathbf{x}-\mathbf{y}|)=\frac{\mu}{2\pi}\int d\mathbf{s}\,\frac{\exp[ik(|\mathbf{x}-\mathbf{s}|+|\mathbf{s}-\mathbf{y}|)]}{|\mathbf{x}-\mathbf{s}|\,|\mathbf{s}-\mathbf{y}|}. \tag{II. 18}$$

Inserting Eq. (II. 18) into the right-hand side of (II. 15) gives

$$\begin{aligned}\frac{d}{dz}\mathrm{Tr}[Vr_0(z)]^2&=2\,\mathrm{Tr}\{r_0(z)[Vr_0(z)]^2\}\\&=\mathrm{Tr}\left(\frac{d}{dz}[Vr_0(z)]^2\right).\end{aligned} \tag{II. 19}$$

These arguments extend to the $n>2$ cases. There Eq. (II. 19) becomes

$$\frac{d}{dz}\mathrm{Tr}[Vr_0(z)]^n=n\,\mathrm{Tr}\{r_0(z)[Vr_0(z)]^n\}. \tag{II. 20}$$

This completes the demonstration of Lemma 1.

Lemma 2: There exist finite k_r and k_i such that for all $|z|>(k_r^2+k_i^2)/2\mu$ the Born series expansion

$$\mathrm{Tr}[r(z)-r_0(z)]=\sum_{n=1}^{\infty}(-1)^n\,\mathrm{Tr}\{r_0(z)[Vr_0(z)]^n\} \tag{II. 21}$$

is valid. The series is uniformly convergent in z.

Proof: As usual, set $z=k^2/2\mu$ and choose z so that it has values $\mathrm{Im}k>k_i=\mu B_2/2\pi$. For z so restricted, the operator Born expansion

$$r(z)-r_0(z)=\sum_{n=1}^{\infty}(-1)^n r_0(z)[Vr_0(z)]^n \tag{II. 22}$$

is valid. It is easy to see that the series (II. 22) is convergent in operator norm for $\mathrm{Im}k>k_i$. A given term in this series has norm

$$\begin{aligned}\|r_0(z)[Vr_0(z)]^n\|&\leq\|r_0(z)\|\,\|Vr_0(z)\|^n\\&\leq(2\mu/k_i^2)\|Vr_0(z)\|_2^n.\end{aligned} \tag{II. 23}$$

For the restricted values of k, $\|Vr_0(z)\|_2$ is less than one. The sum of terms (II. 23) with respect to n then converges absolutely.

Since we know $r(z)-r_0(z)$ is trace class, we can take the trace of Eq. (II. 22) to obtain

$$\mathrm{Tr}[r(z)-r_0(z)]=\mathrm{Tr}\left\{\sum_{n=1}^{\infty}(-1)^n r_0(z)[Vr_0(z)]^n\right\} \tag{II. 24}$$

for $\mathrm{Im}k>k_i$. The series on the right of Eq. (II. 24) suggests we consider the related series

$$\sum_{n=1}^{\infty}(-1)^n\,\mathrm{Tr}\{r_0(z)[Vr_0(z)]^n\}. \tag{II. 25}$$

Introducing the definition of the trace into this expression gives the double sum

$$\sum_n^{\infty}\sum_i^{\infty}(-1)^n(\phi_i,r_0(z)[Vr_0(z)]^n\phi_i). \tag{II. 26}$$

It is easy to demonstrate that this double series is absolutely convergent. Employing the general trace identity (v), we have

$$\begin{aligned}\sum_i^{\infty}|(\phi_i,r_0(z)[Vr_0(z)]^n\phi_i)|&\leq\|r_0(z)[Vr_0(z)]^n\|_1\\&\leq\|r_0(z)\|\,\|Vr_0(z)\|_2^n.\end{aligned} \tag{II. 27}$$

As in Eq. (II. 23), when $\mathrm{Im}k>k_i$ the sum over n of the terms on the right of Eq. (II. 27) converge uniformly in k. This shows that the double series in expression (II. 26) is absolutely convergent. Thus the order of summation may be changed. And so, Eq. (II. 24) may be written in the form given by Eq. (II. 21).

Let us consider the validity of Eq. (II. 21) in a different region of z. Suppose $|\mathrm{Re}k|>k_r$. The trace norm appearing in Eq. (II. 27) may be estimated by

$$\begin{aligned}&\|r_0(z)[Vr_0(z)]^n\|_1\\&\quad=\|r_0(z)|V|^{1/2}A(z)^{n-1}V^{1/2}r_0(z)\|_1\\&\quad\leq\|r_0(z)|V|^{1/2}\|_2^2\|A(z)^2\|_2^{(n-1)/2},\qquad n=\text{odd},\\&\quad\leq\|r_0(z)|V|^{1/2}\|_2^2\|A(z)\|_2\|A(z)^2\|_2^{(n-2)/2},\quad n=\text{even}.\end{aligned} \tag{II. 28}$$

Bounds given in Eqs. (II. 9), (II. 11), and (II. 12) show that the sum over n of $\|r_0(z)[Vr_0(z)]^n\|_1$ converge. Again, the double sum in Eq. (II. 26) is absolutely convergent, and formula (II. 21) is valid for all $|\mathrm{Re}k|>k_r$. In fact the domain specified by $|z|\geq(k_i^2+k_2^2)/2\mu$ lies in the union of $|\mathrm{Re}k|>k_r$ and $|\mathrm{Im}k|>k_i$. So Lemma 2 is proved.

Lemma 3: For all integers $n\geq 2$, $\mathrm{Tr}[Vr_0(z)]^n$ satisfies

$$\lim_{|\mathrm{Re}k|\to\infty} \mathrm{Tr}[Vr_0(z)]^n = 0 \tag{II. 29}$$

for all $\mathrm{Im}k \geq 0$.

Proof: We note first that for $n \geq 2$

$$\mathrm{Tr}[Vr_0(z)]^n = \mathrm{Tr}\,|V|^{1/2}[V^{1/2}r_0(z)|V|^{1/2}]^{n-1}V^{1/2}r_0(z). \tag{II. 30}$$

For $z \in \Pi_\delta$ the operators $A(z)$ and $V^{1/2}r_0(z)$ are Schmidt class and $|V|^{1/2}$ is bounded; thus we may use trace property (iii) to obtain

$$\mathrm{Tr}[Vr_0(z)]^n = \mathrm{Tr}[A(z)^n]. \tag{II. 31}$$

For $n \geq 4$ we have the bounds

$$\begin{aligned} |\mathrm{Tr}[Vr_0(z)]^n| &\leq \|A^2(z)\|_2^{n/2}, && n = \text{even}, \\ &\leq \|A(z)\|_2 \|A^2(z)\|_2^{(n-1)/2}, && n = \text{odd}. \end{aligned} \tag{II. 32}$$

Applying result (II. 11) gives the statement (II. 29) in the lemma.

There remains only the cases $n = 2, 3$ to prove. Consider $n = 2$. Using the integral representation of the trace, we have

$$\mathrm{Tr}[Vr_0(z)]^2 = \left(\frac{\mu}{2\pi}\right)^2 \iint d\mathbf{x}\, d\mathbf{y} \frac{v(\mathbf{x})v(\mathbf{y}) \exp(2ik|\mathbf{x}-\mathbf{y}|)}{|\mathbf{x}-\mathbf{y}|^2}. \tag{II. 33}$$

Because of the condition (II. 1) the nonoscillatory part of the integrand

$$v(\mathbf{x})v(\mathbf{y})|\mathbf{x}-\mathbf{y}|^{-1} \exp(-2\,\mathrm{Im}k|\mathbf{x}-\mathbf{y}|)$$

is L^2 over $(\mathbf{x}, \mathbf{y})$. Thus we can apply the Riemann–Lebesgue lemma to conclude that $\mathrm{Tr}[Vr_0(z)]^2$ vanishes as $|\mathrm{Re}k| \to \infty$. A similar argument works for $\mathrm{Tr}[Vr_0(z)]^3$.

Lemma 4: For all $z \in \Pi_0$ the value of $\mathrm{Tr}[r_0(z)Vr_0(z)]$ is given by

$$\mathrm{Tr}[r_0(z)Vr(z)] = \frac{i}{\pi}\left(\frac{\mu}{2}\right)^{3/2} \frac{1}{\sqrt{z}} \int d\mathbf{x}\, v(\mathbf{x}). \tag{II. 34}$$

Proof: For $z \in \Pi_\delta$, $r_0(z)Vr_0(z)$ is the product of two Schmidt operators, so by trace property (vi),

$$\mathrm{Tr}[r_0(z)Vr_0(z)] = \left(\frac{\mu}{2\pi}\right)^2 \iint d\mathbf{y}\, d\mathbf{x} \frac{v(\mathbf{x}) \exp(2ik|\mathbf{x}-\mathbf{y}|)}{|\mathbf{x}-\mathbf{y}|^2}. \tag{II. 35}$$

For all $\mathrm{Im}k > 0$ the double integral

$$\iint d\mathbf{y}\, d\mathbf{x} \frac{|v(\mathbf{x})| \exp(-2\,\mathrm{Im}k|\mathbf{x}-\mathbf{y}|)}{|\mathbf{x}-\mathbf{y}|^2}$$

exists since $v \in L^1$. Thus, employing the Fubini theorem on interchange of integration, we can write (II. 35) in the form of an iterated integral

$$\mathrm{Tr}[r_0(z)Vr_0(z)] = \left(\frac{\mu}{2\pi}\right)^2 \int d\eta \frac{\exp(2ik|\eta|)}{|\eta|^2} \int d\mathbf{x}\, v(\mathbf{x}), \tag{II. 36}$$

where we have set $\eta = \mathbf{x} - \mathbf{y}$. The integral is trivial and gives Eq. (II. 34). So far the equality (II. 34) is established for $z \in \Pi_\delta$. However, the right-hand side of (II. 34) has Π_0 as its natural domain of analyticity. Thus (II. 34) represents the analytic extension of $\mathrm{Tr}[r_0(z)Vr_0(z)]$ to the domain Π_0.

III. LEVINSON'S THEOREM

In this section we combine the analytic properties of $\mathrm{Tr}[r(z) - r_0(z)]$ established in the previous section with the known features of time delay outlined in the Introduction to complete our derivation of Levinson's theorem. Our proof will require one additional technical assumption about time delay. We assume the existence of the following integral

$$\int_0^\infty dE \left| \mathrm{tr}[q(E)] + \frac{2}{\pi}\left(\frac{\mu}{2}\right)^{1/2} \frac{\tilde{v}}{\sqrt{E}} \right|, \tag{B}$$

where $\tilde{v}$ is defined by

$$\tilde{v} = \int d\mathbf{x}\, v(\mathbf{x}), \quad |\tilde{v}| < B_1. \tag{III. 1}$$

Ideally assumption (B) should be verified directly from the potential property (A). But it would take us far afield to establish (B) in this manner. There are strong physical arguments for believing (B), that will be discussed in the next section.

Consider the function $Q(z)$ defined by

$$Q(z) = \mathrm{Tr}[r(z) - r_0(z) + r_0(z)Vr_0(z)]. \tag{III. 2}$$

We have established that this function is analytic in Π_δ. Bound states of the Hamiltonian H appear as simple poles of the resolvent $r(z)$, with residues that are projection operators onto the bound state eigenfunction space. Physically interesting potentials will always have negative bound state energies. So our formalism will always imply this situation.

Suppose z_0 is some point in Π_δ and C_0 some small circular contour about z_0. Then the Cauchy–Coursat theorem tells one that the integral

$$\oint_{C_0} dz\, Q(z) = 0. \tag{III. 3}$$

Our version of Levinson's theorem is based on this identity. We open up the contour C_0 as indicated in Fig. 1. Now C_0 may be replaced by the contour segments C_Γ, C_δ, and C_i. The contours C_i, which are P in number, encircle the P distinct eigenvalues of the Hamiltonian h. Path C_δ is symmetric about the real axis, always a distance δ away from the positive real axis and ending where the real value of z is equal to Γ. The curve C_Γ is a circle centered about the z plane origin, having a radius equal to $(\Gamma^2 + \delta^2)^{1/2}$.

Because of the behavior of the exact resolvent $r(z)$ in the neighborhood of the eigenfunctions of h, one has

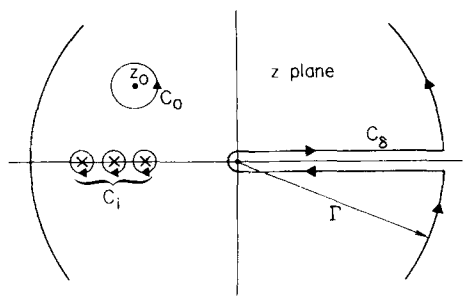


FIG. 1. The complex energy plane.

$$\sum_{i=1}^{P} \oint_{C_i} dz\, Q(z) = 2\pi i N, \tag{III. 4}$$

where N is the total number of bound states of h counting degeneracy. Thus the integral (III. 3) becomes

$$\int_{C_\Gamma + C_\delta} dz\, Q(z) = -2\pi i N. \tag{III. 5}$$

We will now evaluate the double limit

$$\lim_{\Gamma\to\infty}\lim_{\delta\to 0} \int_{C_\Gamma + C_\delta} dz\, Q(z) = -2\pi i N. \tag{III. 6}$$

Consider the C_δ integral first. It may be expanded as

$$\int_{C_\delta} dz\, Q(z) = 2i \int_0^\Gamma d\lambda\, \mathrm{Im}[Q(\lambda + i\delta)] + i\delta \times \int_{3\pi/2}^{\pi/2} d\theta \exp(i\theta) Q[\delta \exp(i\theta)]. \tag{III. 7}$$

Because of the presumption that the Hamiltonian h has no zero energy eigenstates the second integral gives zero in the $\delta \to 0$ limit. Let us study the first integral on the right-hand side of Eq. (III. 7). We shall prove

Lemma 5: For potentials such that (A) and (B) are valid then

$$\lim_{\Gamma\to\infty}\lim_{\delta\to 0} 2\int_0^\Gamma d\lambda\, \mathrm{Im}Q(\lambda + i\delta) = \int_0^\infty dE \left[\mathrm{tr}[q(E)] + \frac{2}{\pi}\left(\frac{\mu}{2}\right)^{3/2} \frac{\tilde{v}}{\sqrt{E}}\right]. \tag{III. 8}$$

Proof: Define $D_1(\Gamma, \delta)$ and $D_2(\Gamma, \delta)$ by the expressions

$$D_1(\Gamma, \delta) = \frac{1}{\pi}\int_0^\Gamma d\lambda \int_0^\infty dE \frac{\delta}{(\lambda - E)^2 + \delta^2} \times \left[\mathrm{tr}[q(E)] + \frac{2}{\pi}\left(\frac{\mu}{2}\right)^{3/2} \frac{\tilde{v}}{\sqrt{E}}\right], \tag{III. 9}$$

$$D_2(\Gamma, \delta) = c_0 \int_0^\Gamma d\lambda \int_0^\infty dE \frac{1}{\pi\sqrt{E}} \frac{\delta}{(\lambda - E)^2 + \delta^2} - c_0 \times \int_0^\Gamma d\lambda\, \mathrm{Re}\frac{1}{\sqrt{\lambda + i\delta}}, \tag{III. 10}$$

where $c_0 = (2/\pi)(\mu/2)^{3/2}\tilde{v}$. Equations (I. 13), (II. 34) together with (III. 2) give

$$2\int_0^\Gamma d\lambda\, \mathrm{Im}Q(\lambda + i\delta) = D_1(\Gamma, \delta) + D_2(\Gamma, \delta) + \int_0^\Gamma d\lambda \left\{\sum_{i=1}^{N} \frac{2\delta}{|E_i + \lambda + i\delta|^2}\right\}. \tag{III. 11}$$

Since E_i are the magnitudes of the negative energy eigenvalues and thus positive, it is obvious that the last integral vanishes when the double limit is taken. So we need only consider the contribution from D_1 and D_2. Consider D_1. Set

$$g(E) = \mathrm{tr}[q(E)] + \frac{2}{\pi}\left(\frac{\mu}{2}\right)^{3/2} \frac{\tilde{v}}{\sqrt{E}} \tag{III. 12}$$

then D_1 is

$$D_1(\Gamma, \delta) = \int_0^\Gamma d\lambda \int_0^\infty dE \frac{\delta g(E)}{\pi[(\lambda - E)^2 + \delta^2]}. \tag{III. 13}$$

Since (B) states that $g(E)$ is L^1 and that $[(\lambda - E)^2 + \delta^2]^{-1}$ is L^1 in λ for all $\delta > 0$, the Fubini theorem allows us to change the order of integrations,

$$D_1(\Gamma, \delta) = \int_0^\infty dE \int_0^\Gamma d\lambda \frac{\delta g(E)}{\pi[(\lambda - E)^2 + \delta^2]}. \tag{III. 14}$$

The integral over $d\lambda$ is elementary and gives us

$$D_1(\Gamma, \delta) = \int_0^\infty dE \frac{g(E)}{\pi}\left(\tan^{-1}\frac{\Gamma - E}{\delta} + \tan^{-1}\frac{E}{\delta}\right). \tag{III. 15}$$

Now the E-dependent integrand is L^1 for all $\delta \geq 0$ and uniformly bounded by $|g(E)|$. The Lebesgue dominated convergence theorem permits us to pass the $\delta \to 0$ limit through the integral to obtain,

$$D_1(\Gamma, 0) = \int_0^\infty dE \frac{g(E)}{\pi} \lim_{\delta\to 0}\left(\tan^{-1}\frac{\Gamma - E}{\delta} + \tan^{-1}\frac{E}{\delta}\right). \tag{III. 16}$$

Using

$$\lim_{\delta\to 0} \tan^{-1}(E/\delta) = \pi/2, \quad \text{all } E > 0,$$
$$\lim_{\delta\to 0} \tan^{-1}[(\Gamma - E)/\delta] = \begin{cases} \pi/2 & \text{all } E < \Gamma, \\ -\pi/2, & \text{all } E > \Gamma, \end{cases} \tag{III. 17}$$

we see that $\tan^{-1}$ functions give us a step function that becomes zero when $E > \Gamma$. Thus

$$D_1(\Gamma, 0) = \int_0^\Gamma dE\, g(E) \tag{III. 18}$$

and

$$\lim_{\Gamma\to\infty}\lim_{\delta\to 0} D_1(\Gamma, \delta) = \int_0^\infty dE\, g(E). \tag{III. 19}$$

A parallel analysis leads one to conclude

$$\lim_{\Gamma\to\infty}\lim_{\delta\to 0} D_2(\Gamma, \delta) = 0. \tag{III. 20}$$

Thus Lemma 5 is proved.

The one remaining integral in relation (III. 5) that we have not yet studied is the C_Γ term. For this integral we have the result:

Lemma 6: For potentials satisfying (A), then

$$\lim_{\Gamma\to\infty} \int_{C_\Gamma} dz\, Q(z) = 0. \tag{III. 21}$$

Proof: Choose $\Gamma > (k_r^2 + k_i^2)/2\mu$. Lemma 2 states that the Born series expansion of $Q(z)$ is uniformly convergent for $z \in C_\Gamma$. Using Eq. (II. 21) to expand $Q(z)$, we can write our integral as

$$\int_{C_\Gamma} dz\, Q(z) = \int_{C_\Gamma} dz \sum_{n=2}^{\infty} (-1)^n \mathrm{Tr}\{r_0(z)[Vr_0(z)]^n\} = \sum_{n=2}^{\infty} (-1)^n \int_{C_\Gamma} dz\, \mathrm{Tr}\{r_0(z)[Vr_0(z)]^n\}. \tag{III. 22}$$

Equation (II. 20) allows us to transform the integrand into an exact differential,

$$\int_{C_\Gamma} dz\, Q(z) = \sum_{n=2}^{\infty} \frac{(-1)^n}{n} \int_{C_\Gamma} dz \frac{d}{dz} \mathrm{Tr}[Vr_0(z)]^n = \sum_{n=2}^{\infty} \frac{(-1)^n}{n}\{\mathrm{Tr}[Vr_0(\Gamma + i0)]^n - \mathrm{Tr}[Vr_0(\Gamma - i0)]^n\}. \tag{III. 23}$$

Estimates (II. 32) imply that this series is uniformly convergent in Γ. Thus the $\Gamma \to \infty$ limit may be passed

through the sum. Lemma 3 shows us each term in Eq. (III. 8) vanishes in the $\Gamma \to \infty$ limit. Thus Eq. (III. 21) is demonstrated.

Combining the conclusions for both Lemmas 5 and 6 together with Eq. (III. 6) gives us:

Theorem: For potentials satisfying condition (A) and the trace of the time delay satisfying (B), then, the following relation holds:

$$\int_0^\infty dE\left[\mathrm{tr}[q(E)] + \frac{2}{\pi}\left(\frac{\mu}{2}\right)^{3/2}\frac{\tilde{v}}{\sqrt{E}}\right] = -2\pi N. \qquad \text{(III. 24)}$$

The positive integer N is the total number of negative energy eigenfunctions of the Hamiltonian h. It is presumed that h has no zero energy resonances or bound states and is free of positive energy bound states.

Under differing circumstances the theorem obtained here has been discussed previously by Buslaev.[12] In Buslaev's version the function $\mathrm{tr}[q(E)]$ is replaced by its logarithmic S-matrix derivative form given in Eq. (I. 11). In the form found by Buslaev there is no recognition that the theorem is a moment property that constrains the energy integral of the time delay. By way of contrast, our basic starting point, the spectral property relation (I. 13), is derived in Ref. 5 using just the definition of the time delay operator. No recognition or use of the S matrix is needed in obtaining Eq. (I. 13).

From a technical perspective the theorem given above is established here for a larger class of potentials than Buslaev's analysis permits. Buslaev requires that the potential $V(\mathbf{r})$ be infinitely many times differentiable. Furthermore, both the potential and all of its derivatives are required to decrease to zero in the limit $|\mathbf{r}| \to \infty$ more rapidly than any power of $|\mathbf{r}|^{-1}$.

IV. ASYMPTOTIC COMPLETENESS AND LEVINSON'S THEOREM

In this section we are concerned with two aspects of our Levinson's theorem. First, we establish how one may derive the result starting from the completeness of the scattering states. Secondly, we give a physical interpretation and explanation of our result. Our aim in this section is to provide some insight into the result obtained above, rather than to supply additional rigorous proofs. Thus, we will use nonrigorous arguments which we believe convincing, even though these arguments tend to lose sight of the exact conditions on the potential for which the analysis is valid.

The derivation given above of Levinson's theorem is based on the spectral property of time delay combined with the analytic features of $\mathrm{Tr}[r(z) - r_0(z)]$. However, in the literature there exists another method of derivation. Jauch[13] established that the usual partial wave form of Levinson's theorem can be obtained from asymptotic completeness and certain properties of the wave operator, $\Omega^{(+)}$.

We adopt Jauch's argument to the case at hand—namely the full amplitude. The mathematical statement of asymptotic completeness is

$$\Omega^{(+)\dagger}\Omega^{(+)} = I, \qquad \text{(IV. 1)}$$

$$\Omega^{(+)}\Omega^{(+)\dagger} = I - P. \qquad \text{(IV. 2)}$$

Here I is the identity operator in $\mathcal{H}$ and P is the projection operator onto the subspace spanned by all eigenfunctions of h. We note that $\mathrm{Tr}P = N$.

The wave operator possesses a well-known representation in terms of the t matrix. Suppose $t(z)$ is the operator satisfying the Lippmann–Schwinger equation

$$t(z) = V - Vr_0(z)t(z). \qquad \text{(IV. 3)}$$

The wave operator may be expanded[13] about the identity,

$$\Omega^{(+)} = I - K, \qquad \text{(IV. 4)}$$

where K is determined by the generalized function

$$\langle \mathbf{p}|K|\mathbf{p}'\rangle = \frac{\langle \mathbf{p}|t(p'^2/2\mu + i0)|\mathbf{p}'\rangle}{\mathbf{p}^2/2\mu - \mathbf{p}'^2/2\mu - i0}. \qquad \text{(IV. 5)}$$

If one forms the commutator $[K^\dagger, K]$, then Eqs. (IV. 1) and (IV. 2) imply,

$$[K^\dagger, K] = P. \qquad \text{(IV. 6)}$$

Levinson's theorem is obtained by taking the trace of Eq. (IV. 6).

On aspect of this approach requires care. The kernel representation of K is a generalized function. As a consequence the trace needs to be computed through a limiting process. It is convenient to introduce a two parameter family of operators on $\mathcal{H}_-$, $\tau(E, E')$, defined by

$$\langle \hat{p}|\tau(E, E')|\hat{p}'\rangle = j(E)\langle E\hat{p}|t(E + i0)|E'\hat{p}'\rangle j(E'), \qquad \text{(IV. 7)}$$

where $|E\hat{p}\rangle$ stands for the element $|p\hat{p}\rangle$ and $p = \sqrt{2\mu E}$. The factor $j(E) = (2\mu^3 E)^{1/4}$. One can easily express the reduced S matrix, $s(E)$, in terms of $\tau(E, E)$, viz.,

$$s(E) = e - 2\pi i\tau(E, E). \qquad \text{(IV. 8)}$$

Here e is the identity on $\mathcal{H}_-$.

Construct now the trace of $[K^\dagger, K]$. Combining Eqs. (IV. 7) and (IV. 5) and the fact that $d\mathbf{p} = j^2(E)\,dE\,d\hat{p}$, we have

$$j(E)\langle E\hat{p}|[K^\dagger, K]|E'\hat{p}\rangle j(E') = -\int dE_1 d\hat{p}_1 \frac{\langle \hat{p}_1|\tau(E_1, E)|\hat{p}\rangle^*\langle \hat{p}_1|\tau(E_1, E')|\hat{p}\rangle - \langle \hat{p}|\tau(E, E_1)|\hat{p}_1\rangle\langle \hat{p}|\tau(E', E_1)|\hat{p}_1\rangle^*}{(E - E_1 - i0)(E_1 - E' - i0)}. \qquad \text{(IV. 9)}$$

Carry out the $d\hat{p}$ integration of both sides of Eq. (IV. 9) and use the adjoint relation

$$\langle \hat{p}|\tau^\dagger(E_1, E)|\hat{p}_1\rangle = \langle \hat{p}|\tau(E_1, E)|\hat{p}\rangle^*. \qquad \text{(IV. 10)}$$

The result is

$$\int d\hat{p}\, j(E)\langle E\hat{p}|[E^\dagger, K]|E'\hat{p}\rangle j(E') = -\int_0^\infty dE_1 \frac{\mathrm{tr}[\tau^\dagger(E_1, E)\tau(E_1, E')] - \mathrm{tr}[\tau(E, E_1)\tau^\dagger(E', E_1)]}{(E - E_1 - i0)(E_1 - E' - i0)}. \qquad \text{(IV. 11)}$$

Here, as before, tr denotes the trace on $\mathcal{H}_-$. The diagonal element of Eq. (IV.11) is obtained by letting $E' \to E$. To carry out this limit let us recall a result established by Jauch.[13] Let $f(E,E')$ and $g(E,E')$ be complex valued functions which are differentiable in their (real) arguments. Then the following formula is valid:

$$\lim_{E'\to E}\int_0^\infty dE_1 \frac{f(E,E_1)g(E_1,E') - g(E,E_1)f(E_1,E')}{(E-E_1-i0)(E_1-E'-i0)} = -i\pi\left[f(E,E)\frac{d}{dE}g(E,E) - g(E,E)\frac{d}{dE}f(E,E)\right]. \quad \text{(IV.12)}$$

If we apply (IV.12) to Eq. (IV.11) and integrate with respect to dE, then the left-hand side of Eq. (IV.12) is $\mathrm{Tr}[K^\dagger, K]$. Thus employing Eq. (IV.6) gives

$$N = -i\pi\int_0^\infty dE\,\mathrm{tr}\left[\tau^\dagger(E,E)\frac{d}{dE}\tau(E,E) - \tau(E,E)\frac{d}{dE}\tau^\dagger(E,E)\right]. \quad \text{(IV.13)}$$

This is Levinson's theorem expressed in terms of the scattering amplitudes. It may be restated in terms of the reduced S matrices, $s(E)$, by utilizing Eq. (IV.8). Simple algebra leads to

$$N = \frac{1}{4\pi}\int_0^\infty dE\left\{\mathrm{tr}\left[is^\dagger(E)\frac{d}{dE}s(E) - i\frac{ds^\dagger(E)}{dE}s(E)\right] + \mathrm{tr}\left[i\frac{d}{dE}[s^\dagger(E) - s(E)]\right]\right\} = \frac{1}{4\pi}\int_0^\infty dE\left\{-2\ \mathrm{tr}[q(E)] + i\frac{d}{dE}\mathrm{tr}[s^\dagger(E) - s(E)]\right\}. \quad \text{(IV.14)}$$

What remains is to understand the behavior of the term $\mathrm{tr}[s^\dagger(E) - s(E)]$. Let us define $t_2(z)$ by

$$t_2(z) = t(z) - V. \quad \text{(IV.15)}$$

Replace $t(E'+i0)$ in Eq. (IV.7) by $t_2(E'+i0)$ and denote the resultant two-parameter operator by $\tau_2(E,E')$. Using Eq. (IV.15) and Eq. (IV.8), one finds

$$i\frac{d}{dE}\mathrm{tr}[s^\dagger(E) - s(E)] = -\frac{4}{\pi}\left(\frac{\mu}{2}\right)^{3/2}\frac{\tilde{v}}{\sqrt{E}} - 8\pi\frac{d}{dE}\mathrm{Re}\,\mathrm{tr}[\tau_2(E,E)]. \quad \text{(IV.16)}$$

Thus Eq. (IV.14) now reduces to,

$$-2\pi N = \int_0^\infty dE\left[\mathrm{tr}[q(E)] + \frac{2}{\pi}\left(\frac{\mu}{2}\right)^{3/2}\frac{\tilde{v}}{\sqrt{E}}\right] + 4\pi\,\mathrm{Re}\,\mathrm{tr}[\tau_2(E,E)]\Big|_{E=0}^{E=\infty}. \quad \text{(IV.17)}$$

This is Levinson's theorem when the last term is zero. The fact that the zero energy on-shell t-matrix is proportional to the scattering length means that $\mathrm{tr}[\tau_2(E,E)]$ behaves like a const$\times\sqrt{E}$ for small E. So, we have $\mathrm{tr}[\tau_2(0,0)] = 0$.

All that is left to consider is the high energy limit of $\mathrm{Re}\,\mathrm{tr}[\tau_2(E,E)]$. Under assumption (A) on the potential, it is well known[14] that

$$\left|\langle \mathbf{p}|\, t(p^2/2\mu + i0)\,|\mathbf{p}\rangle - \langle \mathbf{p}|\, V\,|\mathbf{p}\rangle\right| = \delta(p). \quad \text{(IV.18)}$$

and $\delta(p) \to 0$ as $p \to \infty$. Furthermore, the symmetry properties of the resolvent $r(z)$ under the transformation $p \to -p$ imply that the forward scattering amplitude

$$f(p) = -4\mu\pi^2\langle \mathbf{p}|\, t(p^2/2\mu + i0)\,|\mathbf{p}\rangle \quad \text{(IV.19)}$$

satisfies

$$f^*(p) = f(-p). \quad \text{(IV.20)}$$

Of course, only the forward scattering amplitude is needed to compute $\mathrm{Re}\,\mathrm{tr}[\tau_2(E,E)]$. The symmetry relation (IV.20) means $\mathrm{Re}[f(p)]$ is an even function of p. Thus, at infinity $\mathrm{Re}[f(p)] = O(p^{-2n})$. Now the estimate (IV.18) forces n to be a positive integer. Thus the slowest behavior possible for $\mathrm{Re}[f(p)]$ at infinity is $O(p^{-2})$. This observation combined with the definition (IV.7) of $\tau(E,E')$ implies $\mathrm{Re}\,\mathrm{tr}[\tau_2(E,E)] = O(E^{-1/2})$. Thus the high energy surface term in Eq. (IV.17) vanishes.

The details of this derivation indicate why our theorem must have the term $\tilde{v}$ present. Consider the high energy behavior of $\mathrm{tr}[q(E)]$. For sufficiently high energies we expect that the t matrix will be dominated by the Born term. If we replace the t matrix by v in the expression for the S matrix, then the first order contribution to $\mathrm{tr}[q(E)]$ in powers of potential is

$$\mathrm{tr}[q(E)] \sim -\frac{2}{\pi}\left(\frac{\mu}{2}\right)^{3/2}\frac{\tilde{v}}{\sqrt{E}}. \quad \text{(IV.21)}$$

For this reason $\mathrm{tr}[q(E)]$ is not integrable at infinity with respect to E. The $\mathrm{Tr}[r_0(z)\,Vr_0(z)]$ term in the integrand of the Levinson's theorem exactly cancels this singular behavior of $\mathrm{tr}[q(E)]$. With this singularity subtracted away it is now very reasonable to expect that condition (B) on this time delay is valid.

Since the form of our Levinson's theorem differs from the usual partial wave form, it is intructive to see how the customary result can emerge from the analysis given. This is most easily understood by starting from Eq. (IV.14). When the potential v is spherically symmetric, then the angular momentum operator $\mathbf{L} = \mathbf{x}\times\mathbf{p}$ commutes with h and h_0, so that the reduced S matrix and the time delay operator may be represented by

$$\langle \hat{p}|\, s(E)\,|\hat{p}'\rangle = \sum_{l=0}^\infty \frac{2l+1}{4\pi}\, s_l(E)P_l(\hat{p}\cdot\hat{p}'), \quad \text{(IV.22)}$$

$$\langle \hat{p}|\, q(E)\,|\hat{p}'\rangle = \sum_{l=0}^\infty \frac{2l+1}{4\pi}\, q_l(E)P_l(\hat{p}\cdot\hat{p}'). \quad \text{(IV.23)}$$

Here the S matrix admits the usual phase-shift parametrization, $s_l(E) = \exp[2i\delta_l(E)]$. The corresponding formula for the time delay is $q_l(E) = 2(d/dE)\delta_l(E)$.

Let us compute the contribution of a single partial wave amplitude to Eq. (IV.14). Each bound state has a $2l+1$ degeneracy, so let N_l denote the number of bound states with different energy. Upon substituting Eqs. (IV.22) and (IV.23) into Eq. (IV.14) we have

$$N_l = -\frac{1}{4\pi}\int_0^\infty dE\left\{4\frac{d}{dE}\delta_l(E) - 8\pi\frac{d}{dE}\sin[2\delta_l(E)]\right\}. \quad \text{(IV.24)}$$

The customary phase shift normalization is to set $\delta_l(\infty) = 0$. Thus Eq. (IV.24) becomes

$$\pi N_l = \delta_l(0) - 2\pi\sin[2\delta_l(0)] \quad \text{(IV.25)}$$

and has the solution

$$\pi N_l = \delta_l(0). \qquad \text{(IV. 26)}$$

This is the partial wave Levinson's theorem. For a single partial wave the terms $\mathrm{tr}[q(E)]$ and $(d/dE) \times \mathrm{tr}[s^\dagger(E) - s(E)]$ are individually integrable. For the full amplitude case these terms are separately divergent, but when added together their divergences cancel. The mechanism for changing the behavior of these terms is the infinite sum over partial waves.

We close this section with some general comments about the results found here. One interesting aspect of statement (III. 24) of Levinson's theorem, is that it relates two observables of the scattering system. Both the time delay $\mathrm{tr}[q(E)]$ and the number of bound states N are in principle observable features of the scattering system. One nonintuitive result of the theorem concerns the behavior of time delay when resonances are present. Consider the case when at some energy, E_r, there is a very long-lived resonance. Suppose the potential is slightly perturbed so that the lifetime of the resonance increases but the number of bound states is unaltered. Then Eq. (III. 24) tells us that at energies away from the resonance there must be a corresponding decrease in the time delay since the energy integral is invariant.

A second advantage of this theorem is that it is more general than the usual Levinson's theorem in that it remains valid for scattering from a nonspherically symmetric potential. Furthermore, the general approach given here may obtain Levinson's theorems for the few-body scattering problem. We note that the spectral property of time delay has already been established for the three-body problem.[5]

So far physical applications of this theorem have not been investigated. However, one application is straightforward. The theorem may be used to predict the high temperature behavior of the second virial coefficient for a quantum gas. This will be reported on elsewhere.

ACKNOWLEDGMENTS

One of us (D. B.) would like to thank Professor John Nuttall and the University of Western Ontario for their hospitality. The financial support of the Canada Council is gratefully acknowledged.

*Research supported in part by the Energy Research and Development Administration under Contract No. E(04-3)-515 and by the U.S. National Science Foundation Grant 21101.

†Permanent address: Department of Physics, The University of Manitoba, Winnipeg, Manitoba R3T 2N2, Canada.
‡Permanent address: Institute of Theoretical Physics, Celestijnenlaan, 200D, B-3030 Heverlee, Belgium. Onderzoeker IIKW, Belgium.

[1]D. Bollé and T. A. Osborn, Phys. Rev. D 13, 299 (1976).
[2]J.M. Jauch and J.P. Marchand, Helv. Phys. Acta 40, 217 (1967).
[3]J.M. Jauch, K.B. Sinha, and B.N. Misra, Helv. Phys. Acta 45, 398 (1972).
[4]Ph. A. Martin, Commun. Math. Phys. 47, 221 (1976).
[5]T.A. Osborn and T.Y. Tsang, Ann. Phys. (N.Y.) 101, 119 (1976).
[6]T. Kato, *Perturbation Theory for Linear Operators* (Springer, New York, 1966), Chap. X.
[7]Equation (I.13) of the text may be obtained from Ref. 3 by combining statement (i) of Theorem 2 with the last equation on p. 408.
[8]J. Nowakowski and T.A. Osborn, Nucl. Phys. A 249, 301 (1975).
[9]R.G. Newton, *Scattering Theory of Waves and Particles* (McGraw-Hill, New York, 1966), Sec. 11.
[10]B. Simon, *Quantum Mechanics for Hamiltonians Defined as Quadratic Forms* (Princeton U.P., Princeton, N.J., 1971).
[11]R. Schatten, *Norm Ideals of Completely Continous Operators* (Springer-Verlag, Berlin, 1960).
[12]V.S. Buslaev, Dokl. Akad. Nauk SSSR 143, 1067 (1962) [Sov. Phys. Dokl. 7, 295 (1962)].
[13]J.M. Jauch. Helv. Phys. Acta 30, 143 (1957).
[14]H.M. Nussenzveig, *Causality and Dispersion Relations* (Academic, New York, 1972), Chap. 6.

Spin-frame independent variables in general relativity

K. Wódkiewicz

Department of Physics and Astronomy, University of Rochester, Rochester, New York 14627
and Institute of Theoretical Physics, Warsaw University, 00—681, Warsaw, Poland
(Received 15 March 1976)

The purpose of this paper is to present the spin-frame independent variables in general relativity. The work is based on the fact that the tetrad Newman–Penrose form of Einstein's equations can be put into the Yang–Mills form with the group SL(2,ℂ) as the gauge group. The set of Mandelstam path-dependent dynamical variables for such a theory forms spin-frame independent variables in general relativity. The empty-space field equations for spin-frame independent variables have formally the same form as Maxwell's equations. In addition, the full set of field equations for spin-frame independent variables have formally the same form as the equations of nonlinear electrodynamics.

1. INTRODUCTION

In the formulation of the gravitational canonical formalism or in any effort toward the quantization of this field the main complication is related to the construction of coordinate independent variables. The situation in some respects is similar to the one which we encounter in electrodynamics when we try to formulate the canonical formalism and apply it to the quantization procedure. Due to the gauge invariance of the electromagnetic field any attempt to construct Poisson brackets or equal-time commutators for gauge dependent quantities in a covariant form and in full agreement with the equations of motion fails. The well known solution of the problem is to choose one particular gauge, to form Feynman diagrams, and at the end to prove that all physical quantities like the S matrix are gauge invariant. Due to the renormalization procedure in quantum electrodynamics the last point was not easy to prove.[1] Another possible way of proceeding is to formulate a gauge independent canonical formalism and to quantize the field in a gauge independent way. Such a gauge independent quantization of the electromagnetic field has been known for a long time and is due to Mandelstam.[2]

As a result of the formal similarities between the gravitational field and the electromagnetic field, the literature contains several attempts to build a canonical formalism and to quantize the gravitational field using the methods of classical and quantum electrodynamics.[3] The analogies between the Maxwell and Einstein fields are brought out in a striking way by the spinor formalism. The source free Maxwell equations and the Bianchi identities become, respectively,

$$\partial^{A\dot{C}}\Phi_{AB}=0, \tag{1.1}$$

$$\partial^{A\dot{E}}\Psi_{ABCD}=0. \tag{1.2}$$

A careful investigation of the algebraic structure and of the nonlinearity of the gravitational field indicates a possible way of gaining a better understanding of general relativity in the context of the Yang—Mills field. This similarity is based on the fact that in gravitation, the non-Abelian internal gauge group [SU(2) in the original work of Yang and Mills[4]] is the non-Abelian group of coordinate transformation. As a result, in almost all papers related to the quantization of the Yang—Mills field we can find a paragraph devoted entirely to the gravitational field.[5–8]

In all these papers the difficult problem of quantizing the Yang—Mills field is attacked by different techniques and methods. Nevertheless, one common result emerges: the Feynman diagrams for the Yang—Mills field (obtained for the first time by Feynman himself in his pioneer work on this subject).[5] If we compare the paragraphs related to the gravitational field we discover that each author uses different aspects of gravitational field for the quantization after Yang—Mills. Mandelstam used the path dependent Riemanian tensor $R^{\mu}_{\mu\alpha\beta}(x,P)$,[7] and Fradkin and Tyutin used the Chrystoffel symbol $\Gamma^{\mu}_{\nu\lambda}$ and $g^{\beta}g_{\mu\nu}$ as basic variables.[8] In fact, it is very difficult to compare results due to divergences and to the different gauges used. This situation is due to one basic question. How does one relate the gravitational field to the Yang—Mills field in a compact and general form?

Recently, Carmeli[9] has shown the relationship between the gravitational dynamical variables and the Yang—Mills field variables. He used the spinor—tetrad formalism developed by Newman and Penrose.[10] In the framework of this formalism he was able to show that the spin coefficients and the Riemann tensor play the role of the Yang—Mills potentials and the Yang—Mills fields, respectively. The non-Abelian gauge group is the group SL(2,ℂ) acting in the spin space. However, there is one problem. The Yang—Mills field constructed by Carmeli is not a gauge-independent SL(2,ℂ) invariant object. In contrast to the electromagnetic field (Yang—Mills with Abelian gauge) the field strength is a gauge dependent object and therefore it cannot be used for any gauge independent procedures (canonical formalism or quantization).

The purpose of this paper is to present in the framework of the Newman—Penrose (NP) formalism a set of spin—frame independent variables for the gravitational field.

In Sec. II we briefly introduce the Yang—Mills description of the gravitational field using the Newman—Penrose formalism, and we also establish our own notation. In Sec. III we introduce spin-frame independent fields as path-dependent quantities. The empty-space field equations (1.2), in terms of these new fields, coincide formally with the field equations of Maxwell's electrodynamics. In Sec. III, we generalize the result to the full set of field equations (1.2).

2. THE EMPTY-SPACE FIELD EQUATIONS AND YANG-MILLS FIELDS

The connection between tensors $T^{\mu\nu\cdots}_{\rho\cdots}$ and spinors $T^{A\dot{B}C\dot{D}\cdots}_{E\dot{F}\cdots}$ is[11]

$$T^{A\dot{B}C\dot{D}\cdots}_{E\dot{F}\cdots} = \sigma_\mu^{A\dot{B}}\sigma_\nu^{C\dot{D}}\cdots\sigma^\rho_{E\dot{F}}\cdots T^{\mu\nu\cdots}_{\rho\cdots}. \tag{2.1}$$

The quantities $\sigma_\mu^{A\dot{B}}$ satisfy the relation

$$g_{\mu\nu}\sigma^\mu_{A\dot{B}}\sigma^\nu_{C\dot{D}} = \epsilon_{AC}\epsilon_{\dot{B}\dot{D}} \tag{2.2}$$

where $g_{\mu\nu}$ is the metric tensor and the ϵ's are the Levi-Civita symbols,

$$\epsilon_{AB} = \epsilon^{AB} = \begin{Vmatrix} 0 & 0 \\ -1 & 0 \end{Vmatrix}. \tag{2.3}$$

In the Newman–Penrose formalism we introduce at each point of a curved space–time two basic spinors o^A and ι^A. The spinors o^A and ι^A satisfy the normalization condition

$$o^A\iota_A = 1, \tag{2.4}$$

and the completeness relation

$$\epsilon_{AB} = o_A\iota_B - \iota_B o_A. \tag{2.5}$$

We denote the spin frame by ξ_a^A, where

$$\xi_1^A = o^A, \quad \xi_2^A = \iota^A.$$

The dyadic components of a spinor in a given spin frame following Newman and Penrose are given by[10]

$$T^{a\dot{x}\cdots}_{\cdots} = \xi^a_A\,\xi^{\dot{x}}_{\dot{X}}\cdots T^{A\dot{X}\cdots}_{\cdots}, \tag{2.6}$$

where the rules of lowering or raising spinor indices are

$$\xi^A = \epsilon^{AB}\xi_B, \quad \xi_B = \xi^A\epsilon_{AB}. \tag{2.7}$$

Having a dyad in the spin–space, we can build a tetrad in the vector space,

$$l^\mu = \sigma^\mu_{A\dot{B}}o^A o^{\dot{B}}, \quad n^\mu = \sigma^\mu_{A\dot{B}}\iota^A\iota^{\dot{B}},$$

$$m^\mu = \sigma^\mu_{A\dot{B}}o^A\iota^{\dot{B}}, \quad \bar{m}^\mu = \sigma^\mu_{A\dot{B}}\iota^A o^{\dot{B}}. \tag{2.8}$$

The set of four vectors l^μ, n^μ, m^μ, $\bar{m}^\mu$ form a null tetrad system with the completeness relation

$$g^{\mu\nu} = l^\mu n^\nu + l^\nu n^\mu - m^\mu\bar{m}^\nu - \bar{m}^\mu m^\nu. \tag{2.9}$$

Following the basic idea of the Yang–Mills' fields we introduce a gauge group as follows. At each point of the curved space–time we introduce ξ_a^A to define a spin frame. A spin frame gauge can be defined as an arbitrary way of choosing the orientation of the spin-frame axes at all space–time points. We then demand that all physical processes be invariant under the spin-frame transformation,

$$\xi = S\xi', \tag{2.10}$$

where ξ is a 2×2 complex matrix whose elements are ξ_a^A and $S\in \mathrm{SL}(2,\mathbb{C})$. From the definition of the spin frame, the dynamical variables of the gravitational field can be introduced.[9]

The covariant derivative acting on the spin frame gives

$$\nabla_\mu\xi_a^A = B^b_{a\mu}\xi_b^A. \tag{2.11}$$

Using matrix notation, this equation can be written as

$$\nabla_\mu\xi = B_\mu\xi. \tag{2.12}$$

In the same matrix notation the commutator of the covariant derivatives $\nabla_\mu\nabla_\nu - \nabla_\nu\nabla_\mu$ acting on ξ gives $F_{\mu\nu}$, where

$$F_{\mu\nu} = \nabla_\nu B_\mu - \nabla_\mu B_\nu + [B_\mu, B_\nu]. \tag{2.13}$$

The tetrad components of B_μ and $F_{\mu\nu}$ are given, respectively, by (compare formula 2.1 and 2.6)

$$B_{a\dot{b}} = \xi_a^A\xi_{\dot{b}}^{\dot{B}}\sigma^\mu_{A\dot{B}}B_\mu, \tag{2.14}$$

$$F_{ab\dot{c}\dot{d}} = \xi_a^A\,\xi_b^{\dot{B}}\,\xi_{\dot{c}}^C\,\xi_{\dot{d}}^{\dot{D}}\sigma^\mu_{A\dot{B}}\sigma^\nu_{C\dot{D}}F_{\mu\nu}. \tag{2.15}$$

It was shown in Ref. 9 that the B_μ fields are related to the Newman–Penrose spin coefficients,

$$B_{00} = \begin{pmatrix} \epsilon & -\kappa \\ \pi & -\epsilon \end{pmatrix} \quad B_{01} = \begin{pmatrix} \beta & -\sigma \\ \mu & -\beta \end{pmatrix},$$
$$B_{1\dot{0}} = \begin{pmatrix} \alpha & -\rho \\ \lambda & -\alpha \end{pmatrix}, \quad B_{1\dot{1}} = \begin{pmatrix} \gamma & -\tau \\ \nu & -\gamma \end{pmatrix}, \tag{2.16}$$

and that the $F_{\mu\nu}$ fields are related to the Newman–Penrose tetrad components of the Weyl tensor,

$$F_{0\dot{1}0\dot{1}} = \begin{pmatrix} \psi_1 & -\psi_0 \\ \psi_2 & -\psi_1 \end{pmatrix}, \quad F_{1\dot{1}1\dot{0}} = \begin{pmatrix} \psi_3 & -\psi_2 \\ \psi_4 & -\psi_3 \end{pmatrix},$$
$$F_{1\dot{1}0\dot{0}} = F_{0\dot{1}1\dot{0}} = \begin{pmatrix} \psi_2 & -\psi_1 \\ \psi_3 & -\psi_2 \end{pmatrix}. \tag{2.17}$$

We see that B_μ (the spin coefficients) play the role of Yang–Mills potentials and that $F_{\mu\nu}$ (the Weyl tensor) plays the role of the Yang–Mills fields with the group $\mathrm{SL}(2,\mathbb{C})$ as a gauge group. Under a change of spin frame, $\xi = S\xi'$, the potentials B_μ and the field $F_{\mu\nu}$ are transformed,

$$B'_\mu = SB_\mu S^{-1} - S^{-1}\partial_\mu S, \tag{2.18a}$$

$$F'_{\mu\nu} = SF_{\mu\nu}S^{-1}. \tag{2.18b}$$

As in the Yang–Mills case the simplest Lagrangian density which is invariant under both general coordinate transformations and spin-frame transformations is

$$\mathcal{L} = -\tfrac{1}{4}(-g)^{1/2}\mathrm{Tr}(F_{\mu\nu}F^{\mu\nu}). \tag{2.19}$$

This Lagrangian leads to the following set of field equations:

$$\nabla_\mu F^{\nu\mu} - [B_\mu, F^{\nu\mu}] = 0. \tag{2.20}$$

The tetrad projection of these field equations leads to the Newman–Penrose form of field equation (1.2).[9]

The form of the field equations (2.20), the definition of $F_{\mu\nu}$ (2.13), and the transformations (2.18) under gauge changes indicate the Yang–Mills form to the theory. The important difference is that the gauge group is the group of all possible changes of the spin–frame, the group $\mathrm{SL}(2,\mathbb{C})$.

3. SPIN-FRAME INDEPENDENT DYNAMICAL VARIABLES FOR EMPTY-SPACE FIELD EQUATIONS

In contrast to the Abelian gauge in electrodynamics

given by the group U(1), the field strength $F_{\mu\nu}$ given by Eq. (2.13) is not invariant under SL(2,$\mathbb{C}$) gauge transformations. It means that we cannot build a coordinate independent canonical formalism for the gravitational field based on the Weyl tensor. At this point the analogy between electrodynamics and general relativity breaks down due to the fact that the gauge group of the theory is not Abelian. We need to introduce new objects for the gravitational field which are independent of the particular spin frame chosen.

One can construct spin-frame independent fields $\mathcal{J}_{\mu\nu}$. Such quantities for the Yang—Mills field with the gauge group SU(2) are known (see Ref. 12). Let us now extend this procedure to the field presented in Sec. 2.

Following Mandelstam's path-dependent formulation of quantum electrodynamics we introduce a path dependent matrix,[13]

$$U(x,P)=T\exp\Big(-\int_P^x B_\mu dx^\mu\Big)=T\exp\Big(-\int_P^s B_\mu(s)\frac{dx^\mu}{ds}ds\Big), \tag{3.1}$$

where the linear integrals are evaluated along a certain path P: $x^\mu(s)$. The T operation denotes the s ordering of the B_μ's,

$$T(B_\mu(s_1)B_\mu(s_2))=B_\mu(s_1)B_\mu(s_2)\quad\text{if}\quad s_1>s_2. \tag{3.2}$$

The derivatives of the matrix $U(x,P)$ are given by

$$\nabla_\mu U(P)=-B_\mu U(P). \tag{3.3}$$

Lets define a new gauge-invariant field $\mathcal{J}_{\mu\nu}$ as a path-dependent quantity,

$$\mathcal{J}_{\mu\nu}(P)=U^*(P)F_{\mu\nu}U(P), \tag{3.4}$$

where $F_{\mu\nu}$ is given by Eq. (2.13) and U^* is the Hermitian conjugate of $U(P)$.

Under the gauge transformation (2.18) the gauge-transformed matrix

$$U'=T\exp\Big(-\int_p^x B'_\mu dx^\mu\Big) \tag{3.5}$$

satisfies the following equation:

$$\nabla_\mu U'=-B'_\mu U'=-SB_\mu S^{-1}U'+(\partial_\mu S)S^{-1}U'. \tag{3.6}$$

It is easy to check that the matrix SU satisfies the same equation. Because the two matrices SU and U have the same boundary condition,

$$SU=\big|_{s=-\infty}=U'\big|_{s=-\infty}=1, \tag{3.7}$$

it follows that they are the same,

$$U'=SU\Longrightarrow U'^*=U^*S^{-1}. \tag{3.8}$$

Now we can prove the gauge invariance of $\mathcal{J}_{\mu\nu}(P)$. Under the gauge transformation (2.18), the field $\mathcal{J}_{\mu\nu}(P)$ transforms as follows:

$$\mathcal{J}'_{\mu\nu}(P)=U'^*F'_{\mu\nu}U'=U'^*SF_{\mu\nu}S^{-1}U'=U^*F_{\mu\nu}U=\mathcal{J}_{\mu\nu}(P), \tag{3.9}$$

where we have used the relations (3.8). So we have proved that the path-dependent field $\mathcal{J}_{\mu\nu}(P)$ defined by the formula (3.4) is a spin-frame independent object. In terms of $\mathcal{J}_{\mu\nu}(P)$ we can rewrite the field equations (2.20) and (2.13),

$$\nabla_\mu \mathcal{J}^{\mu\nu}(P)=0, \tag{3.10a}$$

$$\nabla_\lambda \mathcal{J}_{\mu\nu}(P)+\nabla_\nu \mathcal{J}_{\lambda\mu}(P)+\nabla_\mu \mathcal{J}_{\nu\lambda}=0. \tag{3.10b}$$

This set of equations is equivalent to the vacuum Bianchi equations (1.2). The field equations (3.10) obeyed by spin-frame independent fields are linear and coincide formally with the field equations in Maxwell electrodynamics.

4. SPIN-FRAME INDEPENDENT VARIABLES FOR THE FULL FIELD EQUATIONS

The results of Sec. 3 were obtained for the case of empty-space Einstein equations. This means that in the decomposition of the Riemann tensor into its irreducible components we took into account only the Weyl tensor. Now we want to generalize our previous results such that the tracefree part of the Ricci tensor and the Ricci scalar are present. The field,

$$F_{\mu\nu}=\nabla_\nu B_\mu-\nabla_\mu B_\nu+[B_\mu,B_\nu], \tag{4.1}$$

leads to the following tetrad components[9]:

$$F_{0\dot0\dot1\dot0}=\begin{pmatrix}\psi_1 & -\psi_0\\ \psi_2+2\Lambda-\psi_1 & \end{pmatrix},\quad F_{1\dot1\dot1\dot0}=\begin{pmatrix}\psi_2+\phi_{11}-\Lambda & -\psi_1-\phi_{01}\\ \psi_3+\phi_{21} & -\psi_2-\phi_{11}+\Lambda\end{pmatrix},$$

$$F_{1\dot0\dot0\dot0}=\begin{pmatrix}\phi_{10} & -\phi_{00}\\ \phi_{20} & -\phi_{10}\end{pmatrix},\quad F_{1\dot1\dot0\dot1}=\begin{pmatrix}\phi_{12} & -\phi_{02}\\ \phi_{22} & -\phi_{12}\end{pmatrix}, \tag{4.2}$$

$$F_{1\dot1\dot1\dot0}=\begin{pmatrix}\psi_3 & -\psi_2-2\Lambda\\ \psi_4 & -\psi_3\end{pmatrix},\quad F_{1\dot0\dot0\dot1}=\begin{pmatrix}-\psi_2+\phi_{11}+\Lambda & \psi_1-\phi_{01}\\ -\psi_3+\phi_{21} & \psi_2-\phi_{11}-\Lambda\end{pmatrix},$$

where the nine ϕ's describe the tetrad components of the tracefree part of the Ricci tensor, R is the Ricci scalar and $\Lambda=\frac{1}{24}R$.

Using the definition of σ's and the field $F_{\mu\nu}$ we can define the following field:

$$H^{\mu\nu}=\sigma^{\mu C\dot B}\sigma^{\nu A\dot D}\sigma^\alpha_{A\dot B}\sigma^\beta_{C\dot D}F_{\alpha\beta}. \tag{4.3}$$

The full set of gravitational Newman—Penrose equations can be obtained from the Lagrangian density[14]

$$\mathcal{L}=-\tfrac{1}{2}(-g)^{1/2}\mathrm{Tr}\{H^{\mu\nu}(-\tfrac{1}{2}F_{\mu\nu}+\nabla_\mu B_\mu-\nabla_\mu B_\nu+[B_\mu,B_\nu])\}. \tag{4.4}$$

In empty space this Lagrangian becomes the Lagrangian given by the formula (2.19). The Lagrangian density generates the field equations

$$\nabla_\nu H^{\mu\nu}-[B_\nu,H^{\mu\nu}]=0. \tag{4.5}$$

As in the previous case, let's define two new spin-frame invariant fields, $\mathcal{J}_{\mu\nu}$ and $\mathcal{H}_{\mu\nu}$, as path dependent quantities,

$$\mathcal{J}_{\mu\nu}(P)=U^*(P)F_{\mu\nu}U(P), \tag{4.6a}$$

$$\mathcal{H}_{\mu\nu}(P)=U^*(P)H_{\mu\nu}U(P), \tag{4.6b}$$

where the definition of the matrix U is given by (3.1). As in Sec. 3, it is easy to prove that the path-dependent $\mathcal{J}_{\mu\nu}$ and $\mathcal{H}_{\mu\nu}$ are both gauge invariant under the SL(2,$\mathbb{C}$) gauge. In terms of these new fields, the field equations (4.5) and (4.1) can be written as follows:

$$\nabla_\mu \mathcal{H}^{\mu\nu}(P)=0, \tag{4.7a}$$

$$\nabla_\lambda \mathcal{J}_{\mu\nu}(P) + \nabla_\nu \mathcal{J}_{\lambda\mu}(P) + \nabla_\mu \mathcal{J}_{\nu\lambda}(P) = 0, \tag{4.7b}$$

and

$$H^{\mu\nu}(P) = \frac{\delta L}{\delta \mathcal{J}_{\mu\nu}(P)}. \tag{4.7c}$$

This set of equations is equivalent to the full Bianchi equations (1.2). The field equations obeyed by the spin-frame independent fields (4.7) coincide formally with the field equations of nonlinear electrodynamics.

5. CONCLUDING REMARKS

The hope to obtain nontrivial results in quantum gravitation is based on the recent successful investigation of the Yang—Mills field theory. Apart from the discovery of the Feynman diagrams for such a theory, the basic results are due to 't Hooft who proved that the Yang—Mills theory is a renormalizable theory.[15] This fact opened the possibility of investigating higher order Feynman diagrams without divergences. The basic tool in the renormalization procedure is the Slawnov—Taylor identity, which is the generalization of the Ward identity to non-Abelian theories.[16]

As was shown in Sec. 3, the empty-space field equations (1.2) for the path dependent objects $\mathcal{J}_{\mu\nu}(P)$ have the form of the equations of linear electrodynamics. Apart from the hard problem of the interpretation, this theory seems to fit the 't Hooft results. The careful investigation of one loop diagrams and the form of Ward identity can hopefully shed some light on the renormalization of quantum gravitation. This problem will be investigated in a further publication. The example of the full field equations presented in Sec. 4 indicates further complications. Bianchi's equation coupled to matter or to the electromagnetic field for path-dependent quantities has the form of the equations of nonlinear electrodynamics. There exists in the literature some conjectures that such theories are not renormalizable.[17]

The quantization of the gravitational field has still some fundamental difficulties. Even in the simplest case of a Yang—Mills theory based on the group SU(2), there is no proof that the amplitudes or the probabilities are gauge invariant.

The second problem is that the present quantization of the gravitational field is rather a theory for spin two but not for gravity.[17,18] For these reasons a better understanding of the classical theory may give deeper insight into its structure and indicate the direction of future investigations.

ACKNOWLEDGMENTS

The author would like to thank Dr. J.R. Ackerhalt for reading the manuscript.

[1]I. Białynicki-Birula, Phys. Rev. **155**, 1414 (1967).
[2]S. Mandelstam, Ann. Phys. (N.Y.) **19**, 1 (1962).
[3]See, for example, R. Arnowitt, S. Deser, and C.W. Misner, in *Gravitation: An Introduction to Current Research*, edited by L. Witten (Wiley, New York, 1962); A. Komar, Phys. Rev. **134**, 1430B (1964); **164**, 1595 (1967); K. Wódkiewicz, Acta Phys. Polon. B **6**, 509, 521 (1975).
[4]C.N. Yang and R.L. Mills, Phys. Rev. **96**, 191 (1954).
[5]R.P. Feynman, Acta Phys. Polon. **24**, 697 (1963).
[6]B.S. DeWitt, Phys. Rev. **162**, 1195, 1239 (1967); **171**, 1834(E) (1968).
[7]S. Mandelstam, Phys. Rev. **175**, 1580, 1604 (1968).
[8]E.S. Fradkin and I.V. Tyutin, Phys. Rev. D **2**, 2841 (1970); E.S. Fradkin and G.A. Vilkovisky, Phys. Rev. D **8**, 4241 (1973).
[9]M. Carmeli, Gen. Rel. Grav. **5**, 287 (1974) and references cited therein. See also R. Utiyama, Phys. Rev. **101**, 1597 (1956); C.J. Isham, A. Salam, and J. Strathdee, Nuovo Cimento Lett. **5**, 969 (1972).
[10]E.T. Newman and R. Penrose, J. Math. Phys. **3**, 566 (1962).
[11]Latin italic capitals A, B, $\cdots$ are used for spinor indices and take the value $A = 1, 2$, $\dot{B} = \dot{1}, \dot{2}$. Greek letters are used for space—time indices, Latin letters ab denote the tetrad components of the spinor and take the value $a = 1, 2$, $\dot{b} = \dot{1}, \dot{2}$. For a detailed introduction to tetrad—spinor formalism see F.A.E. Pirani, in *Lectures in General Relativity*, Brandeis Summer Institute in Theoretical Physics, edited by S. Deser and K.W. Ford (Prentice-Hall, Englewood Cliffs, N.J., 1965).
[12]I. Białynicki-Birula, Bull. Acad. Polon. Sci. **11**, 135 (1963).
[13]The path integration for gravitational field was introduced first by S. Mandelstam, Ann. Phys. (N.Y.) **19**, 25 (1962). See in this paper the discussion of physical interpretation of the formulation.
[14]M. Carmeli and S.L. Fickler, Phys. Rev. D **5**, 290 (1972).
[15]G. 't Hooft, Nucl. Phys. B **33**, 173 (1971).
[16]See for example, E.S. Abers and B.W. Lee, Phys. Rep. 9C, No. 1 (1973).
[17]S. Deser and P. van Nieuwenhuizen, Phys. Rev. Lett. **32**, 245 (1975); Phys. Rev. D **10**, 401, 411 (1974).
[18]G. 't Hooft and M. Veltman, Ann. Inst. Henri Poincaré, Sec. A **20**, 69 (1974).

Global operator product expansions for free fields of arbitrary mass $m \geq 0$

A. H. Völkel

Department of Mathematics, P.H.B., Berlin, Germany
*and Institut für Theoretische Physik, F.U., Berlin, Germany**
(Received 27 April 1976)

A set (more than countably many) of global operator expansions—"on and off the vacuum"—are proved to hold for free fields of any mass $m \geq 0$. Conformal invariance ($m = 0$) singles out exactly one of them in the case of the "vacuum expansion." There does not exist any termwise conformal covariant expansion "off the vacuum."

I. PRELIMINARY REMARKS

The motivation for studying conformal covariant (QFT) quantum field theory originates from the widespread belief that conformal covariance may play a central role in constructing global operator expansions.[1–9] This program consists of two main parts:

(I) Construction of (composite) field operators with the correct transformation properties under global conformal transformations.

(II) The derivative of global operator expansions by means of these fields "on and off the vacuum."

For both parts some progress has been recently achieved[10–15] respectively.[6–9] Nevertheless, the (unsolved) problems one faces in performing this program in a Minkowski quantum field theory are large and deep enough that for a deeper understanding a detailed investigation in the (technically) simplest possible model of quantum field theory seems to be worthwhile. The simplest such model is, of course, the free field or generalized free field. Even in this case the "off vacuum" expansion is not yet completely understood.[6]

The main problem of Minkowski conformal covariant QFT is the reconciliation of Einstein causality with global conformal transformations. Since the structure of the conformal group in itself is already very complicated, it is of tremendous help to observe that the causality problem is completely understood if it is solved for one single element, namely, the so-called conformal inversions:

$$x \mapsto Rx = -x/x^2, \quad x^2 = (x^0)^2 - \mathbf{x}^2. \tag{I.1}$$

Note that any special conformal transformation may be written as

$$K(b) = RT(b)R, \tag{I.2}$$

where $T(b)$ is a translation by b. Hence we will restrict ourselves to this transformation.

In Sec. II we will very briefly review the generalized free field theory. Section III is devoted to the construction of a set of composite operators (Wick products and their derivatives) for generalized free fields, which is properly covariant with respect to conformal inversions (transformations). Finally in Sec. IV we derive by means of these operators a whole set (more than countably many) of weakly convergent operator product expansions "on and off the vacuum" for any free field theory of arbitrary mass $m \geq 0$—conformally covariant ($m = 0$) or not ($m > 0$).

However, in the "off vacuum" case the coefficients are not tempered distributions anymore. Hence the expansion exists only on a restricted subspace of the Schwartz space[16] S_4 (smearing of the fields in configuration space), for instance, the subspace D_4 of C^∞ functions with compact support.[16,17] Also the weakly convergence has to be understood in the restricted sense of all states from the dense set of functions with compact support in momentum space.

II. REVIEW OF GENERALIZED FREE FIELDS

In this section we want briefly collect the main facts and formulas of generalized free fields, which we need later on. The details and their derivation may be found in Ref. 18. Let $\mathfrak{H}^0$ be the Hilbert space of the complex numbers ψ^0 with the scalar product

$$(\Psi_0, \Phi_0) = \bar{\psi}^0 \phi^0 \tag{II.1}$$

and $\mathfrak{H}^n = L^2(d^n\mu)$ the Hilbert space of all equivalence classes $\psi^n(p_1, \ldots, p_n)$ of complex functions of n four vector variables, which are symmetric under permutations and square integrable with respect to the measure

$$\begin{aligned} d^n\mu(p_1, \ldots, p_n) &= \prod_{i=1}^{n} d^4p_i \theta_+(p_i^2)\mu(p_i^2), \\ \theta_+(p^2) &= \theta(p^0)\theta(p^2). \end{aligned} \tag{II.2}$$

Here $\mu(p^2)$ is some positive tempered measure on the closed positive real axis. For the free field of mass m $\mu(p^2)$ is given by:

$$\mu(p^2) = \delta(p^2 - m^2). \tag{II.3}$$

Now the Hilbert space of the theory is the direct sum of all $\mathfrak{H}^n$

$$\mathfrak{H} = \bigoplus_{n=0}^{\infty} \mathfrak{H}^n. \tag{II.4}$$

The elements Ψ of $\mathfrak{H}$ are sequences

$$\Psi = \{\psi^0, \psi^1(p_1), \psi^2(p_1, p_2), \ldots, \psi^n(p_1, \ldots, p_n), \cdots\} \tag{II.5}$$

with the norm

$$\|\Psi\|^2 = \sum_{n=0}^{\infty} \int d^n\mu(p_1, \ldots, p_n) |\psi^n(p_1, \ldots, p_n)|^2 < +\infty.$$

Local field operators $\varphi(f) = \int d^4x \varphi(x) f(x)$ are introduced in the following way: Let $\mathcal{D}_\varphi$ denote the dense domain in $\mathfrak{H}$, consisting of all *finite* linear combinations of vectors

of form

$$\Psi_n=\{0,\ldots,0,\ \underset{n-1}{}\ \psi^n(p_1,\ldots,p_n),\ \underset{n}{}\ 0,\ \underset{n+1}{}\ 0,\cdots\} \tag{II.6}$$

with $\psi^n(p_1,\ldots,p_n)\in S_{4n}$ being the Schwartz space of all C^∞ functions of strong decrease.[16,17] On $\mathcal{D}_\varphi$ we define for every $f\in S_4$ linear operators $\varphi_j(f)$ $(j=0,1)$ by

$$\varphi_0(f)\Psi_n=\Big\{0,\ldots,0,\ \underset{n}{}\ \frac{1}{\sqrt{n+1}}\ \underset{n+1}{\mathcal{S}}\ \tilde{f}(-p_{n+1})\ \underset{n+1}{}$$

$$\times\psi^n(p_1,\ldots,p_n),\ 0,0,\cdots\}, \tag{II.7}$$

where $\mathcal{S}_{n+1}$ means symmetrization in all $n+1$ vectors p_n, and

$$\varphi_1(f)\Psi_n=\{0,0,\ldots,0,\ \sqrt{n}\int d^1\mu(q)$$

$$\times\tilde{f}(q)\psi^n(p_1,\ldots,p_{n-1},q),0,\cdots\}, \tag{II.8}$$

$$\tilde{f}(q)=[1/(2\pi)^{3/2}]\int d^4x\exp(-ip\cdot x)f(x). \tag{II.9}$$

φ_0 and φ_1 are just the familiar creation and annihilation operators or the positive respectively negative frequency parts of the Wightman field

$$\varphi(f)=\varphi_0(f)+\varphi_1(f). \tag{II.10}$$

They have all properties of a Wightman field except locality or Einstein causality.

Besides the basic local field $\varphi(f)$ and its nonlocal constituents we introduce what we call in the future local composite fields: $:\varphi^l:(f)$. They are nothing else than the Wick products of φ, for instance,

$$:\varphi^2:(x)=\lim_{y\to x}\{\varphi(x)\varphi(y)-(\Psi_0,\varphi(x)\varphi(y)\Psi_0).$$

In close analogy to the basic fields $\varphi(f)$ they may be represented in form of a sum of l nonlocal constituents $:\varphi^l:_j(f)$, $(j=0,1,\ldots,l)$

$$:\varphi^l:(f)\Psi=\sum_{j=0}^{l}:\varphi^l:_j(f)\Psi,\quad \Psi\in\mathcal{D}_\varphi, \tag{II.11}$$

where the fields $:\varphi^l:_j(f)$ give rise to comparatively simple transitions in $\mathfrak{H}$, namely

$$\mathfrak{H}_n\ \xrightarrow{:\varphi^l:_j}\ \begin{cases}\mathfrak{H}_{(n+l-2j)} & \text{for } j\le\min\{(l+n)/2,n\}\\ 0 & \text{for } j>\min\{(l+n)/2,n\}.\end{cases} \tag{II.12}$$

to be compared to the basic fields $\varphi_j(f)$

$$\mathfrak{H}_n\ \xrightarrow{\varphi_j}\ \begin{cases}\mathfrak{H}_{n+1} & \text{for } j=0\\ \mathfrak{H}_{n-1} & \text{for } j=1.\end{cases} \tag{II.13}$$

Explicitly the operators $:\varphi^l:_j(f)$ are defined by

$$:\varphi^l:_j(f)\Psi_n=\{0,\ldots,0,(2\pi)^{-3/2(l-1)}\left(\frac{n!}{(n+l-2j)!}\right)^{1/2}$$

$$\times\int\cdots\int\prod_{i=1}^{j}d^1\mu(q_i)\sum_{k_1<k_2<\cdots<k_{l-j}=1}^{n+l-2j}\frac{1}{j!}$$

$$\times\sum_{P(q_1,\ldots,q_j,-p_{k_1},\ldots,-p_{k_{l-j}})}f\Big(\sum_{i=1}^{j}q_i-\sum_{r=1}^{l-j}p_{k_r}\Big)$$

$$\times\psi^n(q_1,\ldots,q_j,\ p_1,\ldots,\hat{p}_{k_1},\ldots,\hat{p}_{k_{l-j}},\ldots,p_{n+l-2j}),0,0,\cdots\} \tag{II.14}$$

The sum $\sum_{P(\cdots)}$ is over all permutations of the variables $q_1,\ldots,q_j,\ (-p_{k_1}),\ldots,(-p_{k_{l-j}})$, and $\hat{p}$ means that this vector has to be omitted.

By a closer inspection of this complicated expression the property (12) follows at once. It means, for instance, that all $:\varphi^l:_j(f)$ with $j>0$ destroy the vacuum:

$$:\varphi:_j(f)\Psi_0=0\quad\text{for all } j>0. \tag{II.15}$$

Moreover, for $l\ge 2$ all $:\varphi^l:_j(f)$ with $j>1$ destroy in addition the one-particle state:

$$:\varphi^l:_j(f)\Psi_1=0\quad\text{for } 2\le j\le l,\ l\ge 2 \tag{II.15'}$$

and so on.

The definitions above, of course, give only Lorentz scalar composite fields. Besides these we need, however, tensor fields of arbitrary high rank. They are obtained by forming Wick products of derivatives of the basic fields.

Let α denote a sequence $\{\alpha_0,\alpha_1,\alpha_2,\alpha_3\}$ of four nonnegative integers, $|\alpha|=\sum_{i=0}^{3}\alpha_i$ and

$$D^\alpha=\frac{\partial^{|\alpha|}}{\prod_{j=0}^{3}(\partial x_j)^{\alpha_j}}; \tag{II.16}$$

then $D^\alpha\varphi(f)$ is defined by

$$D^\alpha\varphi(f)=(-1)^{|\alpha|}\varphi(D^\alpha f). \tag{II.17}$$

Now the Wick products of these fields are formed in exactly the same way as in the scalar case

$$:\prod_{i=1}^{l}(D^{\alpha^{(i)}}\varphi):(f)\Psi$$

$$=\sum_{j=0}^{l}:\prod_{i=0}^{l}(D^{\alpha^{(i)}}\varphi):_j(f)\Psi \tag{II.18}$$

and the explicit representation of the constituents is obtained from (II.14) by putting the corresponding polynomials in the momenta q_r and $-p_{k_s}$ in between $\sum_{P(\cdots)}$ and $\tilde{f}$:

$$:D^{\alpha(1)}\varphi\times\cdots\times D^{\alpha(l)}\varphi:_j(f)\Psi_n$$

$$=\Big\{0,\ldots,0,(2\pi)^{-3/2(l-1)}\left(\frac{n!}{(n+l-2j)!}\right)^{1/2}\int\cdots\int$$

$$\prod_{i=1}^{j}d^1\mu(q_j)\sum_{k_1<\cdots<k_{l-j}=1}^{n+l-2j}\frac{1}{j!}\Big[\sum_{P(q_1,\ldots,-p_{k_{l-j}})}$$

$$\times(-iq_1)^{\alpha(1)}\cdot\ \cdots\ \cdot(-iq_j)^{\alpha(j)}\cdot(ip_{k_1})^{\alpha(j+1)}\cdot\cdots$$

$$\cdot(ip_{k_{l-j}})^{\alpha(l)}\tilde{f}\Big(\sum_{s=1}^{j}q_s-\sum_{r=1}^{l-j}p_{k_r}\Big)\Big]$$

$$\times\psi^n(q_1,\ldots,q_j,\ p_1,\ldots,\hat{p}_{k_1},\ldots,p_{k_{l-j}},\ldots,p_{n+l-2j}),$$

$$0,0,\cdots\Big\}. \tag{II.19}$$

The symbol $(k)^\alpha$ stands for the product

$$(k^0)^{\alpha_0}(k^1)^{\alpha_1}(k^2)^{\alpha_2}(k^3)^{\alpha_3}.$$

Obviously these composite fields share all the properties (12), (13), and (15) of the scalar ones. In addition by means of the sum $\sum_{P(\cdots)}$ occurring in (19) they are symmetric with respect to any permutation of the $\alpha^{(1)}\cdots\alpha^{(l)}$. Hence if $\{1_p,2_p,\ldots,l_p\}$ denotes any per-

mutation of the index set $\{1, 2, \ldots, l\}$, we have

$$:\prod_{r=1}^{l}(D^{\alpha^{(r)}}\varphi):_f(f) = :\prod_{r=1}^{l}(D^{\alpha^{(r_p)}}\varphi):_f(f). \tag{II.20}$$

This completes the review of generalized free field theories, and we may attack the problem of conformal covariance.

III. CONFORMAL INVERSION COVARIANCE

Since the Hilbert space is a direct sum of direct products of the one particle subspace $\mathfrak{H}^1$, all unitary representations of symmetry transformations have the same structure

$$U = \bigoplus_{n=0}^{\infty} U^n, \quad U^n = \prod_{j=1}^{n} \otimes U^1, \quad n \geq 1. \tag{III.1}$$

Hence we need only construct unitary representations U^1 in $\mathfrak{H}^1$ and U^0 on the vacuum.

Before we consider conformal covariance let us first list the restrictions induced by scale invariance. The existence of unitary operators

$$\left.\begin{aligned} &U_d(\lambda)\Psi_0 = \Psi_0 \\ &(U_d^1(\lambda)\psi^1)(p_1) = \lambda^d \psi^1(\lambda p_1) \end{aligned}\right\}, \quad \lambda > 0, \ d \in \mathbb{R} \tag{III.2}$$

fixes the measure $d^1\mu(p)$, up to a normalization constant $g(d)$ to be

$$d^1\mu(p) = \begin{cases} g(d)\, d^4p\, \theta_+(p^2)(p^2)^{d-2} & \text{for } d > 1 \\ g(1) d^4p\, \delta_+(p^2) & \text{for } d = 1. \end{cases} \tag{III.3}$$

$d < 1$ is forbidden by positivity of the Hilbert space norm. If we chose the normalization constant $g(d)$ to be

$$\begin{aligned} &g(d) = (2\pi)^2 2^{2(1-d)} \Gamma(d-1)^{-1}\Gamma(d)^{-1} && \text{for } d > 1, \\ &g(1) = (2\pi)^2 && \text{for } d = 1, \end{aligned} \tag{III.4}$$

then the Fourier transform of $d^1\mu(p)$ becomes simply

$$\frac{1}{(2\pi)^3}\int d^1\mu(p) \exp(-ipx) = [-(x - i\epsilon)^2]^{-d}. \tag{III.5}$$

In order to get an idea, how the conformal inversions in $\mathfrak{H}^1$ may look, we transform the scalar product into configuration space:

$$(\Phi^1, \Psi^1) = \int d^4x \int d^4y\, \overline{\tilde{\phi}^1(x)}[-(x - y - i\epsilon)^2]^{-d}\tilde{\psi}^1(y) \tag{III.6}$$

with

$$\tilde{\phi}^1(x) = [1/(2\pi)^{5/2}] \int d^4p \exp(-ipx)\phi^1(p). \tag{III.7}$$

Performing the formal substitution

$$R: x \mapsto x_R = -x/x^2 \tag{III.8}$$

in (III.6) and using the distribution identity

$$\begin{aligned} &[-(x - y \pm i\epsilon)^2]^{-d} \\ &= [-(x_R \pm i\epsilon)^2]^{-d}[-(x_R - y_R \pm i\epsilon)^2]^{d}[-(y_R \mp i\epsilon)^2]^{-d} \end{aligned} \tag{III.9}$$

we readily see that the following mapping $U_d^1(R)$ in $\mathfrak{H}^1$ may be a suitable candidate for a unitary representation in $\mathfrak{H}^1$:

$$(U_d^1(R)\psi^1)(x) = [-(x + i\epsilon)^2]^{d-4}\psi^1(-x/x^2). \tag{III.10}$$

In order to prove this, one transforms this relation back to momentum space by means of

$$\begin{aligned} &\pm (2\pi)^3[-(z_R \pm i\epsilon)^2]^d \exp(ipz_R) \\ &= d^4q\, \theta_\pm(q^2) q^0 (q^2)^{(d-2)/2} \exp(iqz) \iint \frac{d^3y\, dr r}{(\mathbf{y}^2 + r^2)^3} \\ &\quad |p^2|^{-(d-2)/2} \mathcal{J}_{d-2}(\sqrt{|p^2|}r)\, \mathcal{J}_{d-2}\left(\frac{\sqrt{q^2}r}{\mathbf{y}^2 + r^2}\right) \\ &\quad \times \exp\{-i[\mathbf{p}\cdot\mathbf{y} - \mathbf{q}\cdot\mathbf{y}(\mathbf{y}^2 + r^2)^{-1}]\}. \end{aligned} \tag{III.11}$$

Definition III.1: Let $\mathfrak{d}$ be the dense domain in $\mathfrak{H}^1$ of all fast decreasing functions:

$$\mathfrak{d} = \{\Psi^1 \in \mathfrak{H}^1 \cap L^1(d^1\mu(q)(q^2)^{(d-2)/2}) \mid [1 + (q^0)^2]^N \psi^1(q) | < +\infty \text{ for all } N \in \mathbb{N}\}.$$

Now one can prove along the same lines as in Ref. 11 the following theorem for $d > 1$. The canonical case $d = 1$ has already been proved in Ref. 11 (Theorem 1).

Theorem III.1: For all $d > 1$ the mappings $\mathfrak{d} \to \mathfrak{H}^1$ with

$$\begin{aligned} &\psi^1(p) \mapsto (U_d^1(R)\psi^1)(p) \\ &= (2\pi)^{-3} \int d^4q\, \theta_+(q^2) q^0 (q^2)^{(d-2)/2} \psi^4(q) \iint \frac{d^3x\, dr r}{(\mathbf{x}^2 + r^2)^3} \\ &\quad \times |p^2|^{-(d-2)/2} \mathcal{J}_{d-2}(\sqrt{|p^2|}r)\, \mathcal{J}_{d-2}\left(\frac{\sqrt{q^2}r}{\mathbf{x}^2 + r^2}\right) \\ &\quad \times \exp\{-i[\mathbf{px} - \mathbf{qx}(\mathbf{x}^2 + r^2)^{-1}]\} \end{aligned} \tag{III.12}$$

define linear isometric and symmetric operators on $\mathfrak{d}$ satisfying the group relation $U_d^1(R)U_d^1(R) = 1$. Hence they possess unique unitary and self-adjoint extensions to $\mathfrak{H}^1$.

The main steps of the proof are these:

(1) Consider the scalar product $(\Phi_1, U_d^1(R)\Psi_1)$ for ϕ^1, $\psi^1 \in \mathfrak{d}$. By means of the reality properties of the Bessel functions and a variable substitution

$$(\mathbf{x}, r) \mapsto (\mathbf{y}, \rho) = (x, r)[\mathbf{x}^2 + r^2]^{-1}$$

in (III.12) one gets at once the symmetry relations

$$\begin{aligned} &(U_d^1(R)\Phi_1; \Psi_1) = (\Phi_1; U_d^1(R)\Psi_1), \\ &(\Phi_1; U_d^1(R)U_d^1(R)\Psi_1) = (U_d^1(R)\Phi_1;\, U_d^1(R)\Psi_1). \end{aligned} \tag{III.13}$$

Hence this scalar product defines a symmetric sequilinear form on $\mathfrak{d}$.

(2) In the second technically much more involved step, one proves the (weak) group relation for all Φ_1, $\Psi_1 \in \mathfrak{d}$:

$$(\Phi_1; U_d^1(R)U_d^1(R)\Psi_1) = (\Phi_1; \Psi_1). \tag{III.14}$$

The last two relations imply the scalar product to define also an *isometric* sesquilinear form on $\mathfrak{d}$, which in turn implies via Schwartz's inequality its boundedness on $\mathfrak{d}$. This proves the theorem.

The details of the last step consist in fooling around with integrals over Bessel functions and will be omitted.

Having established the existence of a unitary representation in $\mathfrak{H}^1$, it is now straightforward to derive the transformation laws of the basic and composite field operators. Since the consequences are nontrivial, let us start with the basic fields $\varphi_\xi(f)$. We have only to calculate the right-hand side of the following equation:

$$(\Phi; U_d(R)\varphi_\zeta(f)\,U_d(R)\Psi) = (U_d(R)\Phi;\ \varphi_\zeta(f)U_d(R)\Psi)$$

for all $\Phi \in \mathcal{D}$ and $\Psi \in \mathcal{D}_\varphi$.

This can immediately be done by means of (II. 7), (II. 8), (III. 11), and (III. 12) using the unitarity and self-adjointness of $U_d(R)$.

The result is

$$U_d(R)\varphi_\zeta(f)U_d(R)\Psi = \varphi_\zeta(V_\zeta^d f)\Psi \tag{III. 15}$$

with

$$(V_\zeta^d f)(x) = \begin{cases} [-(x+i\epsilon)^2]^{d-4} f(-x/x^2) & \text{for } \zeta = 0 \\ [-(x-i\epsilon)^2]^{d-4} f(-x/x^2) & \text{for } \zeta = 1. \end{cases} \tag{III. 16}$$

Again as in the free field case[11] the positive and negative frequency parts transform differently, i.e., there is no local transformation law for the local field $\varphi(f)$ itself. Moreover, since the right-hand side of (III. 16) is not anymore an element of the Schwartz space S_4, from which we started in Chap. II, the conformally covariant fields cannot be tempered. For them we have to reformulate the whole theory, i.e., construct a new test function space, which is invariant under the transformation (III. 16) and Poincaré transformations. Obviously this new space S_4^d cannot share the polynomial decrease at infinity of S_4. However, since we want to construct Lorentz tensors of arbitrary high rank by means of derivatives, it must share the C^∞ properties in S_4. Moreover, it must be a nuclear space in order to be able to construct the Wick products. In one dimension such test function spaces exist depending on the scale parameter d.[19] In higher dimensions, to our knowledge, this problem is yet unsolved. The construction of such spaces goes, of course, beyond the scope of this paper. We shall simply assume the existence of such spaces, and will whenever we restrict ourselves to conformally covariant theories replace S_4 by S_4^d.

Finally we want to mention that the different transformation laws for φ_0 and φ_1 are closely connected to their support properties in momentum space; φ_0 having support only in the closed forward cone $\overline{V}_+$ and φ_1 in the backward cone $\overline{V}_-$. Hence we expect an even more complicated transformation law for our composite fields $:\varphi^l:_j$ since they have also support in spacelike regions of momentum space [see Eq. (II. 14].

Indeed, if we perform the same calculation as above for $:\varphi^l:_j$, we find

$$U_d(R):\varphi^l:_j(f)U_d(R)\Psi = :\varphi^l:_j(V_j^{dl}f)\Psi \tag{III. 17}$$

with

$$(V_j^{dl}f)(x) = [-(x+i\epsilon)^2]^{dl-4+dj}[-(x-i\epsilon)^2]^{dj} f(-x/x^2) \tag{III. 18}$$

and

$$[-(x+i\epsilon)^2]^{dl-4-dj}[-(x-i\epsilon)^2]^{dj} = |x^2|^{dl-4}\begin{cases} 1 & \text{for } x^2 \le 0 \\ \exp[-i\pi d(l-2j)] & \text{for } x^2 > 0 \wedge x^0 > 0 \\ \exp[i\pi d(l-2j)] & \text{for } x^2 > 0 \wedge x^0 < 0. \end{cases} \tag{III. 19}$$

For d not an integer the components $:\varphi^l:_j$ of the Wick products $:\varphi^l:$ transform all differently by a phase factor $\exp\{-i2\pi dj\epsilon(x^0)\theta(x^2)\}$.

This is exactly the transformation behavior suggested by Schroer and Swieca on group theoretical arguments and verified in two-dimensional models.[12] Since the distribution of the powers $(+dj)$ and $(-dj)$ in (18) and (19) is unique only mod 1, (III. 16) is just a special case of (III. 18).

Our final task is to construct traceless Lorentz covariant higher rank tensors, which will bring us new problems. Let us start with the Lorentz vector: $:\partial^\mu\varphi\times\varphi^{l-1}:_j(f)$. Performing the by now familiar calculations, we find

$$\begin{aligned} U_d(R):\partial^\mu\varphi\times\varphi^{l-1}:_j(f)U_d(R)\Psi &= :\partial^\nu\varphi\times\varphi^{l-1}:_j(r_\nu^\mu\cdot(V_j^{dl}f))\Psi \\ &\quad + 2d:\varphi^l:_j(X^\mu\cdot(V_j^{dl}f))\Psi, \end{aligned} \tag{III. 20}$$

where X^μ and $r^{\mu\nu}$ are the multiplication operators in S_4^d by x^μ respectively by

$$r^{\mu\nu}(x) = 2x^\mu x^\nu - x^2 g^{\mu\nu}. \tag{III. 21}$$

Since $\partial^\mu:\varphi^l:_j(f) = l:\partial^\mu\varphi\times\varphi^{l-1}:_j(f)$, it follows from the symmetry (II. 20) that there exists no other Lorentz vector by means of which one could remove the second term in (III. 20).

The situation improves, when we go over to traceless tensors of higher rank. First making the most general ansatz for a traceless second rank tensor,

$$\begin{aligned} T^{\mu\nu}_{dl,j}(f) &= a_1:\partial^\mu\partial^\nu\varphi\times\varphi^{l-1}:_j(f) \\ &\quad + a_2:\partial^\mu\varphi\partial^\nu\varphi\times\varphi^{l-2}:_j(f) + a_3:\varphi^l:_j(\partial^\mu\partial^\nu f) \\ &\quad - \text{Traces}, \end{aligned}$$

and calculating the transformation under $U_d(R)$, we find

Theorem III. 2: There exists one and only one symmetric, traceless tensor of second rank, which transforms convariantly under Poincaré transformations and conformal inversions, namely $(l = 2, 3, 4, \cdots)$

$$\begin{aligned} C^{\mu_1\mu_2}_{dl+2,j}(f) &= 4\Gamma(d+\tfrac{1}{2})\Gamma(d-\tfrac{1}{2})^{-1}\{(d-\tfrac{1}{2}) \\ &\quad\times[:\varphi\overleftrightarrow{\partial}_-^{\mu_1}\overleftrightarrow{\partial}_-^{\mu_2}\varphi\times\varphi^{l-2}:_j(f) - \tfrac{1}{4}g^{\mu_1\mu_2}:\varphi\overleftrightarrow{\partial}_-^{\alpha}\overleftrightarrow{\partial}_{-\alpha}\varphi\times\varphi^{l-2}:_j(f)] \\ &\quad - \tfrac{1}{2}[\varphi\overleftrightarrow{\partial}_+^{\mu_1}\overleftrightarrow{\partial}_+^{\mu_2}\varphi\times\varphi^{l-2}:_j(f) - \tfrac{1}{4}g^{\mu_1\mu_2}:\varphi\overleftrightarrow{\partial}_+^{\alpha}\overleftrightarrow{\partial}_{+\alpha}\varphi\times\varphi^{l-2}:_j(f)]\}. \end{aligned} \tag{III. 22}$$

Here the $\overleftrightarrow{\partial}_\pm$ are defined by:

$$\overleftrightarrow{\partial}_\pm^\mu = \frac{\overrightarrow{\partial}}{\partial x_\mu} \pm \frac{\overleftarrow{\partial}}{\partial x_\mu}. \tag{III. 23}$$

The expression (III. 22) gives us a hint how to construct the higher rank tensors. For the coefficients occurring in (III. 22) are just those of the well known Gegenbauer polynomial of degree 2. Hence, by copying the coefficients from the Gegenbauer polynomials,[20] the correct combinations seem to be

$$\begin{aligned} &C^{\mu_1\cdots\mu_{2n}}_{dl+2n;j}(f) \\ &\quad = :\varphi G^{\mu_1\cdots\mu_{2n}}_{d-1/2}(\overleftrightarrow{\partial}_-;\overleftrightarrow{\partial}_+)\varphi\times\varphi^{l-2}:_j(f), \end{aligned} \tag{III. 24}$$

where $G^{\cdots}_{\cdots}(\overleftrightarrow{\partial}_-;\overleftrightarrow{\partial}_+)$ are the following symmetric, traceless differential monomials of degree $2n$, $n = 1, \frac{3}{2}, 2, \cdots, 2, \cdots,$

$$G^{\mu_1\cdots\mu_{2n}}_{d-1/2}(\overleftrightarrow{\partial}_-;\overleftrightarrow{\partial}_+)$$

$$= \sum_{m=0}^{2n} \frac{(-1)^m \Gamma(d-\frac{1}{2}+2n-m)}{m!\,(2n-2m)!\,\Gamma(d-\frac{1}{2})} 2^{2n-2m}$$

$$\times \Big\{ \mathop{\$}_{\{\mu_r\}} \prod_{j=1}^{2n-2m} \overleftarrow{\partial}_-^{\mu_j} \prod_{s=1}^{2m} \overrightarrow{\partial}_+^{\mu_{2n-2m+s}} - \text{Traces} . \qquad \text{(III. 25)}$$

The symbol $\mathop{\$}_{\{\mu_r\}}$ again means symmetrization with respect to all indices:

$$\mathop{\$}_{\{\mu_r\}} = (2n)!^{-1} \sum_{P(\mu_1 \cdots \mu_{2n})} .$$

In other words the monomials $G^{\cdots}_{\cdots}(\overrightarrow{\partial}_-, \overrightarrow{\partial}_+)$ are constructed in such a way that for any lightlike vector ξ ($\xi^2 = 0$) and two arbitrary Minkowski vectors x, y we have

$$\xi_{\mu_1} \xi_{\mu_2} \cdots \xi_{\mu_{2n}} G^{\mu_1 \cdots \mu_{2n}}_{d-1/2}(x, y)$$

$$= (\xi y)^{2n} C_{2n}^{d-1/2}\left(\frac{(\xi x)}{(\xi y)}\right)$$

with $C_{2n}^{d-1/2}(z)$ being the Gegenbauer polynomial of degree $2n$.

Explicitly we obtain from (II. 19)

$$C^{\mu_1 \cdots \mu_{2n}}_{ld+2n;\,j}(f)\Psi_m = \{0, \ldots, 0, (2\pi)^{-3(l-1)^2} \left(\frac{m!}{(m+2-2j)!}\right)^{1/2}$$

$$\times \int \cdots \int \prod_{i=1}^{j} d^1\mu(q_j) \sum_{k_1 < \cdots < k_{l-j}=1}^{m+l-2j} \frac{1}{j!}$$

$$\times \Bigg[\sum_{P(q_1, \ldots, p_{k_{l-j}})} (-1)^m G^{\mu_1 \cdots \mu_{2m}}_{d-1/2}(\kappa_-, \kappa_+)$$

$$\times f\left(\sum_{s=1}^{j} q_j - \sum_{r=1}^{l-j} p_{k_r}\right)\Bigg]$$

$$\times \psi^n(q_1, \ldots, q_j, p_1, \ldots, \hat{p}_{k_1}, \ldots, \hat{p}_{k_{l-j}}, \ldots p_{n+l-2j}), 0, 0 \cdots\} \qquad \text{(III. 26)}$$

with κ given by

$$\kappa_\pm = \begin{cases} q_1 \pm q_2 & \text{for } j \geq 2 \\ q_1 \mp p_{k_1} & \text{for } j = 1 \\ p_{k_1} \pm p_{k_2} & \text{for } j = 0. \end{cases} \qquad \text{(III. 27)}$$

One can convince oneself by complete induction using the recurrence relations for the Gegenbauer polynomials that these operators are covariant under conformal inversions. A final property of these composite operators may be obtained by a simple calculation from (III. 26) and the representation (III. 2) for scale transformations, namely that the scale dimension of them is $dl + 2n$.

Collecting the results we have

Theorem III. 3: The symmetric, traceless tensors $C^{\mu_1 \cdots \mu_{2n}}_{dl+2n;\,j}(f)$ are of scale dimension $dl + 2n$ and transform under conformal inversions covariantly according to

$$U_d(R) C^{\mu_1 \cdots \mu_{2n}}_{dl+2n;\,j}(f) U_d(R)\Psi$$

$$= C^{\nu_1 \cdots \nu_{2n}}_{dl+2n;\,j}\left(\prod_{s=1}^{2n} r^{\mu_s \nu_s}(V_j^{dl} f)\right)\Psi \qquad \text{(III. 28)}$$

with $\Psi \in \mathcal{D}_\varphi$ and $r^{\mu\nu}(x) = 2x^\mu x^\nu - x^2 g^{\mu\nu}$, $n = 1, \frac{3}{2}, 2, \cdots$.

IV. OPERATOR PRODUCT EXPANSIONS

For the major part of this section we can forget everything we have derived about conformal covariant fields in Sec. III except the definition of the composite operators $C^{\mu_1 \cdots \mu_{2n}}_{2\lambda+2n;\,j}(f)$, i.e., the formulas (III. 22)–(III. 26). They are, of course, independent of conformal invariance and could also be obtained by pure intuition. By means of these operators we are going to prove operator product expansions for free fields of arbitrary mass $m \geq 0$. Hence conformal invariance will serve *only as a bookkeeping* for the composite operators to be used. Only at the very end we come back to proper conformal invariant theories. The global operator expansions rest on the idea that there may exist a complete set of so called composite operators $\{C^{\mu_1 \cdots \mu_n}_{d_n, x_n}(f)\}$, characterized by their Lorentz-tensor structure, their scale dimension, and some other degeneracies, such that at least the matrix elements of the product of two basic fields may be developed into a series of the same matrix elements of the composite fields. Hence in our case we expect (at least formally) something like

$$(\Phi; \varphi(g)\varphi(f)\Psi) = \sum_{r,s=0}^{1} (\Phi; \varphi_r(g)\varphi_s(f)\Psi)$$

$$= (\Psi_0; \varphi_1(g)\varphi_0(f)\Psi_0)(\Phi; \Psi) + \sum_{m=0}^{\infty} \sum_{l=0}^{\infty} \sum_{r,s=0}^{1} \sum_{j=0}^{l} \int d^4z \int d^4x \int d^4y$$

$$\times g(x) f(y) K^{[l\lambda+2m;\,j]r,s}_{\mu_1 \cdots \mu_{2m}}(x - z, y - z)$$

$$\times (\Phi; C^{\mu_1 \cdots \mu_{2m}}_{l\lambda+2m;\,j}(z)\Psi). \qquad \text{(IV. 1)}$$

Due to the Hilbert space structure (direct sum of direct products of the one particle space $\mathfrak{H}^1$) all matrix elements of the first line may be reduced to the following two:

$$(\Phi_2; \varphi(g)\varphi(f)\Phi_0) = (\Phi_2; \varphi_0(g)\varphi_0(f)\Psi_0) ,$$

or $\qquad \text{(IV. 2)}$

$$(\Phi_1; \varphi(g)\varphi(f)\Psi_1) = (\Phi_1; \{\varphi_0(g)\varphi_1(f) + \varphi_1(g)\varphi_0(f)\}\Psi_1).$$

Hence we need to consider (IV. 1) only for these special matrix elements, which consideration also brings considerable simplifications on the right-hand side.

From the condition (II. 12)

$$\mathfrak{H}^n C^{\mu_1 \cdots \mu_{2m}}_{l\lambda+2m;\,j}(f) \begin{cases} \mathfrak{H}^{n+l-2j} & \text{for } j \leq \min\{(l+n)/2, n\} \\ 0 & \text{for } j > \min\{(l+n)/2, n\} \end{cases}$$

we deduce at once that, in the first case, the so-called "vacuum expansion," only the operators with

$$l = 2 \quad \text{and} \quad j = 0 \qquad \text{(IV. 3)}$$

and, in the second case, the so-called "off vacuum" case, only terms with

$$l = 2 \quad \text{and} \quad j = 1 \qquad \text{(IV. 4)}$$

contribute to the sum on the right-hand side of (IV. 1):

$$(\Phi; \varphi(g)\varphi(f)\Psi) = (\Psi_0; \varphi_1(g)\varphi_0(f)\Psi_0)(\Phi, \Psi)$$

$$+ \sum_{m=0}^{\infty} \sum_{r,s=0}^{1} \sum_{j=0}^{1} \int d^4z \int d^4x\, d^4y\, g(x) f(y)$$

$$\times K^{[2\lambda+2m;\,j]r,s}_{\mu_1 \cdots \mu_{2m}}(x - z; y - z)(\Phi; C^{\mu_1 \cdots \mu_{2m}}_{2\lambda+2m;\,j}(z)\Psi). \qquad \text{(IV. 5)}$$

Our task now consists of constructing a set of kernels

$$K^{\cdots}_{\cdots}[g,f](z)=\int d^4x\,d^4y\,g(x)f(y)K^{\cdots}_{\cdots}(x-z,\ y-z) \qquad \text{(IV.6)}$$

such that the z integration can be given a precise mathematical meaning and the sum converges to the left-hand side for the two special matrix elements (IV.2) with Ψ^1, Φ^1 and $\Phi^2\in\mathcal{D}_\varphi$.

Introducing the Fourier transform

$$\hat{K}^{\cdots}_{\cdots}(k,z)=\frac{1}{(2\pi)^{5/2}}\int d^4u\,\exp(iku)K^{\cdots}_{\cdots}(u+z,\ u-z) \qquad \text{(IV.7)}$$

and calculating all matrix elements explicitly by means of (II.7), (II.8), and (III.26), we end up with the following two equations:

(1) "Vacuum expansion":

$$\begin{aligned}
&d^1\mu(p_1)\int d^1\mu(p_2)\,\overline{\phi^2(p_1,p_2)}\,\tilde{g}(-p_1)\tilde{f}(-p_2)\\
&=(2\pi)^{-3/2}\lim_{N\to\infty}\sum_{n=0}^{N}(-1)^n\int d^1\mu(p_1)d^1\mu(p_0)\,\overline{\phi^2(p_1,p_2)}\\
&\quad\times\int d^4q\,\tilde{g}\left(-\frac{p_1+p_2-q}{2}\right)\tilde{f}\left(-\frac{p_1+p_2+q}{2}\right)\\
&\quad\times\int d^4x\,\exp(iqx)\,G^{\mu_1\cdots\mu_{2n}}_{\lambda-1/2}(p_1-p_2,\,p_1+p_2)\\
&\quad\times\hat{K}^{[2\lambda+2n,0]0,0}_{\mu_1\cdots\mu_{2n}}(-p_1+p_2),x).
\end{aligned} \qquad \text{(IV.8)}$$

(2) "Off vacuum expansion":

$$\begin{aligned}
&\int d^1\mu(p_1)\,d^1_\mu(p_2)\,\overline{\phi^1(p_1)}\,\{\tilde{g}(-p_1)\tilde{f}(p_2)+\tilde{g}(p_2)\tilde{f}(-p_1)\}\,\psi^1(p_2)\\
&=(2\pi)^{-3/2}\lim_{N\to\infty}\sum_{n=0}^{N}(-1)^n\int d^1\mu(p_1)\int d^1\mu(p_2)\,\overline{\phi^1(p_1)}\,\psi^1(p_2)\\
&\quad\times\int d^4q\,\tilde{g}\left(-\frac{p_1-p_2-q}{2}\right)\tilde{f}\left(-\frac{p_1-p_2+q}{2}\right)\\
&\quad\times\int d^4x\,\exp(iqx)\,G^{\mu_1\cdots\mu_{2n}}_{\lambda-1/2}(p_1+p_2,\ p_1-p_2)\\
&\quad\times\hat{K}^{[2\lambda+2n,1]+}_{\mu_1\cdots\mu_{2n}}(-(p_1-p_2),x)
\end{aligned} \qquad \text{(IV.9)}$$

with

$$\hat{K}^{[\cdots]0,1}_{\cdots}(k,x)+\hat{K}^{[\cdots]0,1}_{\cdots}(k,x)+\hat{K}^{[\cdots]1,0}_{\cdots}(k,x).$$

Here $G^{\mu_1\cdots\mu_{2n}}_{\lambda-1/2}$ are the polynomials of Gegenbauer type introduced at the end of the last section.

Now the only difference between the right-hand sides of (IV.8) and (IV.9) apart from the wavefunction is the interchange

$$p_+=p_1+p_2 \Longleftrightarrow p_-=p_1-p_2. \qquad \text{(IV.10)}$$

We first solve the less problematic case of the "vacuum expansion" and afterwards look for the changes brought about by this replacement for the "off vacuum expansion."

A. "Vacuum expansion"

The only thing we have to do, is to construct an infinite set of Kernels $\hat{K}^{\cdots}_{\cdots}(k,x)$ such that

$$\begin{aligned}
&(2\pi)^{5/2}\sum_{n=0}^{\infty}(-1)^nG^{\mu_1\cdots\mu_{2n}}_{\lambda-1/2}(p_-;p_+)\hat{K}^{[2\lambda+2n,0]0,0}_{\mu_1\cdots\mu_{2n}}(p_+;x)\\
&=\cos(p_-\circ x),\quad \text{for } p_1\in\overline{V_+},\ p_2\in\overline{V_+},\ x\in\mathbb{M}_4
\end{aligned} \qquad \text{(IV.11)}$$

holds in the sense of tempered distributions.

Now for any Lorentz invariant function $g(p_+;x)$ with the property

$$(p_-\circ x)^2\le g_0(p_+,x)\ \text{ for all } p_1,p_2\in\overline{V_+}\wedge x\in\mathbb{M}_4 \qquad \text{(IV.12)}$$

$\cos(p_-x)$ may be developed into a uniformly convergent series of Gegenbauer polynomials[21,22] in the following way:

$$\begin{aligned}
\cos(p_-x)&=\cos\left(\sqrt{g_0(p_+,x)}\circ\frac{(p_-x)}{\sqrt{g_0(p_+,x)}}\right)\\
&=\frac{\Gamma(\lambda-1/2)2^{\lambda-1/2}}{\sqrt{g_0(p_+,x)}^{\lambda-1/2}}\sum_{n=0}^{\infty}(-1)^n(2n+\lambda-\tfrac{1}{2})\\
&\quad\times\mathcal{J}_{\lambda-1/2+2n}(\sqrt{g_0(p_+,x)})\,C^{\lambda-1/2}_{2n}\left(\frac{p_-x}{\sqrt{g_0(p_+x)}}\right),\quad \lambda>-\tfrac{1}{2}.
\end{aligned} \qquad \text{(IV.13)}$$

Since on the left-hand side of (IV.11) there occur polynomials $G^{\cdots}_{\lambda-1/2}$ closely related to the Gegenbauer polynomials, we may try to reduce our problem to finding solutions of the equations

$$\begin{aligned}
&\{(2\pi)^{5/2}G^{\mu_1\cdots\mu_{2n}}_{\lambda-1/2}(p_-,p_+)\hat{K}^{[2\lambda+2n,0]0,0}_{\mu_1\cdots\mu_{2n}}(-p_+,x)\}\\
&=\frac{\Gamma(\lambda-\frac{1}{2})2^{\lambda-1/2}(2n+\lambda-\frac{1}{2})}{\sqrt{g_0(p_+,x)}^{\lambda-1/2}}\,\mathcal{J}_{2n+\lambda-1/2}(\sqrt{g_0(p_+,x)})\\
&\quad\times C^{\lambda-1/2}_{2n}\left(\frac{p_-\circ x}{\sqrt{g_0(p_+,x)}}\right)\ \text{for all } n\in\mathbf{N}.
\end{aligned} \qquad \text{(IV.14)}$$

This relation suggests the ansatz

$$\begin{aligned}
\hat{K}^{[2\lambda+2n,0]0,0}_{\mu_1\cdots\mu_{2n}}(-p_+,x)&=\Gamma(\lambda-\tfrac{1}{2})(2n+\lambda-\tfrac{1}{2})2^{\lambda-1/2}(2\pi)^{-5/2}\\
&\quad\times(\sqrt{g_0(p_+,x)})^{-\lambda+1/2-2n}\,\mathcal{J}_{\lambda-1/2+2n}(\sqrt{g_0(p_+,x)})\\
&\quad\times\mathfrak{P}^{\lambda+2n}_{\mu_1\cdots\mu_{2n}}(-p_+,x)
\end{aligned} \qquad \text{(IV.15)}$$

with $\mathfrak{P}^{\cdots}_{\cdots}(-p_+,x)$ an arbitrary traceless, symmetric tensor of the form:

$$\begin{aligned}
\mathfrak{P}^{\lambda+2n}_{\mu_1\cdots\mu_{2n}}(-p_+,x)&=x_{\mu_1}\circ\cdots x_{\mu_{2n}}\\
&\quad+\sum_{t=1}^{2n}a^{\lambda,n}_t(p_+,x)\left\{\sum_{\{\mu_1\}}\prod_{j=1}^{2n-t}x_{\mu_j}\prod_{s=1}^{t}(p_+)_{\mu_3}\ -\ \text{Traces}\right\}.
\end{aligned} \qquad \text{(IV.16)}$$

Inserting this ansatz into (IV.14) we end up via (III.25) with one linear relation between the $2n$ coefficients $a^{\lambda,n}_t$

$$\begin{aligned}
&f^{\lambda,n}_{2n}(p_+,p_-)a^{\lambda,n}_{2n}(p_+,x)\\
&\quad+\sum_{i=1}^{2n-1}f^{\lambda,n}_i(p_+,x,p_-)a^{\lambda,n}_i(p_+,x)=f^{\lambda,n}_0(p_+,x,p_-;q_0),
\end{aligned} \qquad \text{(IV.17)}$$

where the $f^{\lambda,n}_i$ are given Lorentz invariant polynomials, homogeneous of degree $2n$.

This relation would admit at least one solution if the coefficients would not depend on p_+^2 and $(p_+\circ p_-)$ as in the case of a free field of mass $m\ge 0$, in which these two variables may be reexpressed in terms of p_+^2 and m^2. In the general case, however, the situation is different.

For instance, if $n=1$, then (IV. 17) reads

$$a_2^{\lambda,1}(p_+,x)\{p_+^2[(\lambda+\tfrac{1}{2})p_-^2+\tfrac{3}{2}p_+^2]-4(\lambda+\tfrac{1}{2})(p_-\cdot p_+)^2\}$$
$$+2a_1^{\lambda,1}(p_+,x)\{(xp_+)[(\lambda+\tfrac{1}{2})p_-^2+\tfrac{3}{2}p_+^2]$$
$$-4(\lambda+\tfrac{1}{2})(xp_-)(p_-p_+)\}$$
$$=2g_0(p_+,x)-2(p_+x)^2+\tfrac{1}{2}x^2p_+^2-(\lambda+\tfrac{1}{2})x^2p_-^2. \qquad \text{(IV. 18)}$$

Now if $d^1\mu(p)$ has a continuous part, there is obviously no solution for the three functions $a_i^{\lambda,1}(p_+,x)$ $(\lambda>-\frac{1}{2})$ and $g_0(p_+,x)$ independent of p_-.

This, however, does not mean that there does not exist a vacuum expansion for our system may not yet be a complete set of operators.[23] The situation improves rapidly if the measure is concentrated on one or more mass hyperbolids. Take, for instance, the case of one mass $m\geq 0$. Then with

$$g_0(p_+,x)=(p_+x)^2-x^2p_+^2\geq(p_-x)^2 \qquad \text{(IV. 19)}$$

Eq. (IV. 18) becomes

$$\left\{a_2^{\lambda,1}(p_+,x)+\frac{x^2}{p_+^2}+2\frac{(xp_+)}{p_+^2}a_1^{\lambda,2}(p_+,x)\right\}$$
$$\times\{(\lambda-1)p_+^2-4m^2(\lambda+\tfrac{1}{2})\}=0 \qquad \text{(IV. 20)}$$

with the most general solution

$$a_2^{\lambda,1}(p_+,x)=\begin{cases}-\dfrac{x^2}{p_+^2}-2\dfrac{(xp_+)}{p_+^2}a_1^{\lambda,1}(p_+,x)\\ \qquad\text{unless } d=1\wedge m=0,\\ \text{arbitrary for } d=1\wedge m=0.\end{cases} \qquad \text{(IV. 21)}$$

The situation remains similar for all higher $n\in\mathbb{N}$ with the result that Eq. (IV. 17) can always be solved for $a_{2n}^{\lambda,n}(p_+,x)$ in terms of the remaining $a_t^{\lambda,n}(p_+,x)$ $(t=1,\ldots,2n-1)$ due to the fact that apart from an overall factor the coefficient of $a_{2n}^{\lambda,n}$ is proportional to $(p_+^2)^{2n}$. A particular simple solution is

$$a_{2s-1}^{\lambda,n}(p_+,x)=0;\ a_{2s}^{\lambda,n}(p_+,x)=\frac{(-1)^s}{2s-1}\left(\frac{x^2}{p_+^2}\right)^s, \qquad \text{(IV. 22)}$$

$s=1,2,\ldots,n,\ n\in\mathbb{N}$.

With our solutions above every term in (IV. 11) is a polynomial bounded continuous function in p_1, p_2, and x. Moreover, since the inequality (IV. 19) holds for all $p_1, p_2\in V_+^m$ and all $x\in\mathbb{M}_4$ the series (IV. 13) respectively (IV. 14) converges uniformly and therefore also in the sense of tempered distributions. Thus we have gained our first result:

Theorem IV. 1 (*"Vacuum expansion"*): Let $\varphi(f)$ be a free scalar field of mass $m\geq 0$. Then for any set of composite field operators

$$\{C_{2\lambda+2n;j}^{\mu_1\cdots\mu_{2n}}(f)\,|\,n\in\mathbb{N}\wedge j=0,1,2\}\ \text{with}\ \lambda>-\tfrac{1}{2}$$

defined in (III. 24)—(III. 26) and any set

$$\{a_t^{\lambda,n}(p_+,x)\,|\,n\in\mathbb{N}\wedge t=1,2,\ldots,2n-1\}$$

of Lorentz invariant, polynomial bounded functions there exists a set of kernels $\{K_{\mu_1\cdots\mu_{2n}}^{[2\lambda+2n,0]0,0}[g,f](z)\,|n\in\mathbb{N}\}$ such that the operator product expansion for $(\Phi,\varphi(g)\varphi(f)\Psi_0)$ converges for all $\Phi\in\mathcal{D}_\varphi$.

B. "Off vacuum expansion"

As already mentioned, the "off vacuum expansion" may be obtained from the "vacuum expansion" (at least formally) by the replacement $p_-\Longleftrightarrow p_+$ in (IV. 11)—(IV. 22). However, the new problems are brought about by the fact that the inequalitites (IV. 12), respectively (IV. 19), break down after the above exchange:

$$(p_+x)\not\geq g_0(p_-,x)$$

for $\qquad \text{(IV. 23)}$

$$g_0(p_-,x)=(p_-x)^2-x^2p_-^2.$$

$g_0(p_-,x)$ can even become negative, i. e., the square root purely imaginary. Hence we have to reinvestigate the convergence of the Neumann series (IV. 13).

Let us introduce the complex variables:

$$\eta=\sqrt{g_0(p_-,x)},\quad z=(p_+\cdot x)/\sqrt{g_0(p_-,x)}. \qquad \text{(IV. 24)}$$

In the center of mass frame of p_+z reads

$$z=x^0/\sqrt{(x^0)^2-\mathbf{x}^2\sin^2\theta},\quad \cos\theta=\mathbf{x}\mathbf{p}_-/|\mathbf{x}|\,|\mathbf{p}_-|. \qquad \text{(IV. 25)}$$

Now the region in the complex plane, for which we need the expansion (IV. 13), is easily obtained and shown in Fig. 1. Note that $g_0(p_-,x)<0$ implies $x^2<0$.

Since for every fixed complex η the function $\cos\eta\cdot z$ is an entire function of z, a theorem of Szegö[22] saves us under certain restrictions on the test functions g and f. It states that if $F(z)$ is analytic on a closed segment $[-1,1]$ of the real axis, then the expansion of $F(z)$ in a Jacobi (Gegenbauer) series converges in the interior of the largest ellipse with foci at ± 1, in which $F(z)$ is regular. The expansion diverges outside this ellipse.

Hence the series (IV. 13) with $p_-\Longleftrightarrow p_+$ and therefore also the analogon of (IV. 11) converges in any bounded region in x of the Minkowski space $\mathbb{M}_4$. This brings about the restriction in f, g to the subspace D_4 (in configuration space) of functions with compact support of S_4 in (IV. 9). Besides this there is another necessary restriction in the wavefunctions $\phi^1(p_1)$ and $\psi^1(p_2)$.

For $g_0(p_-,x)<0$ the kernels $\hat{K}_{\cdots}^{\cdots}(-p_-,x)$ in (IV. 15) contain in contrast to all other cases modified instead of ordinary Bessel functions. The modified ones grow exponentially, i. e.,

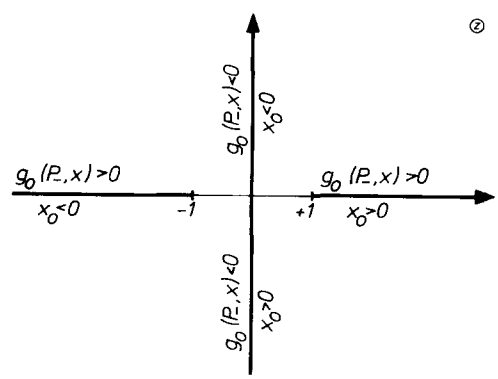


FIG. 1. Domain for the "off vacuum expansion."

$$\hat{K}^{\cdots}_{\cdots}(-p_-,x) \sim \exp\{|\,|(p_-x)^2 - p_-^2x^2\,|^{1/2}\}, \qquad \text{(IV. 26)}$$

which in the rest frame of p_+ becomes

$$\hat{K}^{\cdots}_{\cdots}(-p_-,x) \sim \exp\{|\mathbf{p}_-|\,(|\mathbf{x}^2\cos^2\theta + x^2|)^{1/2}\}.$$

Hence, in order that the individual terms of the "off vacuum expansion" (IV. 9) exist at all, also the wavefunction have to be restricted to the dense subspace D_3 of C^∞ functions with compact support.

However, with these two necessary restrictions, which are deadly for the calculation of the kernels by termwise conformally covariance, all the results from the "vacuum expansion" can be taken over with $p_- \Longleftrightarrow p_+$.

Theorem IV. 2 ("Off Vacuum Expansion"): Let $\varphi(f)$ be a free scalar field of mass $m \geq 0$. Then for any set of composite field operators $\{C^{\mu_1\cdots\mu_{2n}}_{2\lambda+2n;j}(f)\,|\,n\in\mathbb{N}\wedge j=0,1,2\}$ with $\lambda > -\frac{1}{2}$ defined in (III. 24)–(III. 26) and any set $\{a^{\lambda,n}_t(p_-,x)\,|\,n\in\mathbb{N}\wedge t=1,\ldots,2n-1\}$ of Lorentz invariant, polynomial bounded functions there exists for all f, $g\in D_4$ (in configuration space) a set of kernels $\{K^{[2\lambda+2n,0]+}_{\mu_1,\ldots,\mu_{2n}}[g,f](z)\,|\,n\in\mathbb{N}\}$ such that the operator product expansion for $(\Phi, \varphi(g)\varphi(f)\Psi)$ with f, $g\in D_4$ converges for all Φ, $\Psi\in\mathcal{D}_\varphi$ with compact support.

Finally we want to make some remarks concerning conformal (inversion) covariant theories, for which the kernels could at least formally be computed up to a phase by termwise conformal (inversion) covariance of the series.[5]

This covariance first fixes m and λ to be $m=0$ and $\lambda=d=1$. Moreover, for the "vacuum expansion" all functions $a^{\lambda,n}_t(p_+,x)$ are uniquely fixed to be the special solution given in (IV. 22). This is obvious, since all three-point functions are uniquely fixed by conformal (inversion) invariance. Hence conformal (inversion) covariance singles exactly one out from our set given in Theorem IV. 1.

However, the "off vacuum expansion" must disagree, since termwise conformal (inversion) covariance leads to tempered or even better behaved kernels (decreasing exponentially for $x^2<0$), whereas the proof above leads to nontempered kernels, which grow exponentially in certain spacelike directions.

There rises the question whether there exists another set of conformal (inversion) covariant composite operators or other solutions with a different $g_0(p_-,x)$ of (IV. 17)? According to Theorem III. 2 there is one and only one (inversion) covariant second-rank tensor $C^{\mu\nu}_{4;j}(f)$. Moreover, for $d=1$ and $m=0$, there is one and only one solution of Eq. (IV. 18) for $g_0(p_-,x)$, namely

$$g_0(p_-,x) = (p_-x)^2 - p_-^2x^2.$$

Hence, for the "off vacuum expansion," conformal (inversion) covariance can at most serve for finding a complete set of composite operators. For the calculation of the kernels it is completely useless.

A final remark should be made on the consequences of the restrictions in the momentum space wave functions in Theorem IV. 2. According to these restrictions—caused by the asymptotic behavior (IV. 26) of the kernels $\hat{K}^{\cdots}_{\cdots}(-p_-,x)$—no localizable states in configurations space are admitted for the "off vacuum expansion." This in turn kills all duality programs in quantum field theory[24] since locality cannot be applied anymore. Localizable states in configuration space require wavefunctions in momentum space of the type characterized by Jaffe.[25] However, for such wavefunctions our kernels do not exist.

ACKNOWLEDGMENTS

The material of this paper emerged from a collaboration with B. Schroer and J.A. Swieca. I am grateful to both and also to B. Vaziri for many helpful discussions.

*The major part of this work was done during a stay at the Institut für Theoretische Physik, F.U. Berlin. I thank the members of this institute for their hospitality.

[1]S. Ferrara, R. Gatto, and A. Grillo, Phys. Rev. D 5, 3102 (1972); Ann. Phys. (N.Y.) 76, 161 (1973).
[2]L. Bonora, Nuovo Cimento Lett. 3, 548 (1972).
[3]A. A. Migdal and A. M. Polyakov (unpublished).
[4]G. Mack, "Group Theoretical Approach to Conformal Invariant Quantum Field Theory," preprint, Bern (1973).
[5]J. A. Swieca, "Conformal Operator Expansions in the Minkowski Region," preprint Rio de Janeiro (1973).
[6]B. Schroer, J.A. Swieca, and A.H. Völkel, Phys. Rev. D 11, 1509 (1975).
[7]V.K. Dobrev, V.B. Petcova, S.G. Petrova, and I.T. Todorov, "Dynamical Derivation of Vacuum Operator Product Expansions in Euklidean Q. F. T.," preprint, Institute for Advanced Study, New York (1975).
[8]J. Kupsch, W. Rühl, and B.C. Yunni, "Operator Product Expansions in Conformally Convariant Quantum Field Theory, Part I," preprint, Kaiserslautern (1975).
[9]M. Lüscher, "Operator Product Expansion on the Vacuum in Conformal Quantum Field Theory in Two Space-Time Dimensions," preprint, Hamburg, Desy 75/45 (1975).
[10]M. Hortacsu, R. Seiler, and B. Schroer, Phys. Rev. D 5, 2519 (1972).
[11]J.A. Swieca and A.H. Völkel, Commun. Math. Phys. 29, 319 (1973).
[12]B. Schroer, J.A. Swieca, Phys. Rev. D 10, 480 (1974).
[13]M. Lüscher and G. Mack, Commun. Math. Phys. 41, 203 (1975).
[14]J. Kupsch, W. Rühl, and B.C. Yunn, "Conformal Invariance of Quantum Fields in Two-Dimensional Space—Time," preprint, Kaiserslautern (1974).
[15]G. Mack, "All Unitary Ray Representations of the Conformal Group SU(2,2) with Positive Energy," preprint, Hamburg (1975).
[16]L. Schwartz, *Theorié des distributions, Tom. I, II* (Hermann, Paris, 1957).
[17]I. M. Gel'fand and G. E. Schilow, *Verallgemeinerte Funktionen, Bd. I* (VEB, Berlin, 1960).
[18]L. Gårding and A.S. Wightman, Arkiv Fysik 28, (13) (1964).
[19]I. M. Gel'fand, M.I. Graev, and N. Ya. Vilenkin, *Generalized Functions* (Academic, New York, 1966), Vol. 5, Chap. VII.
[20]H. Bateman *et al.*, *Higher Transcendental Functions* (McGraw-Hill, New York, 1953), Vol. II, pp. 174ff.
[21]See Ref. 20, p. 98.
[22]G. Szegö, *Orthogonal Polynomials*, Am. Math. Soc. Colloq. Publ. Vol. XXIII (Am. Math. Soc., New York, 1939), Theorem 9.1.1.
[23]H. Epstein, Nuovo Cimento 27, 886 (1963).
[24]G. Mack, "Duality in Relativistic Quantum Field Theory," Preprint, Hamburg (1975).
[25]A. Jaffe, Phys. Rev. 158, 1454 (1967).

Wigner–Eckart theorem for tensor operators of graded Lie algebras*

L. Mezincescu

Institute of Physics, Theoretical Physics Department, Bucharest, POB 5207, Of. Post. Magurele, Romania
(Received 27 February 1976)

An invariant functional, analog to the group integral associated with a Lie group, is defined for the graded Lie algebras. A sufficient condition for the vanishing of the group volume is given. Orthogonality relations of the matrix elements of the representations are obtained, and the Wigner–Eckart theorem is proved for a class of graded Lie algebras.

1. INTRODUCTION

In connection with the interest in the supersymmetries,[1] much attention has been recently paid to the study of the graded Lie algebras (GLA).[2–5] Pais and Rittenberg[3] in a thorough study of the representations of the GSU(2) [OSp(2/1)] algebra [the graded version of the SU(2) algebra] have brought up the question of a general proof of the Wigner–Eckart (WE) theorem. They have shown its validity in the particular case of superspin $J=\frac{1}{2}$ tensor operators of the GSU(2) algebra.

In this paper we will develop a method which allows a general proof of the WE theorem for the GLA's satisfying the following conditions:

(1) the underlying Lie algebra (LA) is compact,

(2) the reducible representations are fully reducible,

(3) the "group volume" is nonvanishing.

It consists in defining an invariant functional[6] associated with any GLA, which is similar to the group integral for ordinary LA, in the sense that one integrates over the usual commuting parameters and over the anticommuting parameters associated with the generators of nonzero grading. This we do by using the definition of the "integrals" over anticommuting variables given by Berezin[7] and the formulas given by Pakhomov[8] for the change of coordinates in the mixed integrals over commuting and anticommuting variables which we give in Sec. 2.

In Sec. 3 the invariant integral associated with a GLA is defined, the orthogonality relations for the matrix elements of the representations of the GLA's are deduced, and a sufficient condition for the vanishing of the "group volume" is given.

In Sec. 4 for the sake of simplicity we work on the example of GSU(2). We show how all the Clebsch–Gordan coefficients can be determined and prove the WE theorem.

2. PAKHOMOV FORMULAS

Integration over anticommuting variables is defined according to Berezin[7] by the rules

$$\int d\theta_\alpha = 0, \quad \int d\theta_\alpha \theta_\beta = \delta_{\alpha\beta} \tag{2.1}$$

and multiple integrals are defined as products of simple integrals.

For mixed integrals over commuting ξ_i $(i=1,\ldots,n)$ and anticommuting θ_α $(\alpha=1,\ldots,m)$ variables, Pakhomov[8] has given the formulae for the change of coordinates. If the old and new coordinates (ξ,θ) and (ζ,η), respectively, are related by an invertible transformation

$$\xi = g(\zeta,\eta), \quad \theta = h(\zeta,\eta) \tag{2.2}$$

(g is an even and h is an odd element in η), then we have for any integral of an element $f(\xi,\theta)$

$$\int f(\xi,\theta)\, d\xi\, d\theta = \int f(g(\zeta,\eta),h(\zeta,\eta))\, J_{\zeta\eta}(\xi,\theta)\, d\zeta\, d\eta, \tag{2.3}$$

where $J_{\zeta\eta}(\xi,\theta)$ is the Jacobian

$$J_{\zeta\eta}(\xi,\theta) = \frac{D(\xi,\theta)}{D(\zeta,\eta)} = \begin{vmatrix} \frac{Dg}{D\zeta} & -\frac{Dg}{D\eta} \\ \frac{Dh}{D\zeta} & \frac{Dh}{D\eta} \end{vmatrix}, \tag{2.4}$$

which by using the rules of computing the generalized determinants[9] gives

$$J_{\zeta\eta}(\xi,\theta) = \left| \frac{Dg}{D\zeta} \right| \left| \frac{Dh}{D\eta} + \frac{Dh}{D\zeta}\left(\frac{Dg}{D\zeta}\right)^{-1} \frac{Dg}{D\eta} \right|^{-1}, \tag{2.5}$$

where the determinants appearing in (2.5) are usual determinants.

3. INVARIANT FUNCTIONAL

The general form of the commutation relations between generators of a GLA is

$$\begin{aligned} &[Q_i, Q_j] = f_{ij}^k Q_k \\ &[Q_i, V_\alpha] = F_{i\alpha}^\beta V_\beta \\ &\{V_\alpha, V_\beta\} = A_{\alpha\beta}^k Q_k \end{aligned} \tag{3.1}$$

$i=1,\ldots,n$, $\alpha=1,\ldots,m$, where we will consider that the LA generated by Q's is semisimple and compact. In any finite dimensional representation of (3.1) we define a "group" element

$$G = \exp(\theta_\alpha V_\alpha)\exp(i\xi_i Q_i), \tag{3.2}$$

where θ_α is a system of Grassmann anticommuting variables

$$\{\theta_\alpha, \theta_\beta\} = 0 \quad (\alpha=1,\ldots,m). \tag{3.3}$$

ξ_i are c-numbers and summation over repeated indices is implied. We will consider as functions on the "group"

functions $f(G)$ defined on the superspace (ξ, θ). An invariant functional may be defined on these functions with the help of a function $\mu_L(G)$ if:

$$\int dG\, \mu_L(G) f(G'G) = \int dG\, \mu_L(G) f(G), \tag{3.4}$$

where $\mu_L(G)$ stands for invariance to the left and similarly with $\mu_R(G)$ for invariance to the right,

$$dG = d^n\xi\, d^m\theta, \tag{3.5}$$

and

$$G' = \exp(\eta_\alpha V_\alpha)\exp(i\zeta_i Q_i), \tag{3.6}$$

η_α being another system of anticommuting variables.

Using (1.3) the relation (3.4) is equivalent to

$$\mu_L(G) = \left| \frac{D(GG')}{D(G')} \right|^{-1}_{G'=E} \tag{3.7}$$

and similarly for $\mu_R(G)$, with G and G' in (3.7) interchanged. If the underlying LA of (3.1) is compact we have from (3.7) and (1.5)

$$\mu_L(G) = \mu_R(G) \left| \frac{\partial \alpha_\rho(\theta,\eta)}{\partial \eta_\gamma} \right|^{-1} \times \left| \frac{\partial \alpha_\chi(\eta,\theta)}{\partial \eta_\epsilon} + \frac{\partial R_{\chi\kappa}(\zeta)}{\partial \zeta_e}\theta_\kappa \frac{\partial \beta_e(\eta,\theta)}{\partial \eta_\epsilon} \right|_{\eta,\zeta=0}, \tag{3.8}$$

where $\alpha_\rho(\theta,\eta)$, $\beta_e(\theta,\eta)$, and $R_{\chi\kappa}(\zeta)$ are defined by

$$\exp(\theta_\rho V_\rho)\exp(\eta_\kappa V_\kappa) = \exp[\alpha_\rho(\theta,\eta)V_\rho]\exp[i\beta_k(\theta,\eta)Q_k],$$
$$\alpha_\rho(\theta,\eta) = \theta_\rho + \eta_\rho + \cdots,$$
$$\beta_k(\theta,\eta) = \theta A_k \eta + \cdots, \tag{3.9}$$
$$\exp(-i\zeta_i Q_i)\, V_\alpha \exp(i\zeta_i Q_i) = R_{\beta\alpha}(\zeta)\, V_\beta, \tag{3.10}$$

respectively.

The relation (3.8) could be further simplified using the group properties of the composition law (3.9). Indeed, the inverse of (3.9) can be taken in two ways: either by interchanging the order of exponentials in (3.9) and changing the signs of the exponents, or by changing $\theta \to -\eta$ in (3.9) and using (3.10) to commute the Q term to the left. In this way we obtain

$$\beta_k(\theta,\eta) = -\beta_k(-\eta,\theta), \tag{3.11}$$

$$R(\beta_k(\theta,\eta))\alpha(-\eta,-\theta) = -\alpha(\theta,\eta). \tag{3.12}$$

Differentiating (3.12) with respect to η and using (3.11) and the fact that the part linear in η of $\alpha(\theta,\eta)$ is an even function of θ, we get

$$\left.\frac{\partial \alpha_\rho(\theta,\eta)}{\partial \eta_\gamma}\right|_{\eta=0} = \left.\frac{\partial \alpha_\rho(\eta,\theta)}{\partial \eta_\gamma} + \frac{\partial R_{\rho\kappa}(\zeta)}{\partial \zeta_e}\theta_\kappa \frac{\partial \beta_e(\eta,\theta)}{\partial \eta_\gamma}\right|_{\eta,\zeta=0}. \tag{3.13}$$

Thus, if the underlying LA of the GLA (3.1) is compact, the invariant forms to the left and right are equal

$$\mu_L(G) = \mu_R(G). \tag{3.14}$$

The functional (3.4) is also invariant with respect to changing G into G^{-1} in $f(G)$. These properties of the functional (3.4) are independent of the way (3.2) of parametrizing the elements, as can be seen by changing the variables in (3.4) and using the formulae (1.3).

We remark that unlike the ordinary Lie groups, in the case of the GLA the normalization integral (group volume)

$$\int dG\, \mu(G) \tag{3.15}$$

may vanish. This is the case whenever for the GLA we have irreducible representations with the number of even dimensions equal to the number of odd dimensions or when the number of even generators in (3.1) equals the number of odd generators.

To show this let us consider a finite dimensional irreducible representation of the GLA (3.1). The representation vectors are labeled by $a(\alpha)$, where α is the grade of the vectors which may be 0 or 1 and a is a complete system of quantum numbers. Let us consider the matrix elements

$$T(G)_{a(\alpha),b(\beta)} = [\exp(\theta V)\exp(i\xi \mathbf{Q})]_{a(\alpha),b(\beta)}. \tag{3.16}$$

Two such matrix elements anticommute whenever they are both odd, i.e., $\alpha+\beta$ is odd. By the standard technique, taking into consideration the fact that the matrix elements may anticommute, we get

$$\int dG\, \mu(G)\, T_{a(\alpha),b(\beta)}(G)\, T_{c(\gamma),d(\rho)}(G^{-1}) = (-)^\beta \lambda \delta_{c(\gamma),b(\beta)}\delta_{a(\alpha),d(\rho)} \tag{3.17}$$

which is the orthogonality relation for the matrix elements (3.16). We see that, unlike the ordinary compact Lie groups, (3.17) is not positively defined.

From (3.17) we get

$$\int dG\, \mu(G) = \lambda(N_e - N_0), \tag{3.18}$$

where $N_{e(0)}$ is the dimension of the even (odd) subspace in the irreducible representation of the algebra (3.1). In particular a sufficient condition for the vanishing of (3.18) is $N_e = N_0$. This is the case when the number of Q's equals the number of V's because then the adjoint representation is of this type (provided it is irreducible).

4. WIGNER-ECKART THEOREM

The WE theorem can be demonstrated for the GLA for which the CG decomposition theorem of direct products of irreducible representations holds, provided that (3.18) does not vanish.

In what follows we shall restrict ourselves to the simple example of the GSU(2) algebra. The commutation relations are

$$[Q_m, Q_n] = i\epsilon_{mnp}Q_p,$$
$$[Q_m, V_\alpha] = \tfrac{1}{2}(\tau^m)_{\beta\alpha}V_\alpha, \tag{4.1}$$
$$\{V_\alpha, V_\beta\} = \tfrac{1}{2}(c\tau^m)_{\alpha\beta}Q_m,$$

where $m = 1,2,3$, $\alpha = 1,2$, $c = \left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right)$, and τ^m are the Pauli matrices.

The functions α_ρ and β_m which appear in (3.9) are

$$\alpha(\theta,\eta) = \theta + \eta - \tfrac{1}{8}(\theta c\eta)\theta + \tfrac{1}{4}(\theta c\eta)\eta$$
$$\beta_m(\theta,\eta) = -(i/4)\theta c\tau_m\eta. \tag{4.2}$$

From (3.7) we obtain

$$\mu(G)=\tfrac{1}{2}(\sin^2\xi/\xi^2)(1+\tfrac{1}{8}\theta c\theta), \tag{4.3}$$

where $\xi=(\sum \xi_i^2)^{1/2}$ and ξ_i are the exponential parameters of the SU(2) group.

As shown by Pais and Rittenberg[3] the GSU(2) algebra has irreducible representations characterized by an integer or half-integer quantum number J called superspin, the spin content I being J and $J-\frac{1}{2}$. Consequently a vector of the representation is labeled by JII_3.

In the canonical basis the operators V_α have the following matrix elements:

$$(V^J_\alpha)_{II_3,MM_3}=[-\alpha\sqrt{J-2\alpha I_3}\;\delta_{IJ}\delta_{MJ-1/2} -\tfrac{1}{2}\sqrt{J+2\alpha I_3+\tfrac{1}{2}}\,\delta_{IJ-1/2}\delta_{MJ}]\,\delta_{I_3+\alpha,M_3}. \tag{4.4}$$

The orthogonality relation (2.17) is now

$$\int dG\,\mu(G)\,T^J_{II_3,MM_3}(G)\,T^{J'}_{M'M'_3,I'I'_3}(G^{-1}) =(-)^{4(J^2+M^2)}\delta_{JJ'}\delta_{II'}\,\delta_{I_3I'_3}\,\delta_{MM'}\,\delta_{M_3M'_3} \tag{4.5}$$

where we have used

$$T^J_{II_3,MM_3}(G)\,T^{J'}_{I'I'_3,M'M'_3}(G) =(-)^{4(I+M)(I'+M')}\,T^{J'}_{I'I'_3,M'M'_3}(G)\,T^J_{II_3,MM_3}(G). \tag{4.6}$$

The relation (4.5) can be obtained directly by integrating the left hand side with the form (4.4) for the matrix elements of the SU(2) and using the orthogonality relations for the matrix elements of the SU(2) group representations.

We can go further and using the CG theorem[3]

$$J\otimes J'=|J-J'|\oplus|J-J'|+\tfrac{1}{2}\oplus\cdots\oplus J+J' \tag{4.7}$$

deduce relations in which under the integral sign there appear three matrix elements.

We have

$$(-)^{4M(I'+M')}\int dG\,\mu(G)\,T^J_{II_3,MM_3}(G)\,T^{J'}_{I'I'_3,M'M'_3}(G)\,T^j_{mm_3,ii_3}(G^{-1}) =(-)^{4(j^2+m^2)}\langle JII_3\,J'I'I'_3\,|\,jii_3\rangle\langle JMM_3\,J'M'M'_3|\,jmm_3\rangle, \tag{4.8}$$

where $\langle JII_3\,J'I'I'_3\,|\,jii_3\rangle$ is the CG coefficient for the vector $|\,jii_3\rangle$ in the direct product of the representations J and J'. Formula (4.8) can be used to determine all the CG coefficients, as in the case of compact Lie groups.

The tensor operators are defined as a set of operators $V^j_{ii_3}$ which acting on a space of a representation $T(G)$ of the algebra (4.1) have the property

$$T(G)\,V^j_{ii_3}\,T(G^{-1})=T^j_{ii_3,mm_3}(G)\,V^j_{mm_3}. \tag{4.9}$$

For these operators we have the WE theorem for the matrix elements of $V^j_{ii_3}$:

$$\langle J'I'I'_3\alpha'\,|\,V^j_{ii_3}\,|\,JII_3\alpha\rangle =\langle jii_3\,JII_3\,|\,J'I'I'_3\rangle\langle J'\alpha'\,||\,V^j\,||\,J\alpha\rangle \tag{4.10}$$

where

$$\langle J'\alpha'\,||\,V^j\,||\,J\alpha\rangle =\sum_{mm_3MM_3M'M'_3}(-)^{4(J'^2+M'^2)} \times\langle jmm_3\,JMM_3\,|\,J'M'M'_3\rangle\langle J'M'M'_3\alpha'\,|\,V^j_{mm_3}\,|\,JMM_3\alpha\rangle \tag{4.11}$$

are the reduced matrix elements of the operators $V^j_{ii_3}$.

This can be easily shown by putting to the left and to the right of $V^j_{ii_3}$ in (4.10) the product $T(G^{-1})T(G)$. Using (4.9) two of the operators $T(G)$ are exchanged by the matrix element $T^j_{ii_3,mm_3}(G)$. Then one acts to the right on the state vector with the remaining products of operators and gets two more matrix elements $T^J_{II_3,MM_3}(G)$ and $T^{J'}_{I'I'_3,M'M'_3}(G)$. Finally integrating over the "group" and making use of (4.8) we obtain (4.10).

5. CONCLUSIONS

We have shown that for GLA there exist invariant forms which can be used to abstract in a simple way properties of the algebra such as the CG coefficients and the WE theorem. The invariant form (3.17) is not positively defined. We gave a sufficient condition for the vanishing of the "group volume" but it would be interesting to have also a necessary condition.

ACKNOWLEDGMENTS

The author is very indebted to Dr. M. Bundaru for very useful discussions and a critical reading of the manuscript. He also wishes to express deep gratitude to Dr. I. Raszillier and Dr. L. Bányai for a careful reading of the manuscript.

*Work performed under contract with the State Committee for Nuclear Energy of Romania.

[1]For a thorough list of references see B. Zumino, *Proceedings of the International Conference on High Energy Physics* (The Research Council, Rutherford Laboratory, 1974), I-245; A. Salam and J. Strathdee, Phys. Rev. D **11**, 1521 (1975); L. Corwin, O. Ne'eman, and S. Sternberg, Rev. Mod. Phys. **47**, 573 (1975); V.I. Ogievetsky and L. Mezincescu, Usp. Fiz. Nauk **117**, 637 (1975).

[2]V.G. Katz, Funkts Anal. Pril. **9**, 91 (1975) (in Russian).

[3]A. Pais and V. Rittenberg, J. Math. Phys. **16**, 2062 (1975).

[4]P.G.O. Freund and I. Kaplansky, "Simple Supersymmetries," University of Chicago preprint EFI 75-49 (1975).

[5]W. Nahm, V. Rittenberg, and M. Scheunert, "The classification of graded Lie algebras," Bonn University preprint-HE-75-16 (1975).

[6]The question of defining a group integral for GLA appears in F.A. Berezin and G.I. Katz, Mat. Sbornik **82**, 124 (1970) [Math. USSR Sbornik **11**, 311 (1970)].

[7]The integrals over anticommuting parameters are defined in F.A. Berezin, *The Method of Second Quantization* (Academic, New York, 1966).

[8]B.F. Pakhomov, Mat. Zametki **16**, 65 (1974) (in Russian).

[9]R. Arnowitt, P. Nath, and B. Zumino, Phys. Lett. B **56**, 81 (1975).

Wightman distributions on conformal space*

D. H. Mayer

Institut des Hautes Etudes Scientifiques,† 91440 Bures-sur-Yvette, France
(Received 24 February 1976)

We show that every tempered distribution $T_n \in \mathcal{S}'(M_4^n)$ which is the boundary value of a function $f_n(z)$ holomorphic in the field theoretic tube domain $\mathcal{T}_4^n$ can be uniquely continued to a distribution $\tilde{T}_n$ on the universal covering space $\tilde{M}_4^n$ of the conformally compactified Minkowski space M_c^{4n}. It can be shown that $\tilde{T}_n$ is the boundary value of a function $\tilde{f}_n$ holomorphic in a certain domain $\tilde{\mathcal{T}}_4^n$ of the complex manifold $\mathbb{C}\tilde{M}_4^n = (\mathbb{C}\times\Sigma_3)^n$, where Σ_3 denotes the affine complex three-dimensional unit sphere.

INTRODUCTION

On the way to construct a globally conformal invariant quantum field theory it has become clear that the so called conformal superworld $\tilde{M}_4$, which is nothing else but the universal covering manifold of the compactified Minkowski space M_c^4, can perhaps play the role of an underlying manifold of space and time on which such a theory can live.[1-4].

One can be lead to such a conclusion by several recent results concerning the structure of this manifold: We know that the manifold $\tilde{M}_4$ admits a global causal structure which is invariant under the smooth action of the universal covering group $\widetilde{SO(2,4)}$ of the conformal group $SO(2,4)/Z_2 \cong C_{15}(M_4)$.[2-5] Furthermore, it was shown in Refs. 6 and 7 that the unitary representations of this universal covering group are those which appear in the transformation laws of quantized fields under the conformal group.

Generalizing this result, Lüscher and Mack[1] proved that in every weakly conformal invariant Wightman field theory,[8] the Hilbert space of physical states carries an unitary representation of the universal covering group $\widetilde{SO(2,4)}$. They also could show that the Wightman functions of such a theory can be analytically continued to a domain of holomorphy which has as a real boundary the space $\tilde{M}_4$.

In answering the question if there exist on $\tilde{M}_4$ fields as operator valued distributions it is necessary to know if these analytically continued Wightman functions have on $\tilde{M}_4$ boundary values in the sense of distributions on the space $\tilde{M}_4$. This is the problem we will be considering in this paper.

In contrast to the authors of Ref. 1, we do not use the Euclidean version of a Wightman field theory which would mean working with Euclidean Green functions, but consider the Wightman distributions and the Wightman functions themselves. As is well known, these are tempered distributions and boundary values of the holomorphic Wightman functions.

We consider therefore the subclass of tempered distributions on $\mathbb{R}^{4n}$ which are boundary values of functions which are holomorphic in the so-called field theoretic tube domain $\mathcal{T}_4^n$. For this class of distributions one can show that they can be uniquely extended to distributions on the compact manifold M_c^{4n}. By embedding M_c^{4n} into a complex compact manifold $\mathbb{C}M_c^{4n}$ we are able to show that these extended distributions are boundary values of functions holomorphic in a certain tube domain $\mathcal{T}_4^n$, the real boundary of which is exactly M_c^{4n}.

The well-known principle of "recollement des morceaux" allows us to project these distributions on M_c^{4n} back to the universal covering manifold $\tilde{M}_4^n$ via the canonical mapping π_M: $\tilde{M}_4 \to M_c^4$. The distributions we get in this way on $\tilde{M}_4^n$ are then invariant under the action of the group of deck transformations[2] on the manifold $\tilde{M}_4^n$.

Using then a certain complexification $(\mathbb{C}\times\sum_3)^n$ of the manifold $\tilde{M}_4^n$, we can also show that the distributions we have constructed on $\tilde{M}_4^n$ above are boundary values of holomorphic functions on the complex manifold $(\mathbb{C}\times\sum_3)^n$.

As a first application of these considerations we get the result that the Wightman distributions $W_n(\xi_1,\ldots,\xi_n)$, which are the boundary values of holomorphic functions in exactly the tube domain $\mathcal{T}_4^n$, can be uniquely extended to distributions on the manifold $\tilde{M}_4^n$ which are boundary values of holomorphic functions on certain domains $\tilde{\mathcal{T}}_4^n$ of the complex manifold $(\mathbb{C}\times\sum_3)^n$.

We treat these problems in the following way: In Sec. I we very briefly repeat the definitions of the conformally compactified Minkowski space M_c^4 and its universal covering space $\tilde{M}_4$. We construct a topological isomorphism between the space $\mathcal{S}(M_4)$ and a closed subspace $\mathcal{D}^\kappa(M_c^4)$ of the space $\mathcal{D}(M_c^4)$, which is the space of all C^∞ complex valued functions on M_c^4.

In Sec. II we show how a first class of distributions on M_4 can be uniquely extended to distributions on the manifold M_c^4. The problem of boundary values of holomorphic functions on the manifold $\mathbb{C}M_c^4$ is also discussed.

In Sec. III we show how every distribution on M_c^4 determines a unique distribution on the universal covering manifold $\tilde{M}_4$ which is invariant under the group of deck transformations of the covering $\pi_M: \tilde{M} \to M_c^4$. Furthermore, we show that the distributions on $\tilde{M}_4$ which we get in this way are boundary values of holomorphic functions on the complex manifold $\mathbb{C}\tilde{M}_4$.

In Appendix A we repeat the notion of an equicontinuous set, and in Appendices B, C, D, and E we give the proofs of some lemmas and inequalities which are the main steps for getting the indicated results.

I. THE RELATION BETWEEN THE SPACE $\mathcal{S}(M_4)$ AND THE SPACE $\mathcal{D}(M_c^4)$

Let us briefly recall the definitions of the different

spaces we are dealing with in this paper. We denote by M_4 the four-dimensional pseudo-Euclidean Minkowski space endowed with the Lorentz metric g

$$g=\sum_{i,j=0}^{3} g_{ij}\,dx_i\otimes dx_j, \tag{1}$$

where $g_{ij}=0$ for $i\neq j$ and $g_{00}=-g_{11}=-g_{22}=-g_{33}=1$. We denote the points of M_4 by $x=(x^0,x^1,x^2,x^3)=(x^0,\mathbf{x})$. The compactification of the space M_4 we are interested in is the space M_c^4 as described, for instance, in Ref. 9. This space is diffeomorphic to $(S_1\times S_3)/Z_2$, where S_n denotes the n-dimensional unit sphere and Z_2 the discrete group $\{1,-1\}$. For the following it is convenient to look at the manifold M_c^4 as a closed subset of the five-dimensional real projective space $\mathbb{P}^5$.[10] If we introduce for the elements $[\eta]\in\mathbb{P}^5$ the projective coordinates $(\eta^0,\eta^1,\ldots,\eta^5)$, the space M_c^4 can be described as follows:

$$M_c^4=\{[\eta]\in\mathbb{P}^5:\eta_0^2-\eta_1^2-\eta_2^2-\eta_3^2-\eta_4^2+\eta_5^2=0\}. \tag{2}$$

Minkowski space M_4 is then diffeomorphic to the complement of the intersection of M_c^4 with any projective hyperplane of $\mathbb{P}^5$, especially if we take for this the hyperplane $\eta^4-\eta^5=0$. From this it follows immediately that M_4 can be densely embedded into M_c^4. If we define an open set $U_\kappa\subset M_c^4$ by

$$U_\kappa:=\{[\eta]\in M_c^4:\kappa=\eta^4-\eta^5\neq 0\},$$

and a mapping φ_κ: $U_\kappa\to M_4$ by

$$\varphi_\kappa:[\eta]\mapsto(\eta^0/\kappa,\eta^1/\kappa,\eta^2/\kappa,\eta^3/\kappa), \tag{3}$$

then the mapping (3) is a diffeomorphism of U_κ onto M_4 and φ_κ^{-1} gives a C^∞ embedding of M_4 into the manifold M_c^4. For the definition of a basis for a complete atlas $\mathfrak{A}$ on M_c^4 which makes M_c^4 a differentiable manifold see, for instance, Ref. 10.

The universal covering space of M_c^4 is denoted by $\tilde{M}_4$ and it is clear that it is diffeomorphic to the space $\mathbb{R}\times S_3$. For the elements of $\tilde{M}_4$ we write $\tilde{x}=(\tau,\mathbf{n})$, that is

$$\tilde{M}_4=\left\{\tilde{x}=(\tau,\mathbf{n}):\tau\in\mathbb{R},\mathbf{n}=(n^1,\ldots,n^4):\sum_{i=1}^{4}n_i^2=1\right\}. \tag{4}$$

The canonical mapping of $\tilde{M}_4$ onto M_c^4 is denoted by π_M and is given by

$$\pi_M:\tilde{x}\to[(\cos\tau,\mathbf{n},\sin\tau)]. \tag{5}$$

Let us next consider the function spaces $\mathcal{S}(M_4)$ and $\mathcal{D}(M_c^4)$. The space $\mathcal{S}(M_4)$ is the well-known Schwartz space of C^∞ functions on M_4 which vanish at infinity together with all their derivatives faster than any power of $|x|^{-1}$, where $|x|=(x_0^2+x_1^2+x_2^2+x_3^2)^{1/2}$ together with the usual topology on it.[11]

The space $\mathcal{D}(M_c^4)$ is the space of all C^∞ functions on the compact manifold M_c^4. The topology on $\mathcal{D}(M_c^4)$ is defined in the following way[12]: Let $f_n\in\mathcal{D}(M_c^4)$ be a sequence. We say that f_n converges in $\mathcal{D}(M_c^4)$ to 0 iff for every chart (U,φ) of M_c^4 and every compact set $K\subset\varphi(U)\subset\mathbb{R}^4$ the sequence $f_n\circ\varphi^{-1}$ converges, together with all its derivatives, uniformly on K to 0.

Of special interest for us is the following subspace $\mathcal{D}^\kappa(M_c^4)$ of the space $\mathcal{D}(M_c^4)$:

$$\mathcal{D}^\kappa(M_c^4):=\{f\in\mathcal{D}(M_c^4):f\text{ vanishes together with all its derivatives on the intersection of the hyperplane }\kappa=0\text{ with }M_c^4\}.$$

We assume $\mathcal{D}^\kappa(M_c^4)$ carries the induced topology of $\mathcal{D}(M_c^4)$. In generalizing a well-known result of Schwartz[11] we can prove the following theorem:

Theorem I: The space $\mathcal{S}(M_4)$ is topologically isomorphic to the space $\mathcal{D}^\kappa(M_c^4)$.

For the proof of this theorem we need some lemmas which we will state and prove first.

Lemma 1: Let $f_n\in\mathcal{D}(M_c^4)$ be a sequence which converges to $g\in\mathcal{D}(M_c^4)$. Then the set $H=\{f_n\}\cup g$ is uniformly equicontinuous.

Proof: Because H is a compact subset of $\mathcal{D}(M_c^4)$ it follows from the theorem of Ascoli (see Appendix A) that H is equicontinuous, but because M_c^4 is compact, every equicontinuous set $H\subset C(M_c^4)$ is uniformly equicontinuous.[13]

Lemma 2: Let $f_n\in\mathcal{D}^\kappa(M_c^4)$ be a sequence which converges to 0 and let P be any C^∞ differential operator on M_c^4. Then the sequence Pf_n also converges in $\mathcal{D}^\kappa(M_c^4)$ to 0.

Proof: Because P is a C^∞ differential operator on M_c^4 it follows[14] that there exists for every chart (U_i,φ_i) of M_c^4, a C^∞ differential operator P_i on $\varphi_i(U_i)$ such that

$$(Pf_n)\circ\varphi_i^{-1}=P_i(f_n\circ\varphi_i^{-1}) \tag{6}$$

on the open set $\varphi_i(U_i)\subset\mathbb{R}^4$. Now P_i has the following general form in the local coordinates $x=(x^0,x^1,x^2,x^3)$ on $\varphi_i(U_i)$:

$$P_i=\sum_{r=0}^{\infty}a_r D^r, \tag{7}$$

where the a_r are locally finite functions from $C^\infty(\varphi_i(U_i))$, and D^r is the familiar abbreviation

$$D^r:=\frac{\partial^r}{\partial x_0^{r_0}\partial x_1^{r_1}\partial x_2^{r_2}\partial x_3^{r_3}},\quad r=\sum_{i=0}^{3}r_i. \tag{8}$$

Now if K is any compact subset of $\varphi_i(U_i)$, it follows immediately from the definition of the topology in $\mathcal{D}^\kappa(M_c^4)$ that the sequence $(Pf_n)\circ\varphi_i^{-1}$ converges [because of (6)] uniformly to 0 on K together with all its derivatives. That the functions Pf_n are again elements of the space $\mathcal{D}^\kappa(M_c^4)$, is clear from the definition of this space.

The content of Lemma 2 can also be expressed by saying that every C^∞ differential operator P on M_c^4 is a continuous mapping of $\mathcal{D}(M_c^4)$ onto itself, such that the subspace $\mathcal{D}^\kappa(M_c^4)$ is left invariant. Let us now prove Theorem I.

Proof of Theorem I: First we define a mapping h: $\mathcal{D}^\kappa(M_c^4)\to\mathcal{S}(M_4)$ which, as we will show, has the desired properties. Let φ_κ be the diffeomorphism defined in (3) and let f be any element from $\mathcal{D}^\kappa(M_c^4)$. The mapping h is then defined as follows[11]:

$$h:\bar{f}\to f:=\bar{f}\circ\varphi_\kappa^{-1}. \tag{9}$$

On the other hand, if $f\in\mathcal{S}(M_4)$, we define a mapping h^{-1} as

$$h^{-1}: f \to \bar{f} := \begin{cases} f \circ \varphi_\kappa & \text{on } U_\kappa, \\ 0 & \text{on } M_c^4 \setminus U_\kappa = \complement U_\kappa. \end{cases} \tag{10}$$

We have to show that the function f in (9) is an element of $\mathcal{S}(M_4)$. Clearly f is from $C^\infty(M_4)$. Next we have to show that for every $m \in \mathbb{N}$, every multiindex $r = (r_0, r_1, r_2, r_3)$, and for every $\epsilon > 0$ there exists a number $N = N(m, r, \epsilon)$ such that

$$|x|^{2m} |D^r f(x)| < \epsilon \quad \text{for all } x \in M_4 \text{ with } |x| > N.$$

We have shown in Ref. 10 that for all $i = 0, 1, 2, 3$ the differential operator $(\partial/\partial x)^i$ on $\varphi_\kappa(U_\kappa) \cong M_4$ can be continued to a well-defined C^∞ differential operator P^i on M_c^4. Therefore, the differential operator D^r also defines a C^∞ differential operator P^r on M_c^4. Lemma 2 then shows us that $P^r \bar{f} \in \mathcal{D}^\kappa(M_c^4)$ and also that the function

$$\bar{f}_1 := \frac{(\eta_0^2 + \eta_1^2 + \eta_2^2 + \eta_3^2)^m}{\kappa^{2m}} P^r \bar{f} \tag{11}$$

is an element of $\mathcal{D}^\kappa(M_c^4)$ for every $m \in \mathbb{N}$ and every multiindex r. Denote by $V^\delta(\kappa)$ the following open set of M_c^4:

$$V^\delta(\kappa) := \{[\eta] \in M_c^4 : \operatorname{dist}([\eta], \complement U_\kappa) < \delta\}. \tag{12}$$

Because any function $\bar{f} \in \mathcal{D}^\kappa(M_c^4)$ is uniformly continuous on M_c^4 there exists for $\epsilon > 0$ a $\delta > 0$ such that

$$|\bar{f}([\eta_1]) - \bar{f}([\eta_2])| < \epsilon \quad \text{for all } [\eta_1], [\eta_2] \in M_c^4 \text{ with } \operatorname{dist}([\eta_1], [\eta_2]) < \delta.$$

Applying this to the set $V^\delta(\kappa)$, we get

$$|\bar{f}([\eta])| < \epsilon \quad \text{for all } [\eta] \in V^\delta(\kappa),$$

because $\bar{f}([\eta]) = 0$ for $[\eta] \in \complement U_\kappa$.

It is clear that for all $x \in M_4$, the points $[\eta] = \varphi_\kappa^{-1}(x) \in U_\kappa$ are contained in the open set $V^\delta(\kappa)$ if $|x| > N(\delta)$, where $N(\delta)$ depends on δ and therefore on ϵ. So we get for all $x \in M_4$ with $|x| > N(\delta)$, and for the function $\bar{f}_1$ from (11),

$$|\bar{f}_1 \circ \varphi_\kappa^{-1}(x)| < \epsilon.$$

But this reads, when we insert definition (3) of φ_κ and use (6),

$$|x|^{2m} |D^r f(x)| < \epsilon \quad \text{for all } |x| > N(\delta) = N(\epsilon).$$

Let us next show that the function $\bar{f}$ defined in (10) is an element of $\mathcal{D}^\kappa(M_c^4)$. For this we have only to show that $\bar{f}$ is C^∞ on $\complement U_\kappa$ because it is trivially C^∞ on U_κ, and that all derivatives of $\bar{f}$ vanish on $\complement U_\kappa$. The same reasoning as above shows us that

$$\left| \left(\frac{\eta_0^2 + \cdots + \eta_3^2}{\kappa^2} \right)^m P^r \bar{f}([\eta]) \right| < \epsilon \quad \text{for all } [\eta] \in V^\delta(\kappa) \cap U_\kappa, \tag{13}$$

with $V^\delta(\kappa)$ as defined above. If we now use the fact that relation (13) is also true for all C^∞ differential operators P_α^r on M_c^4, which are obtained as extensions of the local differential operators D_α^r in the different local charts $(\varphi_\alpha, U_\alpha)$ (see for instance Ref. 10), we get from (13) in the local chart $(\varphi_\alpha, U_\alpha)$, the following inequality:

$$|a(x_\alpha) D_\alpha^r (\bar{f} \circ \varphi_\alpha^{-1})(x_\alpha)| < \epsilon \quad \text{for all } x_\alpha \in \varphi_\alpha(V^\delta(\kappa) \cap U_\kappa \cap U_\alpha),$$

where $a(x_\alpha)$ is some C^∞ function on $\varphi_\alpha(V^\delta(\kappa) \cap U_\kappa \cap U_\alpha)$.

Putting $\bar{f}$ and all its derivatives equal to 0 on $\complement U_\kappa$ therefore defines a C^∞ function on M_c^4. This shows that the mapping h is a one–one mapping. In a second step we have to show that h is even a homeomorphism. Consider therefore a sequence $\bar{f}_n \in \mathcal{D}^\kappa(M_c^4)$ which converges to 0. From Lemma 2 we also get that the sequence $[(\eta_0^2 + \cdots + \eta_3^2)/\kappa^2]^m P \bar{f}_n$ converges to 0 in $\mathcal{D}^\kappa(M_c^4)$ for all $m \in \mathbb{N}$ and for all C^∞ differential operators P on M_c^4. Because of Lemma 1 the set H of functions $\bar{f}_n^{r,m} := [(\eta_0^2 + \cdots + \eta_3^2)/\kappa^2]^m P^r \bar{f}_n$, where P^r is the differential operator on M_c^4 which on $\varphi_\kappa(U_\kappa)$ is just D^r, is a uniformly equicontinuous set. Therefore, there exists for $\epsilon > 0$, a $\delta > 0$ such that

$$|\bar{f}_n^{r,m}([\eta])| < \epsilon \quad \text{for all } [\eta] \in V^\delta(\kappa) \text{ and for all } n. \tag{14}$$

Because $\complement V^\delta(\kappa)$ is closed and therefore compact we have $\varphi_\kappa(\complement V^\delta(\kappa))$ is compact in $\varphi_\kappa(U_\kappa)$. Therefore, there exists a number $N_0 = N_0(\epsilon)$ such that

$$|\bar{f}_n^{r,m} \circ \varphi_\kappa^{-1}(x)| < \epsilon \quad \text{for all } n > N_0 \text{ and for all } x \in \varphi_\kappa(\complement V^\delta(\kappa)), \tag{15}$$

because $\bar{f}_n^{r,m} \to 0$ in $\mathcal{D}^\kappa(M_c^4)$.

For $x \notin \varphi_\kappa(\complement V^\delta(\kappa))$ we have $\varphi_\kappa^{-1}(x) \in V^\delta(\kappa)$ and therefore from (14),

$$|\bar{f}_n^{r,m} \circ \varphi_\kappa^{-1}(x)| < \epsilon \quad \text{for all } n \in \mathbb{N} \text{ and for all } x \notin \varphi_\kappa(\complement V^\delta(\kappa)). \tag{16}$$

Inserting now the definition of the function $\bar{f}_n^{r,m}$ into the two relations (15) and (16) gives

$$|x|^{2m} |D^r(\bar{f}_n \circ \varphi_\kappa^{-1})(x)| < \epsilon \quad \text{for all } x \in M_4 \text{ and for all } n > N_0, \tag{17}$$

which shows that $\bar{f}_n \circ \varphi_\kappa^{-1} = f_n$ converges to 0 in $\mathcal{S}(M_4)$.

On the other hand, let $f_n \in \mathcal{S}(M_4)$ be a sequence which converges to 0 in $\mathcal{S}(M_4)$, that means (17) is true. From this we immediately get

$$|P_\alpha^r \bar{f}_n([\eta])| < \epsilon \quad \text{for all } n > N_0 \text{ and all } [\eta] \in U_\kappa,$$

where P_α^r is any C^∞ differential operator on M_c^4 which is the global extension of any differential operator D_α^r on the local chart $\varphi_\alpha(U_\alpha)$. Because $\bar{f}_n$ and all their derivatives vanish on $\complement U_\kappa$ we have even

$$|P_\alpha^r \bar{f}_n([\eta])| < \epsilon \quad \text{for all } n > N_0 \text{ and for all } [\eta] \in M_c^4, \tag{18}$$

but (18) written in the local coordinates of $\varphi_\alpha(U_\alpha)$ gives

$$|D_\alpha^r(\bar{f}_n \circ \varphi_\alpha^{-1})(x_\alpha)| < \epsilon \quad \text{for all } x_\alpha \in \varphi_\alpha(U_\alpha) \text{ and for all } n > N_0.$$

This concludes the proof of Theorem I.

II. THE RELATION BETWEEN THE SPACE $\mathcal{S}'(M_4)$ AND THE SPACE $\mathcal{D}'(M_c^4)$

Theorem I of the last section allows us to identify the space $\mathcal{S}(M_4)$ with the closed subspace $\mathcal{D}^\kappa(M_c^4) \subset \mathcal{D}(M_c^4)$. We can therefore also identify the space $\mathcal{S}'(M_4)$ with the space $\mathcal{D}^{\kappa\prime}(M_c^4)$ via the mapping h^*

$$(h^* T)(\bar{f}) := T(h\bar{f}) \tag{19}$$

for $T \in \mathcal{S}'(M_4)$ and $\bar{f} \in \mathcal{D}^\kappa(M_c^4)$ with $h\bar{f}$ defined in (9).

In complete analogy to (19) we can define the inverse mapping $h^{*-1} = h^{-1*}$. Now if $T \in \mathcal{S}'(M_4)$, we get $h^* T$

$\in \mathcal{D}^{\kappa\prime}(M_c^4)$. Because $\mathcal{D}^\kappa(M_c^4)$ is a closed subspace of $\mathcal{D}(M_c^4)$, we can extend the definition of h^*T via the Hahn–Banach theorem to the whole space $\mathcal{D}(M_c^4)$. We call the distribution we get in this way on M_c^4, an extension of T onto the manifold M_c^4 and denote it by $\bar{T}$. It is however clear that this extension is not unique because we can add to $\bar{T}$ any distribution from $\mathcal{D}'(M_c^4)$ which has its support on the space $\complement U_\kappa$. The aim of this section is to find a class of tempered distributions on M_4 which have the property that they allow for a unique extension to a distribution on M_c^4. It turns out that this class contains all distributions which are the boundary values of holomorphic functions in certain domains which on the other hand, are also of great interest from the physical point of view.

Let

$$\begin{aligned} \mathcal{T}_4 &:= M_4 + iV_+ \\ &= \{z = x + iy : x \in M_4, y \in V_+ \\ &= \{y \in M_4 : y_0^2 - \mathbf{y}^2 > 0 \text{ and } y^0 > 0\}\}. \end{aligned} \tag{20}$$

Consider further the set

$$\mathcal{H}(\mathcal{T}_4) := \{f : \mathcal{T}_4 \to \mathbb{C},\ f \text{ holomorphic in } \mathcal{T}_4\}.$$

Definition: A function $f \in \mathcal{H}(\mathcal{T}_4)$ is an element of the space $\mathcal{S}'_{1oc}(\mathcal{T}_4)$ iff $\lim_{z \to \partial} \mathcal{T}_4 f(z) = T$ is a tempered distribution on the boundary $\partial\mathcal{T}_4$ of the domain $\mathcal{T}_4$. The limit is understood in the topology of $\mathcal{S}'(\mathbb{C}^4)$.

It has been shown by Martineau[15] that $f \in \mathcal{S}'_{1oc}(\mathcal{T}_4)$ iff there exist constants $C, \alpha, \beta \geq 0$ such that

$$|f(z)| \leq C(1 + |z|^2)^\alpha \operatorname{dist}(z, \partial\mathcal{T}_4)^{-\beta} \text{ for all } z \in \mathcal{T}_4. \tag{21}$$

Let us next construct a complexification $\mathbb{C}M_c^4$ of the real manifold M_c^4. For this we only have to complexify the space M_c^4 as defined in (2) to a space $\mathbb{C}M_c^4$ as follows:

$$\mathbb{C}M_c^4 := \{[\xi] : [\xi] \in \mathbb{CP}^5 : \xi_0^2 - \xi_1^2 - \xi_2^2 - \xi_3^2 - \xi_4^2 + \xi_5^2 = 0\}. \tag{22}$$

In (22), $\mathbb{CP}^5$ denotes the complex compact five-dimensional projective space. It follows therefore that $\mathbb{C}M_c^4$ is also a compact space, it is even a complex manifold.[16]

On $\mathbb{C}M_c^4$ we introduce the following complex structure. We denote by

$$\begin{aligned} &\kappa := \xi^4 - \xi^5, \quad \lambda := \xi^4 + \xi^5, \quad \tilde{\kappa} := \xi^1 - \xi^0, \\ &\tilde{\lambda} := \xi^1 + \xi^0, \quad \gamma := \xi^2 - \xi^5, \end{aligned} \tag{23}$$

for any point $[\xi] = [(\xi^0, \xi^1, \ldots, \xi^5)] \in \mathbb{C}M_c^4$. We use for the space $\mathbb{C}M_c^4$ the same letters $\kappa, \lambda, \cdots$ as we did in Ref. 10 for M_c^4, but hope that there is no confusion. Let us further denote by Z the set $Z := \{\kappa, \lambda, \tilde{\kappa}, \tilde{\lambda}, \gamma\}$. For every $\beta \in Z$ we introduce the following open subset U_β of $\mathbb{C}M_c^4$:

$$U_\beta := \{[\xi] \in \mathbb{C}M_c^4 : \beta \neq 0\}. \tag{24}$$

All these definitions are straightforward complex versions of the respective real cases in M_c^4. With the above definitions one verifies that

$$\mathbb{C}M_c^4 = \bigcup_{\beta \in Z} U_\beta.$$

To define the complex structure we still have to give the local homeomorphisms φ_β on the open sets U_β. First we define the mapping φ_κ: $U_\kappa \to \mathbb{C}^4$ as follows:

$$\varphi_\kappa : [\xi] \to (\xi^0/\kappa, \xi^1/\kappa, \xi^2/\kappa, \xi^3/\kappa). \tag{25}$$

Then φ_κ is a biholomorphic mapping of the set U_κ onto $\mathbb{C}^4$. Next we introduce certain transformations on $\mathbb{C}M_c^4$ which we use to define the other homeomorphisms φ_β. Let $[\xi] \in \mathbb{C}M_c^4$ be given in the projective coordinates by $[(\xi^0, \ldots, \xi^5)]$.

$$\begin{aligned} &\textit{Definition: } I_\lambda : [\xi] \mapsto [\xi'] := [(-\xi^0, \xi^1, \xi^2, \xi^3, \xi^4, -\xi^5)], \\ &I_{\tilde{\kappa}} : [\xi] \mapsto [\xi'] := [(\xi^5, \xi^4, \xi^2, \xi^3, \xi^1, \xi^0)], \\ &I_\chi := I_\lambda \circ I_{\tilde{\kappa}}, \\ &I_\gamma : [\xi] \mapsto [\xi'] := [(\xi^0, \xi^1, \xi^4, \xi^3, \xi^2, \xi^5)]. \end{aligned} \tag{26}$$

It is clear from the relations (24) and (26) that for all $\beta \in Z$ we have

$$I_\beta^2 = 1 \quad (\text{when we set } I_\kappa = \mathbf{1}_{\mathbb{C}M_c^4})$$

and

$$I_\beta(U_\kappa) = U_\beta. \tag{27}$$

Then we define the homeomorphisms $\varphi_\beta : U_\beta \to \mathbb{C}^4$ by

$$\varphi_\beta := \varphi_\kappa \circ I_\beta. \tag{28}$$

It is easy to see that all the transformations I_β, $\beta \in Z$, are elements of the group $SO(2,4)/Z_2$ which acts in a completely analogous way on the manifold $\mathbb{C}M_c^4$ as it acts on M_c^4.[2]

The charts (U_β, φ_β), $\beta \in Z$, define the basis of an atlas for the complex manifold $\mathbb{C}M_c^4$. The mappings φ_β then become biholomorphic mappings from $U_\beta \to \mathbb{C}^4$. Because $\varphi_\kappa^{-1} : \mathbb{C}^4 \to U_\kappa$ gives a biholomorphic embedding of $\mathbb{C}^4$ into the manifold $\mathbb{C}M_c^4$, and $\bar{U}_k = \mathbb{C}M_c^4$, the manifold $\mathbb{C}M_c^4$ is a compactification of the space $\mathbb{C}^4$ on which the group $SO(2,4)/Z_2$ acts as a group of biholomorphic transformations.

It is interesting to note that the space $\mathbb{C}M_c^4$ is simply connected. We shall prove this in Appendix C.

Consider now the image of the field theoretic tube domain $\mathcal{T}_4$ under the mapping φ_κ^{-1} in $\mathbb{C}M_c^4$. It is clear that

$$\mathcal{J}_4 := \varphi_\kappa^{-1}(\mathcal{T}_4) \tag{29}$$

is contained in U_κ, but it is even contained in the intersection of all the charts U_β, $\beta \in Z$, as we want to show next.

Lemma 3: The open set $\mathcal{J}_4$ is contained in $\cap_{\beta \in Z} U_\beta$ and is invariant under the action of the conformal group $SO(2,4)/Z_2$ on $\mathbb{C}M_c^4$.

Proof: It is known[17] that the domain $\mathcal{T}_4 \subset \mathbb{C}^4$ is invariant under the action of the group $C_{15}(M_4)$ on $\mathbb{C}^4$. This group is even a subgroup of the group of automorphisms of the domain $\mathcal{T}_4$. It is also clear, for instance from Ref. 18, that the action of $SO(2,4)/Z_2$ on $\mathbb{C}M_c^4$ gives locally on U_κ exactly the action of $C_{15}(M_4)$ on $\mathbb{C}^4$. Therefore, $\mathcal{J}_4$ is invariant under the action of $SO(2,4)/Z_2$. Because $\mathcal{J}_4 \subset U_\kappa$ and $w\mathcal{J}_4 = \mathcal{J}_4$ for all $w \in SO(2,4)/Z_2$, we get

$$\mathcal{J}_4 = I_\beta \mathcal{J}_4 \subset I_\beta U_\kappa = U_\beta \text{ for all } \beta \in Z, \tag{30}$$

because of (27). Therefore, $\mathcal{J}_4 \subset U_\beta$ for all $\beta \in Z$.

By using the explicit definitions (25) and (20) we get for the set $\mathcal{T}_4$, in terms of the projective coordinates ξ:

$$\mathcal{T}_4=\{[\xi]\in\mathbb{C}M_c^4:\text{If } v=(v^0,v^1,v^2,v^3)$$

$$\text{with } v^j=\mathrm{Im}(\xi^i\kappa^*),\ \text{then } v\in V_+\}. \tag{31}$$

For the boundary $\partial\mathcal{T}_4$, we get the set

$$\partial\mathcal{T}_4=\{[\xi]\in\mathbb{C}M_c^4: v^jv_j=0,\ v^0\geq 0\}. \tag{32}$$

Therefore, the real manifold M_c^4 belongs to the boundary $\partial\mathcal{T}_4$.

If we denote by $H(\mathcal{T}_4)$ the set of all holomorphic functions on the domain $\mathcal{T}_4$, we say that $\tilde{f}\in H(\mathcal{T}_4)$ belongs to the class $\mathcal{S}'_{\mathrm{loc}}(\mathcal{T}_4)$ iff the function $\tilde{f}$ has on the boundary $\partial\mathcal{T}_4$ a boundary value in the sense of distributions. To be more precise, this means the following: For any point $[\xi]\in\partial\mathcal{T}_4$, and for every chart (U,φ) of $\mathbb{C}M_c^4$ with $[\xi]\in U$ we have

$$\tilde{f}\circ\varphi^{-1}\in\mathcal{S}'_{\mathrm{loc}}(\varphi(U\cap\mathcal{T}_4)). \tag{33}$$

Because the property of having a boundary value in the sense of distributions is a local property,[15] we then get by using the principle of "recollement des morceaux" that a function $\tilde{f}$ with the property (33) has a boundary value in the sense of distributions on the boundary $\partial\mathcal{T}_4$.

To prove the main theorem of this section, namely that every tempered distribution T on M_4 which is the boundary value of a holomorphic function in the tube domain $\mathcal{T}_4$ can be uniquely continued to a distribution $\bar{T}$ on M_c^4 which again is the boundary value of a holomorphic function on $\mathcal{T}_4$, we need the following lemma.

Lemma 4: Let $\mathrm{Aut}(\mathcal{T}_4)$ be the group of automorphisms of the tube domain $\mathcal{T}_4$. For every $f\in\mathcal{S}'_{\mathrm{loc}}(\mathcal{T}_4)$ we have

$$wf\in\mathcal{S}'_{\mathrm{loc}}(\mathcal{T}_4)$$

where

$$wf(z):=f(w^{-1}z)\ \text{ for all } w\in\mathrm{Aut}(\mathcal{T}_4).$$

Proof: It is known that the group $\mathrm{Aut}(\mathcal{T}_4)$ is generated by the Poincaré transformations, the dilatations, and the inversion R, which is defined by

$$Rz=-z/z^2\ \text{ with } z^2=z_0^2-z_1^2-z_2^2-z_3^2,\quad z\in\mathcal{T}_4. \tag{34}$$

Because the lemma is trivially true for all Poincaré transformations and dilatations we only have to consider the case where $w=R$. Because R as defined in (34) is a biholomorphic mapping of $\mathcal{T}_4$, we have $Rf\in H(\mathcal{T}_4)$. Then we have to show that there exist constants $C,\alpha,\beta\geq 0$ such that

$$|Rf(z)|\leq C(1+|z|^2)^{\alpha}\mathrm{dist}(z,\partial\mathcal{T}_4)^{-\beta}.$$

Now a trivial geometrical consideration shows that the distance of a point $z=x+iy\in\mathcal{T}_4$ from the boundary $\partial\mathcal{T}_4$ is given by

$$\mathrm{dist}(z,\partial\mathcal{T}_4)=c(y^0-|\mathbf{y}|), \tag{35}$$

where c is some constant.

Because $f\in\mathcal{S}'_{\mathrm{loc}}(\mathcal{T}_4)$, there exist constants $C',\alpha',\beta'\geq 0$ with

$$|f(z')|\leq C'(1+|z'|^2)^{\alpha'}(y^{0\prime}-|\mathbf{y}'|)^{-\beta'}$$

$$\text{for all } z'=x'+iy'\in\mathcal{T}_4. \tag{36}$$

We now have to find some relation between $y^{0\prime}-|\mathbf{y}'|$ and $y^0-|\mathbf{y}|$ on the one hand, and $|z'|$ and $|z|$ on the other hand, where $z'=x'+iy'=Rz$. In Appendix B we will prove the following two inequalities:

$$y^{0\prime}-|\mathbf{y}'|\geq\tfrac{1}{5}(y^0-|\mathbf{y}|)^3|z|^{-4} \tag{37}$$

and

$$|z'|\leq 6|z|^3(y^0-|\mathbf{y}|)^{-4}. \tag{38}$$

Using the two relations (37) and (38) we get

$$|f(z')|\leq C(1+|z|^2)^{4\alpha'+2\beta'}\,\mathrm{dist}(z,\partial\mathcal{T}_4)^{-3\beta'-8\alpha'}$$

$$\text{for all } z\in\mathcal{T}_4, \tag{39}$$

with C a constant which is determined in a straightforward way. Because $f(z')=Rf(z)$ for $z'=Rz$, the lemma is proven.

Now we are able to prove the main theorem.

Theorem II: Let $f(z)\in\mathcal{S}'_{\mathrm{loc}}(\mathcal{T}_4)$. Then the function $\tilde{f}:=f\circ\varphi_\kappa$ is an element of $\mathcal{S}'_{\mathrm{loc}}(\mathcal{T}_4)$.

Proof: Because φ_κ is a biholomorphic mapping of $U_\kappa\to\varphi_\kappa(U_\kappa)=\mathbb{C}^4$ with $\varphi_\kappa(\mathcal{T}_4)=\mathcal{T}_4$, we get that $\tilde{f}\in H(\mathcal{T}_4)$. Now let $[\xi]$ be any point in $\partial\mathcal{T}_4$. Then there exists a $\beta\in Z$ such that $[\xi]\in U_\beta$. Let us look at $\varphi_\beta(U_\beta\cap\mathcal{T}_4)$. Because $\mathcal{T}_4\subset\cap_{\beta\in Z}U_\beta$ we have $\varphi_\beta(U_\beta\cap\mathcal{T}_4)=\varphi_\beta(\mathcal{T}_4)$, but with (28) and Lemma 3 we get

$$\varphi_\beta(\mathcal{T}_4)=\varphi_\kappa\circ I_\beta(\mathcal{T}_4)=\varphi_\kappa(\mathcal{T}_4)=\mathcal{T}_4. \tag{40}$$

Therefore, the image of $\mathcal{T}_4$ under the different mappings φ_β, $\beta\in Z$, is the same, namely the tube domain $\mathcal{T}_4\subset\mathbb{C}^4$. Because the different transformations I_β are biholomorphic transformations on the domain $\mathcal{T}_4$, we also get that

$$\tilde{f}\circ\varphi_\beta^{-1}=f\circ\varphi_\kappa\circ I_\beta\circ\varphi_\kappa^{-1} \tag{41}$$

is a holomorphic function on $\varphi_\beta(\mathcal{T}_4)=\mathcal{T}_4$.

The mapping $\varphi_\kappa\circ I_\beta\circ\varphi_\kappa^{-1}:\mathbb{C}^4\to\mathbb{C}^4$ is an element of the conformal group $C_{15}(M_4)$. Therefore, with the definition of Lemma 4 we can also write the function $\tilde{f}\circ\varphi_\beta^{-1}$ as wf with $w=\varphi_\kappa\circ I_\beta\circ\varphi_\kappa^{-1}$ an element of $C_{15}(M_4)$. Lemma 4 gives us then, that $\tilde{f}\circ\varphi_\beta^{-1}=wf\in\mathcal{S}'_{\mathrm{loc}}(\mathcal{T}_4)$ $=\mathcal{S}'_{\mathrm{loc}}(\varphi_\beta(U_\beta\cap\mathcal{T}_4))$, but this proves the theorem.

As an immediate consequence we get that for every $T\in\mathcal{S}'(M_4)$ such that $T=\lim_{z\to\partial}\mathcal{T}_4f(z)$, we have a unique extension to a distribution $\bar{T}$ from $\mathcal{D}'(M_c^4)$ which is defined as

$$\bar{T}=\lim_{[\xi]\to\partial\mathcal{T}_4}f\circ\varphi_\kappa([\xi]).$$

III. EXTENSION OF DISTRIBUTIONS FROM $\mathcal{D}'(M_c^4)$ TO DISTRIBUTIONS FROM $\mathcal{D}'(\tilde{M}_4)$

Let us consider the canonical projection π_M as defined in (5) and let us assume that $\tilde{M}_4$ is given such a differentiable structure that π_M is a local diffeomorphism. Then give a distribution $\bar{T}$ on the space M_c^4. Then $\bar{T}$ determines in a unique way a distribution $\tilde{T}$ on $\tilde{M}_4$.[19] Namely, consider any point $\tilde{x}\in\tilde{M}_4$. Then there exists an open neighborhood $\tilde{U}(\tilde{x})$ such that $\pi_{M|\tilde{U}}$ is a diffeomorphism of $\tilde{U}$ onto $\pi_M(\tilde{U})$. Denote by $\bar{T}_{|U}$ the restriction of the distribution $\bar{T}$ to the open set $U:=\pi_M(\tilde{U})$. The diffeomorphism $\pi_{M|\tilde{U}}$ induces a topological isomorphism of the space $\mathcal{D}(\tilde{U})$ onto the space $\mathcal{D}(U)$ in the following way:

$$\pi_{M|\tilde{U}}(\tilde{f}) := \tilde{f} \circ \pi^{-1}_{M|\tilde{U}} \text{ for } \tilde{f} \in \mathcal{D}(\tilde{U}). \tag{42}$$

The mapping $\pi^{-1}_{M|\tilde{U}}$ also then induces a mapping of the space $\mathcal{D}'(U)$ onto the space $\mathcal{D}'(\tilde{U})$, namely

$$(\pi^{-1}_{M|\tilde{U}} \overline{T}_{|U}) \tilde{f} := \overline{T}_{|U}(\pi_{M|\tilde{U}}(\tilde{f}))$$

$$= \overline{T}_{|U}(\tilde{f} \circ \pi^{-1}_{M|\tilde{U}}) \text{ for } \tilde{f} \in \mathcal{D}(\tilde{U}). \tag{43}$$

In this way we get for every point $\tilde{x} \in \tilde{M}_4$ and every admissible neighborhood $\tilde{U}$ of $\tilde{x}$ (admissible in the sense that $\pi_{M|\tilde{U}}$ is a diffeomorphism), a distribution $\tilde{T}_{\tilde{U}} := \pi^{-1}_{M|\tilde{U}} \overline{T}_{|U}$. Because these distributions trivially coincide for $\tilde{U} \cap \tilde{V} \neq \phi$, $\tilde{U}$, $\tilde{V}$ admissible neighborhoods in $\tilde{M}_4$, that means

$$\tilde{T}_{\tilde{U}} = \tilde{T}_{\tilde{V}} \text{ on } \tilde{U} \cap \tilde{V}, \tag{44}$$

we get from the principle of "recollement des morceaux" a unique distribution $\tilde{T}$ on $\tilde{M}_4$ such that for all admissible open sets $\tilde{U}$ in $\tilde{M}_4$ we have

$$\tilde{T}_{|\tilde{U}} = \pi^{-1}_{M|\tilde{U}} \overline{T}_{|U}, \tag{45}$$

where $U = \pi_M(\tilde{U})$.

We want to show that the above constructed distribution $\tilde{T}$ on $\tilde{M}_4$ is invariant under the action of the group $\Gamma(\tilde{M}_4, \tilde{M}^4_c)$ of deck transformations of the covering π_M: $\tilde{M}_4 \to M^4_c$. From Ref. 2 we know that every element $\gamma \in \Gamma(\tilde{M}_4, M^4_c)$ has the property

$$\pi_M \circ \gamma = \pi_M. \tag{46}$$

To show the invariance of $\tilde{T}$ under every $\gamma \in \Gamma(\tilde{M}_4, M^4_c)$ we have only to show that for every admissible open set $\tilde{U} \subset \tilde{M}_4$ we have

$$(\gamma \tilde{T})_{|\tilde{U}} = \tilde{T}_{|\tilde{U}}, \tag{47}$$

because this implies immediately that $\gamma \tilde{T} = \tilde{T}$ on the whole space $\tilde{M}_4$. Applying (45) we get for every $\tilde{f} \in \mathcal{D}(\tilde{U})$,

$$(\gamma \tilde{T})_{|\tilde{U}}(\tilde{f}) = \tilde{T}_{|\gamma^{-1}(\tilde{U})}(\gamma^{-1} \tilde{f}) = \tilde{T}_{|\gamma^{-1}(\tilde{U})}(\tilde{f} \circ \gamma),$$

where we have used the relations (42) and (43).

Using (45) we get

$$\tilde{T}_{|\gamma^{-1}(\tilde{U})}(\tilde{f} \circ \gamma) = \pi^{-1}_{M|\gamma^{-1}(\tilde{U})} \overline{T}_{|U}(\tilde{f} \circ \gamma) = \overline{T}_{|U}(\tilde{f} \circ \gamma \circ \pi^{-1}_{M|\gamma^{-1}\tilde{U}})$$

$$= \overline{T}_{|U}(\tilde{f} \circ \pi^{-1}_{M|\tilde{U}}) = \pi^{-1}_{M|\tilde{U}} \overline{T}_{|U}(\tilde{f}) = \tilde{T}_{|\tilde{U}}(\tilde{f}). \tag{48}$$

Having established the existence of a distribution $\tilde{T}$ on $\tilde{M}_4$ given a distribution $\overline{T}$ on M^4_c, we want to show that $\tilde{T}$ can be obtained as the boundary value of a holomorphic function when $\overline{T}$ is the boundary value of a function holomorphic in the domain $\mathcal{T}_4$. In order to achieve this, we first have to embed the manifold $\tilde{M}_4$ into a complex manifold $\mathbb{C}\tilde{M}_4$. Because of (4), a natural choice is the following:

$$\mathbb{C}\tilde{M}_4 := \mathbb{C} \times \Sigma_3, \tag{49}$$

where Σ_3 denotes the complex affine three-dimensional unit sphere that means $\Sigma_3 = \{z \in \mathbb{C}^4 : z_1^2 + \cdots + z_4^2 = 1\}$.

If we remember the definition of the canonical mapping π_M in (5) we can immediately extend this mapping to a mapping

$$\pi^{\mathbb{C}}_M : \mathbb{C}\tilde{M}_4 \to \mathbb{C}M^4_c \tag{50}$$

defined by

$$\pi^{\mathbb{C}}_M : \tilde{z} \mapsto (\cos z, \mathbf{m}, \sin z), \tag{51}$$

where $\tilde{z} \in \mathbb{C}\tilde{M}_4$ is given by $\tilde{z} = (z, \mathbf{m})$ with $z \in \mathbb{C}$ and $\mathbf{m} \in \Sigma_3$.

It is clear that $\pi^{\mathbb{C}}_M$ defined in this way is not a covering map because $\pi^{\mathbb{C}}_M$ is not a mapping onto the manifold $\mathbb{C}M^4_c$. This can be seen from the fact that for all $[\xi] \in \pi^{\mathbb{C}}_M(\mathbb{C}\tilde{M}_4)$ we have $\xi_0^2 + \xi_5^2 = 1$, whereas in $\mathbb{C}M^4_c$ there are also points with $\xi_0^2 + \xi_5^2 = 0$.

Let us therefore consider the complex manifold $(\Sigma_1 \times \Sigma_3)/Z_2$ which is a submanifold of $\mathbb{C}M^4_c$. This manifold has the property that $\pi_1(\Sigma_1 \times \Sigma_3) = \mathbb{Z}$. This follows from the fact that Σ_n is diffeomorphic to the tangent bundle $T(S_n)$ of the real n-dimensional unit sphere. So we get as a special case that Σ_1 is diffeomorphic to the two-dimensional cylinder in $\mathbb{R}^3$ and therefore $\pi_1(\Sigma_1) = \mathbb{Z}$.

The mapping

$$\pi^{\mathbb{C}}_M : \mathbb{C} \times \Sigma_3 \to (\Sigma_1 \times \Sigma_3)/Z_2 \tag{52}$$

is a universal covering map. It is, as one can see from (51), a holomorphic mapping from the complex manifold $\mathbb{C} \times \Sigma_3$ onto the complex manifold $(\Sigma_1 \times \Sigma_3)/Z_2$ which is locally even biholomorphic.[20]

Let us next investigate the relation between $(\Sigma_1 \times \Sigma_3)/Z_2$ and the domain $\mathcal{T}_4$ in $\mathbb{C}M^4_c$. For this we define the set

$$N_0 := \{[\xi] \in \mathbb{C}M^4_c : [\xi] \notin (\Sigma_1 \times \Sigma_3)/Z_2\}. \tag{53}$$

For $[\xi] \in N_0$ we have $\xi_0^2 + \xi_5^2 = 0$ or $\xi_0 = \pm i\xi_5$. We are interested in the set $N_0 \cap \mathcal{T}_4$. Because $\mathcal{T}_4 \subset U_\kappa$ we can assume $\kappa = 1$ and therefore

$$\xi^4 - \xi^5 = 1 \text{ or } \xi^4 \pm i\xi^0 = 1. \tag{54}$$

Writing $\xi^\mu = \eta^\mu + i\zeta^\mu$ we get for (54),

$$\eta^4 \mp \zeta^0 = 1 \text{ and } \zeta^4 = \mp \eta^0, \tag{55}$$

and therefore

$$\eta^4 = 1 \pm \zeta^0 \text{ and } \zeta^4 = \mp \eta^0.$$

Because $\xi_0^2 + \xi_5^2 = 0$ we also have $\xi_1^2 + \cdots \xi_4^2 = 0$, but this means

$$\eta_1^2 + \cdots + \eta_4^2 = \zeta_1^2 + \cdots + \zeta_4^2 \text{ and } \eta_1\zeta_1 + \cdots + \eta_4\zeta_4 = 0. \tag{56}$$

With (25) we get for $z = x + iy = \varphi_\kappa([\xi])$, $y = (\zeta^0, \zeta^1, \zeta^2, \zeta^3)$ and therefore $|y|^2 = \zeta_0^2 + \cdots + \zeta_3^2$.

With (55) this reads

$$|y|^2 = (1 - \eta^4)^2 + \eta_4^2 + \eta_1^2 + \eta_2^2 + \eta_3^2 - \zeta_4^2. \tag{57}$$

As we will show in Appendix E, we have

$$\eta_1^2 + \eta_2^2 + \eta_3^2 \geq \zeta_4^2 \tag{58}$$

and therefore from (57) it follows that

$$|y|^2 \geq (1 - \eta^4)^2 + \eta_4^2 \geq \tfrac{1}{2}.$$

This shows that we have

$$\varphi_\kappa(N_0 \cap \mathcal{T}_4) \subset \{z \in T_4 : |y|^2 \geq \tfrac{1}{2}\}.$$

This, on the other hand, implies that the set $(\Sigma_1 \times \Sigma_3)/Z_2 \cap \mathcal{T}_4$ has the property

$$\varphi_\kappa((\Sigma_1 \times \Sigma_3)/Z_2 \cap \mathcal{T}_4) \supset \{z \in T_4 : |y|^2 < \tfrac{1}{2}\}. \tag{59}$$

In the coordinates ξ, this means

$$(\Sigma_1\times\Sigma_3)/Z_2\cap\mathcal{T}_4\supset\{[\xi]\in\mathcal{T}_4: v_0^2+\cdots+v_3^2<\tfrac{1}{2}\kappa\kappa^*\}. \tag{60}$$

Since the mapping $\pi_M^{\mathbb{C}}$ as defined in (51) is a local biholomorphic mapping of $\mathbb{C}\tilde{M}_4$ onto $(\Sigma_1\times\Sigma_3)/Z_2$ there exists for every point $\tilde{x}\in M_4\subset\mathbb{C}\tilde{M}_4$ a neighborhood $\tilde{U}(\tilde{x})$ in $\mathbb{C}\tilde{M}_4$ such that $\pi^{\mathbb{C}}_{M|\tilde{U}}$ is a biholomorphic mapping of $\tilde{U}(\tilde{x})$ onto $\pi_M^{\mathbb{C}}(\tilde{U})\subset(\Sigma_1\times\Sigma_3)/Z_2$, which is a neighborhood of the point $\pi_M^{\mathbb{C}}(\tilde{x})=\pi_M(\tilde{x})\in M_c^4$.

Let us denote by $U:=\pi_M^{\mathbb{C}}(\tilde{U})\cap\mathcal{T}_4$ the intersection of $\pi_M^{\mathbb{C}}(\tilde{U})$ with the tube domain $\mathcal{T}_4$. It is clear that then $\pi_M^{\mathbb{C}}(\tilde{x})\in\partial U$. If there is given a holomorphic function $\bar{f}$ on $\mathcal{T}_4$, then $\bar{f}_{|U}$ is also a holomorphic function on U. But then also the function $f_{\tilde{V}}$ defined as

$$\tilde{f}_{\tilde{V}}:=\bar{f}\circ\pi^{\mathbb{C}}_{M|\tilde{V}}, \tag{61}$$

where $\tilde{V}=\tilde{V}_{\tilde{x}}:=\pi^{\mathbb{C}^{-1}}_{M|\tilde{U}}(U)$ is the preimage of U in the open set $\tilde{U}$, is a holomorphic function on $\tilde{V}_{\tilde{x}}$ and the point $\tilde{x}$ is an element of the boundary $\partial\tilde{V}_{\tilde{x}}$ of the open set $\tilde{V}_{\tilde{x}}$. In this way we can construct for every point $\tilde{x}\in\tilde{M}_4$ an open set $\tilde{V}_{\tilde{x}}\subset\mathbb{C}\tilde{M}_4$ such that for a given holomorphic function $\bar{f}$ on $\mathcal{T}_4$ there exists a function $f_{\tilde{V}}$ which is holomorphic on $\tilde{V}_{\tilde{x}}$ and which fulfills the relation (61). Furthermore, the point $\tilde{x}$ belongs to $\partial\tilde{V}_{\tilde{x}}$. It is easily seen that for any two $\tilde{V}$ and $\tilde{W}$ such that $\tilde{V}\cap\tilde{W}\neq\phi$ the functions $f_{\tilde{V}}$ and $f_{\tilde{W}}$ are identical on $\tilde{V}\cap\tilde{W}$.

Therefore, the two functions are analytic continuations of each other and there exists a holomorphic function $\tilde{f}$ on $\tilde{V}\cup\tilde{W}$ such that $\tilde{f}_{|\tilde{V}}=\tilde{f}_{\tilde{V}}$ and $\tilde{f}_{|\tilde{W}}=f_{\tilde{W}}$.

Let us next introduce the union of all the sets $\tilde{V}_{\tilde{x}}$ in $\mathbb{C}\tilde{M}_4$,

$$\tilde{\mathcal{T}}_4:=\bigcup_{\tilde{x}\in\tilde{M}_4}\tilde{V}_{\tilde{x}}. \tag{62}$$

By construction, $\tilde{\mathcal{T}}_4$ is an open set of the manifold $\mathbb{C}\tilde{M}_4$ and because it is connected it is a domain. It also follows immediately that there exists for every function $\bar{f}$ holomorphic in $\mathcal{T}_4$, a holomorphic function $\tilde{f}$ on $\tilde{\mathcal{T}}_4$ such that

$$\tilde{f}_{|\tilde{V}}=\bar{f}\circ\pi^{\mathbb{C}}_{M|\tilde{V}} \tag{63}$$

for all $\tilde{V}\subset\tilde{\mathcal{T}}_4$ such that $\pi^{\mathbb{C}}_{M|\tilde{V}}$ is a biholomorphic map. It is also clear from the construction that $\tilde{M}_4\subset\partial\tilde{\mathcal{T}}_4$.

We can also show that the function $\tilde{f}$ obeying the relation (63) is invariant under the action of the group of deck transformations of the covering $\pi_M^{\mathbb{C}}:\mathbb{C}\tilde{M}_4\to(\Sigma_1\times\Sigma_3)/Z_2$. If we denote this group by $\Gamma^{\mathbb{C}}$ we see that the group $\Gamma^{\mathbb{C}}$ is isomorphic to the group $\Gamma=\Gamma(\tilde{M}_4,M_c^4)$ and that its action restricted to $\tilde{M}_4$ is exactly the same as that of Γ.

If therefore $\tilde{x}\in\tilde{M}_4$ and $\tilde{V}_{\tilde{x}}\subset\tilde{\mathcal{T}}_4$ is an admissible neighborhood of $\tilde{x}$, then also $\gamma\tilde{V}_{\tilde{x}}$ is an admissible neighborhood for $\gamma\tilde{x}$ and we have $\gamma\tilde{x}\in\partial\gamma\tilde{V}_{\tilde{x}}$ for every $\gamma\in\Gamma^{\mathbb{C}}$. To show the invariance of the function $\tilde{f}$ under the action of the group $\Gamma^{\mathbb{C}}$ we have to show that

$$\tilde{f}_{|\tilde{V}}=(\gamma\tilde{f})_{|\tilde{V}} \tag{64}$$

for every $\gamma\in\Gamma^{\mathbb{C}}$ and every admissible open set $\tilde{V}$. In (64) the function $\gamma\tilde{f}$ is defined as follows:

$$(\gamma\tilde{f})(\tilde{y}):=\tilde{f}(\gamma^{-1}\tilde{y})\quad\text{for all }\tilde{y}\in\tilde{\mathcal{T}}_4\text{ and all }\gamma\in\Gamma^{\mathbb{C}}. \tag{65}$$

To show the property (64) let $\tilde{V}$ be any admissible open set and let $\tilde{y}\in\tilde{V}$. Then $\gamma^{-1}\tilde{y}\in\gamma^{-1}\tilde{V}$ and therefore because of (63) and property (46), which for $\gamma\in\Gamma^{\mathbb{C}}$ reads

$$\pi_M^{\mathbb{C}}\circ\gamma=\pi_M^{\mathbb{C}},$$

we have

$$\tilde{f}(\gamma^{-1}\tilde{y})=\bar{f}\circ\pi_M^{\mathbb{C}}(\gamma^{-1}\tilde{y})=\bar{f}\circ\pi_M^{\mathbb{C}}(\tilde{y})=\tilde{f}(\tilde{y}).$$

Therefore, we get $\tilde{f}_{|\tilde{V}}=(\gamma\tilde{f})_{|\tilde{V}}$, a property which is in complete analogy to the relation (47) for the distribution $\tilde{T}$ on $\tilde{M}_4$.

We want to finally show in this section that if the function $\bar{f}$ is an element of the set $\mathcal{S}'_{\mathrm{loc}}(\mathcal{T}_4)$, then also the function $\tilde{f}$ which is holomorphic on $\tilde{\mathcal{T}}_4$ has a boundary value $\tilde{T}$ on the manifold $\tilde{M}_4$ in the sense of distributions on $\tilde{M}_4$ which has the property (45), where $\bar{T}$ is the boundary value of the function $\bar{f}$ on M_c^4.

For this we write the relation (63) as follows:

$$\tilde{f}_{|\tilde{V}}=\pi^{\mathbb{C}}_{M|\tilde{V}}{}^{-1}\bar{f}_{|V}, \tag{66}$$

where $V=\pi_M^{\mathbb{C}}(\tilde{V})$.

If we now perform the limit $\tilde{z}\in\tilde{V}$ going to the real boundary of $\tilde{V}$, the point $\pi_M^{\mathbb{C}}(\tilde{z})$ approaches the real boundary B of $V\subset\mathcal{T}_4$ which is a subset of the real manifold M_c^4. If therefore W is an open subset contained in B we know, because $\bar{f}\in\mathcal{S}'_{\mathrm{loc}}(\mathcal{T}_4)$, that the right-hand side of (66) tends to $\pi_{M|\tilde{W}}^{-1}\bar{T}_{|W}$, W is the image of a certain open subset $\tilde{W}$ of the real boundary of the set $\tilde{V}$ under the mapping π_M. But this is exactly the definition of the distribution $\tilde{T}$ on the open subset $\tilde{W}\subset\tilde{M}_4$ as given in (45).

Summarizing our results, we have shown that every tempered distribution $T\in\mathcal{S}'(M_4)$ which is the boundary value of a function f holomorphic in the tube domain $\mathcal{T}_4$, can be uniquely extended to a distribution $\tilde{T}$ on the manifold $\tilde{M}_4$ which is again the boundary value of a function $\tilde{f}$ holomorphic in $\tilde{\mathcal{T}}_4$, which is invariant under the action of the group of deck transformations of the covering $\pi_M^{\mathbb{C}}:\mathbb{C}\tilde{M}_4\to(\Sigma_1\times\Sigma_3)/Z_2$. This also induces invariance of the distribution T under the action of the group of deck transformations of the covering $\pi_M:\tilde{M}_4\to M_c^4$.

It is straightforward that all that we have done for the space M_4 can be immediately written down and with almost no change for any direct product $M_4^n=M_4\times\cdots\times M_4$. The tube domain $\mathcal{T}_4$ then becomes the domain $\mathcal{T}_4^n=\mathcal{T}_4\times\cdots\times\mathcal{T}_4$ and all steps can be repeated for this domain.

As a result of this, we then get that the Wightman distributions $W_n(\xi_1,\ldots,\xi_n)$ where $\xi_i=x_{i+1}-x_i$, which are tempered distributions in $\mathcal{S}'(M_4^n)$ and boundary values of functions holomorphic in the tube domain $\mathcal{T}_4^{n\,2)}$, can be uniquely extended to distributions in $\mathcal{D}'(M_4^n)$ which are on M_4^n boundary values of functions holomorphic in the domain $\tilde{\mathcal{T}}_4^n$.

ACKNOWLEDGMENTS

I am very much indebted to Professors N. Kuiper and L. Michel for the kind hospitality extended to me at the I.H.E.S. I want to thank Professors L. Michel,

H. Epstein, H. Kastrup, and M. Oka for helpful discussions. This work was done under a fellowship of the Deutsche Forschungsgemeinschaft.

APPENDIX A: THEOREM OF ASCOLI

In this Appendix we shall repeat some definitions and the theorem of Ascoli which we applied in Sec. I of this paper.

Let F be a Banach space with norm $\| \ \|_F$ and E a metric space. We denote by $C_F(E)$ the space of all continuous mappings from E into F. Let H be a subset of $C_F(E)$.

Definition: H is called equicontinuous on the space E if for every $x \in E$ and for every $\epsilon > 0$ there exists a $\delta > 0$ such that $\|f(x) - f(y)\|_F < \epsilon$ for all $y \in E$ with $\text{dist}(x, y) < \delta$ and for all $f \in H$.

We call H uniformly equicontinuous if there exists for every $\epsilon > 0$ a $\delta > 0$ such that $\|f(x) - f(y)\|_F < \epsilon$ for all $x, y \in E$ with $\text{dist}(x, y) < \delta$ and all $f \in H$.

The theorem of Ascoli gives a simple criterion for H to be an equicontinuous set.[13]

Theorem (Ascoli): Let F be a Banach space, E a compact metric space, $H \subset C_F(E)$. If H is relatively compact then H is equicontinuous.

APPENDIX B: A LEMMA

Lemma: Let $\mathcal{T}_4$ be the field theoretic tube domain and let $f(z)$ be a holomorphic function on $\mathcal{T}_4$. Assume there exist constants $C', \alpha', \beta' \geq 0$ such that $|f(z)| \leq C'(1 + |z|^2)^{\alpha'} \ \text{dist}(z, \partial\mathcal{T}_4)^{-\beta'}$. Then there exist constants $C, \alpha, \beta \geq 0$ such that

$$|Rf(z)| \leq C(1 + |z|^2)^{\alpha} \, \text{dist}(z, \partial\mathcal{T}_4)^{-\beta}$$

where $Rf(z) = f(Rz)$.

Proof: Let $z = x + iy$ with $y \in V_+$. Then we have $\text{dist}(z, \partial\mathcal{T}_4) = c(y^0 - |\mathbf{y}|)$, where c is some constant we are not interested in. Now let $z' = Rz = x' + iy'$. Using the definition of the operator R we get

$$x' = \frac{-x(x^2 - y^2) - 2y(x \cdot y)}{(x^2 - y^2)^2 + 4(x \cdot y)^2}, \quad y' = \frac{-y(x^2 - y^2) + 2x(x \cdot y)}{(x^2 - y^2)^2 + 4(x \cdot y)^2}. \tag{B1}$$

For the component $y^{0\prime}$, this reads

$$y^{0\prime} = \frac{-y^0(x^2 - y^2) + 2x^0(x \cdot y)}{(x^2 - y^2)^2 + 4(x \cdot y)^2}. \tag{B2}$$

Let us first look at the numerator N of the expression (B2),

$$N = -y^0(x^2 - y^2) + 2x^0(xy).$$

This can be written as

$$N = y^0 y^2 + y^0 x_0^2 + y^0 |\mathbf{x}|^2 - 2x^0 \mathbf{x} \cdot \mathbf{y}. \tag{B3}$$

For the space part of the vector y', we get from (B1)

$$\mathbf{y}' = \frac{-\mathbf{y}(x^2 - y^2) + 2\mathbf{x}(x \cdot y)}{(x^2 - y^2)^2 + 4(x \cdot y)^2}. \tag{B4}$$

Let us again look at the numerator $\mathbf{N}$ of expression (B4). It can be written as

$$\mathbf{N} = \mathbf{y} y^2 + \mathbf{M}, \tag{B5}$$

where $\mathbf{M} = 2\mathbf{x}x^0y^0 - 2\mathbf{x}(\mathbf{x} \cdot \mathbf{y}) - \mathbf{y}x_0^2 + \mathbf{y}|\mathbf{x}|^2$.

We claim that the following inequality is true:

$$y^0 x_0^2 + y^0 |\mathbf{x}|^2 - 2x^0 \mathbf{x} \cdot \mathbf{y} \geq |\mathbf{M}|. \tag{B6}$$

This is true because this inequality is equivalent to the following:

$$(y_0^2 - |\mathbf{y}|^2)(x_0^4 + |\mathbf{x}|^4 - 2|\mathbf{x}|^2 x_0^2) \geq 0,$$

which is true because $y_0^2 > |\mathbf{y}|^2$.

Therefore, we get from (B6) and (B3),

$$y^{0\prime} - |\mathbf{y}| \geq \frac{y^0 y^2 - |\mathbf{y}| y^2 + (y^0 x_0^2 + y^0 |\mathbf{x}|^2 - 2x^0 \mathbf{x} \cdot \mathbf{y} - |\mathbf{M}|)}{(x^2 - y^2)^2 + 4(x \cdot y)^2},$$

and therefore

$$y^{0\prime} - |\mathbf{y}'| \geq \frac{(y^0 - |\mathbf{y}|)y^2}{(x^2 - y^2)^2 + 4(x \cdot y)^2}. \tag{B7}$$

Because $y^2 = y_0^2 - |\mathbf{y}|^2 \geq (y^0 - |\mathbf{y}|)^2$ and $(x^2 - y^2)^2 + 4(x \cdot y)^2 \leq 5|z|^4$ we get

$$y^{0\prime} - |\mathbf{y}'| \geq \tfrac{1}{5}|z|^{-4}(y^0 - |\mathbf{y}|)^3. \tag{B8}$$

To get an upper bound for $|z'|$ we only have to look at the denominator D of the expressions in (B1),

$$D = (x^2)^2 + (y^2)^2 - 2x^2y^2 + 4(x \cdot y)^2. \tag{B9}$$

Because

$$(x \cdot y)^2 \geq x^2 y^2 \tag{B10}$$

for all $z = x + iy$ with $y^2 > 0$ and $y^0 > 0$, we get for (B9),

$$D \geq (y^2)^2 \geq (y^0 - |\mathbf{y}|)^4. \tag{B11}$$

From (B11) and (B1) we then get

$$|z'| \leq 6|z|^3 (y^0 - |\mathbf{y}|)^{-4}. \tag{B12}$$

Putting together the inequalities (B8) and (B12) gives

$$(1 + |z'|^2)^{\alpha'} (y^{0\prime} - |\mathbf{y}'|)^{-\beta'} \leq C(1 + |z|^2)^{4\alpha' + 2\beta'} (y^0 - |\mathbf{y}|)^{-3\beta' - 8\alpha'},$$

where $C = 2^{8\alpha' + 1} \ 5^{\beta'}$. But this proves the lemma.

APPENDIX C: THE MANIFOLD $\mathbb{C}M_c^4$

We want to prove that the complex manifold $\mathbb{C}M_c^4$ defined in (22) is simply connected. It turns out that $\mathbb{C}M_c^4$ is a special case of a more general class of projective varieties which all share this property. The proof I will now give is essentially due to Oka.

Theorem: Let $\mathbb{C}^{n+2} = \{z = (z_0, \ldots, z_{n+1}), \ z_i \in \mathbb{C}\}$ and let be $\mathbb{C}\mathbb{P}^{n+1} = \{[z] = [(z_0, \ldots, z_{n+1})]\}$ be the $(n+1)$-dimensional complex projective space. Let $f: \mathbb{C}^{n+2} \to \mathbb{C}$ be a homogeneous polynomial. Let V be the following subset of $\mathbb{C}\mathbb{P}^{n+1}$: $V := \{[z] \in \mathbb{C}\mathbb{P}^{n+1} : f([z]) = 0\}$. Then V is simply connected.

Proof: Consider the Hopf fibering[22] $\pi: S^{2n+3} \to \mathbb{C}\mathbb{P}^{n+1}$, where π is the following mapping:

$$\pi: z \to [z] = [(z_0, \ldots, z_{n+1})], \tag{C1}$$

where $z \in S^{2n+3}$; that means $z = (z_0, \ldots, z_{n+1})$ fulfills the equation

$$|z_0|^2 + \ldots + |z_{n+1}|^2 = 1. \tag{C2}$$

From the definition (C1) of the mapping π it follows that

all points $e^{i\varphi}z \in S^{2n+3}$, where $0 \leq \varphi \leq 2\pi$, are mapped under π onto the same point $[z] \in \mathbb{C}\mathbb{P}^{n+1}$. Therefore, the fiber of the above map is just the one-dimensional unit sphere S_1. It is known that $(S^{2n+3}, \mathbb{C}\mathbb{P}^{n+1}, S_1, \pi)$ defines a fiber bundle. Next consider the set W defined as

$$W := \{z \in \mathbb{C}^{2n+2} : f(z) = 0\}. \tag{C3}$$

That means W is the kernel of the polynomial f in $\mathbb{C}^{2n+2}$. If we then define the set K by

$$K := W \cap S^{2n+3}, \tag{C4}$$

we get a restricted fibering[23]

$$\pi_{|K} : K \to V, \tag{C5}$$

whose fiber again is S_1, because f was homogeneous in z. If we now write down the exact homotopy sequence[24] for the fibering $\pi : K \xrightarrow{S_1} V$, we get

$$\begin{aligned} &\cdots \to \pi_k(S_1) \to \pi_k(K) \to \pi_k(V) \to \pi_{k-1}(S_1) \\ &\quad \to \pi_{k-1}(K) \to \pi_{k-1}(V) \cdots, \\ &\cdots \to \pi_1(S_1) \to \pi_1(K) \to \pi_1(V) \to \pi_0(S_1). \end{aligned} \tag{C6}$$

It is known from the work of Milnor[25] on singular points on complex hypersurfaces that the space K is $(n-1)$ connected which means[24] that

$$\pi_k(K) = 0 \quad \text{for all } k \leq n-1, \tag{C7}$$

where $\pi_k(K)$ denotes the kth homotopy group of the space K. Because S_1 is path connected we have $\pi_0(S_1) = 0$.

Therefore, we get from (C6) that

$$0 \to \pi_1(V) \to 0 \tag{C8}$$

is an exact sequence, but from this it immediately follows that $\pi_1(V) = 0$, which means V is simply connected if $n \geq 2$.

Because $\mathbb{C}M_c^4$ is just the space V for $n=4$, and f is the homogeneous polynomial $f(z) = z_0^2 - z_1^2 - \cdots - z_4^2 + z_5^2$, we get the desired result.

APPENDIX D: PRINCIPLE OF "RECOLLEMENT DES MORCEAUX"[26]

Let M be a differentiable manifold and let $\{U_\lambda, \lambda \in L\}$ be an open covering of M. Let us assume that for every $\lambda \in L$ there exists a distribution T_λ on U_λ with the property that for every pair $\lambda, \mu \in L$ the restrictions of T_λ and T_μ to the open set $U_\lambda \cap U_\mu$ are equal. Then there exists an unique distribution T on M such that for every $\lambda \in L$ the restriction of T to U_λ is exactly the distribution T_λ.

APPENDIX E: AN INEQUALITY

Let $w = (w_1, \ldots, w_n)$, $v = (v_1, \ldots, v_n) \in \mathbb{R}^n$ be two vectors with $|w|^2 = |v|^2$ and $w \cdot v = w_1 v_1 + \cdots + w_n v_n = 0$. Then we have $w_n^2 \leq v_1^2 + \cdots + v_{n-1}^2$.

Proof: Assume $w_n^2 > v_1^2 + \cdots + v_{n-1}^2$, then also $v_n^2 > w_1^2 + \cdots + w_{n-1}^2$, but then also because of the Schwarz inequality, the following is true:

$$v_n^2 w_n^2 > (v_1^2 + \cdots + v_{n-1}^2)(w_1^2 + \cdots + w_{n-1}^2) \geq (v_1 w_1 + \cdots + v_{n-1} w_{n-1})^2.$$

But this is a contradiction to $w \cdot v = 0$.

*Work supported by the Deutsche Forschungsgemeinschaft.
†On leave of absence from RWTH Aachen.

[1] M. Lüscher and G. Mack, Commun. Math. Phys. **41**, 203 (1975).
[2] D. H. Mayer, "Conformal Invariant Causal Structures on Pseudo-Riemannian Manifolds," preprint, Aachen (1974).
[3] T. H. Go, Commun. Math. Phys. **41**, 157 (1975).
[4] I. T. Todorov, CERN TH 1697 (1973).
[5] I. E. Segal, Bull. Am. Math. Soc. **77**, 958 (1971).
[6] J. Swieca and A. H. Völkel, Commun. Math. Phys. **29**, 319 (1973).
[7] B. Schroer and J. Swieca, "Conformal Transformations for Quantized Fields," preprint, Rio de Janeiro (1974).
[8] M. Hortaçsu, B. Schroer, and R. Seiler, Phys. Rev. D **5**, 2518 (1972).
[9] T. H. Go, H. A. Kastrup, and D. H. Mayer, "Properties of Dilatations and Conformal Transformations in Minkowski Space," Rep. Math. Phys. **6**, 395 (1974).
[10] T. H. Go and D. H. Mayer, Rep. Math. Phys. **5**, 187 (1974).
[11] L. Schwartz, *Théorie des Distributions* (Herman, Paris, 1966), p. 233–36.
[12] J. Dieudonné, *Eléments d'Analyse* (Gauthier-Villars, Paris, 1970), Vol. III, p. 230.
[13] J. Dieudonné, *Elêments d'Analyse* (Gauthier-Villars, Paris, 1969), Vol. I, p. 142–45.
[14] L. Hörmander, *Linear Partial Differential Operators* (Springer-Verlag, Berlin, 1963), p. 29.
[15] A. Martineau, "Distributions et valeurs au bord de fonctions holomorphes," in *Prepubl. de la Recherche Coopérative sur Programme n°* 25 (Centre National de la Recherche Scientifique, Strasbourg, 1967).
[16] R. Gunning and H. Rossi, *Analytic Functions of Several Complex Variables* (Prentice-Hall, Englewood Cliffs, New Jersey, 1965), p. 149, 151.
[17] V. S. Vladimirov, *Methods of the Theory of Functions of Several Complex Variables* (Cambridge, Mass., 1966).
[18] H. A. Kastrup, Ann. Phys. **9**, 388 (1962).
[19] See Ref. 12, p. 239.
[20] I. R. Shafarevich, *Basic Algebraic Geometry* (Springer-Verlag, Berlin, 1974), p. 380.
[21] R. F. Streater and A. S. Wightman, *PCT, Spin and Statistics, and All That* (Benjamin, New York, 1964), p. 114.
[22] P. Hilton, *An Introduction to Homotopy Theory* (Cambridge U. P., Cambridge, 1961), p. 53.
[23] M. Oka, "On the Cohomology Structure of Projective Varieties," to appear in Proc. Int. Conf. on Manifolds and Related Topics in Topology, Tokyo (1973).
[24] S. Hu, *Homotopy Theory* (Academic, New York, 1959), p. 152.
[25] J. Milnor, "Singular Points of Complex Hypersurfaces," in Ann. Math. Studies 61 (Princeton U. P., 1969), p. 45.
[26] See Ref. 12, p. 238.

Approximating functions with a given singularity

Pierre Hillion

Institut Henri Poincaré, Paris, France
(Received 17 February 1976; revised manuscript received 27 April 1976)

In this paper, assuming that one knows one of the singularities s_1 of the function $f(z)$ and its power series expansion on a domain D of the complex plane, we introduce some sequences of Gammel–Baker generalized Padé approximants with the same type of singularity s_1. Two examples are given: One concerns the convergence acceleration of approximations for functions with a logarithmic singularity; in the other one, a physical application to optical polarizability is discussed.

1. INTRODUCTION

Let us consider the power series $\sum_{n=0}^{\infty} f_n z^{-n}$, convergent for every $z \in D_f$ in the complex plane $\mathbb{C}$. In this paper, we assume that one of the singularities of the function $f(z)$ with the Taylor expansion $\sum_0^{\infty} f_n z^{-n}$ on D_f is known and we use this extra information to build sequences of Gammel–Baker generalized Padé approximants with the same type of singularities. We also give an example of approximating functions able to cope with two known singularities of $f(z)$.

2. DEFINITION OF THE NEW APPROXIMATING FUNCTIONS

A. Convolution and approximation

Let us consider the following equality:

$$f(z) = k(z) \circledast g(z) = \frac{1}{2\pi i} \int_{\Gamma_k} t^{-1} k(t) g(z/t)\, dt\,, \tag{1}$$

where Γ_k is a contour enclosing all the singularities of the function $t^{-1}k(t)$. We now prove that Eq. (1) defines the complex convolution of the kernel $k(t)$ with the function $g(t)$.

Lemma 1: If

(i) $k(t) = \sum_{n=0}^{\infty} k_n t^{-n}$

$$\text{for } t \in D_k = \{t: |t| > \sigma_k = \lim_{n\to\infty} \sup(|k_n|)^{1/n}\}\,, \tag{2}$$

(ii) $g(t) = \sum_{n=0}^{\infty} g_n t^{-n}$

$$\text{for } t \in D_g = \{t: |t| > \sigma_g = \lim_{n\to\infty} \sup(|g_n|)^{1/n}\}, \tag{3}$$

then

$$f(t) = \sum_{n=0}^{\infty} k_n g_n t^{-n}$$

$$\text{for } t \in D_f = \{t: |t| > \sigma_f = \lim_{n\to\infty} \sup(|k_n g_n|)^{1/n}\}. \tag{4}$$

Indeed, from (1), D_k is inside the domain with boundary Γ_k, and including the origin of the z^{-1} plane, we can then deform the contour of integration in (1) so that it lies in D_k and it is now possible to choose z such that for $t \in \Gamma_k$, we have $z/t \in D_g$. Under these conditions, we can substitute the series (2) and (3) in (1) and this gives

$$f(z) = \frac{1}{2\pi i} \int_{\Gamma_k} t^{-1} \left(\sum_{n=0}^{\infty} k_n t^{-n}\right)\left(\sum_{i=0}^{\infty} g_i z^{-i} t^{i}\right) dt$$

$$= \sum_{n=0}^{\infty} k_n g_n z^{-n}$$

and, from the Cauchy criterion, this series is convergent for $z \in D_f$.

Let $g^{n,m}(t)$ be a Padé approximant of the power series $\sum_{j=0}^{\infty} g_j t^{-j}$ and let $\tilde{f}^{n,m}(z)$ be the following expression:

$$\tilde{f}^{n,m}(z) = \frac{1}{2\pi i} \int_{\Gamma_k} t^{-1} k(t) g^{n,m}(z/t)\, dt\,. \tag{5}$$

Lemma 2: $\tilde{f}^{n,m}(z)$ is a form of Gammel–Baker approximant[1] for a product kernel. Making a partial fraction decomposition of $g^{n,m}(z/t)$ in (5) gives, for z sufficiently large,

$$\tilde{f}^{n,m}(z) = \sum_{i=1}^{n} \alpha_i k(z/\beta_i) \tag{6}$$

and the result follows by analytic continuation to smaller z. We give now some examples of (1) and (5) for different kernels $k(z)$.

Example 1. $k_1(t) = (1 - t^{-1})^{-1}$, $f_1(z) = g(z)$,

$$\tilde{f}_1^{n,m}(z) = g^{n,m}(z)\,, \tag{7a}$$

Example 2. $k_2(t) = t(t^{-1})^{-2}$, $f_2(z) = \left[\frac{d}{dt} g\left(\frac{z}{t}\right)\right]_{t=1}$,

$$\tilde{f}_2^{n,m}(z) = \left[\frac{d}{dt} g^{n,m}\left(\frac{z}{t}\right)\right]_{t=1}, \tag{7b}$$

Example 3. $k_3(t) = \frac{t(t+1)}{(t-1)^3}$, $f_3(z) = \frac{1}{2}\left[\frac{d^2}{dt^2}(t+1) g\left(\frac{z}{t}\right)\right]_{t=1}$,

$$\tilde{f}_3^{n,m}(z) = \frac{1}{2}\left[\frac{d^2}{dt^2}(t+1) g^{n,m}\left(\frac{z}{t}\right)\right]_{t=1}. \tag{7c}$$

The Padé approximants correspond to the particular kernel $k_1(t)$. It is easy to build similar examples but the following ones run in opposite direction since they require integrations rather than derivatives.

Example 4. $k_4(t) = \log(1 - t^{-1})^{-1}$, $f_4(z) = \int_z^{\infty} t^{-1} g(t)\, dt$,

$$\tilde{f}_4^{n,m}(z) = \int_z^{\infty} t^{-1} g^{n,m}(t)\, dt\,, \tag{8a}$$

Example 5. $k_5(t)=\int_t^\infty p\log\frac{p}{1-p}dp$,

$$f_5(z)=\int_z^\infty q^{-1}\int_q^\infty t^{-1}g(t)\,dtdq,$$

$$\tilde f_5^{n,m}(z)=\int_z^\infty q^{-1}\int_q^\infty t^{-1}g^{n,m}(t)\,dtdq, \qquad (8b)$$

Example 6. $k_6(t)=t^\beta\int_t^\infty\frac{p^{-\beta}}{p-1}dp$, $f_6(z)=z^\beta\int_z^\infty t^{-(1+\beta)}g(t)\,dt$,

$$\tilde f_6^{n,m}(z)=z^\beta\int_z^\infty t^{-(1+\beta)}g^{n,m}(t)\,dt,\quad \beta>0. \qquad (8c)$$

Lemma 3: If $k(t)$ has, at the point $t=1$, a singularity of the s type, then every simple pole of $g^{n,m}(t)$ gives a singularity of $\tilde f^{n,m}(z)$ of the s type. Indeed, we can also write (1) in the form (provided that the integral along the infinite circle is zero)

$$f(z)=g(z)\circledast k(z)=\frac{1}{2\pi i}\int_{\Gamma_g}t^{-1}g(t)k(z/t)\,dt$$

and, in the same way, we have

$$\tilde f^{n,m}(z)=\frac{1}{2\pi i}\int_{\Gamma_{g^{n,m}}}t^{-1}g^{n,m}(t)k(z/t)\,dt.$$

The result then follows from the partial fraction expansion of $g^{n,m}(t)$. The six previous examples give a good illustration of Lemma 3. For the properties of the Gammel–Baker generalized Padé approximants see Ref. 1.

B. New approximating function

Let $\sum_{n=0}^\infty f_nz^{-n}$ be the power series expansion on D_f of a function $f(z)$ which has a singularity s_1 at the point z_1 of the complex plane and let $k(z)$ be a kernel with the same type of singularity at the point $z=1$ and such that $k(z)=\sum_{n=0}^\infty k_nz^{-n}$ for $z\in D_k$. Now, if we choose for $g(z)$, the function with the power series expansion $\sum_{n=0}^\infty f_nk_n^{-1}z^{-n}$, then from (1) and (4) we have $f(z)=k(z)\circledast g(z)$ and from (5), the function $\tilde f^{n,m}(z)=(1/2\pi i)\int_{\Gamma_k}t^{-1}\times k(t)g^{n,m}(z/t)\,dt$, where $g^{n,m}(z)$ is a Padé approximant of $g(z)$, is an approximation of $f(z)$ which according to Lemma 3 has a singularity of s_1 type at every z_i where z_i is a simple pole of $g^{n,m}(z)$.

Lemma 4: If $k(z)=f(z)$, then $\tilde f^{n,m}(z)=f(z)$ for every integer n,m with $n\neq0$. Indeed, for $k(z)=f(z)$, we have $g(z)=(1-z^{-1})^{-1}$ and as it is known, for this particular $g(z)$, $g^{n,m}(z)=g(z)\ \forall\ n,m,\ n\neq0$.

The case where the known singularity of $f(z)$ is a pole of multiplicity $m=1,2,\ldots$ is not difficult and the examples in (7) show the type of kernel (2) to use. From now, we consider functions $f(z)$ with one essential singularity and kernels of the type

$$k_\beta(t)=t^\beta\int_t^\infty\frac{p^{-\beta}}{p-1}dp,\quad \beta>0,\quad k_0(t)=\log(1-t^{-1})^{-1},$$

where β is an arbitrary positive real number.

Using (8a), (8c), and the previous definition of $g(z)$, it is easy to prove that the approximants $\tilde f^{n,m}(z)$ are

$$\tilde f_\beta^{n,m}(z)=\int_0^1\left[\frac{\partial}{\partial t}\left(t^\beta f\left(\frac{z}{t}\right)\right)\right]^{n,m}dt+\delta_{\beta0}f_0,\quad \beta\geq0, \qquad (9)$$

where the integrand denotes the $[n,m]$ Padé approximant of the power series expansion of the derivative $(\partial/\partial t)\times[t^\beta f(z/t)]$, $\delta_{\beta0}$ being the Kronecker symbol.

Remark: Let us assume that one knows r singularities $s_1,s_2,\ldots,s_r$, of $f(z)$ and that one can write the power series $f(z)=\sum_{n=0}^\infty f_nz^{-n}$ as $f(z)=\sum_{i=1}^r f_i(z)$ with $f_i(z)=\sum_{k=0}^\infty f_{k,i}z^{-k}$ such that $f_i(z)$ has a singularity of the s type, then the previous results can be applied to every component $f_i(z)$ of $f(z)$. We will give an example of such a possibility with $r=2$ in Sec. 4.

3. APPLICATION TO CONVERGENCE ACCELERATION

A. Numerical tests

We consider two particular functions with a logarithmic singularity and we compute from (9) $\tilde f_\beta^{1,0}(z)$ and $\tilde f_\beta^{1,1}(z)$. Let us begin with

$$f(z)=z^2\{(1-z^{-1})\log(1-z^{-1})+z^{-1}\}=\sum_{n=1}^\infty\frac{1}{n(n+1)}z^{-(n-1)}.$$

The logarithmic singularity of $f(z)$ suggests that we take $\beta=1$ in (9), so we have

$$f^{1,0}(z)=\frac{1}{2}\left(1-\frac{1}{3z}\right)^{-1},\quad \left[\frac{\partial}{\partial t}\left(tf\left(\frac{z}{t}\right)\right)\right]^{1,0}=\frac{1}{2}\left(1-\frac{2t}{3z}\right)^{-1},$$

$$\tilde f_1^{1,0}(z)=-\frac{3z}{4}\log\left(1-\frac{2}{3z}\right),$$

$$f^{1,1}(z)=\frac{1/2-1/12z}{1-1/2z},\quad \left[\frac{\partial}{\partial t}\left(tf\left(\frac{z}{t}\right)\right)\right]^{1,1}=\frac{1}{2}\frac{1-t/12z}{1-3t/4z},$$

$$\tilde f_1^{1,1}(z)=\frac{1}{18}\left[1-\frac{32z}{3}\log\left(1-\frac{3}{4z}\right)\right].$$

Table I allows us to compare the Padé approximants and the approximants in (9) with $f(z)$ for some values of z^{-1}.

We now consider

$$f(z)=\frac{(1-z^{-2})\log[(1-z^{-1})/(1+z^{-1})]+2z^{-1}}{4z^{-3}}$$

$$=\sum_{n=1}^\infty\frac{1}{(2n-1)(2n+1)}z^{-2(n-1)},$$

still with $\beta=1$, it becomes

$$f^{1,0}(z)=\frac{1}{3}\left(1-\frac{1}{5z^2}\right)^{-1},$$

$$\tilde f_1^{1,0}(z)=\frac{1}{6}\left(\frac{5}{3}\right)^{1/2}\log\left(\frac{1+(3/5)^{1/2}z^{-1}}{1-(3/5)^{1/2}z^{-1}}\right),$$

$$f^{1,1}(z)=\frac{\frac{1}{3}-8/105z^2}{1-3/7z^2},$$

$$\tilde f_1^{1,1}(z)=\frac{4}{75}+\frac{7}{50}\left(\frac{7}{5}\right)^{1/2}\log\left(\frac{1+(5/7)^{1/2}z^{-1}}{1-(5/7)^{1/2}z^{-1}}\right).$$

These results (see Table II) are consistent with the conjecture of the previous section.

TABLE I.

z^{-1}	$f^{1,0}(z)$	$\tilde{f}_1^{1,0}(z)$	$f^{11}(z)$	$\tilde{f}_1^{1,1}(z)$	$f(z)$
1	0.750	0.829	0.833	0.877	1
−1	0.375	0.383	0.388	0.387	0.386
−2	0.300	0.318	0.333	0.327	0.324
−1	0.600	0.608	0.638	0.612	0.613

B. Numerical approximation of Eq. (9)

The integral in (9) is of the type $If = \int_0^1 x^r f(x)\,dx$ ($r = \beta - 1$) and it requires in most cases a numerical approximation that we now discuss.

Let us consider the following integral:

$$I^{r,s}f = \int_0^1 x^r(1-x)^s f(x)\,dx\,, \quad r, s > -1.$$

As it is well known,[2] an approximation $I_d^{r,s}f$ of $I^{r,s}f$ of degree $d = 2n-1$ (that is exact for the polynomials of degree $d \le 2n-1$) is

$$I_d^{r,s}f = \sum_{i=1}^{n} W_{i,n}^{r,s} f(t_{i,n}^{r,s})\,, \quad d = 2n-1\,, \tag{10}$$

where:

1. The abscissas $t_{i,n}^{r,s}$ are the real roots in the interval $(0,1)$ of the Jacobi polynomial $H_n^{r,s}(x)$ of degree n, such that

$$\int_0^1 x^r(1-x)^s H_n^{r,s}(x) H_m^{r,s}(x)\,dx = \delta_{nm}\,,$$

with[3]

$$H_n^{r,s}(x) = N_n^{r,s} x^{-r}(1-x)^{-s}\frac{d^n}{dx^n}\left[x^{r+n}(1-x)^{s+n}\right],$$

$$N_n^{r,s} = B^{-1/2}(r+1, s+1) \times\left(\frac{(r+s+1)(r+s+2)\cdots(r+s+n)}{(r+1)(r+2)\cdots(r+n)(s+1)(s+2)\cdots(s+n)}\right)^{1/2} \times\left(\frac{r+s+1+2n}{n!}\right)^{1/2}.$$

2. The weights $W_{i,n}^{r,s}$ ($i = 1, 2, \ldots, n$) are the Christoffel numbers

$$W_{i,n}^{r,s} = 1\Big/\sum_{\nu=0}^{n-1}\left(H_\nu^{r,s}(t_{i,n}^{r,s})\right)^2\,, \quad i = 1, 2, \ldots, n.$$

For instance, for $r = \beta$, $s = 1$, we have

$$H_0^{\beta,1}(x) = [(\beta+1)(\beta+2)]^{1/2}\,,$$

$$H_1^{\beta,1}(x) = \left(\frac{(\beta+2)(\beta+4)}{2}\right)^{1/2}[\beta+1-(\beta+3)x]\,,$$

$$H_2^{\beta,1}(x) = \left(\frac{(\beta+3)(\beta+6)}{3!}\right)^{1/2}\{x^2[(\beta+1)(\beta+2)+6(\beta+2)+6] - 2x(\beta+2)(\beta+4)+(\beta+1)(\beta+2)\}\,,$$

and this gives

$$t_{1,1}^{\beta,1} = (\beta+1)(\beta+3)^{-1}\,, \quad W_{1,1}^{\beta,1} = (\beta+1)^{-1}(\beta+2)^{-1}\,, \tag{11a}$$

$$t_{1,2}^{\beta,1} = \frac{(\beta+2)(\beta+4)+[3(\beta+2)(\beta+4)]^{1/2}}{(\beta+1)(\beta+2)+6(\beta+2)+6}\,,$$

$$t_{2,2}^{\beta,1} = \frac{(\beta+2)(\beta+4)-[3(\beta+2)(\beta+4)]^{1/2}}{(\beta+1)(\beta+2)+6(\beta+2)+6}\,, \tag{11b}$$

$$W_{i,2}^{\beta,1} = \{(\beta+1)(\beta+2) + (\beta+1)(\beta+2)[\beta+1-(\beta+3)t_{i,2}^{\beta,1}]^2/2\}^{-1}\,, \quad i = 1, 2.$$

We hope to publish in the near future a program for computing any $t_{i,n}^{r,s}$ and $W_{i,n}^{r,s}$ (see Ref. 4). In many cases, it is useful to include in the approximation of $I^{r,s}f$ the values of the function at the ends of the interval.

Lemma 5: The following expression including $f(0)$ is of degree $d = 2n$:

$$I_{2n}^{r,s}f = W_{0,n}^{r+1,s}f(0) + \sum_{i=1}^{n}\frac{W_{i,n}^{r+1,s}}{t_{i,n}^{r+1,s}} f(t_{i,n}^{r+1,s})\,, \tag{12}$$

$$W_{0,n}^{r+1,s} = B(r+1, s+1) - \sum_{i=1}^{n}\frac{W_{i,n}^{r+1,s}}{t_{i,n}^{r+1,s}}\,. \tag{12'}$$

[$B(r,s)$ denotes the usual beta function.] Indeed, we have, assuming $f \in C^\infty(0,1)$,

$$I^{r,s}f = \int_0^1 x^r(1-x)^s f(x)\,dx = \sum_{j=0}^{\infty} B(r+j+1, s+1) f^{(j)}(0)/j!\,, \tag{13}$$

and from (10),

$$I_{2n-1}^{r,s}f = \sum_{j=0}^{\infty}[f^{(j)}(0)/j!]\sum_{k=1}^{n} W_{k,n}^{r,s}(t_{k,n}^{r,s})^j\,. \tag{13'}$$

The comparison between (13) and (13') gives

$$\sum_{k=1}^{n} W_{k,n}^{r,s}(t_{k,n}^{r,s})^j = B(r+j+1, s+1)\,, \quad j = 0, 1, 2, \ldots, 2n-1\,. \tag{14}$$

Now, from (12) it follows that

$$I_{2n}^{r,s}f = W_{0,n}^{r,s}f(0) + \sum_{j=0}^{\infty}\frac{f^{(j)}(0)}{j!}\sum_{k=1}^{n} W_{k,n}^{r+1,s}(t_{k,n}^{r+1,s})^{j-1}. \tag{15}$$

The right-hand side of (14) being invariant under the transformation $r \mapsto r+1$, $j \mapsto j-1$, we have the equalities

$$\sum_{k=1}^{n} W_{k,n}^{r+1,s}(t_{k,n}^{r+1,s})^{j-1} = B(r+j+1, s+1)\,, \quad j = 1, 2, \ldots, 2n. \tag{14'}$$

The comparison between (13) and (15), taking into account (12'), (14') completes this lemma. The assumption $f \in C^\infty(0,1)$ was made only in order to simplify the proof.

Lemma 6: The following approximation including $f(1)$ is of degree $d = 2n$.

TABLE II.

z^{-1}	$f^{1,0}(z)$	$\tilde{f}_1^{1,0}(z)$	$f^{1,1}(z)$	$\tilde{f}_1^{1,1}(z)$	$f(z)$
±1	0.416	0.450	0.444	0.464	0.500
$\pm\frac{1}{2}$	0.350 87	0.352 00	0.351 68	0.352 05	0.352 08

$$I_{2n}^{r,s}f = W_{0,n}^{r,s+1}f(1) + \sum_{i=1}^{n} \frac{W_{i,n}^{r,s+1}}{1 - t_{i,n}^{r,s+1}} f(t_{i,n}^{r,s+1}) , \tag{16}$$

$$W_0^{r,s+1} = B(r+1, s+1) - \sum_{i=1}^{n} \frac{W_{i,n}^{r,s+1}}{1 - t_{i,n}^{r,s+1}} . \tag{16'}$$

The proof is the same as for Lemma 5 either by using a Taylor series expansion near $x = 1$ or by changing x into $(1 - x)$ and by interchanging r and s.

Lemma 7: The following approximation including $f(0)$ and $f(1)$ is of degree $d = 2n + 1$:

$$I_{2n+1}^{r,s} f = W_{0,n;0}^{r+1,s+1} f(0) + W_{0,n;1}^{r+1,s+1} f(1) + \sum_{i=1}^{n} \frac{W_{i,n}^{r+1,s+1}}{t_{i,n}^{r+1,s+1}(1 - t_{i,n}^{r+1,s+1})} f(t_{i,n}^{r+1,s+1}) , \tag{17}$$

$$W_{0,n;0}^{r+1,s+1} + W_{0,n;1}^{r+1,s+1} = B(r+1, s+1) - \sum_{k=1}^{n} \frac{W_{k,n}^{r+1,s+1}}{t_{k,n}^{r+1,s+1}(1 - t_{k,n}^{r+1,s+1})} , \tag{17'}$$

$$W_{0,n;1}^{r+1,s+1} = B(r+1, s+1) - \sum_{k=1}^{n} \frac{W_{k,n}^{r+1,s+1}}{1 - t_{k,n}^{r+1,s+1}} . \tag{17''}$$

The equality of both Taylor expansions (17) and (13) for $j = 0$, $j = 1$ is trivial, taking into account (17') and (17''); for $j \geq 2$, the coefficient of $f^{(j)}(0)/j!$ in (17) is

$$W_{0,n;1}^{r+1,s+1} + \sum_{k=1}^{n} \frac{W_{k,n}^{r+1,s+1}(t_{k,n}^{r+1,s+1})^{j-1}}{1 - t_{k,n}^{r+1,s+1}}$$

$$= B(r+2, s+1) - \sum_{k=1}^{n} W_{k,n}^{r+1,s+1} \frac{(1 - t_{k,n}^{r+1,s+1})^{j-1}}{1 - t_{k,n}^{r+1,s+1}}$$

$$= B(r+2, s+1) - \sum_{l=0}^{j-2} \sum_{k=1}^{n} W_{k,n}^{r+1,s+1} (t_{k,n}^{r+1,s+1})^{l}$$

$$= B(r+2, s+1) - \sum_{l=0}^{j-2} B(r+l+2, s+2) ,$$

$$j = 2, 3, \ldots, 2n+1 ,$$

where we used relations (14). The result follows then from the equality.

$$B(r+2, s+1) - \sum_{l=0}^{j-2} B(r+l+2, s+2) = B(r+j+1, s+1) ,$$

$$j = 2, 3, \ldots, 2n+1$$

is easy to prove, since

$$\int_0^1 x^{r+1}(1-x)^s dx$$

$$= \int_0^1 x^{r+j}(1-x)^s dx + \sum_{l=0}^{j-2} \int_0^1 x^{r+l+1}(1-x)^s dx$$

$$- \sum_{l=0}^{j-2} \int_0^1 x^{r+l+2}(1-x)^s dx$$

$$= \int_0^1 x^{r+j}(1-x)^s dx + \sum_{l=0}^{j-2} \int_0^1 x^{r+l+1}(1-x)^{s+1} dx .$$

We now come back to (9) which can be written $\int_0^1 t^{\beta} \varphi^{n,m}(z/t)\, dt$. Since, for z^{-1} little enough, the integrand differs from a polynomial to the order $n+m+1$, we have to take a quadrature formula of degree $d > n+m+1$ so as not to change the order of approximation.

As an illustration, let us consider the approximant $\tilde{f}^{1,1}(z)$ in Example 1 of Sec. 3 for $z^{-1} = 1$,

$$\tilde{f}^{1,1}(1) = \frac{1}{6} \int_0^1 \frac{12 - t}{4 - 3t}\, dt = 0.8770 . \tag{18}$$

Formula (17) for $r = s = 0$, $n = 1$ ($t_{1,1}^{1,1} = \frac{1}{2}$, $W_{1,1}^{1,1} = 1/6$, $W_{0,1,0}^{11} = W_{0,1,1}^{11} = 1/6$) is $I_3^{00}f = (1/6)f(0) + (1/6)f(1) + (2/3)f(\frac{1}{2})$, and applied to (18) it becomes $\tilde{f}^{1,1}(1) = 0,9$. For $n = 2$, we have

$$I_5^{0,0}f = 0.0834[f(0) + f(1)] + 0.4166[f(0.2764) + f(0.7236)]$$

and we obtain $\tilde{f}^{1,1}(1) = 0.8793$.

4. APPLICATION TO OPTICAL POLARIZABILITY

We do not intend here to make a contribution to the theory of optical polarizability, but only to use this theory in order to show an interesting application of the new approximants considered in the previous sections. All the data are borrowed from a paper by Langshoff and Karplus.[5]

A. Summary of the problem

The dynamic dipole polarizability of an atomic system is defined by the Kramers–Heisenberg formula (P means the Cauchy principal value),

$$\alpha(z) = P \int_0^{\infty} \frac{df/d\epsilon}{\epsilon^2 - z^2}\, d\epsilon ,$$

$df/d\epsilon = \sum_{n=0}^{\infty} f_{n0}\delta(\omega_{n0} - \epsilon) + dg/d\epsilon$, $dg/d\epsilon$ corresponding to the continuum part of the spectrum with a threshold at ω_{∞}. Thus, on the real positive axis, $\alpha(z)$ has poles for $0 < \omega < \omega_{\infty}$ and a cut for $\omega \geq \omega_{\infty}$. Let μ_k be the Cauchy moments

$$\mu_k = \int_0^{\infty} \epsilon^{-(2k+2)} \frac{df}{d\epsilon}\, d\epsilon = \sum_{n=1}^{\infty} \frac{f_{n0}}{\omega_{n0}^{2k+2}} + \int_{\omega_{\infty}}^{\infty} \frac{dg/d\epsilon}{\omega(\epsilon)^{2k+2}}\, d\epsilon$$

for $\omega < \omega_0$. $\alpha(\omega^2)$ has the power series expansion $\alpha(\omega^2) = \sum_{k=0}^{\infty} \mu_k \omega^{2k}$ and the physics allows us to compute a finite number of moments μ_k; the problem is then to approximate $\alpha(z)$ with these ones. Of course, the Padé approximants are an interesting tool for solving this problem and a thorough discussion is given in Ref. 5.

In order to use the results of the previous sections and to find approximations with poles and a cut on the positive real axis we write $\alpha(\omega^2) = \alpha_1(\omega^2) + \alpha_2(\omega^2)$ with

$$\alpha_1(\omega^2) = \sum_{j=0}^{\infty} k_j \mu_j \omega^{2j} , \quad \alpha_2(\omega^2) = \sum_{j=0}^{\infty} (1 - k_j)\mu_j \omega^{2j}$$

and we consider the approximations $\tilde{\alpha}^{n,m}(\omega^2) = \alpha_1^{n,m}(\omega^2) + \tilde{\alpha}_2^{n,m}(\omega^2)$, where $\alpha_1^{n,m}(\omega^2)$ is a usual Padé approximant of the power series $\alpha_1(\omega^2)$ and $\tilde{\alpha}_2^{n,m}(\omega^2)$ is an approximation of $\alpha_2(\omega^2)$ obtained with the help of (9). Of course, we have to determine the $(n+m+1)$ parameters k_i and the exponent β from $(n+m+2)$ experimental data. As we shall see, this is not always possible.

The first literal expressions are

1. $n=1,\ m=0$

$$\tilde{\alpha}^{1,0}(\omega^2)=\frac{k_0\mu_0}{1-(k_1\mu_1/k_0\mu_0)\omega^2}+(1-k_0)\mu_0\beta\int_0^1\frac{t^{\beta-1}\,dt}{1-[(1-k_1)/(1-k_0)](\mu_1/\mu_0)[(\beta+1)/\beta]\omega^2 t}\,, \tag{19}$$

with

$$\frac{k_0\mu_0}{k_1\mu_1}=\omega_{10}^2\,,\quad \frac{k_0^2\mu_0^2}{k_1\mu_1}=f_{10}\,,\quad \frac{\beta}{\beta+1}\,\frac{\mu_0}{\mu_1}\,\frac{1-k_0}{1-k_1}=\omega_\infty^2\,, \tag{19'}$$

2. $n=1,\ m=1$

$$\tilde{\alpha}^{1,1}(\omega^2)=\omega_{10}^2\,\frac{k_0\mu_0+(k_1\mu_1-k_0\mu_0\omega_{10}^{-2})\omega^2}{\omega_{10}^2-\omega^2}+\omega_\infty^2\int_0^1 t^{\beta-1}\,\frac{\beta(1-k_0)\mu_0+[(1-k_1)(1+\beta)\mu_1-(1-k_0)\beta\mu_0\omega_\infty^{-2}]\omega^2 t}{\omega_\infty^2-\omega^2 t}\,dt \tag{20}$$

with

$$\frac{k_1\mu_1}{k_2\mu_2}=\omega_{10}^2\,,\quad k_1\mu_1\omega_{10}^4=f_{10}\,,\quad \frac{1-k_1}{1-k_2}\,\frac{\mu_1}{\mu_2}\,\frac{\beta+1}{\beta+2}=\omega_\infty^2\,. \tag{20'}$$

The f_{n0} and ω_{n0} are respectively the oscillator strengths and transition frequencies of the real system.

3. $n=2,\ m=1$.

Here we put

$$g_i=\mu_i k_i\,,\quad x_i=\mu_i(1-k_i)\,,\quad i=0,1,2,3\,.$$

$$\tilde{\alpha}^{2,1}(\omega^2)=\frac{g_0+(b_1g_0+g_1)\omega^2}{1+b_1\omega^2+b_2\omega^4}+\int_0^1 t^{\beta-1}\,\frac{\beta x_0+[\beta C_1x_0+(\beta+1)x_1]\omega^2 t}{1+C_1\omega^2 t+C_2\omega^4 t^2}\,dt \tag{21}$$

with

$$b_1=\frac{g_1g_2-g_0g_3}{g_0g_2-g_1^2}=-\frac{\omega_{10}^2+\omega_{20}^2}{\omega_{10}^2\omega_{20}^2}\,,$$

$$b_2=\frac{g_1g_3-g_2^2}{g_0g_2-g_1^2}=\frac{1}{\omega_{10}^2\omega_{20}^2}\,,$$

$$C_1=\frac{(\beta+1)(\beta+2)x_1x_2-\beta(\beta+3)x_0x_3}{\beta(\beta+2)x_0x_2-(\beta+1)x_1^2}\,,$$

$$C_2=\frac{(\beta+1)(\beta+3)x_1x_3-(\beta+2)^2x_2^2}{\beta(\beta+2)x_0x_2-(\beta+1)^2x_1^2}\,.$$

The following conditions determine the four parameters k_i:

$$g_0=\frac{f_{10}}{\omega_{10}^2}+\frac{f_{20}}{\omega_{20}^2}\,,\quad g_1=\frac{f_{10}}{\omega_{10}^4}+\frac{f_{20}}{\omega_{20}^4}\,, \tag{21'}$$

$$g_2=-(b_2g_0+b_1g_1)\,,\quad g_3=-(b_2g_1+b_1g_2)$$

while β is a solution of the equation

$$C_2\omega_\infty^4+C_1\omega_\infty^2+1=0\,. \tag{21''}$$

These results need the following comments:

1. Since the literal expressions of $\tilde{\alpha}^{n,m}(\omega^2)$ are to be known (because of the existence of the k_i parameters), the method is practically limited to the approximants of lower order ($m=0,1,2$, $n=0,1,2$) but if the conjecture of the second paragraph is valid, these approximations could be better than the approximations with the usual Padé approximants of higher order. We are already sure, provided that the system of k_i,β, parameters has a solution to obtain exactly some (one or two) transition frequencies and the corresponding strengths of the real system.

2. Let us assume that the poles of the integrand in $\tilde{\alpha}_2^{n,m}(\omega^2)$ are simple, then a necessary condition that the system of $(m+n+2)$ parameters k_i,β has one and only one solution, is $m=n-1$ for $\alpha_2^{n,m}(\omega^2)$ has n poles (identified with transition frequencies) and this leads to $2n$ relations by identifying the corresponding residues with the oscillator strengths. Taking into account the condition for obtaining ω_∞, we then have $2n+1=m+n+2$, that is $m=n-1$.

3. Of course, $\beta>0$, but besides, β must be such that the branching points of $\tilde{\alpha}_2^{n,m}(\omega^2)$ exist only for $\omega\geq\omega_\infty$.

B. Polarizability of the atomic hydrogen

The data borrowed from Ref. 5 are

$$\mu_0=4.5\,,\quad \mu_1=26.5833\,,\quad \mu_2=172.188\,,\quad \mu_3=1162.09\,,$$

$$\omega_{10}=0.3750\,,\quad f_{10}=0.4162\,,$$

$$\omega_{20}=0.4444\,,\quad f_{20}=0.079\,11\,,\quad \omega_\infty=0.5\,.$$

The solution of system (19′) is then

$$k_0=0.657\,69\,,\quad k_1=0.791\,71\,,\quad \beta=8.866\,44\,,$$

and this leads to

$$\tilde{\alpha}^{1,0}(\omega^2)=\frac{0.4162}{(0.3750)^2-\omega^2}+3.4143\int_0^1\frac{t^{7.866}}{0.25-\omega^2 t}\,dt\,. \tag{22}$$

The solution of (21′) is

$$k_0=0.746\,715\,317\,,\quad k_1=0.868\,013\,921\,,$$

$$k_2=0.928\,801\,24\,,\quad k_3=0.960\,574\,952\,,$$

and Eq. (21″) has two roots, $\beta_1=76.156\,821\,20$ and

$\beta_2 = 1.906\,814\,694$. For $\beta = \beta_2$, the equation $1 + C_1\omega^2 + C_2\omega^4 = 0$ has a root lower than ω_∞, so we must take $\beta = \beta_1$ and for this choice, we have

$$C_1 = -5.635\,122\,061\,, \quad C_2 = 6.540\,488\,32\,,$$

$$\beta x_0 = 86.802\,103\,16\,, \quad \beta C_1 x_0 + (\beta + 1)x_1 = -218.426\,053\,6\,.$$

Finally it becomes

$$\tilde{\alpha}^{2,1}(\omega^2) = \frac{0.4162}{(0.3750)^2 - \omega^2} + \frac{0.079\,11}{(0.4444)^2 - \omega^2} + \int_0^1 t^{75.15} \frac{86.80 - 218.92\omega^2 t}{1 - 5.635\omega^2 t + 6.540\omega^4 t^2}\, dt\,. \tag{23}$$

We now compare $\tilde{\alpha}^{1,0}(\omega^2)$ with $\alpha^{2,1}(\omega^2)$ which, according to Ref. 5, represents exactly $\alpha(\omega^2)$ for $\omega < \omega_{10}$. Computing with the help of (16) for $n = 2$, some easy calculations lead to the approximation

$$\tilde{\alpha}_4^{1,0}(\omega^2) = \frac{0.4162}{(0.3750)^2 - \omega^2} + 3.4143\left\{\frac{0.030\,676\,7}{0.25 - \omega^2} + \frac{0.072\,479\,2}{0.25 - 0.889\,58\omega^2} + \frac{0.009\,634\,4}{0.25 - 0.644\,07\omega^2}\right\}. \tag{22'}$$

A look at the following table shows that the agreement is pretty good:

ω	0	0.1	0.2	0.3	0.35	0.3750
$\tilde{\alpha}_4^{1,0}(\omega^2)$	4.5	4.7840	5.935	10.367	25.731	∞
$\alpha^{2,1}(\omega^2)$	4.5	4.7842	5.941	10.521	24.349	139.8

On the imaginary axis, $y^2 = -\omega^2$, it is also easy to see that results are very good.

5. CONCLUSION

To conclude, we discuss the advantages and the drawbacks of both approximants $f^{n,m}(z)$ and $\tilde{f}^{n,m}(z)$ for a function $f(z)$ whose singularity is known.

1. The Padé approximants $f^{n,m}(z)$, relative to $\tilde{f}^{n,m}(z)$, are easy to compute but they are only meromorphic functions and they do not include any information on the known singularity of $f(z)$.

2. In the opposite direction, the approximants $\tilde{f}^{n,m}(z)$ convey this information which intuitively ought to make them a better approximation, but they are more difficult to compute and in most cases they cannot be computed exactly. Besides, it seems difficult to state some useful convergence criteria.

Thus, the choice between $f^{n,m}(z)$ and $\tilde{f}^{n,m}(z)$ depends on whether one wants to pay the price in introducing extra information in the usual approximations $f^{n,m}(z)$. It seems that it could be interesting to pay this price for solving some problems in theoretical physics.

Of course, when no singularity of $f(z)$ is known, the choice of the Padé approximants is the best one.

ACKNOWLEDGMENTS

The author is grateful to a referee for pointing out that Eq. (5) is a form of the Gammel–Baker approximant.

[1] G. A. Baker, in *Essentials of Padé Approximant* (Academic, New York, 1975).

[2] P. Davis and P. Rabinowitz, *Numerical Integration* (Blaisdell, New York, 1967).

[3] F. Oberhettinger and W. Magnus, *Formulas and Theorems for the Functions of Mathematical Physics* (Chelsea, New York, 1954).

[4] A. H. Stroud and Don Secrest, *Gaussian Quadrature Formulas* (Prentice Hall, Englewood Cliffs, New Jersey, 1966).

[5] P. W. Langhoff and M. Karplus, in *The Padé Approximants in Theoretical Physics*, edited by G. Baker and J. L. Gammel (Academic, New York, 1970).

A gauge invariant formulation of quantum electrodynamics using local currents*

R. Menikoff and D. H. Sharp

Theoretical Division, Los Alamos Scientific Laboratory, University of California, Los Alamos, New Mexico 87545
(Received 9 August 1976)

In this paper we formulate nonrelativistic quantum electrodynamics in a local and manifestly gauge invariant manner. This is accomplished by using the electromagnetic field strengths, rather than potentials, to describe the electromagnetic field and local currents, rather than canonical fields, to describe the matter. The exponentiated currents and field strengths form a group, whose representations can be studied using the Gel'fand–Vilenkin formalism. The currents and electromagnetic field strengths can be represented on a physical Hilbert space having positive norm. (The necessity for an indefinite metric does not arise here.) Furthermore, the classical equations of motion hold as operator equations on this Hilbert space. In this formulation, the requirement of gauge invariance is essentially replaced by imposing the Maxwell initial value eqution, which in turn lead to constraints on systems of Gel'fand–Vilenkin multipliers.

I. INTRODUCTION

In the standard formulation of quantum electrodynamics, one begins by using potentials $\mathbf{A}(\mathbf{x}, t)$ and $\phi(\mathbf{x}, t)$ to represent the electric and magnetic fields,

$$\mathbf{E}(\mathbf{x}, t) = -\frac{\partial \mathbf{A}(\mathbf{x}, t)}{\partial t} - \nabla\phi(\mathbf{x}, t), \tag{1.1}$$

$$\mathbf{B}(\mathbf{x}, t) = \nabla \times \mathbf{A}(\mathbf{x}, t); \tag{1.2}$$

then the potentials are quantized. The choice of potentials is not unique. Classically, the same $\mathbf{E}(\mathbf{x})$ and $\mathbf{B}(\mathbf{x})$ fields result from any pair of potentials related by the transformation

$$\begin{aligned} \phi'(\mathbf{x}, t) &= \phi(\mathbf{x}, t) - \frac{\partial \chi(\mathbf{x}, t)}{\partial t}, \\ \mathbf{A}'(\mathbf{x}, t) &= \mathbf{A}(\mathbf{x}, t) + \nabla\chi(\mathbf{x}, t), \end{aligned} \tag{1.3}$$

where χ can be a rather general function of $\mathbf{x}$ and t. This fact expresses the gauge invariance of electrodynamics, which is a necessary requirement on the theory since $\mathbf{E}(\mathbf{x})$ and $\mathbf{B}(\mathbf{x})$ are physical observables while $\mathbf{A}(\mathbf{x})$ and $\phi(\mathbf{x})$ are not.

In quantum theory the situation is more complicated, as can be illustrated by considering the following two specific choices for the vector potential. In the Coulomb gauge $[(\nabla \cdot \mathbf{A})(\mathbf{x}) = 0]$, the commutation relations of $\mathbf{A}(\mathbf{x})$ with $\mathring{\mathbf{A}}(\mathbf{x})$ have the nonlocal form[1]

$$[A_i(\mathbf{x}, t),\ \mathring{A}_j(\mathbf{y}, t)] = i\delta_{ij}\delta(\mathbf{x} - \mathbf{y}) - \frac{i}{4\pi}\,\partial_{x_i}\partial_{y_j}\frac{1}{|\mathbf{x} - \mathbf{y}|}. \tag{1.4}$$

Furthermore, $\mathbf{A}(\mathbf{x})$ turns out not to be the space component of a Lorentz covariant 4-vector field. In the Lorentz gauge $[(\nabla \cdot \mathbf{A})(\mathbf{x}) + \partial\phi(\mathbf{x})/\partial t = 0]$, on the other hand, the commutation relations among the components of the free field $A_\mu(x) = (\mathbf{A}, \phi)$ are local[1]

$$[A_\mu(\mathbf{x}, x_0), A_\nu(\mathbf{y}, y_0)] = -2g_{\mu\nu}\int \frac{d^4k}{(2\pi)^3}\exp[ik\cdot(x - y)]\,\theta(k_0)\,\delta(k^2), \tag{1.5}$$

and $A_\mu(x)$ is a Lorentz covariant 4-vector. However, the operators must be represented on a linear space having an indefinite metric. It is then necessary to select states with positive norm to form a Hilbert space.[2]

This state of affairs is inconvenient for general discussions of gauge invariance. One must require the physical observables $[\mathbf{E}(\mathbf{x}), \mathbf{B}(\mathbf{x})]$, acting on the physical Hilbert space, to be the same in different gauges, but one cannot in general represent two different vector potentials on the same Hilbert space. Thus, it would be useful to have a manifestly gauge invariant formulation of quantum electrodynamics.

Manifestly gauge invariant formulations of quantum electrodynamics have been discussed previously by a number of authors. For example, DeWitt[3] and Mandelstam[4] showed how quantum electrodynamics can be written in a formally gauge invariant way by introducing path dependent fields. These, however, result in a nonlocal quantum field theory.[2]

An alternative approach was outlined by one of the present authors who showed,[5] again on a formal level, that the electrodynamics of charged scalar mesons could be written in a *local* and manifestly gauge invariant way if the mesons were described using local currents. In exploring this idea further here, we shall consider nonrelativistic particles interacting with an electromagnetic field,[6,7] because in this case much more is known about the properties of local currents.

We begin by considering a single species of charged particle. A system of spinless, nonrelativistic particles can be described[8] using the number density of particles $\rho(\mathbf{x})$ and the particle flux density $\mathbf{J}(\mathbf{x})$. We use $\mathbf{E}(\mathbf{x})$ and $\mathbf{B}(\mathbf{x})$, rather than potentials, to describe the electromagnetic field. The fields $\rho(\mathbf{x})$, $\mathbf{J}(\mathbf{x})$, $\mathbf{E}(\mathbf{x})$, and $\mathbf{B}(\mathbf{x})$ are physical observables and hence gauge invariant. This set of operators generates a closed but nonlinear algebra when commuted at equal times. The interaction appears explicitly in the algebra. The Hamiltonian that gives the operator form of the classical equations of motion is formally the same as the free Hamiltonian, with the matter part of the free Hamiltonian written in terms of currents. Thus in this model what changes in passing from the free to the interacting theory is the

structure of the equal time algebra of observables, rather than the form of the Hamiltonian. This model thus provides an even more striking example than Haag's theorem[9] of the dependence of the equal-time algebra on the interaction, for in this case the structure of the algebra itself changes rather than just the representation that must be employed.

Since the equal-time algebra of $\rho(\mathbf{x})$, $\mathbf{J}(\mathbf{x})$, $\mathbf{E}(\mathbf{x})$, and $\mathbf{B}(\mathbf{x})$ is nonlinear, it is not a Lie algebra. A Lie algebra results if the variable $\mathbf{J}(\mathbf{x})$ is replaced by the total momentum density of the system,

$$\mathbf{P}(\mathbf{x}) = m\mathbf{J}(\mathbf{x}) + \tfrac{1}{2}(\mathbf{E}\times\mathbf{B} - \mathbf{B}\times\mathbf{E})(\mathbf{x}). \tag{1.6}$$

The Lie algebra of $\rho(\mathbf{x})$, $\mathbf{P}(\mathbf{x})$, $\mathbf{E}(\mathbf{x})$, and $\mathbf{B}(\mathbf{x})$, in contrast to the nonlinear algebra of $\rho(\mathbf{x})$, $\mathbf{J}(\mathbf{x})$, $\mathbf{E}(\mathbf{x})$, and $\mathbf{B}(\mathbf{x})$, does not contain the coupling constant or other explicit vestige of the interaction, which instead reappears in the Hamiltonian. In order to obtain the correct commutation relations for this Lie algebra, the initial value equations

$$\begin{aligned}&(\nabla\cdot\mathbf{E})(\mathbf{x}) = e\rho(\mathbf{x}),\\ &(\nabla\cdot\mathbf{B})(\mathbf{x}) = 0\end{aligned} \tag{1.7}$$

must be imposed at a fixed time. The equations of motion insure that Eqs. (1.7) hold at all subsequent times. The Lie algebra can be exponentiated to form a group whose representations can be described in terms of a measure and a set of multipliers using the Gel'fand–Vilenkin formalism.[10,11] When one works with the group rather than the algebra, the initial value equations (1.7) get replaced by a set of constraints on the Gel'fand–Vilenkin multipliers. These constraints on the multipliers are the conditions which replace gauge invariance when the theory is formulated in a manifestly gauge invariant fashion. Finally, we note that this algebra can be generalized to include a magnetic charge (monopole).

When considering more than one species of charged particle, one must introduce currents $\rho_i(\mathbf{x})$ and $\mathbf{J}_i(\mathbf{x})$ to describe each of the i different species of particle. These fields, together with $\mathbf{E}(\mathbf{x})$ and $\mathbf{B}(\mathbf{x})$, generate a closed nonlinear algebra under equal time commutation. A Lie algebra can be generated if each variable $\mathbf{J}_i(\mathbf{x})$ is replaced by

$$\mathbf{J}_i(\mathbf{x}) + (e_i/m_i c)\,\rho_i(\mathbf{x}) \int (d\mathbf{y}/4\pi)\,|\mathbf{x}-\mathbf{y}|^{-1}\nabla\times\mathbf{B}(\mathbf{y}). \tag{1.8}$$

However, this algebra is nonlocal. The quasilocal operator

$$\mathbf{A}(\mathbf{x}) = \int (d\mathbf{y}/4\pi)\,|\mathbf{x}-\mathbf{y}|^{-1}\nabla\times\mathbf{B}(\mathbf{y}) \tag{1.9}$$

behaves like the vector potential in Coulomb gauge, in that $\mathbf{A}(\mathbf{x})$ and $\mathring{\mathbf{A}}(\mathbf{x})$ commute as in Eq. (1.4) and

$$\mathbf{B}(\mathbf{x}) = \nabla\times\mathbf{A}(\mathbf{x}), \quad \mathbf{E}_\perp(\mathbf{x}) = -\frac{\partial\mathbf{A}(\mathbf{x})}{\partial t}, \quad (\nabla\cdot\mathbf{A})(\mathbf{x}) = 0. \tag{1.10}$$

However, since $\mathbf{A}(\mathbf{x})$ is computed from $\mathbf{B}(\mathbf{x})$ here, this quantity is also gauge invariant. Thus, when dealing with several species of charged particle, one seems to have the choice of working with gauge invariant local fields which generate a nonlinear algebra or quasilocal gauge invariant fields which generate a Lie algebra. For several species of particles we do not know the analog, if any, of the variables $\rho(\mathbf{x})$, $\mathbf{P}(\mathbf{x})$, $\mathbf{E}(\mathbf{x})$, and $\mathbf{B}(\mathbf{x})$ which generate a local Lie algebra in the case of a single type of charged particle.

This paper is organized as follows. In Sec. II we briefly review the description of systems of nonrelativistic particles using local currents $\rho(\mathbf{x})$ and $\mathbf{J}(\mathbf{x})$. In Sec. III the representation theory for the free electromagnetic field is described using the variables $\mathbf{E}(\mathbf{x})$ and $\mathbf{B}(\mathbf{x})$.[12] In Sec. IV, the results of Secs. II and III are combined to give a manifestly gauge invariant formulation of nonrelativistic quantum electrodynamics. The Gel'fand–Vilenkin representation theory for this model is discussed, particularly with regard to the question of multipliers and gauge invariance. In Sec. V we discuss possible generalizations of these results to the case where several species of charged particle are considered. We end with some concluding remarks in Sec. VI.

II. NONRELATIVISTIC QUANTUM MECHANICS AND LOCAL CURRENTS

We employ the number density of particles, $\rho(\mathbf{x})$, and the particle flux density, $\mathbf{J}(\mathbf{x})$, as variables to describe a system of identical, spinless nonrelativistic particles. In this section we briefly summarize the properties of the local currents which will be used in later sections.[13]

Our starting point is the local current algebra generated by commutation of the operators $\rho(\mathbf{x})$ and $\mathbf{J}(\mathbf{x})$ at equal times. This is given by[8] (in this section we take $\hbar = m = 1$):

$$\begin{aligned}&[\rho(\mathbf{x}), \rho(\mathbf{y})] = 0,\\ &[\rho(\mathbf{x}), J_k(\mathbf{y})] = -i\partial_{x_k}[\delta(\mathbf{x}-\mathbf{y})\,\rho(\mathbf{x})],\\ &[J_j(\mathbf{x}), J_k(\mathbf{y})] = -i\partial_{x_k}[\delta(\mathbf{x}-\mathbf{y})\,J_j(\mathbf{x})]\\ &\qquad\qquad + i\partial_{y_j}[\delta(\mathbf{x}-\mathbf{y})\,J_k(\mathbf{y})].\end{aligned} \tag{2.1}$$

The dynamics is determined by a Hamiltonian which, for particles interacting via a two-body potential, is given by[8,13]

$$\begin{aligned}H = {}&\tfrac{1}{8}\int d^3x\, K_i^*(\mathbf{x})[1/\rho(\mathbf{x})]\,K_i(\mathbf{x})\\ &+ \tfrac{1}{2}\int d^3x \int d^3y : \rho(\mathbf{x})\rho(\mathbf{y}) : V(\mathbf{x}-\mathbf{y}),\end{aligned} \tag{2.2}$$

where $\mathbf{K}(\mathbf{x}) = \nabla\rho(\mathbf{x}) + 2i\mathbf{J}(\mathbf{x})$. The local current is conserved

$$\frac{\partial\rho(\mathbf{x})}{\partial t} = i[H, \rho(\mathbf{x})] = -\nabla\cdot\mathbf{J}(\mathbf{x}), \tag{2.3}$$

as follows from Eqs. (2.1)–(2.2). In this way of formulating nonrelativistic quantum mechanics, the problem is to find a representation of the current algebra (2.1) in which the Hamiltonian (2.2) is a well-defined operator. This question has been studied in Ref. 14.

To discuss the representation theory for the local currents it is convenient to study the group formed by exponentiating the currents. To do this, one introduces the unitary operators[11,13]

$$U(f) = \exp[i\rho(f)]$$

and (2.4)

$$V(\phi_t^g) = \exp[itJ(g)],$$

where $\rho(f) = \int d^3x\, \rho(x) f(x)$ and $J(g) = \int d^3x\, J(x) \cdot g(x)$. Also, $\phi_t^g(x)$ denotes the flow for time t by the vector field g. It is defined by the equation

$$\frac{\partial \phi_t^g(x)}{\partial t} = g \circ \phi_t^g(x), \tag{2.5}$$

with $\phi_{t=0}^g(x) = x$ and where "$\circ$" stands for composition, i.e., $g \circ \phi(x) = g(\phi(x))$. It can then be shown that U and V satisfy the group multiplication law:

$$\begin{aligned} &U(f_1)U(f_2) = U(f_1 + f_2), \\ &V(\phi)U(f) = U(f \circ \phi)V(\phi), \\ &V(\phi_1)V(\phi_2) = V(\phi_2 \circ \phi_1). \end{aligned} \tag{2.6}$$

The representation theory for this group has been worked out by Goldin,[11] using the Gel'fand–Vilenkin formalism.[10] We will introduce the specifics of the representation theory as needed in the following sections.

III. THE FREE ELECTRIC AND MAGNETIC FIELDS

In this section we discuss the representation theory for the free electric and magnetic fields, using a formalism which works directly with the operator field strengths $E(x)$ and $B(x)$, rather than potentials.

A. Representation of the free electromagnetic fields

We take as our starting point the familiar equal-time commutation relations among the components of the electromagnetic field[1] (we again choose $c = \hbar = 1$):

$$[E_i(x), E_j(y)] = 0, \tag{3.1}$$

$$[B_i(x), B_j(y)] = 0, \tag{3.2}$$

$$[E_i(x), B_j(y)] = i\epsilon_{ijk}\partial_{y_k}\delta(x-y). \tag{3.3}$$

In this section we shall construct a representation of the algebra (3.1)–(3.3) which accommodates the free Hamiltonian

$$H = \tfrac{1}{2}\int d^3x : (E^2 + B^2)(x) : \tag{3.4}$$

and which is compatible with the initial value equations

$$\nabla \cdot E(x) = 0, \quad \nabla \cdot B(x) = 0 \tag{3.5}$$

and the condition

$$H\Omega = 0, \tag{3.6}$$

where Ω is the vacuum state. It is assumed that Ω is a cyclic vector for the representation.

To construct this representation, we begin by writing the Hamiltonian (3.4) in the factored form:

$$H = \tfrac{1}{2}\int d^3x \left[B(x) + i(2\pi^2)^{-1}\nabla \times \int d^3y\, \frac{E(y)}{|x-y|^2} \right]^\dagger \cdot \left[B(x) + i(2\pi^2)^{-1}\nabla \times \int d^3y'\, \frac{E(y')}{|x-y'|^2} \right]. \tag{3.7}$$

To show that this expression for the Hamiltonian is equivalent to that given in Eq. (3.4), we multiply out the two factors in Eq. (3.7). The result is

$$H = \tfrac{1}{2}\int d^3x \left\{ B(x) \cdot B(x) + i(2\pi^2)^{-1} \times \left[B_i(x), \left(\nabla \times \int d^3y\, \frac{E(y)}{|x-y|^2} \right)_i \right] + \left((2\pi^2)^{-1} \nabla \times \int d^3y\, \frac{E(y)}{|x-y|^2} \right) \cdot \left((2\pi^2)^{-1}\nabla \times \int d^3y'\, \frac{E(y')}{|x-y'|^2} \right) \right\}. \tag{3.8}$$

Evaluation of the commutator gives an (infinite) c-number which takes account of the normal ordering indicated in Eq. (3.4). To evaluate the last term, we introduce the Fourier transform[15] of the electric field, $\tilde{E}_i(k)$. Then the last term can be written:

$$\begin{aligned} &\frac{1}{2}\int \frac{d^3k}{(2\pi)^3} [\hat{k} \times \tilde{E}(k)^\dagger] \cdot [\hat{k} \times \tilde{E}(k)] \\ &= \frac{1}{2}\int \frac{d^3k}{(2\pi)^3}\, \tilde{E}_i(k)^\dagger (\delta_{ij} - \hat{k}_i\hat{k}_j) E_j(k) \\ &= \frac{1}{2}\int d^3x\, E_\perp(x) \cdot E_\perp(x), \end{aligned} \tag{3.9}$$

where $\hat{k} = k/|k|$ and $E_\perp(x)$ denotes the transverse part of the electric field. Thus the right-hand side of Eq. (3.7) is equal to

$$H = \tfrac{1}{2}\int d^3x \{B^2(x) + E_\perp^2(x) + (c\text{-number})\}, \tag{3.10}$$

which is equivalent to Eq. (3.4) in view of the constraint $\nabla \cdot E = 0$, which implies that $E = E_\perp$.

We have chosen the zero of energy so that the vacuum state Ω has zero energy. Equation (3.6), together with the fact that the integrand of Eq. (3.7) is positive at each point, implies that

$$\left[B(x) + i(2\pi^2)^{-1}\nabla \times \int d^3y\, \frac{E(y)}{|x-y|^2} \right] \Omega = 0. \tag{3.11}$$

This condition will lead to a functional equation which determines the generating functional

$$L(f) = (\Omega, \exp[iE(f)]\Omega), \tag{3.12}$$

where $E(f) = \int d^3x\, E(x) \cdot f(x)$. To derive this equation, we need to use the fact that

$$(\Omega, B(x) \exp[iE(f)]\Omega) = \tfrac{1}{2}(\nabla \times f)(x)\, L(f). \tag{3.13}$$

Equation (3.13) follows from time reversal invariance and Eq. (3.3), and is proved in Appendix A. Now, taking the inner product of Eq. (3.11) with the state $\exp[-iE(f)]\Omega$, we obtain the following functional differential equation for $L(f)$:

$$\tfrac{1}{2}(\nabla \times f)(x)\, L(f) + i(2\pi^2)^{-1}\nabla \times \int d^3y\, |x-y|^{-2}\, \frac{1}{i}\, \frac{\delta}{\delta f(y)}\, L(f) = 0. \tag{3.14}$$

This equation has the solution[16,17]

$$L(f) = \exp\left[-\tfrac{1}{4}\int [d^3k/(2\pi)^3]\, k \tilde{f}_i^*(k)(\delta_{ij} - k_ik_j/k^2)\tilde{f}_j(k) \right], \tag{3.15}$$

where $k = |k|$ and $\tilde{f}_i(k)$ is the Fourier transform of $f_i(x)$. By taking functional derivatives of this expression for $L(f)$ and using Eqs. (3.3) and (3.11), one can obtain all the equal time n-point functions involving products of

$\mathbf{E}(\mathbf{x})$ and $\mathbf{B}(\mathbf{x})$. Furthermore, it is not difficult to adapt standard arguments about Gaussian functionals to show that $L(\mathbf{f})$ is a positive functional in the sense of Bochner and that it is continuous in the topology of the test function space (e.g., Schwartz space $\mathcal{S}$). Obviously, $L(0) = 1$. Thus, by the GNS construction, Eq. (3.15) determines a representation of the free electromagnetic field.

Remarks: (1) The constraint $(\nabla \cdot \mathbf{E})(\mathbf{x}) = 0$ has been incorporated into the Hamiltonian, by replacing $\mathbf{E}$ with $\mathbf{E}_\perp$. Equation (3.11) implies that

$$(\nabla \cdot \mathbf{B})(\mathbf{x})\,\Omega = 0. \tag{3.16}$$

Once the Lorentz invariance of the theory is established, Eq. (3.16) and the Reeh–Schlieder theorem will imply that

$$\nabla \cdot \mathbf{B}(\mathbf{x}) = 0. \tag{3.17}$$

(2) There is a direct and simple correspondance between the above results and those obtained using the vector potential in Coulomb gauge. For example, using the standard expression for the vector potential for the free field in Coulomb gauge[1]

$$\mathbf{A}(\mathbf{x}, t) = \int \frac{d^3k}{[2\omega(2\pi)^3]^{1/2}} \sum_{\lambda=1,2} \epsilon(\mathbf{k}, \lambda)[a(\mathbf{k}, \lambda) \times \exp(-ik \cdot x) + a^\dagger(\mathbf{k}, \lambda) \exp(ik \cdot x)], \tag{3.18}$$

where $\omega = |\mathbf{k}|$ and $k \cdot x = \omega t - \mathbf{k} \cdot \mathbf{x}$, one can show that the Fourier transform of the quantity $[\mathbf{B}(\mathbf{x}) + i(2\pi^2)^{-1} \nabla \times \int d^3y\, \mathbf{E}(\mathbf{y})\, |\mathbf{x} - \mathbf{y}|^{-2}]$ is

$$i(2\pi)^{3/2}(2\omega)^{1/2}\, \hat{k} \times \sum_{\lambda=1,2} \epsilon(\mathbf{k}, \lambda)\, a(\mathbf{k}, \lambda) \tag{3.19}$$

and the Hamiltonian, Eq. (3.7), is then just

$$H = \int d^3k \omega \sum_{\lambda=1,2} a^\dagger(\mathbf{k}, \lambda)\, a(\mathbf{k}, \lambda).$$

B. The dynamics of the free electromagnetic field

We next indicate how to determine the dynamics, in particular the structure of the time dependent n-point functions, of the free electromagnetic field. We start with the equations of motion for the fields:

$$\dot{\mathbf{E}}(\mathbf{x}, t) = i[H, \mathbf{E}(\mathbf{x}, t)] = (\nabla \times \mathbf{B})(\mathbf{x}, t), \tag{3.20}$$

$$\dot{\mathbf{B}}(\mathbf{x}, t) = i[H, \mathbf{B}(\mathbf{x}, t)] = -(\nabla \times \mathbf{E})(\mathbf{x}, t). \tag{3.21}$$

Using them, it is easy to obtain the explicit time dependence of $\mathbf{E}$ and $\mathbf{B}$:

$$\tilde{\mathbf{E}}(\mathbf{k}, t) = \cos(\omega t)\, \tilde{\mathbf{E}}(\mathbf{k}) + \sin(\omega t)\, \hat{k} \times \tilde{\mathbf{B}}(\mathbf{k}), \tag{3.22}$$

and

$$\mathbf{B}(\mathbf{k}, t) = \cos(\omega t)\, \tilde{\mathbf{B}}(\mathbf{k}) - \sin(\omega t)\, \hat{k} \times \tilde{\mathbf{E}}(\mathbf{k}). \tag{3.23}$$

To compute the time dependent n point functions, it is useful to introduce the generating functional

$$L_1(\mathbf{f}, \mathbf{g}) = (\Omega, \exp i[E(\mathbf{f}) + B(\mathbf{g})]\Omega), \tag{3.24}$$

where

$$E(\mathbf{f}) = \int d^4x\, \mathbf{E}(\mathbf{x}, t) \cdot \mathbf{f}(\mathbf{x}, t)$$

and

$$B(\mathbf{g}) = \int d^4x\, \mathbf{B}(\mathbf{x}, t) \cdot \mathbf{g}(\mathbf{x}, t).$$

One can construct an explicit expression for $L_1(\mathbf{f}, \mathbf{g})$ by expanding the exponential in Eq. (3.24), using Eqs. (3.22)–(3.23) to relate the fields at time t to their value at $t = 0$, using our previous results to evaluate these n-point functions and finally resuming the series. The result is

$$L_1(\mathbf{f}, \mathbf{g}) = \exp\{-\tfrac{1}{4} \int [d^3k/(2\pi)^3]\omega[\tilde{f}(\mathbf{k}, \omega) - \hat{k} \times \tilde{g}(\mathbf{k}, \omega)]_i^* \times [\delta_{ij} - k_i k_j / k^2][\tilde{f}(\mathbf{k}, \omega) - \hat{k} \times \tilde{g}(\mathbf{k}, \omega)]_j\}. \tag{3.25}$$

It is useful to write Eq. (3.25) in a manifestly Lorentz invariant form, since one can then avoid the straightforward but tedious task of demonstrating directly the Lorentz invariance of Eq. (3.25). To do this, we introduce a second rank tensor $T_{\mu\nu}$ whose components are formed from the test functions $\mathbf{f}$ and $\mathbf{g}$ as follows:

$$T_{\mu\nu} = \begin{pmatrix} 0 & -f_1 & -f_2 & -f_3 \\ f_1 & 0 & g_3 & -g_2 \\ f_1 & -g_3 & 0 & g_1 \\ f_3 & g_2 & -g_1 & 0 \end{pmatrix}. \tag{3.26}$$

Then one can write

$$E(\mathbf{f}) + B(\mathbf{g}) = \tfrac{1}{2} \int d^4x\, F^{\mu\nu}(\mathbf{x}, t)\, T_{\mu\nu}(\mathbf{x}, t),$$

where $F^{\mu\nu}(\mathbf{x}, t)$ is the electromagnetic field tensor and Eq. (3.25) can be written in the form

$$L_1(T) = \exp\left\{-\tfrac{1}{4} \int \frac{d^3k}{\omega}\, \tilde{T}_{\mu\nu}(\mathbf{k}, \omega)^* k^\nu g^{\mu\beta} k^\alpha \tilde{T}_{\alpha\beta}(\mathbf{k}, \omega)\right\}. \tag{3.27}$$

Hence, the representation of the free electromagnetic fields which we have constructed is Lorentz invariant.

Finally, we note that the commutation relations between $\mathbf{E}$ and $\mathbf{B}$ at unequal times[1] can be computed from Eqs. (3.1)–(3.3) and (3.22)–(3.23). These commutation relations, together with the generating functional (3.25), enable one to compute all the time dependent n-point functions.

Remarks: (1) The theory has been divided into two parts; determining the representation of the equal-time algebra and calculating the dynamics. The representation defines the Hilbert space on which the fields act at a fixed time, while the dynamics governs their time evolution. In order for the dynamics of a given physical system to be well defined, we must select a representation of the equal-time algebra which is compatible with the Hamiltonian. To discuss Lorentz invariance it is necessary to know both the representation and the dynamics.

(2) In the free theory which we have discussed, polynomials of $E(\mathbf{f})$ acting on the vacuum are dense. This is because the operators $E(\mathbf{f})$ form a maximal commuting set of observables. For the theory to be manifestly Lorentz covariant one would need the Hilbert space to be defined by polynomials of $F^{\mu\nu}(T_{\mu\nu})$ acting on the vacuum. But these states are overcomplete and constraints would be needed. This is compatible with the theorem[2] that a local manifestly Lorentz invariant formulation of quantum electrodynamics using a 4-vector potential requires an indefinite metric.

IV. QUANTUM ELECTRODYNAMICS FOR A SINGLE SPECIES OF CHARGED PARTICLE

In this section we combine the local currents with the electric and magnetic field operators to describe the electromagnetic interaction of a system of charged particles in a manifestly gauge invariant way. We first consider one species of particle with electric charge e and mass m.[18] The case of more than one species of particle is discussed in Sec. V.

A. The local current algebra of the operators $\rho(\mathbf{x})$, $\mathbf{J}(\mathbf{x})$, $\mathbf{E}(\mathbf{x})$, and $\mathbf{B}(\mathbf{x})$

We use the canonical fields as a heuristic to motivate the form of the local current algebra and the Hamiltonian. The number density of particles $\rho(\mathbf{x})$, the flux density of particles $\mathbf{J}(\mathbf{x})$, and the electric and magnetic fields are given by

$$\rho(\mathbf{x}) = \psi^\dagger(\mathbf{x})\,\psi(\mathbf{x}),$$

$$\mathbf{J}(\mathbf{x}) = (2im)^{-1}\left[\psi^\dagger(\mathbf{x})\left(\hbar\nabla - \frac{ie}{c}\mathbf{A}(\mathbf{x})\right)\psi(\mathbf{x}) - \left(\hbar\nabla + \frac{ie}{c}\mathbf{A}(\mathbf{x})\right)\psi^\dagger(\mathbf{x})\,\psi(\mathbf{x})\right],$$

$$\mathbf{E}(\mathbf{x}) = -\nabla A_0(\mathbf{x}) - \frac{1}{c}\frac{\partial \mathbf{A}}{\partial t},$$

$$\mathbf{B}(\mathbf{x}) = (\nabla \times \mathbf{A})(\mathbf{x}).$$

These quantities are all physical observables and hence gauge invariant. Note that the mass density is $m\rho(\mathbf{x})$, the charge density is $e\rho(\mathbf{x})$, the electric current density is $e\mathbf{J}(\mathbf{x})$, and the particles' momentum density is $m\mathbf{J}(\mathbf{x})$. From the commutation relations for the canonical fields

$$[\psi(\mathbf{x}), \psi^\dagger(\mathbf{y})]_\pm = \delta(\mathbf{x}-\mathbf{y}),$$

$$[A_\mu(\mathbf{x}), \dot{A}_\nu(\mathbf{y})] = -c^2 i\hbar g_{\mu\nu}\delta(\mathbf{x}-\mathbf{y}),$$

one obtains the following equal-time commutation relations among the operators ρ, $\mathbf{J}$, $\mathbf{E}$, and $\mathbf{B}$:

$$[\rho(\mathbf{x}), J_i(\mathbf{y})] = -i\left(\frac{\hbar}{m}\right)\frac{\partial}{\partial x_i}[\delta(\mathbf{x}-\mathbf{y})\rho(\mathbf{x})], \tag{4.1}$$

$$[J_i(\mathbf{x}), J_k(\mathbf{y})] = i\left(\frac{\hbar}{m}\right)\frac{\partial}{\partial y_i}[\delta(\mathbf{x}-\mathbf{y})J_k(\mathbf{y})] - i\left(\frac{\hbar}{m}\right)\frac{\partial}{\partial x_k}[\delta(\mathbf{x}-\mathbf{y})J_i(\mathbf{x})] + i\left(\frac{\hbar e}{m^2 c}\right)\epsilon_{ikj}\rho(\mathbf{x})B_j(\mathbf{x})\delta(\mathbf{x}-\mathbf{y}), \tag{4.2}$$

$$[E_i(\mathbf{x}), B_k(\mathbf{y})] = i(c\hbar)\epsilon_{ikj}\frac{\partial}{\partial y_j}\delta(\mathbf{x}-\mathbf{y}), \tag{4.3}$$

$$[J_i(\mathbf{x}), E_k(\mathbf{y})] = i\left(\frac{\hbar e}{m}\right)\rho(\mathbf{x})\,\delta_{ik}\,\delta(\mathbf{x}-\mathbf{y}), \tag{4.4}$$

with all other commutators vanishing.[19] Thus, under equal-time commutation the local currents and the components of the electromagnetic field form a closed algebra. However, the algebra is nonlinear owing to the $\rho(\mathbf{x})\,\mathbf{B}(\mathbf{x})$ term which appears in Eq. (4.2), and hence it is not a Lie algebra. Another important property of the algebra is that the interaction and the coupling constant, e, enter explicitly in the commutation relations (4.2) and (4.4). As a result, the algebra (4.1)–(4.4) for the interacting theory is different from that of the free theory and hence a different representation of the local currents and fields is required. Furthermore, different representations are needed for different values of the coupling constant.

Remark: If the particles were interacting with external electric and magnetic fields, then $\mathbf{E}(\mathbf{x})$ and $\mathbf{B}(\mathbf{x})$ would be c-number fields and the local currents $\rho(\mathbf{x})$ and $\mathbf{J}(\mathbf{x})$ would form a Lie algebra.

We next consider the Hamiltonian. In terms of the canonical fields it is given by

$$H = \frac{\hbar^2}{2m}\int d^3x\left(\nabla + i\frac{e}{\hbar c}\mathbf{A}(\mathbf{x})\right)\psi^\dagger(\mathbf{x}) \cdot \left(\nabla - i\frac{e}{\hbar c}\mathbf{A}(\mathbf{x})\right)\psi(\mathbf{x}) + \tfrac{1}{2}\int d^3x\,(\mathbf{E}^2+\mathbf{B}^2)(\mathbf{x}).$$

By using the local currents, this can be written as

$$H = \frac{\hbar^2}{8m}\int d^3x\,K_i^\dagger(\mathbf{x})\frac{1}{\rho(\mathbf{x})}K_i(\mathbf{x}) + \tfrac{1}{2}\int d^3x\,(\mathbf{E}^2+\mathbf{B}^2)(\mathbf{x}), \tag{4.5}$$

where

$$\mathbf{K}(\mathbf{x}) = \nabla\rho(\mathbf{x}) + (2im/\hbar)\,\mathbf{J}(\mathbf{x}). \tag{4.5'}$$

At this point, one can forget about the canonical fields and take the local currents, their commutation relations [Eqs. (4.1)–(4.4)] and Eq. (4.5) for the Hamiltonian as defining the theory.

As one check on the consistency of this description, one can verify that Eqs. (4.1)–(4.5) lead to the correct operator form of the classical equations of motion. These are given by

$$\partial\rho(\mathbf{x})/\partial t = (i/\hbar)[H, \rho(\mathbf{x})] = \nabla\cdot\mathbf{J}(\mathbf{x}) \tag{4.6}$$

(current conservation),

$$(1/c)\,\partial\mathbf{B}(\mathbf{x})/\partial t = i(1/c\hbar)[H, \mathbf{B}(\mathbf{x})] = -(\nabla\times\mathbf{E})(\mathbf{x}), \tag{4.7}$$

$$(1/c)\,\partial\mathbf{E}(\mathbf{x})/\partial t = i(1/c\hbar)[H, \mathbf{E}(\mathbf{x})] = -(e/c)\,\mathbf{J}(\mathbf{x}) + (\nabla\times\mathbf{B})(\mathbf{x})$$

(the Maxwell's equations governing time evolution),

(4.8)

$$m\dot{J}_i(\mathbf{x}) = i(m/\hbar)[H, J_i(\mathbf{x})] = e\rho(\mathbf{x})\,E_i(\mathbf{x}) + (e/c)[\mathbf{J}(\mathbf{x})\times\mathbf{B}(\mathbf{x})]_i - (\hbar^2/m)\,\partial\theta_{ij}/\partial x_j, \tag{4.9}$$

where

$$\theta_{ij}(\mathbf{x}) = K_i^\dagger(\mathbf{x})\frac{1}{\rho(\mathbf{x})}K_j(\mathbf{x}) + K_j^\dagger(\mathbf{x})\frac{1}{\rho(\mathbf{x})}K_i(\mathbf{x}) - 2\partial^2\rho(\mathbf{x})/\partial x_i\,\partial x_j \tag{4.10}$$

(conservation of momentum).

The first two terms on the right-hand side of Eq. (4.9) comprise the Lorentz force, while the last term describes convection. The other two Maxwell equations, $\nabla\cdot\mathbf{E} = e\rho$ and $\nabla\cdot\mathbf{B} = 0$, must be imposed as initial value equations in order for the total momentum operator to

be identifiable as the generator of translations in space. This point is discussed in detail in the next subsection.

An interesting aspect of this formulation is that the Hamiltonian (4.5) is, formally, the sum of the free Hamiltonian for the particles plus the free Hamiltonian for the electric and magnetic fields [compare Eqs. (2.2) and (3.4)]. The only place where the interaction appears explicitly is in the equal-time algebra (4.1)–(4.4).

B. The local Lie algebra of the operators $\rho(\mathbf{x})$, $\mathbf{P}(\mathbf{x})$, $\mathbf{E}(\mathbf{x})$, and $\mathbf{B}(\mathbf{x})$

To discuss the representation theory of the local currents, it is convenient to make a change of variables so as to obtain a nuclear Lie algebra. For this purpose we introduce the total momentum density:

$$\mathbf{P}(\mathbf{x}) = m\mathbf{J}(\mathbf{x}) + (1/2c)\,[\mathbf{E}(\mathbf{x})\times\mathbf{B}(\mathbf{x}) - \mathbf{B}(\mathbf{x})\times\mathbf{E}(\mathbf{x})]. \tag{4.11}$$

Next we determine the equal-time algebra generated by $\rho(\mathbf{x})$, $\mathbf{P}(\mathbf{x})$, $\mathbf{E}(\mathbf{x})$, and $\mathbf{B}(\mathbf{x})$. A straightforward but tedious calculation shows that the commutation relations involving $\mathbf{P}(\mathbf{x})$ are given by:

$$[\rho(\mathbf{x}), P_i(\mathbf{y})] = -i\hbar\,\frac{\partial}{\partial x_i}\,[\delta(\mathbf{x}-\mathbf{y})\,\rho(\mathbf{x})],$$

$$\begin{aligned}[P_i(\mathbf{x}), P_k(\mathbf{y})] &= i\hbar\,\frac{\partial}{\partial y_i}\,[\delta(\mathbf{x}-\mathbf{y})\,P_k(\mathbf{y})] - i\hbar\,\frac{\partial}{\partial x_k}\,[\delta(\mathbf{x}-\mathbf{y})\,P_i(\mathbf{x})] \\ &\quad + i\left(\frac{\hbar}{c}\right)\epsilon_{ikj}\{[\nabla\cdot\mathbf{E}(\mathbf{x}) - e\rho(\mathbf{x})]\,B_j(\mathbf{x}) + \nabla\circ\mathbf{B}(\mathbf{x})\,E_j(\mathbf{x})\} \\ &\quad \times\delta(\mathbf{x}-\mathbf{y}),\end{aligned}$$

$$\begin{aligned}[E_i(\mathbf{x}), P_k(\mathbf{y})] &= -i\hbar\,\frac{\partial}{\partial x_k}\,[E_i(\mathbf{x})\,\delta(\mathbf{x}-\mathbf{y})] - i\hbar\delta_{ik}\,\frac{\partial}{\partial y_j}\,[E_j(\mathbf{y})\,\delta(\mathbf{x}-\mathbf{y})] \\ &\quad + i\hbar\delta_{ik}[\nabla\cdot\mathbf{E}(\mathbf{x}) - e\rho(\mathbf{x})]\,\delta(\mathbf{x}-\mathbf{y}),\end{aligned}$$

$$\begin{aligned}[B_i(\mathbf{x}), P_k(\mathbf{y})] &= -i\hbar\,\frac{\partial}{\partial x_k}\,[B_i(\mathbf{x})\,\delta(\mathbf{x}-\mathbf{y})] - i\hbar\delta_{ik}\,\frac{\partial}{\partial y_j}\,[B_j(\mathbf{y})\,\delta(\mathbf{x}-\mathbf{y})] \\ &\quad + i\hbar\delta_{ik}\nabla\cdot\mathbf{B}(\mathbf{x})\,\delta(\mathbf{x}-\mathbf{y}).\end{aligned}$$

For a consistent physical interpretation, we require

$$\mathbf{P} = \int d^3x\,\mathbf{P}(\mathbf{x})$$

to be the total momentum, i.e., the generator of space translations. The commutator of $\mathbf{P}$ with any local operator $O(\mathbf{x})$ must therefore have the form

$$[O(\mathbf{x}), \mathbf{P}] = -i\hbar\nabla O(\mathbf{x}).$$

In order for this equation to be consistent with the above commutation relations, it is necessary to impose the constraints

$$(\nabla\circ\mathbf{E})(\mathbf{x}) = e\rho(\mathbf{x}) \tag{4.12}$$

and

$$(\nabla\circ\mathbf{B})(\mathbf{x}) = 0. \tag{4.13}$$

These constraints can be interpreted as initial value equations since it can be shown that if Eqs. (4.12) and (4.13) hold at a fixed time t, then they hold for all time, as a consequence of the equations of motion (4.6)–(4.8). The connection between the initial value equations and the gauge invariance of quantum electrodynamics when formulated using potentials will be discussed in the next subsection.

After imposing the initial value equations one obtains the following nuclear Lie algebra:

$$[\rho(\mathbf{x}), P_i(\mathbf{y})] = -i\hbar\,\frac{\partial}{\partial x_i}\,[\delta(\mathbf{x}-\mathbf{y})\,\rho(\mathbf{x})], \tag{4.14}$$

$$\begin{aligned}[P_i(\mathbf{x}), P_j(\mathbf{y})] &= i\hbar\,\frac{\partial}{\partial y_i}\,[\delta(\mathbf{x}-\mathbf{y})\,P_j(\mathbf{y})] - i\hbar\,\frac{\partial}{\partial x_j}\,[\delta(\mathbf{x}-\mathbf{y})\,P_i(\mathbf{x})],\end{aligned} \tag{4.15}$$

$$[E_i(\mathbf{x}), B_j(\mathbf{y})] = i(c\hbar)\,\epsilon_{ijk}\,\frac{\partial}{\partial y_k}\,\delta(\mathbf{x}-\mathbf{y}), \tag{4.16}$$

$$\begin{aligned}[E_i(\mathbf{x}), P_j(\mathbf{y})] &= -i\hbar\,\frac{\partial}{\partial x_j}\,[E_i(\mathbf{x})\,\delta(\mathbf{x}-\mathbf{y})] - i\hbar\delta_{ij}\,\frac{\partial}{\partial y_k}\,[E_k(\mathbf{y})\,\delta(\mathbf{x}-\mathbf{y})],\end{aligned} \tag{4.17}$$

$$\begin{aligned}[B_i(\mathbf{x}), P_j(\mathbf{y})] &= -i\hbar\,\frac{\partial}{\partial x_j}\,[B_i(\mathbf{x})\,\delta(\mathbf{x}-\mathbf{y})] - i\hbar\delta_{ij}\,\frac{\partial}{\partial y_k}\,[B_k(\mathbf{y})\,\delta(\mathbf{x}-\mathbf{y})],\end{aligned} \tag{4.18}$$

with all other commutators vanishing.

Remark: It can easily be checked that these commutation relations are consistent with the initial value equations and that they satisfy the Jacobi identity.

We can also describe the dynamics in terms of these variables. For example, the Hamiltonian is still given by Eq. (4.5), but with $\mathbf{K}(\mathbf{x})$ given by

$$\mathbf{K}(\mathbf{x}) = \nabla\rho(\mathbf{x}) + (2i/\hbar)\{\mathbf{P}(\mathbf{x}) - (1/2c)(\mathbf{E}\times\mathbf{B} - \mathbf{B}\times\mathbf{E})(\mathbf{x})\}, \tag{4.19}$$

instead of (4.5′). The interaction now appears in the Hamiltonian, as usual, instead of in the algebra. Note that the coupling constant e appears only in the initial value equation (4.12). If we choose the charge density $e\rho(\mathbf{x})$ as our variable instead of the number density $\rho(\mathbf{x})$, then the coupling constant would appear only in the Hamiltonian. Henceforth, we use units in which $1 = c = \hbar = m$.

C. Gauge invariance

At this point we discuss the relationship between the formulation of electrodynamics using local currents and that using potentials, particularly with regard to the role of gauge invariance. In either case one obtains Maxwell's equations, but they come about in different ways.

In terms of the $\mathbf{E}$ and $\mathbf{B}$ fields, the dynamics is determined by two first order equations of motion, Eqs. (4.7) and (4.8). In addition, two initial value equations,

(4.12) and (4.13), must be imposed in order for the theory to accommodate a representation of the translation group. Together, Eqs. (4.7)—(4.8) and (4.12)—(4.13) comprise Maxwell's equations.

Alternatively, one can write Maxwell's equations in the form

$$\partial_\mu F^{\mu\nu} = j^\nu, \tag{4.20}$$

$$\partial_\mu {}^*F^{\mu\nu} = 0, \tag{4.21}$$

where ${}^*F^{\mu\nu}$ is the dual of $F^{\mu\nu}$. Equation (4.21) will be satisfied identically if $F^{\mu\nu}$ is written as the curl of a 4-vector potential A^μ; $F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu$. However, the transformation from $F^{\mu\nu}$ to A^μ is not unique. To determine A^μ uniquely, a constraint, or gauge condition, must be imposed.[20] The equation of motion (second order in time for **A**) is then equivalent to Eq. (4.20). Thus, Eq. (4.20) is the dynamical equation when the theory is expressed in terms of potentials.

To summarize:

(1) When working with **E** and **B**, Eqs. (4.7) and (4.8) are dynamical, while Eqs. (4.12) and (4.13) are initial value equations (constraints).

(2) When working with the potentials, Eqs. (4.8) and (4.12) are dynamical. Equations (4.7) and (4.13) are identically satisfied, but gauge invariance is required.

Thus, the condition that the **E** and **B** fields satisfy initial value equations replaces the requirement of gauge invariance in the formulation of electrodynamics using potentials. However, the dynamical equations are not the same in the two formulations. Note that the Coulomb gauge is an exceptional case. In Coulomb gauge, Eq. (4.12) is not a dynamical equation as it is in other gauges, but is instead an initial value equation.

D. The Lie group of the exponentiated currents and fields

Before applying the representation theory we need to obtain the group formed by exponentiating the commutation relations (4.14)—(4.18). Since the subalgebra formed by $\rho(\mathbf{x})$ and $\mathbf{P}(\mathbf{x})$ is the same as the free $\rho(\mathbf{x})$, $\mathbf{J}(\mathbf{x})$ algebra, we can use the results of Sec. II to handle this part of the problem. For the exponential of $\mathbf{P}(\mathbf{x})$ we introduce the operator

$$U(\phi_t^{\mathbf{g}}) = \exp[itP(\mathbf{g})],$$

where $\phi_t^{\mathbf{g}}$ is the flow corresponding to the vector field $\mathbf{g}(\mathbf{x})$, defined by Eq. (2.5). Then the multiplication law for the ρ, $\mathbf{P}$ subgroup is given by[11]

$$\begin{aligned} &\exp[i\rho(f_1)]\exp[i\rho(f_2)] = \exp[i\rho(f_1+f_2)], \\ &U(\phi)\exp[i\rho(f)] = \exp[i\rho(f\circ\phi)]U(\phi), \\ &U(\phi_1)U(\phi_2) = U(\phi_2\circ\phi_1). \end{aligned} \tag{4.22}$$

It is not difficult to obtain the group associated with the subalgebra formed by $\mathbf{E}(\mathbf{x})$ and $\mathbf{B}(\mathbf{x})$. The multiplication law is simply

$$\exp[iE(\mathbf{g})]\exp[iE(\mathbf{h})] = \exp[iE(\mathbf{g}+\mathbf{h})],$$

$$\exp[iB(\mathbf{g})]\exp(iB(\mathbf{h})] = \exp[iB(\mathbf{g}+\mathbf{h})],$$

$$\exp[iE(\mathbf{g})]\exp[iB(\mathbf{h})]$$

$$= \exp[-i\int d^3x\, \mathbf{g}(\mathbf{x})\cdot(\nabla\times\mathbf{h})(\mathbf{x})]\exp[iB(\mathbf{h})]\exp[iE(\mathbf{g})]. \tag{4.23}$$

Since $\rho(\mathbf{x})$ commutes with $\mathbf{E}(\mathbf{x})$ and $\mathbf{B}(\mathbf{x})$, we also have

$$\begin{aligned} &\exp[iE(\mathbf{g})]\exp[i\rho(f)] = \exp[i\rho(f)]\exp[iE(\mathbf{g}))], \\ &\exp[iB(\mathbf{g})]\exp[i\rho(f)] = \exp[i\rho(f)]\exp[iB(\mathbf{g}))]. \end{aligned} \tag{4.24}$$

Finally, we need to know how to multiply the group element $U(\phi)$ together with $\exp[iE(\mathbf{g})]$ or $\exp[iB(\mathbf{g})]$. In Appendix B, we show that these elements multiply according to the rule

$$U(\phi)\exp[iE(\mathbf{g})] = \exp[iE(\mathbf{W}_\phi(\mathbf{g}))]U(\phi)$$

and (4.25)

$$U(\phi)\exp[iB(\mathbf{g})] = \exp[iB(\mathbf{W}_\phi(\mathbf{g}))]U(\phi),$$

where

$$\mathbf{W}_\phi(\mathbf{g})(\mathbf{x}) = (\nabla\phi_k)(\mathbf{x})(g_k\circ\phi)(\mathbf{x}). \tag{4.26}$$

To summarize, an arbitrary group element can be written as

$$\Gamma(f,\mathbf{g},\mathbf{h},\phi) = \exp[i\rho(f)]\exp[iE(\mathbf{g})]\exp[iB(\mathbf{h})]U(\phi), \tag{4.27}$$

and the law for multiplying any two such elements is given by

$$\begin{aligned} &\Gamma(f_1,\mathbf{g}_1,\mathbf{h}_1,\phi_1)\Gamma(f_2,\mathbf{g}_2,\mathbf{h}_2,\phi_2) \\ &\quad= \exp[i\int d^3x\, \mathbf{W}_{\phi_1}(\mathbf{g}_2)(\mathbf{x})\cdot(\nabla\times\mathbf{h}_1)(\mathbf{x})] \\ &\qquad\times\Gamma(f_1+f_2\circ\phi_1, \mathbf{g}_1+\mathbf{W}_{\phi_1}(\mathbf{g}_2), \mathbf{h}_1+\mathbf{W}_{\phi_1}(\mathbf{h}_2), \phi_2\circ\phi_1). \end{aligned} \tag{4.28}$$

Remark: From the multiplication law above it can be checked that

$$U(\hat{n},\theta) = \exp[i\theta\hat{n}\cdot\int d^3x\, \mathbf{x}\times\mathbf{P}(\mathbf{x})]$$

is formally [since the function $\mathbf{g}(\mathbf{x}) = \mathbf{x}$ is not in the space of test functions] a unitary operator for a rotation by an angle θ about the $\hat{n}$ axis. For example,

$$U(\hat{n},\theta)\exp[iE(\mathbf{g})]U^\dagger(\hat{n},\theta) = \exp[iE(\mathbf{g}_R)],$$

where $(g_R)_i(\mathbf{x}) = R_{ij}g_j(R^{-1}\mathbf{x})$ and R is the matrix for a rotation by an angle θ about the $\hat{n}$ axis.

E. The Gel'fand-Vilenkin representation theory

We are now ready to describe the representation of the local currents $\rho(\mathbf{x})$ and $\mathbf{P}(\mathbf{x})$ and the fields $\mathbf{E}(\mathbf{x})$ and $\mathbf{B}(\mathbf{x})$ in terms of a measure and a set of multipliers using the Gel'fand—Vilenkin representation theory.[10] Here we shall summarize the results of this theory as they apply in the present case. (See Ref. 11 for proofs in the case of the ρ, $\mathbf{J}$ algebra).

We begin by considering the generating functional

$$L(\mathbf{g}) = (\Omega, \exp[iE(\mathbf{g})]\Omega), \tag{4.29}$$

where Ω is a cyclic vector for the representation, which we will always identify with the ground state of the Hamiltonian [Eqs. (4.5) and (4.19)]. One can write $L(\mathbf{g})$

as the Fourier transform of a positive measure μ on $\mathcal{S}'$, the real continuous dual of the space of test functions. Thus,

$$L(\mathbf{g}) = \int_{\mathcal{S}'} d\mu(\mathbf{G}) \exp[i(\mathbf{G}, \mathbf{g})], \tag{4.30}$$

where $(\mathbf{G}, \mathbf{g})$ means the functional $\mathbf{G}$ evaluated at $\mathbf{g}(\mathbf{x})$. The measure μ can be used to define a Hilbert space $\mathcal{H} = L^2_\mu(\mathcal{S}')$. The group elements act in the following way on this Hilbert space. Let $\Psi(\mathbf{G}) \in \mathcal{H}$. Then:

(1) The element $\exp[iE(\mathbf{g})]$ is represented as multiplication by $\exp[i(\mathbf{G}, \mathbf{g})]$, i.e.,

$$(\exp[iE(\mathbf{g})]\Psi)(\mathbf{G}) = \exp[i(\mathbf{G}, \mathbf{g})]\Psi(\mathbf{G}). \tag{4.31}$$

(2) The operator $\rho(f)$ is defined by Eq. (4.12), or $e\rho(f) = -E(\nabla f)$. As a result

$$(\exp[ie\rho(f)]\Psi)(\mathbf{G}) = \exp[-i(\mathbf{G}, \nabla f)]\Psi(\mathbf{G}). \tag{4.32}$$

(3) Next we have

$$(U(\phi)\Psi)(\mathbf{G}) = \chi_\phi(\mathbf{G})\Psi(\phi^*\mathbf{G})[d\mu(\phi^*\mathbf{G})/d\mu(\mathbf{G})]^{1/2}, \tag{4.33}$$

where ϕ^* is a map from $\mathcal{S}'$ into $\mathcal{S}'$ defined by

$$(\phi^*\mathbf{G}, \mathbf{g}) = (\mathbf{G}, \tilde{\phi}\mathbf{g}) \tag{4.34}$$

with

$$(\tilde{\phi}g)_j(\mathbf{x}) = (\partial_j\phi_k)(\mathbf{x})(g_k \circ \phi)(\mathbf{x}).$$

Also, $\chi_\phi(\mathbf{G})$ is a multiplier for $U(\phi)$ and $d\mu(\phi^*\mathbf{G})/d\mu(\mathbf{G})$ is the Radon–Nikodym derivative. For this derivative to exist, it is necessary that the measure μ be quasi-invariant.

(4) Finally,

$$(\exp[iB(\mathbf{h})]\Psi)(\mathbf{G}) = Z_\mathbf{h}(\mathbf{G})\Psi(\mathbf{h}^*\mathbf{G})[d\mu(\mathbf{h}^*\mathbf{G})/d\mu(\mathbf{G})]^{1/2}, \tag{4.35}$$

where $\mathbf{h}^*$ is another map from $\mathcal{S}'$ into $\mathcal{S}'$, here defined by

$$(\mathbf{h}^*\mathbf{G}, \mathbf{g}) = (\mathbf{G} + \nabla \times \mathbf{h}, \mathbf{g}), \tag{4.36}$$

and $Z_\mathbf{h}(\mathbf{G})$ is a multiplier for $B(\mathbf{h})$.

The multipliers χ_ϕ and $Z_\mathbf{h}$ are complex valued functions of modulus one. For the group multiplication law (4.28) to be obeyed, the multipliers must satisfy

$$\begin{aligned}
&\chi_{\phi_1}(\mathbf{G})\chi_{\phi_2}(\phi_1^*\mathbf{G}) = \chi_{\phi_2\circ\phi_1}(\mathbf{G}),\\
&Z_{\mathbf{h}_1}(\mathbf{G})Z_{\mathbf{h}_2}(\mathbf{h}_1^*\mathbf{G}) = Z_{\mathbf{h}_1+\mathbf{h}_2}(\mathbf{G}),\\
&\chi_\phi(\mathbf{G})Z_\mathbf{h}(\phi^*\mathbf{G}) = Z_{\tilde{\phi}\mathbf{h}}(\mathbf{G})\chi_\phi((\tilde{\phi}\mathbf{h})^*\mathbf{G}).
\end{aligned} \tag{4.37}$$

While Eqs. (4.37) follow from the general representation theory, an additional constraint on the multipliers follows from the initial value equation $(\nabla \cdot \mathbf{B})(\mathbf{x}) = 0$. To satisfy $(\nabla \cdot \mathbf{B})(\mathbf{x}) = 0$, we require

$$\exp[iB(\nabla f)] = I. \tag{4.38}$$

From Eq. (4.35) and the relation $(\nabla f)^*\mathbf{G} = \mathbf{G}$, we see that Eq. (4.38) can be satisfied if and only if

$$Z_{\nabla f}(\mathbf{G}) = 1. \tag{4.39}$$

The multiplier law (4.37) then requires that

$$Z_\mathbf{h}(\mathbf{G}) = Z_{\mathbf{h}_\perp}(\mathbf{G}), \tag{4.40}$$

where $\mathbf{h}_\perp$ is the transverse part of $\mathbf{h}$.

Thus, in the representation theory of the local currents the initial value equations, which as we saw in Sec. IV C replace gauge invariance, have the following effect. The equation $\nabla \cdot \mathbf{E}(\mathbf{x}) = e\rho(\mathbf{x})$ is used to define $\rho(\mathbf{x})$. The action of the operator $\exp[i\rho(f)]$ can be expressed entirely in terms of a multiplier. The equation $\nabla \cdot \mathbf{B} = 0$ is equivalent to a constraint on the multipliers.

Remark: The algebra of ρ, $\mathbf{P}$, $\mathbf{E}$, and $\mathbf{B}$ is the same for either Bose or Fermi particles. Representations of the algebra corresponding to systems of bosons or fermions are distinguished by the choice of the multipliers χ_ϕ.[11,21] In the next subsection we show that the multipliers $Z_\mathbf{h}$ are related to magnetic charge.

F. Magnetic charge

In this subsection we show how magnetic charges can be incorporated into the current algebra formulation. We suppose each particle has a magnetic charge (monopole) q in addition to an electric charge e and mass m. The current algebra is given by Eqs. (4.1)–(4.4) with the following modifications:

$$\begin{aligned}
[J_k(\mathbf{x}), J_k(\mathbf{y})] = {}& i\left(\frac{\hbar}{m}\right)\frac{\partial}{\partial y_i}[\delta(\mathbf{x}-\mathbf{y})J_k(\mathbf{y})]\\
&- i\left(\frac{\hbar}{m}\right)\frac{\partial}{\partial x_k}[\delta(\mathbf{x}-\mathbf{y})J_i(\mathbf{x})]\\
&+ i\left(\frac{\hbar}{m^2c}\right)\epsilon_{ikj}\rho(\mathbf{x})[eB_j(\mathbf{x}) - qE_j(\mathbf{x})]\delta(\mathbf{x}-\mathbf{y})
\end{aligned} \tag{4.41}$$

and

$$[J_i(\mathbf{x}), B_k(\mathbf{y})] = i(\hbar q/m)\rho(\mathbf{x})\delta_{ik}\delta(\mathbf{x}-\mathbf{y}). \tag{4.42}$$

The Hamiltonian is still given by Eq. (4.5), but the equations of motion are modified as follows:

$$\begin{aligned}
(1/c)\partial\mathbf{B}(\mathbf{x})/\partial t &= i(1/c\hbar)[H, \mathbf{B}(\mathbf{x})]\\
&= -(q/c)\mathbf{J}(\mathbf{x}) - \nabla \times \mathbf{E}(\mathbf{x})
\end{aligned} \tag{4.43}$$

and

$$\begin{aligned}
m\dot{J}_k(\mathbf{x}) &= i(m/\hbar)[H, J_k(\mathbf{x})]\\
&= e\rho(\mathbf{x})E_k(\mathbf{x}) + q\rho(\mathbf{x})B_k(\mathbf{x}) + (1/2c)[\mathbf{J} \times (e\mathbf{B} - q\mathbf{E})\\
&\quad - (e\mathbf{B} - q\mathbf{E}) \times \mathbf{J}]_k(\mathbf{x}) - (\hbar^2/m)\partial_j\theta_{kj}.
\end{aligned} \tag{4.44}$$

We may again note that the interaction appears explicitly in the current algebra, rather than in the Hamiltonian.

The Lie algebra formed by $\rho(\mathbf{x})$, $\mathbf{P}(\mathbf{x})$, $\mathbf{E}(\mathbf{x})$, and $\mathbf{B}(\mathbf{x})$ is the same as before, Eqs. (4.14)–(4.18). However, in order to interpret $\mathbf{P}$ as the total momentum, i.e., the generator of space translations, we must now impose the initial value equations

$$\begin{aligned}
&(\nabla \cdot \mathbf{E})(\mathbf{x}) = e\rho(\mathbf{x})\\
&\text{and}\\
&(\nabla \cdot \mathbf{B})(\mathbf{x}) = q\rho(\mathbf{x}).
\end{aligned} \tag{4.45}$$

The Lie algebra again contains no explicit dependence on the interaction, which reappears in the Hamiltonian.

The coupling constants e and q appear explicitly only in the initial value equations (4.45).

The representation theory follows along the lines indicated in the previous section, except that, to satisfy the initial value equation, we now require

$$\exp[iB(\nabla f)] = \exp[-iq\rho(f)] = \exp[i(q/e)E(\nabla f)]. \tag{4.46}$$

This imposes a constraint on the multiplier $Z_{\mathbf{h}}$, namely,

$$Z_{\mathbf{h}}(\mathbf{G}) = \exp[i(q/e)(\mathbf{G}, \mathbf{h}_{\parallel})]\, Z_{\mathbf{h}_{\perp}}(\mathbf{G}), \tag{4.47}$$

where $\mathbf{h}_{\parallel}$ and $\mathbf{h}_{\perp}$ are, respectively, the longitudinal and transverse parts of $\mathbf{h}$.

Remarks: (1) The charge quantization condition for e and q must result from the requirement of rotational invariance, as it does in the usual formulation of the magnetic monopole in nonrelativistic quantum mechanics.[22] This is similar to the way in which translation invariance imposes constraints (the initial value equations) on the electric and magnetic fields.

(2) Since we are dealing with a single species of particle, the algebra with magnetic charge can be obtained formally from the algebra with only an electric charge by performing a duality rotation on $\mathbf{E}$, $\mathbf{B}$, e, and q.[23] If this transformation could be unitarily implemented, then we could obtain a solution to the multiplier equations with the magnetic charge constraint, Eq. (4.47), from a solution with the constraint Eq. (4.40), corresponding to no magnetic charge, which has at least one solution ($\chi_{\phi} = 1$, $Z_{\mathbf{h}} = 1$). However, Strocchi[24] has shown in the relativistic field theory framework used in Ref. 2 (the potential is local and a Lorentz 4-vector) that the duality rotation is not unitarily implementable. Furthermore, the existence of magnetic charges conflicts with the possibility of formulating the theory in terms of local Wightman fields (unless there are two kinds of photons). This raises the question of whether, in our formulation of nonrelativistic quantum electrodynamics, the multiplier equations for the magnetic charge case have any solution. This should not be confused with the question of the consistency of the description of magnetic monopoles in nonrelativistic quantum mechanics. That theory is consistent; however, the electromagnetic fields are not quantized.

V. QUANTUM ELECTRODYNAMICS FOR SEVERAL SPECIES OF CHARGED PARTICLE

To deal with several species of particles (having mass m_i and charge e_i) we need to introduce the particle number density $\rho^i(\mathbf{x})$ and the particle flux density $\mathbf{J}^i(\mathbf{x})$ for each of the i species of particle. The currents $\rho^i(\mathbf{x})$ and $\mathbf{J}^i(\mathbf{x})$ together with the fields $\mathbf{E}$ and $\mathbf{B}$ generate an equal-time algebra as follows. The currents associated with different species of particle commute. For a given species, the commutation relations for the currents are the same as we have obtained for a single species, Eqs. (4.1) and (4.2), but with the appropriate mass and charge. The commutation relations between $\mathbf{E}$ and $\mathbf{B}$ are still given by Eq. (4.3). The commutation relation between $\mathbf{J}^i$ and $\mathbf{E}$ is given by Eq. (4.4), but with e_i, m_i, and $\rho^i(\mathbf{x})$ appearing on the right-hand side.

The Hamiltonian, given by

$$H = \sum_i \tfrac{1}{8}(\hbar^2/m_i) \int d^3x \sum_{j=1}^{3} K_j^i(\mathbf{x})^{\dagger}[1/\rho^i(\mathbf{x})] K_j^i(\mathbf{x}) + \tfrac{1}{2}\int d^3x\, (\mathbf{E}^2 + \mathbf{B}^2)(\mathbf{x}), \tag{5.1}$$

with

$$\mathbf{K}^i(\mathbf{x}) = \nabla\rho^i(\mathbf{x}) + (2im_i/\hbar)\,\mathbf{J}^i(\mathbf{x}), \tag{5.2}$$

leads to the appropriate equations of motion: Maxwell's equations (4.7) and (4.8) but with $e\mathbf{J}(\mathbf{x})$ replaced by $\sum_i e_i \mathbf{J}^i(\mathbf{x})$, and current conservation (4.6) and the force equation (4.9) for each species of particle. The interaction appears explicitly in the equal-time algebra, and the Hamiltonian is the sum of the (formally) free Hamiltonians for each species of particles and the Hamiltonian for the electromagnetic field.

The nonlinearity in the equal-time algebra can be eliminated by replacing each current $\mathbf{J}^i(\mathbf{x})$ by

$$\tilde{\mathbf{J}}^i(\mathbf{x}) = \mathbf{J}^i(\mathbf{x}) + (e_i/m_i c)\,\rho^i(\mathbf{x})\,\mathbf{A}(\mathbf{x}), \tag{5.3}$$

where the quasilocal field $\mathbf{A}(\mathbf{x})$ is given by

$$\mathbf{A}(\mathbf{x}) = (4\pi)^{-1}\int d^3y\, |\mathbf{x}-\mathbf{y}|^{-1}\nabla\times\mathbf{B}(\mathbf{y}). \tag{5.4}$$

Then $\rho^i(\mathbf{x})$ and $\mathbf{J}^i(\mathbf{x})$ satisfy the same commutation relations as do $\rho(\mathbf{x})$ and $\mathbf{J}(\mathbf{x})$ for a single species of particle *without* electromagnetic interaction, Eq. (2.1). However, the commutation relation of $\tilde{\mathbf{J}}^i(\mathbf{x})$ with $\mathbf{E}(\mathbf{x})$ is nonlocal:

$$[\tilde{J}_k^i(\mathbf{x}), E_j(\mathbf{y})] = \left(-\frac{i\hbar e_i}{4\pi m_i}\right)\rho^i(\mathbf{x})\,\frac{\partial^2}{\partial y_j \partial y_k}\,|\mathbf{x}-\mathbf{y}|^{-1}. \tag{5.5}$$

The quantity $\mathbf{A}(\mathbf{x})$ has the same properties as the vector potential in Coulomb gauge, namely,

$$(\nabla\cdot\mathbf{A})(\mathbf{x}) = 0, \quad (\nabla\times\mathbf{A})(\mathbf{x}) = \mathbf{B}(\mathbf{x}), \quad -\frac{1}{c}\frac{\partial\mathbf{A}(\mathbf{x})}{\partial t} = \mathbf{E}_{\perp}(\mathbf{x}),$$

and has the commutation relations given by Eq. (1.5). Thus, one has the choice of using a local but nonlinear algebra or a linear but nonlocal algebra.

Remark: The interaction of a magnetic charge q_i associated with each species of particle can be introduced in the same manner as in Sec. IV F. When one does this, one is lead by formal calculations to a result that appears to violate the Jacobi identity, namely,

$$[J^1(\mathbf{f}), [J^2(\mathbf{g}), J^2(\mathbf{h})]] + \circ\circ\circ + \cdots = \int d^3x\, \rho^1(\mathbf{x})\,\rho^2(\mathbf{x})\,\mathbf{f}(\mathbf{x})\cdot(\mathbf{g}\times\mathbf{h})(\mathbf{x})(e_1q_2 - e_2q_1). \tag{5.6}$$

A similar situation arises in the usual quantum mechanical description of electric and magnetic charges using a singular vector potential.[25] In that case the problem can be resolved by carefully considering the domain on which the (unbounded) operators can be applied. We believe similar considerations would apply in the present case.

VI. CONCLUDING REMARKS

We have presented a formulation of nonrelativistic quantum electrodynamics using local currents and the

electromagnetic field. The formulation is local and gauge invariant. We expect that a similar formulation of relativistic quantum electrodynamics is possible, and in fact previous work[5] on the electrodynamics of charged scalar mesons supports this view.

At this point we briefly discuss the relationship between this approach and one which uses a local Lorentz covariant potential. The potential $A_\mu(x)$ is defined on a Hilbert space $\mathcal{H}$. An indefinite metric must be introduced to obtain a unitary representation of the Poincaré group. Subspaces $\mathcal{H}'$ and $\mathcal{H}''$ are defined on which the metric is positive and zero respectively. The physical Hilbert space is given by the quotient space $\mathcal{H}_{\text{phys}} = \mathcal{H}'/\mathcal{H}''$. Maxwell's equations then hold between matrix elements of states in $\mathcal{H}_{\text{phys}}$. It is our belief that by formulating the theory using the **E** and **B** fields and local currents, we are describing $\mathcal{H}_{\text{phys}}$ directly. Furthermore, the equations of motion, including Maxwell's, are to be interpreted as operator equations. As far as we can tell, these results do not violate any known theorem requiring the use of an indefinite metric.[26–28]

The local currents are fields which carry zero charge. Hence, an irreducible representation of the equal-time algebra describes a fixed charge sector. We believe the formal expression for the Hamiltonian, Eq. (4.5), can be given a well-defined meaning in more than one irreducible representation of the algebra and that these inequivalent representations correspond to different charge sectors. In Appendix C we give a simple example to illustrate how this can work. The condition that the ground state be translation invariant will select out the zero charge sector. For the charge sectors, the Hamiltonian will not have a unique ground state. A reducible representation of the equal-time algebra may then be needed to obtain a representation of the translation group.

When using a covariant potential there arises the question of which operators are physical. There are several possible definitions of "physical" operators.[2,29] In the present formulation, the local currents and the electromagnetic fields appear to be the natural candidates for a complete set of observable fields.

ACKNOWLEDGMENTS

The authors wish to offer warm thanks to F. Rohrlich F. Strocchi, and A. S. Wightman for their valued comments and suggestions. We would also like to acknowledge the hospitality of the Aspen Center for Physics, where some of the work reported here was carried out.

APPENDIX A: DERIVATION OF EQ. (3.13)

In this appendix we prove that

$$(\Omega, \mathbf{B}(\mathbf{x}) \exp[iE(\mathbf{f})]\Omega) = \tfrac{1}{2}(\nabla\times\mathbf{f})(\mathbf{x})(\Omega, \exp[iE(\mathbf{f})]\Omega). \tag{A1}$$

First, we need to show that

$$\exp[iE(\mathbf{f})]\,\mathbf{B}(\mathbf{x}) = [\mathbf{B}(\mathbf{x}) - \nabla\times\mathbf{f}(\mathbf{x})]\exp[iE(\mathbf{f})]. \tag{A2}$$

This may be proved by using the relationship

$$\exp(A)\,\mathbf{B}\exp(-A) = \sum_{n=1}^{\infty} \frac{1}{n!}(\mathrm{ad}A)^n\,\mathbf{B}, \tag{A3}$$

where $(\mathrm{ad}A)\,\mathbf{B} = [A, \mathbf{B}]$, and the commutation relation (3.3).

We need to assume there is an antiunitary time reversal operator T having the properties that

$$\begin{aligned} &T\mathbf{E}(\mathbf{x})\,T^* = \mathbf{E}(\mathbf{x}),\\ &T\mathbf{B}(\mathbf{x})\,T^* = -\mathbf{B}(\mathbf{x}),\\ &T\Omega = \Omega. \end{aligned} \tag{A4}$$

Time reversal invariance then implies that

$$\begin{aligned} (\Omega, \mathbf{B}(\mathbf{x})\exp[iE(\mathbf{f})]\,\Omega) &= -(\mathbf{B}(\mathbf{x})\exp[-iE(\mathbf{f})]\,\Omega, \Omega)\\ &= -(\Omega, \exp[iE(\mathbf{f})]\,\mathbf{B}(\mathbf{x})\Omega). \end{aligned} \tag{A5}$$

Finally, substituting Eq. (A2) into Eq. (A5), we obtain Eq. (A1).

APPENDIX B: DERIVATION OF EQ. (4.25)

In this appendix we prove that

$$U(\phi)\exp[iE(\mathbf{g})] = \exp[iE(\mathbf{W}_\phi(\mathbf{g}))]\,U(\phi), \tag{B1}$$

where

$$\mathbf{W}_\phi(\mathbf{g})(\mathbf{x}) = (\nabla\phi_k)(\mathbf{x})(g_k\circ\phi)(\mathbf{x}). \tag{B2}$$

We start by considering the quantity:

$$U(\phi_t^{\mathbf{h}}) = \exp[itP(\mathbf{h})], \tag{B3}$$

where $\phi_t^{\mathbf{h}}$ is the flow for time t by the vector field $\mathbf{h}(\mathbf{x})$, defined by Eq. (2.5). By using Eqs. (A3) and (4.17) it can be shown that

$$\exp[itP(\mathbf{h})]\,E(\mathbf{g})\exp[-itP(\mathbf{h})] = E(\exp[tW]\,\mathbf{g}), \tag{B4}$$

where

$$(W\mathbf{g})_i(\mathbf{x}) = (\mathbf{h}\cdot\nabla)\,g_i(\mathbf{x}) + (\partial_i\mathbf{h})\cdot\mathbf{g}(\mathbf{x}). \tag{B5}$$

It then follows that

$$\exp[itP(\mathbf{h})]\exp[iE(\mathbf{g})] = \exp[iE(\exp(tW)\mathbf{g})]\exp[iP(\mathbf{h})]. \tag{B6}$$

We next introduce the function

$$\mathbf{S}(\mathbf{x}, t) = \exp(tW)\,\mathbf{g}(\mathbf{x}), \tag{B7}$$

which evidently satisfies the differential equation

$$\frac{d\mathbf{S}(\mathbf{x}, t)}{dt} = W\mathbf{S}(\mathbf{x}, t), \tag{B8}$$

with the boundary condition

$$\mathbf{S}(\mathbf{x}, 0) = \mathbf{g}(\mathbf{x}).$$

Since Eq. (B8) has a unique solution, we can prove that

$$\mathbf{S}(\mathbf{x}, t) = [\nabla(\phi_t^{\mathbf{h}})_k](\mathbf{x})(g_k\circ\phi_t^{\mathbf{h}})(\mathbf{x}) \tag{B9}$$

by showing that the right-hand side of Eq. (B9) satisfies Eq. (B8). Using the chain rule and Eq. (2.5), we can obtain from Eq. (B9) that

$$\frac{dS_\alpha(\mathbf{x}, t)}{dt} = \partial_\alpha(\phi_t^{\mathbf{h}})_\beta(\mathbf{x})[(\partial_\beta h_k)\,g_k + (\mathbf{h}\cdot\nabla)\,g_\beta]\circ\phi_t^{\mathbf{h}}(\mathbf{x}). \tag{B10}$$

Then, from Eqs. (B5) and (B9) we obtain

$$(W\mathbf{S})_\alpha(\mathbf{x}, t) = h_j(\mathbf{x})\partial_\alpha\partial_j[(\phi_t^{\mathbf{h}})_\beta(\mathbf{x})](g_\beta\circ\phi_t^{\mathbf{h}})(\mathbf{x})$$

$$+\partial_\alpha[(\phi_t^h)_\beta(x)][(h\circ\nabla)(\phi_t^h)_k(x)][(\partial_k g_\beta)\circ\phi_t^h(x)]$$
$$+(\partial_\alpha h_j)(x)\,\partial_j(\phi_t^h)_\beta(x)(g_\beta\circ\phi_t^h)(x). \tag{B11}$$

At this point we need to prove the following:

Lemma: $h_j(x)\,\partial_j(\phi_t^h)_k(x)=(h_k\circ\phi_t^h)(x)$.

Proof: The flow $\phi_t^h(x)$ represents the position of a particle which starts at point x and moves in the velocity field $h(x)$ for a time t. As a result, it is easy to see that Eq. (2.5) and its associated boundary condition implies that

$$\phi_{t+\epsilon}^h(x)=\phi_t^h(\phi_\epsilon^h(x)). \tag{B12}$$

Taking the derivative of (B12) with respect to ϵ, one obtains

$$\frac{d}{d\epsilon}\,\phi_{t+\epsilon}^h(x)=\left(\frac{d}{d\epsilon}\,\phi_\epsilon^h(x)\circ\nabla\right)(\phi_t^h)(\phi_\epsilon^h(x)). \tag{B13}$$

Setting $\epsilon=0$ and using Eq. (2.5), we obtain the lemma.

After some manipulation and using the lemma, Eq. (B11) can be reduced to Eq. (B10). Combining Eqs. (B6) and (B9), we obtain Eq. (B1) when $V(\phi)$ is given by Eq. (B3). We then extend this result to all $V(\phi)$.

APPENDIX C: THE HARMONIC OSCILLATOR HAMILTONIAN EXPRESSED IN TERMS OF BILINEARS

Formally, the local currents can be defined as bilinear expressions in the canonical fields. A one-dimensional prototype of this situation is provided by the harmonic oscillator in which x and and p are analogous to the canonical fields and the bilinears $S=x^2$ and $\dot{S}=xp+px$ are analogous to the local currents. We use this example to illustrate how the formal expression for the Hamiltonian in terms of bilinears can be defined as an operator in two different ways.

The canonical operators x and $p=(1/i)(d/dx)$ are defined on the Hilbert space $\mathcal{H}=L^2(-\infty,\infty)$. The harmonic oscillator Hamiltonian

$$H=\tfrac{1}{2}(p^2+x^2)-\tfrac{1}{2} \tag{C1}$$

has a discrete spectrum

$$H\psi_n=n\psi_n,$$

where $n=0,1,2,\cdots$. The eigenfunctions with even n are symmetric functions

$$\psi_{2n}(x)=\psi_{2n}(-x),$$

while those with odd n are antisymmetric

$$\psi_{2n+1}(x)=-\psi_{2n+1}(-x).$$

The operators S and $\dot{S}$ satisfy the equal-time algebra

$$[S,\mathring{S}]=4iS. \tag{C2}$$

The harmonic oscillator can be written formally in terms of S and $\dot{S}$ as

$$H=\tfrac{1}{8}(\mathring{S}-i)(1/S)(\dot{S}+i)+\tfrac{1}{2}S-\text{const}. \tag{C3}$$

This expression can be given a well-defined meaning in two ways in an irreducible representation of the $S,\mathring{S}$ algebra, as follows.

The Hilbert space $\mathcal{H}$ can be decomposed into an even and odd parity subspace;

$$\mathcal{H}=\mathcal{H}_+\oplus\mathcal{H}_-,$$

where

$$\mathcal{H}_+=\{\psi\in\mathcal{H};\ \psi(x)=\psi(-x)\},$$

and

$$\mathcal{H}_-=\{\psi\in\mathcal{H};\ \psi(x)=-\psi(-x)\}.$$

The operators S and $\dot{S}$ leave each subspace invariant, and, restricted to each subspace, form an irreducible representation of the equal-time algebra. Since each subspace is isomorphic to the Hilbert space $\mathcal{H}'=L^2(0\ \infty)$, these representations are unitarily equivalent. $\mathcal{H}'$ can be spanned by either $\{\psi_{2n}\}$ or $\{\psi_{2n+1}\}$, $n=0,1,2,\cdots$, and each ψ_n is an eigenfunction of the harmonic oscillator Hamiltonian. As a result, we can define two different Hamiltonians on $\mathcal{H}'$;

(1) H_0 defined by $H_0\psi_{2n}=2n\psi_{2n}$

and

(2) H_1 defined by $H_1\psi_{2n+1}=2n\psi_{2n+1}$.

The formal expression for the Hamiltonian in terms of S and $\mathring{S}$, Eq. (C3), can be given a well-defined meaning corresponding to either H_0 or H_1, by expressing it in the factored form

$$H_0=\tfrac{1}{8}[\mathring{S}-i(2S-1)]^\dagger(1/S)[\dot{S}-i(2S-1)] \tag{C4}$$

and

$$H_1=\tfrac{1}{8}[\mathring{S}-i(2S-3)]^\dagger(1/S)[\dot{S}-i(2S-3)]. \tag{C5}$$

Alternatively, these representations of the S, $\dot{S}$ algebra could have been determined without reference to the underlying canonical operators x and p through the use of a generating functional

$$L(\alpha)=(\Omega,\exp[i\alpha S]\Omega)$$

and the ground state conditions which follow from the Hamiltonians (C4) and (C5):

$$[\dot{S}-i(2S-1)]\,\Omega_0=0, \tag{C6}$$

$$[\mathring{S}-i(2S-3)]\,\Omega_1=0. \tag{C7}$$

Remark: In $\mathcal{H}'$ the operator x is self-adjoint, but p is not. Thus the canonical operators do not necessarily exist in an irreducible representation of the bilinear algebra, even though the bilinears and the algebra they satisfy were abstracted from the canonical theory.

In the simple one-dimensional case considered here the representations of S and $\dot{S}$ are unitarily equivalent but the Hamiltonian can be defined in two ways. For the local currents we believe the formal expression for the Hamiltonian can be given a well-defined meaning in the inequivalent representations which correspond to different charge sectors.

Finally, we note that a similar problem is encountered in distinguishing bosons from fermions when local currents are used in the formulation of nonrelativistic quantum mechanics. In this case the local current algebra and the formal expression for the Hamiltonian are the same, but unitarily inequivalent representations

of the current algebra are associated with different particle statistics,[11,14] and the Hamiltonian must be defined differently in each of the different representations.

*Work supported by the Energy Research and Development Administration.

[1]See, for example, W. Heitler, *The Quantum Theory of Radiation* (Oxford U.P., Oxford 1954), 3rd ed.

[2]For an excellent recent review of these and related matters see F. Strocchi and A.S. Wightman, J. Math. Phys. **15**, 2198 (1974).

[3]B. DeWitt, Phys. Rev. **125**, 2189 (1962).

[4]S. Mandelstam, Ann. Phys. (N.Y.) **19**, 1 (1962).

[5]D.H. Sharp, Phys. Rev. **165**, 1867 (1968).

[6]Strocchi and Wightman, Ref. (2), have recently used local currents to formulate the problem of a single nonrelativistic particle interacting with an external electromagnetic field in a gauge invariant way. In this paper they also present an argument which in our opinion refutes the claim that the Aharonov—Bohm Effect[7] is one which depends directly on the potentials $\mathbf{A}(\mathbf{x})$ and $\phi(\mathbf{x})$ rather than the field strengths $\mathbf{E}(\mathbf{x})$ and $\mathbf{B}(\mathbf{x})$. This is an important point, because if the Aharonov—Bohm Effect did have that implication, it would contradict the idea that quantum electrodynamics could be written in a local and manifestly gauge invariant form.

[7]Y. Aharonov and D. Bohm, Phys. Rev. **115**, 485 (1959); **123**, 1511 (1961).

[8]R.F. Dashen and D.H. Sharp, Phys. Rev. **165**, 1857 (1968).

[9]R. Haag, Dan. Mat. Fys. Medd. **29**, No. 12 (1955). An excellent discussion of this theorem and its implications can be found in A.S. Wightman, "Relativistic Dynamics of Quantized Fields," in *High Energy Electromagnetic Interactions and Field Theory*, edited by M. Lêvy (Gordon and Breach, New York, 1967), Sec. VI.

[10]I. Gel'fand and N. Vilenkin, *Generalized Functions* (Academic, New York, 1964), Vol. 4.

[11]G.A. Goldin, J. Math. Phys. **12**, 462 (1971).

[12]A completely consistent formulation of the quantum electrodynamics of the free electromagnetic field in terms of $\mathbf{E}(\mathbf{x})$ and $\mathbf{B}(\mathbf{x})$ has been given long ago by P. Jordan and W. Pauli, Z. Phys. **47**, 151 (1928). However, the treatment of the problem which we give in Sec. III differs from theirs, particularly in its emphasis on the generating functional $L(f)$.

[13]Brief reviews of the representation theory of local current algebras, the definition of $1/\rho(\mathbf{x})$ and related subjects can be found in G.A. Goldin and D.H. Sharp, "Lie Algebras of Local Currents and Their Representations," in *Group Representations in Mathematics and Physics, Battelle Seattle* 1969 *Rencoutres*, edited by V. Bargmann (Springer, New York, 1970), pp. 300—11, and D.H. Sharp, "What we have learned about representing local nonrelativistic current algebras," in *Local Currents and Their Applications*, edited by D.H. Sharp and A.S. Wightman (North-Holland, Amsterdam, 1974), pp. 85—98.

[14]R. Menikoff and D.H. Sharp, J. Math. Phys. **16**, 2341 (1975).

[15]In defining Fourier transforms, we adopt the convention

$$f(\mathbf{x}) = \int \frac{d^3p}{(2\pi)^3}\, e^{i\mathbf{p}\cdot\mathbf{x}} \tilde{f}(\mathbf{p}) \quad \text{and} \quad f(\mathbf{p}) = \int d^3x\, e^{-i\mathbf{p}\cdot\mathbf{x}} f(\mathbf{x}).$$

[16]In order for Eq. (3.14) to uniquely determine the solution Eq. (3.15), it must be supplemented with suitable boundary conditions. We will not go into this point here. For a discussion of the boundary conditions which uniquely determine the solutions of a functional differential equation, similar to Eq. (3.14), which defines the representations of a nonrelativistic current algebra, see Ref. 17.

[17]G.A. Goldin, J. Grodnik, R.T Powers, and D.H. Sharp, J. Math. Phys. **15**, 88 (1974).

[18]We shall work in Heaviside—Lorentz units. For a useful comparison of different systems of units see the Appendix to J.D. Jackson, *Classical Electrodynamics* (Wiley, New York, 1962).

[19]The algebra given by Eqs. (4.1)—(4.4) with the commutation relations replaced by Poisson brackets also describes a relativistic fluid coupled to the electromagnetic field, when the variables are suitably interpreted. A derivation can be found in I. Bialynicki-Birula and Z. Iwinski, Rep. Math. Phys. **4**, 139 (1973).

[20]In Ref. 2 it has been shown for free quantum electrodynamics that the well-known gauges are related by a generalized gauge transformation. However, gauge invariance has not been proven in general.

[21]A.A. Dicke and G.A. Goldin (private communication).

[22]C.A. Hurst, Ann. Phys. (N.Y.) **50**, 51 (1968).

[23]D. Zwanziger, Phys. Rev. **176**, 1489 (1968).

[24]F. Strocchi (unpublished).

[25]D. Zwanziger, Phys. Rev. **176**, 1480 (1968).

[26]F. Strocchi, Phys. Rev. **162**, 1429 (1967).

[27]F. Strocchi, Phys. Rev. D **2**, 2334 (1970).

[28]R. Ferrari, L. Picasso, and F. Strocchi, Comm. Math. Phys. **35**, 25 (1974).

[29]D. Zwanziger, Phys. Rev. D **14**, 2570 (1976).

Meta universe

Sushil K. Srivastava and Mahendra P. Pathak

Department of Mathematics, University of Gorakhpur, Gorakhpur, India
(Received 14 April 1976)

Foster and Ray have pointed out that since tachyons are particles having energy and momentum, they should hence contribute to the gravitational field through the energy–momentum tensor. Following them, a spacelike metric for the tachyon dust model has been proposed, and the problem of condensation of tachyons, after a slight perturbation of the model, has been examined. The tetrad technique has been frequently used for the purpose.

1. INTRODUCTION

A. Background

Around the turn of the century shortly before Einstein had published his revolutionary paper on the special theory of relativity, Sommerfeld[1] examined the problem of accelerating particles to velocities greater than that of light $c(=3\times10^{10}$ cm/sec) and concluded that at such velocities particles would behave in a patently absurd fashion as they would be accelerated upon loss of energy But at that time physicists were not prepared to accept this absurd notion. Instead, Einstein's theory remedied this state of affairs, by using the mass-variation equation, $m=m_0/(1-v^2/c^2)^{1/2}$, which shows that m increases infinitely as v approaches c. Thus the velocity of light was a barrier.

This argument was too strong to assail. Yet in the sixties Bilaniuk, Deshpande, and Sudarshan[2] reexamined the problem studied by Sommerfeld and proposed the possibility of the existence of particles moving faster than light. They called these, meta particles.[3] Some other physicists, notably Terletskii[4] and Feinberg,[5] also reiterated the same conclusion. In 1967, Feinberg suggested that these superluminal particles be called tachyons.[5]

Later on, Foster and Ray[6] argued that since tachyons have energy and momentum, they should contribute to the gravitational field through the energy–momentum tensor T_{ij}. Gott[7] has also given the parallel argument that since tachyons have appeared in the context of the special theory of relativity hence they should also appear in the general theory of relativity.

Following these arguments, in the present paper, we have proposed a tachyon dust metric different from that proposed by Foster and Ray.[6] When this metric is used in the Einstein's field equations, using an energy–momentum tensor for dust, solutions exist for spacelike 4-velocities. Like Foster and Ray,[6] we also interpret it as a solution for tachyon dust.

In Sec. 2, Part A, we choose the surface $x^3=\text{const}$ as a hypersurface, such that spacelike geodesics along $(x^1, x^2, t)=\text{const}$ will be orthogonal to it. Thus we have chosen a coordinate system in which coordinates x^1, x^2, and t will be constant for each tachyon, i.e., we assume the motion of a tachyon along the x^3 axis only and its velocity vanishing along the x^1, x^2, and t axes. Since the meta universe (universe filled with tachyons) will be a spacelike universe, hence on the basis of orthogonality of $x^3=\text{const}$ and $(x^1, x^2, t)=\text{const}$, we assume a spacelike metric

$$ds^2=e^{\mu}\{(dx^1)^2+(dx^2)^2\}+e^{\nu}(dx^3)^2-e^{\mu}(dt)^2,$$

where $\mu=\mu(x^3, t)$ and $\nu=\nu(x^3, t)$.

On substituting the metric tensor of this metric into Einstein's field equations, we obtain a tachyon dust metric in the form

$$ds^2=\frac{R^2(x^3)}{(1+t)^2}\{(dx^1)^2+(dx^2)^2\}+(dx^3)^2-\frac{R^2(x^3)}{(1+t)^2}(dt)^2.$$

In Sec. 2, Part B, we have obtained the momentum flux[6] of the tachyon dust as

$$\rho=8C^2/27\eta^2,$$

where η is a spacelike quantity.

In Sec. 3, we have derived the spacelike counterpart of Chaudhuri's equation[8] which represents the Einstein's field equations in a simple form. Here we have used a projection tensor H_{ij} to project a quantity from $x^3=\text{const}$ to $(x^1, x^2, t)=\text{const}$. This tensor has been defined as $H_{ij}=g_{ij}-u_iu_j$ in the notation section. In Sec. 4, we have considered the small perturbations of the proposed spacelike metric by the tetrad technique given by Ellis and Stewart[9] and Ellis.[10,11] We have frequently used the method of perturbations proposed by Johri.[12] In the last section, we have made an attempt to interpret the results. Here we have given the idea of condensation of tachyons regardless of whether or not we actually find tachyons existing. In the whole paper, we have treated tachyons as real particles rather than as hypothetical particles.

B. Notation

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

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

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

where u^a are the spacelike 4-velocities for the tachyon fluid so that $u^au_a=1$. The acceleration of the fluid is

$$\dot{u}_i=u_{i;j}u^j, \tag{1.2}$$

where the dot denotes $\partial/\partial t$.

The velocity gradient may be further split up as

$$u_{i;j}=w_{ij}+\sigma_{ij}+\tfrac{1}{3}\theta H_{ij}-\dot{u}_i u_j,$$

where $\theta=u^i_{;i}$ is the expansion scalar, $\sigma_{ij}=u_{(i;j)}+\dot{u}_{(i}u_{j)}-\frac{1}{3}\theta H_{ij}$ is the trace-free shear tensor, and $w_{ij}=u_{[i;j]}+u_{[i}\dot{u}_{j]}$ is the vorticity tensor.

Here H_{ij} is a tensor which projects a quantity from $x^3=\text{const}$ to $(x^1, x^2, t)=\text{const}$ defined by

$$H_{ij}=g_{ij}-u_iu_j, \tag{1.3}$$

i.e., $H_{ij}u^j=0$, $H^i_i=3$.

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

The Ricci rotation coefficients are defined by

$$\Gamma_{abc}=e_a\cdot\nabla_b e_c=e^i_a e_{ci;j}e^j_b$$

so that

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

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

The Lie derivative of e_b with respect to e_a is

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

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

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

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

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

$$\begin{aligned}R_{bd}&=\partial_d\Gamma^c_{cb}-\partial_c\Gamma^c_{db}-\Gamma^c_{cg}\Gamma^g_{db}+\Gamma^g_{cb}\Gamma^c_{gd}\\&=-(\Lambda-\tfrac{1\over2}\rho)H_{bd}-(\Lambda+\tfrac{1}{2}\rho)u_bu_d.\end{aligned}\tag{1.4}$$

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

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

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

$$u^a=\delta^a_3,\quad u_a=\delta^3_a.$$

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

$$\rho'+\rho\theta=0, \tag{1.6}$$

where $\theta=\theta_1+\theta_2+\theta_4$.

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

Perturbation of (1.6) gives

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

Therefore

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

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

2. THEORY

A. Tachyon dust metric

Suppose a space–time has the metric

$$ds^2=e^\mu\{(dx^1)^2+(dx^2)^2\}+e^\nu(dx^3)^2-e^\mu(dx^4)^2, \tag{2.1}$$

where $\mu=\mu(x^3, t)$, $\nu=\nu(x^3, t)$, and $x^4=t$.

Now by Dingle's formula,[13] we have Einstein's field equation (1.1) as

$$\tfrac{3}{4}(\mu')^2+\mu''+\ddot{\mu}/2e^\mu=\Lambda, \tag{2.2a}$$

$$\tfrac{1}{4}(\dot{\mu})^2/e^\mu+\ddot{\mu}/e^\mu-\tfrac{3}{4}(\mu')^2=\rho+\Lambda, \tag{2.2b}$$

$$\tfrac{3}{4}(\mu')^2+\mu''+\tfrac{1}{4}\dot{\mu}^2/e^\mu=\Lambda, \tag{2.2c}$$

$$\tfrac{1}{2}\mu'\dot{\nu}-\dot{\mu}'=0. \tag{2.2d}$$

Here $\dot{\mu}=\partial u/\partial t$ and $\mu'=\partial\mu/\partial x^3$.

Integrating Eq. (2.2d) partially with respect to t, we have

$$\log\mu'=\tfrac{1}{2}\nu+\log(2R'/R),$$

where R depends on x^3 only, or

$$\mu'=(2R'/R)e^{(1/2)\nu}.$$

Integrating again, we have

$$\mu=\chi(t)+2\int(R'/R)e^{\nu/2}\,dx^3.$$

Now putting $\nu=0$, we have

$$\mu=\chi(t)+2\int(R'/R)dx^3=\chi(t)+2\log R. \tag{2.3}$$

Equating Eqs. (2.2a) and (2.2c),

$$\tfrac{3}{4}(\mu')^2+\mu''+\ddot{\mu}/2e^\mu=\tfrac{3}{4}(\mu')^2+\mu''+\tfrac{1}{4}[(\dot{\mu})^2/e^\mu],$$

which gives

$$\ddot{\mu}=\tfrac{1}{2}(\dot{\mu})^2. \tag{2.4}$$

Equations (2.3) and (2.4) give the partial differential equation

$$\ddot{\chi}=\tfrac{1}{2}\dot{\chi}^2,$$

which possesses the solution

$$\chi=-2\log(t+2\alpha)\beta.$$

Therefore by Eq. (2.3),

$$\mu=\log[R(x^3)/\beta(t+2\alpha)]^2.$$

Since α, β are arbitrary constants, hence a coordinate system can be assumed such that $2\alpha=1=\beta$.

Thus the metric (2.1) reduces to

$$ds^2=\frac{R^2(x^3)}{(1+t)^2}\{(dx^1)^2+(dx^2)^2\}+(dx^3)^2-\frac{R^2(x^3)}{(1+t)^2}(dt)^2. \tag{2.5}$$

B. Momentum flux of the tachyon dust

Let us consider the metric (2.5) as the metric of the background model. The nonvanishing tetrad components corresponding to the components of the fundamental tensor in the line element (2.5) are given by

$$(e_1{}^1)_{t=0}=(e_2{}^2)_{t=0}=(e_4{}^4)_{t=0}=1/R.$$

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

$$[\theta_\nu]_{t=0} = -\gamma^\nu_{3\nu} = R'/R = \theta_0 \quad \text{(say)}$$

and the other components of γ^a_{bc} vanish.

The tetrad field equations (1.4) for a pressure-free tachyon fluid are

$$\begin{aligned} &D\theta_1 + \theta_1(\theta_1+\theta_2+\theta_4) = \Lambda - \rho/2, \\ &D\theta_2 + \theta_2(\theta_1+\theta_2+\theta_4) = \Lambda - \rho/2, \\ &D(\theta_1+\theta_2+\theta_4) + (\theta_1^2+\theta_2^2+\theta_4^2) = -\Lambda - \rho/2, \\ &D\theta_4 + \theta_4(\theta_1+\theta_2+\theta_4) = -\Lambda + \rho/2. \end{aligned} \tag{2.6}$$

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

Substituting $\theta_1 = \theta_2 = \theta_4 = \theta_0$ in (2.6), we have

$$\begin{aligned} &5D\theta_0 + 9\theta_0^2 = \Lambda - \tfrac{3}{2}\rho, \\ &D\theta_0 + 3\theta_0^2 = -\Lambda + \rho/2, \end{aligned} \tag{2.7}$$

which are easily integrable (for $\Lambda = 0$) giving

$$R^{9/4} = Cx^3 + A, \tag{2.8}$$

where C and A are integration constants.

This yields the momentum flux for the tachyon dust as

$$\rho = 8C^2/27\eta^2, \tag{2.9}$$

where $\eta = Cx^3 + A$.

Here η is interpreted as the proper path for tachyons. It plays the same role as proper time for tardyons (particles moving slower than light), because proper time becomes imaginary[14] for tachyons. In the present paper, we will call η a proper distance which is a spacelike quantity. Instead, we find that the space–time is singular on the hypersurface $\eta = 0$.

3. SPACELIKE COUNTERPART OF RAYCHAUDHURI'S EQUATION

In general relativity, arbitrary vector fields obey Ricci identity, hence

$$u_{a;d;c} - u_{a;c;d} = R_{abcd}u^a. \tag{3.1}$$

Now multiplying (3.1) by u^d, we get

$$u_{a;d;c}u^d - u_{a;c;d}u^d = R_{abcd}u^a u^d,$$

or

$$(u_{a;c})' - u'_{a;c} + u_{a;d}u^d_c + R_{abcd}u^a u^d = 0. \tag{3.2}$$

The three velocities in $x^3 = \text{const}$ can be given by the tensor defined[11] as

$$v_{ab} = H^c_a H^d_b u_{c;d}. \tag{3.3}$$

Equations (3.2) and (3.3) give

$$H^c_a H^d_b (v_{cd})' - u'_a u'_b - H^c_a H^d_b u'_{c;d} + v_{ad}v^d_b + R_{acbd}u^c u^d = 0. \tag{3.4}$$

Splitting up the tensor v_{ab} into symmetric and antisymmetric parts we have

$$v_{ab} = \theta_{ab} + w_{ab}, \tag{3.5}$$

where $\theta_{ab} = \theta_{(ab)}$ and $w_{ab} = w_{[ab]}$, i.e., $v_{(ab)} = \theta_{(ab)}$ and $v_{[ab]} = w_{[ab]}$.

Splitting θ_{ab} we have

$$\theta_{ab} = \sigma_{ab} + \tfrac{1}{3}\theta H_{ab},$$

where $\sigma^a_a = 0$, $\sigma_{ab} = \sigma_{(ab)}$, $\sigma_{ab}u^b = 0$, and $\theta = u^a_{;a}$.

Thus

$$v_{ab} = w_{ab} + \sigma_{ab} + \tfrac{1}{3}\theta H_{ab}. \tag{3.6}$$

Here θ_{ab} is the expansion tensor, w_{ab} is the vorticity tensor which describes the rotation, and σ_{ab} is the shear tensor which gives distortion.

Now multiplying (3.4) by g^{ab} we have

$$g^{ab}H^c_a H^d (v_{cd})' - g^{ab}u'_a u'_b - g^{ab}H^c_a H^d_b \dot{u}_{c;d} + g^{ab}v_{ad}v^d_b + g^{ab}R_{abcd}u^c u^d = 0,$$

or

$$\begin{aligned} g^{ab}H^c_a H^d_b(\sigma_{cd} + \tfrac{1}{3}\theta H_{cd} + w_{cd})' - g^{ab}u'_a u'_b - g^{ab}H^c_a H^d_b \dot{u}_{c;d} \\ + g^{ab}(\sigma_{ad} + \tfrac{1}{3}\theta H_{ad} + w_{ad})(\sigma^d_b + \tfrac{1}{3}\theta H^d_b + w^d_b) \\ + g^{ab}R_{abcd}u^c u^d = 0, \end{aligned}$$

or

$$\begin{aligned} g^{ab}H^c_a H^d_b(\sigma'_{cd} + \tfrac{1}{3}\theta' H_{cd} + w'_{cd}) - g^{ab}u'_a u'_b - g^{ab}H^c_a H^d_b u'_{c;d} \\ + (\sigma^d_b + \tfrac{1}{3}\theta H^d_b + w^d_b)^2 + R_{cd}u^c u^d = 0. \end{aligned}$$

This equation gives

$$\theta' - u'^a_{;a} + 2(\sigma^2 + \tfrac{1}{6}\theta^2 - \tfrac{1}{2}w^{ad}w_{ad}) + R_{cd}u^c u^d = 0. \tag{3.7}$$

From Einstein's field equation (1.1) we can compute

$$R_{cd} = (T_{cd} - \tfrac{1}{2}Tg_{cd}) + \Lambda g_{cd}. \tag{3.8}$$

Now Eq. (3.7) with the help of (3.8) reduces to

$$\theta' - u'^a_{;a} + 2(\sigma^2 + \tfrac{1}{6}\theta^2 - \tfrac{1}{2}w^{ad}w_{ad}) + T_{cd}u^c u^d - \tfrac{1}{2}Tg_{cd}u^c u^d + \Lambda g_{cd}u^c u^d = 0. \tag{3.9}$$

In general, the energy–momentum tensor T_{cd} is given as

$$T_{cd} = \underset{\sim}{\mu} u_c u_d + (q_c u_d + u_c q_d) + pH_{cd} + \pi_{cd}, \tag{3.10}$$

where q_a is the energy flux relative to u^a (which represent processes such as diffusion and heat conduction), p is the isotropic pressure, π_{ab} is the anisotropic matter pressure (due to processes such as viscosity, and $\underset{\sim}{\mu}$ is the total relativistic energy of matter measured by u^a given by the rotation $\underset{\sim}{\mu} = \rho(1+\epsilon)$, where ϵ is the specific internal momentum flux of the tachyon fluid.

From (3.10)

$$T = \underset{\sim}{\mu} + 3p \ (\text{since } u^a u_a = +1). \tag{3.11}$$

Now Eqs. (3.9) and (3.11) after some adjustments give

$$\theta' + \tfrac{1}{3}\theta^2 + 2(\sigma^2 - w^2) + \tfrac{1}{2}(\underset{\sim}{\mu} - 3p) + \Lambda = 0. \tag{3.12}$$

This is the spacelike counterpart of Raychaudhuri's equation[8] which gives the field equations in the simplest form.

4. PERTURBATIONS OF MOMENTUM FLUX IN THE PROPOSED TACHYON DUST MODEL

Since dust is characterized as pressure-free fluid, hence $p = 0$. Moreover, for simplicity, we take σ, w, and the cosmical constant Λ also vanishing. Hence in this case Eq. (3.12) reduces to

$$\theta' + \tfrac{1}{3}\theta^2 + \tfrac{1}{2}\rho = 0. \quad (4.1)$$

After a slight perturbation, Eq. (4.1) reduces to

$$D^2\theta + 2\theta_0 D\theta + \tfrac{1}{2}D\rho = 0, \quad (4.2)$$

where D has its previous meaning.

Now with the help of (1.8), (4.2) reduces to

$$D^2K + 2\theta_0 DK - \tfrac{1}{2}\rho K = 0,$$

or

$$D^2K + (2R'/R)DK - \tfrac{1}{2}\rho K = 0.$$

From Eq. (2.7) it becomes

$$D^2K + \frac{8C}{9(Cx^3+A)}DK - \frac{4C^2}{27(Cx^3+A)^2}K = 0. \quad (4.3)$$

Now substituting $Cx^3 + A = \eta$, Eq. (4.3) reduces to

$$\frac{\partial^2 K}{\partial \eta^2} + \frac{8}{9\eta}\frac{\partial K}{\partial \eta} - \frac{4}{27\eta^2}K = 0, \quad (4.4)$$

which has the solution

$$K = \alpha_1 \eta^{11/36} + \alpha_2 \eta^{-7/36}, \quad (4.5)$$

where α_1 and α_2 are integration constants.

Thus the differential equation (4.4) gives two solutions

$$K_1 = \alpha_1 \eta^{11/36}, \quad (4.6)$$

and

$$K_2 = \alpha_2 \eta^{-7/36}. \quad (4.7)$$

5. DISCUSSION

The momentum flux $\rho(=8C^2/27\eta^2)$ shows that the momentum flux of the meta universe at a certain time will decrease with the increase of the proper distance. Since there will be homogeneity in time, this process will continue forever. Further, the first solution (4.6) of the perturbation equation of momentum flux shows that the K $(=\delta\rho/\rho)$ will increase algebraically as $\eta^{11/36}$ while the second solution (4.7) shows its decrease as $\eta^{-7/36}$.

Let us suppose that like big-bang singularity (for tardyons) there exists an undiscovered singular point from which tachyons are emitting by some unknown process, and being superluminal particles they move away from the point source. Therefore, it is obvious that the momentum flux of the meta universe formed by them will go on decreasing as being a process forever. Let us further assume that there exists a region at some distance from the point source of tachyons assumed above, such that its momentum flux is $\rho + \delta\rho$ against ρ. Now the solution (4.6) shows that the momentum flux of this particular region will go on increasing as its proper distance from the point source will increase, and thus condensation will occur in this region. The interpretation of solution (4.7) is too tedious to further interpret.

ACKNOWLEDGMENTS

The authors are deeply indebted to Dr. V. B. Johri, Department of Mathematics, University of Gorakhpur, for reading this paper and for his stimulating discussions and important suggestions. Financial assistance received from the University Grants Commission and State Council of Science and Technology are also acknowledged.

[1]A. Sommerfeld, Kgl. Akad. Wet. Amsterdam Proc. **8**, 346 (1906).
[2]O. M. P. Bilaniuk, V. K. Deshpande, and E. C. G. Sundarshan, Am. J. Phys. **30**, 718 (1962).
[3]O. M. P. Bilaniuk and E. C. G. Sudarshan, Phys. Today **22**, 43—51 (May 1969).
[4]Ya. P. Terletskii, *Parodoksy Teorii Otnesitelnosti* (Nauka, Moscow, 1966) [English transl., *Paradoxo in the Theory of Relativity* (Plenum, New York, 1968)].
[5]G. Feinberg, Phys. Rev. **98**, 1089 (1967).
[6]J. C. Foster, Jr., and J. R. Ray, J. Math. Phys. **13**, 979 (1972).
[7]J. R. Gott, III, Nuovo Cimento B 22, 49—69 (1974).
[8]A. K. Raychaudhuri, Phys. Rev. **98**, 1123 (1955).
[9]G. F. R. Ellis and J. M. Stewart, J. Math. Phys. **9**, 1072 (1968).
[10]G. F. R. Ellis, J. Math. Phys. **8**, 1171 (1967).
[11]G. F. R. Ellis, "Lecture on Relativistic Cosmology" in *General Relativity and Cosmology*, edited by R. K. Sachs (Academic, New York, 1971).
[12]V. B. Johri, Prog. Math. 3, 42—56 (1969).
[13]G. C. McVittie, *General Relativity and Cosmology* (Chapmann and Hall, London, 1956), Vol. 4.
[14]H. K. Wimmel, Lett. Nuovo Cimento 2, 363—67 (1971).

Statistical theory of effective electrical, thermal, and magnetic properties of random heterogeneous materials. VII. Comparison of different approaches

Motoo Hori

Department of Applied Physics, Tokyo Institute of Technology, Meguroku, Tokyo 152, Japan
(Received 13 September 1976)

A general analysis is given for the theoretical evaluation of the effective permittivity of random heterogeneous materials that are statistical homogeneous. The perturbative and variational formulations presented in Papers I–VI in this series of work are reconstructed from a slightly different point of view and compared with recent approaches developed by other authors. Perturbation expansions for the effective permittivity are derived through electric field, electric displacement, Lorentz field, and T matrix. The validity of various approximate solutions involving the effective-medium approximation and the cumulant expansion method is discussed with the aid of a diagrammatic representation of the perturbation series. It is confirmed that, at the present stage, the cumulant theory is the best approximation for a three-dimensional system, while the effective-medium theory is the best for a two-dimensional system. The meaning and applicability of variational approaches are also reviewed.

1. INTRODUCTION

In earlier papers,[1–6] henceforth referred to as I–VI, a systematic study was made on the effective permitivity of inhomogeneous continuum media where local permittivity varies randomly from point to point. Various theoretical methods including the perturbation expansion, the variational approach, the effective-medium (EM) approximation, and the percolation theory were used to analyze the overall properties of two-dimensional (2D) and three-dimensional (3D) disordered materials that are statistically homogeneous. Although our formulations were carried out in the language of the dielectric constant, all the results obtained also hold for other physical constants such as magnetic permeability, electrical and thermal conductivity, and diffusion constant.

The statistical theory of heterogeneous materials traces back to Brown[7,8] and has attracted frequent attention thereafter. Even in recent years more than a few researchers have attempted perturbation or variational approaches[9–18] more or less different from ours. The purpose of the present paper is to compare our formulations with other existing theories and to check the validity of the approximate methods. With this aim in mind, we shall modify the previous treatments so as to facilitate the comparison and discuss their applicability from a unified point of view. The physical meaning in connection with the problem of electron localization in disordered systems are clarified in another publication,[19] which also presents a summary of the results of this article.

In Sec. 2 we explain basic concepts and notations used in the present formalism. Sections 3, 4, and 5 give perturbation expansions in terms of the electric field, the electric displacement, and the Lorentz field, respectively, while Sec. 6 is devoted to the so-called T matrix expansions. Finally, in Sec. 7, we review derivations of upper and lower bounds on the effective permittivity by means of variational principles.

It should be noted that our discussion is restricted to continuum mixtures. There is an alternative approach based on a lattice model, say a random network of resistors. Kirkpatrick[20,21] applied the EM approximation to bond percolation on a 2D square and 3D cubic lattice, and Watson and Leath[22] performed a similar treatment for the site model. Stinchcombe[23,24] and Essam *et al.*[25] investigated conduction in a disordered Bethe lattice (Cayley tree) that can be dealt with exactly. More recently, Blackman[26] developed a perturbation theory of conductivity in square bond networks, which proved to have much in common with our results. As suggested by Kirkpatrick[21] and Blackman,[26] in general, the continuum model shows a direct correspondence to the bond model rather than to the site model. A detailed study of this problem will appear elsewhere.

2. DEFINITIONS AND FUNDAMENTAL EQUATIONS

We consider a 3D (or 2D) heterogeneous material with spatially fluctuating permittivity $\epsilon_{ij}(\mathbf{r})$ whose volume V (or area S) is eventually brought to infinity. Assume that the medium is statistically homogeneous and postulate an ergodic hypothesis that the ensemble average can be replaced by the spatial average. For a multiphase material, then, the ensemble average of $\epsilon_{ij}(\mathbf{r})$ is expressed as

$$\langle \epsilon_{ij}(\mathbf{r}) \rangle = \sum_{\alpha} \epsilon_{\alpha,ij} v_{\alpha}, \tag{2.1}$$

where $\epsilon_{\alpha,ij}$ and v_{α} are the permittivity and the volume fraction of the αth phase, respectively.

The electric displacement $D_i(\mathbf{r})$ is related to the electric field $E_i(\mathbf{r})$ by

$$D_i(\mathbf{r}) = \epsilon_{ij} E_j(\mathbf{r}). \tag{2.2}$$

The effective permittivity of a statistically homogeneous material is defined as

$$\langle D_i(\mathbf{r}_i) \rangle = \int d\mathbf{r}_2 \epsilon_{ij}^{\mathrm{eff}}(\mathbf{r}_{12}) \langle E_j(\mathbf{r}_2) \rangle. \tag{2.3}$$

Here the integral is extended over the whole space V or S and $\mathbf{r}_{12}$ denotes the relative position $\mathbf{r}_2 - \mathbf{r}_1$. For simplicity we shall confine ourselves to a random system subjected to a constant-average electric field. In

this case Eq. (2.3) reduces to

$$\langle D_i(\mathbf{r}_1)\rangle = \epsilon^*_{ij}\langle E_j(\mathbf{r}_1)\rangle, \quad \epsilon^*_{ij} = \int d\mathbf{r}_2 \epsilon^{\mathrm{eff}}_{ij}(\mathbf{r}_{12}). \tag{2.4}$$

The problem is to determine the effective permittivity ϵ^*_{ij} or $\epsilon^{\mathrm{eff}}_{ij}$ from the statistical information about the random field $\{\epsilon_{ij}(\mathbf{r})\}$.

Let us separate the permittivity tensor $\epsilon_{ij}(\mathbf{r})$ into constant and fluctuating parts as

$$\epsilon_{ij}(\mathbf{r}) = \epsilon^0_{ij} + \delta\epsilon_{ij}(\mathbf{r}) = \epsilon_0\delta_{ij} + \delta\epsilon_{ij}(\mathbf{r}); \tag{2.5}$$

then the equation governing the electrostatic potential $\Phi(\mathbf{r})$ becomes

$$\epsilon_0 \frac{\partial^2\Phi(\mathbf{r})}{\partial x_i^2} + \frac{\partial}{\partial x_i}\left(\delta\epsilon_{ij}(\mathbf{r})\frac{\partial\Phi(\mathbf{r})}{\partial x_j}\right) = 0. \tag{2.6}$$

Introducing a Green's function $g(\mathbf{r})$ that satisfies

$$\epsilon_0 \frac{\partial^2 g(\mathbf{r})}{\partial x_i^2} + \delta(\mathbf{r}) = 0, \tag{2.7}$$

we obtain

$$\Phi(\mathbf{r}_1) = \Phi_0(\mathbf{r}_1) + \int d\mathbf{r}_2\, g(\mathbf{r}_{12}) \frac{\partial}{\partial x_{2,i}} \times \left(\delta\epsilon_{ij}(\mathbf{r}_2)\frac{\partial\Phi(\mathbf{r}_2)}{\partial x_{2,j}}\right), \tag{2.8}$$

where $\Phi_0(\mathbf{r}_1)$ depends only on ϵ^0_{ij} and the boundary conditions. The explicit forms of the free-space Green's function $g(\mathbf{r})$ in the 3D and 2D cases are, respectively,

$$g(\mathbf{r}) = \frac{1}{4\pi\epsilon_0 r}, \tag{2.9}$$

$$g(\mathbf{r}) = \frac{1}{2\pi\epsilon_0}\log\frac{1}{r}. \tag{2.10}$$

Differentiation of Eq. (2.8) with respect to $x_{1,i}$ yields

$$E_i(\mathbf{r}_1) = \langle E_i(\mathbf{r}_1)\rangle + \int d\mathbf{r}_2 \frac{\partial g(\mathbf{r}_{12})}{\partial x_{1,i}} \frac{\partial}{\partial x_{2,j}}[\delta\epsilon_{jk}(\mathbf{r}_2)E_k(\mathbf{r}_2)], \tag{2.11}$$

which is equivalent to Eq. (2.8) in I or Eq. (2.6) in V. After integrating by parts we have

$$E_i(\mathbf{r}_1) = E^0_i(\mathbf{r}_1) + \int d\mathbf{r}_2\, G_{ij}(\mathbf{r}_{12})\delta\epsilon_{jk}(\mathbf{r}_2)E_k(\mathbf{r}_2). \tag{2.12}$$

Here $E^0_i(\mathbf{r}_1)$ is independent of the random variable $\delta\epsilon_{ij}(\mathbf{r}_1)$ and

$$G_{ij}(\mathbf{r}_{12}) = -\frac{\partial^2 g(\mathbf{r}_{12})}{\partial x_{1,i}\partial x_{2,j}}. \tag{2.13}$$

For a 3D material, the Green's function tensor $G_{ij}(\mathbf{r})$ is given by

$$G_{ij}(\mathbf{r}_{12}) = -\frac{1}{3\epsilon_0}\delta_{ij}\delta(\mathbf{r}_{12}) + P\frac{1}{4\pi\epsilon_0 r_{12}^3} \times \left(3\frac{x_{12,i}}{r_{12}}\frac{x_{12,j}}{r_{12}} - \delta_{ij}\right). \tag{2.14}$$

The letter P indicates that the integration around a singular point $\mathbf{r}_{12} = 0$ is taken in the sense of the principal value. Similarly, the 2D Green's function tensor has the form

$$G_{ij}(\mathbf{r}_{12}) = -\frac{1}{2\epsilon_0}\delta_{ij}\delta(\mathbf{r}_{12}) + P\frac{1}{2\pi\epsilon_0 r_{12}^2} \times \left(2\frac{x_{12,i}}{r_{12}}\frac{x_{12,j}}{r_{12}} - \delta_{ij}\right). \tag{2.15}$$

Our previous formulation developed in I–V started the iteration process with Eq. (2.11), but Dederichs and Zeller[9–11] founded their approach on Eq. (2.12). As a matter of fact, the latter makes closer contact with the standard methods of the Green's function theory in quantum mechanics. In operator notation we may write Eq. (2.12) shortly as

$$\mathbf{E} = \mathbf{E}^0 + \mathbf{G}\delta\epsilon\mathbf{E}. \tag{2.16}$$

The product $\mathbf{AB}$ of the two operators $\mathbf{A}$ and $\mathbf{B}$ means

$$[\mathbf{AB}(\mathbf{r}_1, \mathbf{r}_2)]_{ij} = \int d\mathbf{r}_3 A_{ik}(\mathbf{r}_1, \mathbf{r}_3)A_{kj}(\mathbf{r}_3, \mathbf{r}_2). \tag{2.17}$$

In particular, the permittivity operator $\delta\epsilon$ is defined by

$$\delta\epsilon_{ij}(\mathbf{r}_1, \mathbf{r}_2) = \delta\epsilon_{ij}(\mathbf{r}_1)\delta(\mathbf{r}_{12}), \tag{2.18}$$

while the unit operator implies

$$1_{ij}(\mathbf{r}_1, \mathbf{r}_2) = \delta_{ij}\delta(\mathbf{r}_{12}). \tag{2.19}$$

By the use of a T matrix, Eq. (2.16) is transformed into

$$\mathbf{E} = \mathbf{E}^0 + \mathbf{GTE}^0, \quad \mathbf{T} = \delta\epsilon + \delta\epsilon\mathbf{GT}. \tag{2.20}$$

Averaging Eqs. (2.16) and (2.20) we find

$$\langle\mathbf{T}\rangle = \delta\epsilon^* + \delta\epsilon^*\mathbf{G}\langle\mathbf{T}\rangle, \tag{2.21}$$

with $\delta\epsilon^* = \epsilon^* - \epsilon^0$.

We can derive similar expressions for the electric displacement $\mathbf{D}$. Let $\gamma_{ij}(\mathbf{r})$ be the inverse matrix of $\epsilon_{ij}(\mathbf{r})$ and define γ^*_{ij} by

$$\langle E_i(\mathbf{r})\rangle = \langle\gamma_{ij}(\mathbf{r})D_j(\mathbf{r})\rangle = \gamma^*_{ij}\langle D_j(\mathbf{r})\rangle. \tag{2.22}$$

If we put

$$\delta\gamma_{ij}(\mathbf{r}) = \gamma_{ij}(\mathbf{r}) - \gamma_0\delta_{ij}, \quad D^0_i(\mathbf{r}) = \gamma_0 E^0_i(\mathbf{r}), \tag{2.23}$$

it follows from Eq. (2.12) that

$$D_i(\mathbf{r}_1) = D^0_i(\mathbf{r}_1) + \int d\mathbf{r}_2\, \Gamma_{ij}(\mathbf{r}_{12})\delta\gamma_{jk}(\mathbf{r}_2)D_k(\mathbf{r}_2), \tag{2.24}$$

where

$$\gamma_0\Gamma_{ij}(\mathbf{r}_{12}) = -\delta_{ij}\delta(\mathbf{r}_{12}) - \epsilon_0 G_{ij}(\mathbf{r}_{12}). \tag{2.25}$$

It is easily seen that for a 3D material

$$\Gamma_{ij}(\mathbf{r}_{12}) = -\frac{2}{3\gamma_0}\delta_{ij}\delta(\mathbf{r}_{12}) - P\frac{1}{4\pi\gamma_0 r_{12}^3} \times\left(3\frac{x_{12,i}}{r_{12}}\frac{x_{12,j}}{r_{12}} - \delta_{ij}\right), \tag{2.26}$$

while for a 2D material

$$\Gamma_{ij}(\mathbf{r}_{12}) = -\frac{1}{2\gamma_0}\delta_{ij}\delta(\mathbf{r}_{12}) - P\frac{1}{2\pi\gamma_0 r_{12}^2} \times \left(2\frac{x_{12,i}}{r_{12}}\frac{x_{12,j}}{r_{12}} - \delta_{ij}\right). \tag{2.27}$$

Thus, the operator equation corresponding to Eq. (2.16)

is

$$\mathbf{D}=\mathbf{D}^0+\Gamma\delta\gamma\mathbf{D}. \tag{2.28}$$

For convenience we are concerned mainly with an inhomogeneous material having isotropic permittivity $\epsilon(\mathbf{r})$ at point $\mathbf{r}$. It is straightforward to extend the treatments to locally anisotropic materials such as random polycrystals. For such systems we refer the reader, for example, to Molyneux[27,28] and Dederichs and Zeller.[9-11] When the medium is statistically homogeneous and locally isotropic, the n-point moment of $\delta\epsilon(\mathbf{r})$ becomes

$$\langle\delta\epsilon(\mathbf{r}_1)\delta\epsilon(\mathbf{r}_2)\cdots\delta\epsilon(\mathbf{r}_n)\rangle=\langle\delta\epsilon(0)\delta\epsilon(\mathbf{r}_{12})\cdots\delta\epsilon(\mathbf{r}_{1n})\rangle$$

$$=\langle(\delta\epsilon)^n\rangle h(\mathbf{r}_{12},\mathbf{r}_{23},\ldots,\mathbf{r}_{n-1,n}). \tag{2.29}$$

As the n-point cumulant or semi-invariant we have

$$\langle\delta\epsilon(\mathbf{r}_1)\delta\epsilon(\mathbf{r}_2)\cdots\delta\epsilon(\mathbf{r}_n)\rangle_c=\langle(\delta\epsilon)^n\rangle_c f(\mathbf{r}_{12},\mathbf{r}_{23},\ldots,\mathbf{r}_{n-1,n}), \tag{2.30}$$

$f(\mathbf{r}_{12},\mathbf{r}_{23},\ldots,\mathbf{r}_{n-1,n})$ being called the normalized n-point correlation function. In Appendix A it is shown that the correlation function $f(\mathbf{r}_{12},\mathbf{r}_{23},\ldots,\mathbf{r}_{n-1,n})$ [not $h(\mathbf{r}_{12},\mathbf{r}_{23},\ldots,\mathbf{r}_{n-1,n})$] for a two-phase mixture is independent of the phase permittivities.

Hereafter we shall restrict ourselves to the behavior of a symmetric cell material. The cell model was first proposed by Miller[29,30] and discussed in detail by Brown[15] and in VI. The most reasonable definition of a symmetric cell material is as follows:

(i) The space is completely covered by nonoverlapping cells within which the material property is constant;

(ii) cells are distributed in a manner such that the material is statistically homogeneous;

(iii) the material property of a cell is statistically independent of that of any other cell;

(iv) the material property of a cell is statistically independent of the geometrical distribution (shape and arrangement) of cells.

Clearly, the assumptions made by Dederichs and Zeller[9-11] are equivalent to the above four postulates. Furthermore, a completely random material may be regarded as a limiting case of a symmetric cell material in which cells have spherical shape and infinitesimal size. Note that the completely random material is not only statistically homogeneous but also statistically isotropic.

For a symmetric cell material we obtain

$$\epsilon(\mathbf{r})=\sum_\alpha\epsilon_\alpha(\mathbf{r})=\sum_\alpha\epsilon^\alpha\xi_\alpha(\mathbf{r}),\quad\sum_\alpha\xi_\alpha(\mathbf{r})=1, \tag{2.31}$$

where $\xi_\alpha(\mathbf{r})$ is an indicator in the sense that $\xi_\alpha(\mathbf{r})=1$ or 0 according as $\mathbf{r}$ lies inside or outside the cell α. The permittivity ϵ^α of the αth cell is independent of $\xi_\alpha(\mathbf{r})$ as well as of ϵ^β for $\beta\neq\alpha$. It turns out that

$$\left\langle\sum_\alpha\xi_\alpha(\mathbf{r}_1)\xi_\alpha(\mathbf{r}_2)\right\rangle=P(\mathbf{r}_1,\mathbf{r}_2), \tag{2.32}$$

$$\left\langle\sum_\alpha\xi_\alpha(\mathbf{r}_1)\xi_\alpha(\mathbf{r}_2)\xi_\alpha(\mathbf{r}_3)\right\rangle=P(\mathbf{r}_1,\mathbf{r}_2,\mathbf{r}_3), \tag{2.33}$$

$$\left\langle\sum_\alpha\xi_\alpha(\mathbf{r}_1)\xi_\alpha(\mathbf{r}_2)\xi_\alpha(\mathbf{r}_3)\xi_\alpha(\mathbf{r}_4)\right\rangle=P(\mathbf{r}_1,\mathbf{r}_2,\mathbf{r}_3,\mathbf{r}_4), \tag{2.34}$$

$$\left\langle\sum_{\alpha\neq\beta}\sum\xi_\alpha(\mathbf{r}_1)\xi_\alpha(\mathbf{r}_2)\xi_\beta(\mathbf{r}_3)\xi_\beta(\mathbf{r}_4)\right\rangle=P(\mathbf{r}_1,\mathbf{r}_2;\mathbf{r}_3,\mathbf{r}_4), \tag{2.35}$$

$$\left\langle\sum_{\alpha\neq\beta}\sum\xi_\alpha(\mathbf{r}_1)\xi_\beta(\mathbf{r}_2)\xi_\beta(\mathbf{r}_3)\xi_\alpha(\mathbf{r}_4)\right\rangle=P(\mathbf{r}_1,\mathbf{r}_4;\mathbf{r}_2,\mathbf{r}_3), \tag{2.36}$$

$$\left\langle\sum_{\alpha\neq\beta}\sum\xi_\alpha(\mathbf{r}_1)\xi_\beta(\mathbf{r}_2)\xi_\alpha(\mathbf{r}_3)\xi_\beta(\mathbf{r}_4)\right\rangle=P(\mathbf{r}_1,\mathbf{r}_3;\mathbf{r}_2,\mathbf{r}_4), \tag{2.37}$$

and so on. Here, for instance, $P(\mathbf{r}_1,\mathbf{r}_2,\mathbf{r}_3,\mathbf{r}_4)$ designates the probability that the four points $\mathbf{r}_1,\mathbf{r}_2,\mathbf{r}_3,\mathbf{r}_4$ are in the same cell, and $P(\mathbf{r}_1,\mathbf{r}_2;\mathbf{r}_3,\mathbf{r}_4)$ signifies the probability that two pairs of points, $(\mathbf{r}_1,\mathbf{r}_2)$ and $(\mathbf{r}_3,\mathbf{r}_4)$, are in two different cells. It has been stated in III that for a completely random material

$$P(\mathbf{r}_1,\mathbf{r}_2)=\delta_{\mathbf{r}_{12}}, \tag{2.38}$$

$$P(\mathbf{r}_1,\mathbf{r}_2,\mathbf{r}_3)=\delta_{\mathbf{r}_{12}}\delta_{\mathbf{r}_{23}}, \tag{2.39}$$

$$P(\mathbf{r}_1,\mathbf{r}_2,\mathbf{r}_3,\mathbf{r}_4)=\delta_{\mathbf{r}_{12}}\delta_{\mathbf{r}_{23}}\delta_{\mathbf{r}_{34}}, \tag{2.40}$$

$$P(\mathbf{r}_1,\mathbf{r}_2;\mathbf{r}_3,\mathbf{r}_4)=\delta_{\mathbf{r}_{12}}\delta_{\mathbf{r}_{34}}(1-\delta_{\mathbf{r}_{23}}), \tag{2.41}$$

$$P(\mathbf{r}_1,\mathbf{r}_4;\mathbf{r}_2,\mathbf{r}_3)=\delta_{\mathbf{r}_{14}}\delta_{\mathbf{r}_{23}}(1-\delta_{\mathbf{r}_{12}}), \tag{2.42}$$

$$P(\mathbf{r}_1,\mathbf{r}_3;\mathbf{r}_2,\mathbf{r}_4)=\delta_{\mathbf{r}_{13}}\delta_{\mathbf{r}_{24}}(1-\delta_{\mathbf{r}_{12}}). \tag{2.43}$$

3. PERTURBATION EXPANSIONS IN TERMS OF THE ELECTRIC FIELD

We will first seek perturbation series for $\delta\epsilon^*$ in terms of the electric field $\mathbf{E}$. Successive substitution of Eq. (2.16) in itself leads to

$$\mathbf{E}=\sum_{n=0}^{\infty}(\mathbf{G}\delta\epsilon)^n\mathbf{E}^0, \tag{3.1}$$

whence

$$\langle\mathbf{E}\rangle=\sum_{n=0}^{\infty}\langle\mathbf{G}\delta\epsilon)^n\rangle\mathbf{E}^0, \tag{3.2}$$

$$\delta\epsilon^*\langle\mathbf{E}\rangle=\langle\delta\epsilon\mathbf{E}\rangle=\sum_{n=0}^{\infty}\langle\delta\epsilon(\mathbf{G}\delta\epsilon)^n\rangle\mathbf{E}^0. \tag{3.3}$$

Eliminating $\mathbf{E}^0$ from these equations we find

$$\mathbf{E}=[\mathbf{1}+\mathbf{G}(\delta\epsilon-\langle\delta\epsilon\rangle)+\mathbf{G}(\delta\epsilon\mathbf{G}\delta\epsilon-\delta\epsilon\mathbf{G}\langle\delta\epsilon\rangle$$

$$-\langle\delta\epsilon\mathbf{G}\delta\epsilon\rangle+\langle\delta\epsilon\rangle\mathbf{G}\langle\delta\epsilon\rangle)+\cdots]\langle\mathbf{E}\rangle$$

$$\equiv\sum_{n=0}^{\infty}\mathbf{E}^{(n)}, \tag{3.4}$$

$$\delta\epsilon^*=\langle\delta\epsilon\rangle+[\langle\delta\epsilon\mathbf{G}\delta\epsilon\rangle-\langle\delta\epsilon\rangle\mathbf{G}\langle\delta\epsilon\rangle]$$

$$+[\langle\delta\epsilon\mathbf{G}\delta\epsilon\mathbf{G}\delta\epsilon\rangle-\langle\delta\epsilon\mathbf{G}\delta\epsilon\rangle\mathbf{G}\langle\delta\epsilon\rangle$$

$$-\langle\delta\epsilon\rangle\mathbf{G}\langle\delta\epsilon\mathbf{G}\delta\epsilon\rangle+\langle\delta\epsilon\rangle\mathbf{G}\langle\delta\epsilon\rangle\mathbf{G}\langle\delta\epsilon\rangle]+\cdots$$

$$\equiv\sum_{n=1}^{\infty}\delta\epsilon^{(n)}. \tag{3.5}$$

For example, the second-order term appearing in

Eq. (3.5) is explicitly

$$\delta\epsilon_{ij}^{(2)} = \int d\mathbf{r}_2\, G_{ij}(\mathbf{r}_{12})[\langle\delta\epsilon(\mathbf{r}_1)\delta\epsilon(\mathbf{r}_2)\rangle - \langle\delta\epsilon(\mathbf{r}_1)\rangle\langle\delta\epsilon(\mathbf{r}_2)\rangle]. \tag{3.6}$$

Especially when $\epsilon_0 = \langle\epsilon\rangle$, we set $\epsilon' = \delta\epsilon$ and $\epsilon'^* = \delta\epsilon^*$; then Eq. (3.5) simplifies

$$\epsilon'^* = \langle\epsilon' G\epsilon'\rangle + \langle\epsilon' G\epsilon' G\epsilon'\rangle + [\langle\epsilon' G\epsilon' G\epsilon' G\epsilon'\rangle - \langle\epsilon' G\epsilon'\rangle G\langle\epsilon' G\epsilon'\rangle] + \cdots, \tag{3.7}$$

which is nothing but Eq. (2.28) of I.

Now we shall replace the many-point moments by the corresponding cumulants; that is,

$$\langle\delta\epsilon(\mathbf{r}_1)\rangle = \langle\delta\epsilon(\mathbf{r}_1)\rangle_c, \tag{3.8}$$

$$\langle\delta\epsilon(\mathbf{r}_1)\delta\epsilon(\mathbf{r}_2)\rangle = \langle\delta\epsilon(\mathbf{r}_1)\delta\epsilon(\mathbf{r}_2)\rangle_c + \langle\delta\epsilon(\mathbf{r}_1)\rangle_c\langle\delta\epsilon(\mathbf{r}_2)\rangle_c, \tag{3.9}$$

$$\begin{aligned}\langle\delta\epsilon(\mathbf{r}_1)\delta\epsilon(\mathbf{r}_2)\delta\epsilon(\mathbf{r}_3)\rangle &= \langle\delta\epsilon(\mathbf{r}_1)\delta\epsilon(\mathbf{r}_2)\delta\epsilon(\mathbf{r}_3)\rangle_c + \langle\delta\epsilon(\mathbf{r}_1)\delta\epsilon(\mathbf{r}_2)\rangle_c \\ &\times\langle\delta\epsilon(\mathbf{r}_3)\rangle_c + \langle\delta\epsilon(\mathbf{r}_1)\delta\epsilon(\mathbf{r}_3)\rangle_c\langle\delta\epsilon(\mathbf{r}_2)\rangle_c \\ &+ \langle\delta\epsilon(\mathbf{r}_2)\delta\epsilon(\mathbf{r}_3)\rangle_c\langle\delta\epsilon(\mathbf{r}_1)\rangle_c + \langle\delta\epsilon(\mathbf{r}_1)\rangle_c\langle\delta\epsilon(\mathbf{r}_2)\rangle_c\langle\delta\epsilon(\mathbf{r}_3)\rangle_c,\end{aligned} \tag{3.10}$$

and so on. Insertion of these equations into Eq. (3.5) yields

$$\delta\epsilon_{ij}^{(1)} = \langle\delta\epsilon(\mathbf{r}_1)\rangle_c, \tag{3.11}$$

$$\delta\epsilon_{ij}^{(2)} = \int d\mathbf{r}_2\, G_{ij}(\mathbf{r}_{12})\langle\delta\epsilon(\mathbf{r}_1)\delta\epsilon(\mathbf{r}_2)\rangle_c, \tag{3.12}$$

$$\begin{aligned}\delta\epsilon_{ij}^{(3)} &= \int d\mathbf{r}_2 \int d\mathbf{r}_3\, G_{ik}(\mathbf{r}_{12})G_{kj}(\mathbf{r}_{23}) \\ &\times[\langle\delta\epsilon(\mathbf{r}_1)\delta\epsilon(\mathbf{r}_2)\delta\epsilon(\mathbf{r}_3)\rangle_c + \langle\delta\epsilon(\mathbf{r}_1)\delta\epsilon(\mathbf{r}_3)\rangle_c\langle\delta\epsilon(\mathbf{r}_2)\rangle_c], \\ &\equiv \delta\epsilon_{ij}^{(3,1)} + \delta\epsilon_{ij}^{(3,2)},\end{aligned} \tag{3.13}$$

$$\begin{aligned}\delta\epsilon_{ij}^{(4)} &= \int d\mathbf{r}_2 \int d\mathbf{r}_3 \int d\mathbf{r}_4\, G_{ik}(\mathbf{r}_{12})G_{kh}(\mathbf{r}_{23})G_{hj}(\mathbf{r}_{34}) \\ &\times[\langle\delta\epsilon(\mathbf{r}_1)\delta\epsilon(\mathbf{r}_2)\delta\epsilon(\mathbf{r}_3)\delta\epsilon(\mathbf{r}_4)\rangle_c \\ &+ \langle\delta\epsilon(\mathbf{r}_1)\delta\epsilon(\mathbf{r}_2)\delta\epsilon(\mathbf{r}_4)\rangle_c\langle\delta\epsilon(\mathbf{r}_3)\rangle_c \\ &+ \langle\delta\epsilon(\mathbf{r}_1)\delta\epsilon(\mathbf{r}_3)\delta\epsilon(\mathbf{r}_4)\rangle_c\langle\delta\epsilon(\mathbf{r}_2)\rangle_c \\ &+ \langle\delta\epsilon(\mathbf{r}_1)\delta\epsilon(\mathbf{r}_4)\rangle_c\langle\delta\epsilon(\mathbf{r}_2)\delta\epsilon(\mathbf{r}_3)\rangle_c \\ &+ \langle\delta\epsilon(\mathbf{r}_1)\delta\epsilon(\mathbf{r}_3)\rangle_c\langle\delta\epsilon(\mathbf{r}_2)\delta\epsilon(\mathbf{r}_4)\rangle_c \\ &+ \langle\delta\epsilon(\mathbf{r}_1)\delta\epsilon(\mathbf{r}_4)\rangle_c\langle\delta\epsilon(\mathbf{r}_2)\rangle_c\langle\delta\epsilon(\mathbf{r}_3)\rangle_c] \\ &\equiv \sum_{m=1}^{6}\delta\epsilon_{ij}^{(4,m)}.\end{aligned} \tag{3.14}$$

Higher-order terms can be derived in like manner.

A convenient way of representing various contributions to $\delta\epsilon^*$ is to employ cumulant diagrams explained in IV and V. Cumulant diagrams consist of cross vertices, dashed vertical lines (interaction lines), and solid horizontal lines (propagators). Diagrams of this type are constructed in accordance with the following rules:

(i) Represent the points $\mathbf{r}_1, \mathbf{r}_2, \cdots$ by means of nodes on the horizontal base line;

(ii) assign the Green's function $G(\mathbf{r}_{i,i+1})$ to a propagator connecting the points $\mathbf{r}_i$ and $\mathbf{r}_{i+1}$;

(iii) allot the s-point cumulant $\langle\delta\epsilon(\mathbf{r}_i)\delta\epsilon(\mathbf{r}_j)\cdots\rangle_c$ to a cross vertex from which s interaction lines emanate;

(iv) take the operator product of all the Green's functions and cumulants described above.

The prescription for drawing third-order diagrams is illustrated in Fig. 1 and a diagram equation corresponding to Eq. (3.5) is presented in Fig. 2. By careful inspection of these diagrams we observe that the effective permittivity $\delta\epsilon^*$ is nothing but the sum of all possible proper diagrams, which is usually called the mass operator or the proper self-energy part.

In order to calculate the leading terms of Eq. (3.5), we need only to utilize the results given in I–V. In fact, the second-order term $\delta\epsilon_{ij}^{(2)}$ is rewritten by partial integration as

$$\begin{aligned}\delta\epsilon_{ij}^{(2)} &= \int d\mathbf{r}_2 \frac{\partial g(\mathbf{r}_{12})}{\partial x_{1,i}} \frac{\partial\langle\delta\epsilon(\mathbf{r}_1)\delta\epsilon(\mathbf{r}_2)\rangle_c}{\partial x_{2,j}} \\ &= -\langle(\delta\epsilon)^2\rangle_c \int d\mathbf{r}_{12} \frac{\partial g(\mathbf{r}_{12})}{\partial x_{12,i}} \frac{\partial f(\mathbf{r}_{12})}{\partial x_{12,j}},\end{aligned} \tag{3.15}$$

which implies Eq. (2.30) of I or (2.23) of V. Thus it follows that for any statistically isotropic material

$$\delta\epsilon^{(2)} = -\frac{1}{d}\frac{\langle(\delta\epsilon)^2\rangle_c}{\epsilon_0} \tag{3.16}$$

and for a symmetric cell material composed of uniformly oriented ellipsoids or ellipses

$$\delta\epsilon_{ij}^{(2)} = -L_{ij}\langle(\delta\epsilon)^2\rangle_c/\epsilon_0, \tag{3.17}$$

where d is the dimensionality of the medium and L_{ij} is the depolarization or demagnetization tensor of cells. Similarly we get

$$\delta\epsilon_{ij}^{(3,1)} = L_{ik}L_{kj}\langle(\delta\epsilon)^3\rangle_c/\epsilon_0^2, \tag{3.18}$$

$$\delta\epsilon_{ij}^{(3,2)} = L_{ij}\langle(\delta\epsilon)^2\rangle_c\langle\delta\epsilon\rangle_c/\epsilon_0^2, \tag{3.19}$$

which reduce for a spherical-cell material to

$$\delta\epsilon^{(3,1)} = \frac{1}{d^2}\frac{\langle(\delta\epsilon)^3\rangle_c}{\epsilon_0^2}, \tag{3.20}$$

$$\delta\epsilon^{(3,2)} = \frac{1}{d}\frac{\langle(\delta\epsilon)^2\rangle_c\langle\delta\epsilon\rangle_c}{\epsilon_0^2} \tag{3.21}$$

Referring to the calculations in III and V we are able to determine fourth-order terms for completely random materials as

$$\delta\epsilon^{(4,1)} = -\frac{1}{d^3}\frac{\langle(\delta\epsilon)^4\rangle_c}{\epsilon_0^3}, \tag{3.22}$$

$$\delta\epsilon^{(4,2)} = -\frac{1}{\epsilon_0}\delta\epsilon^{(3,1)}\delta\epsilon^{(1)} = -\frac{1}{d^2}\frac{\langle(\delta\epsilon)^3\rangle_c\langle\delta\epsilon\rangle_c}{\epsilon_0^3}, \tag{3.23}$$

$$\delta\epsilon^{(4,3)} = -\frac{1}{\epsilon_0}\delta\epsilon^{(3,1)}\delta\epsilon^{(1)} = -\frac{1}{d^2}\frac{\langle(\delta\epsilon)^3\rangle_c\langle\delta\epsilon\rangle_c}{\epsilon_0^3}, \tag{3.24}$$

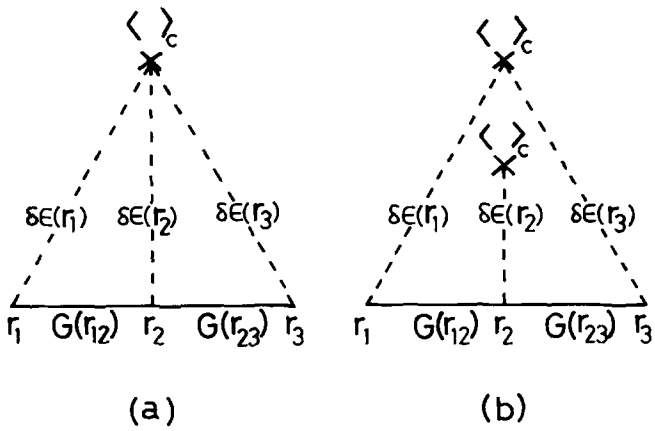


FIG. 1. Prescription for drawing third-order diagrams. The diagram (a) represents $\delta\epsilon^{(3,1)} = \langle\delta\epsilon G\delta\epsilon\delta\epsilon\rangle_c$ and the diagram (b) represents $\delta\epsilon^{(3,2)} = \langle\delta\epsilon G\langle\delta\epsilon\rangle_c G\delta\epsilon\rangle_c$.

$$\delta\epsilon^{(4,4)} = -\frac{1}{\epsilon_0}\delta\epsilon^{(2)}\delta\epsilon^{(2)} = -\frac{1}{d^2}\frac{\langle(\delta\epsilon)^2\rangle_c^2}{\epsilon_0^3}, \tag{3.25}$$

$$\delta\epsilon^{(4,5)} = \begin{cases} -\dfrac{1}{d^3}\dfrac{\langle(\delta\epsilon)^2\rangle_c^2}{\epsilon_0^3} & \text{in 1D or 2D,} \\ 0 & \text{in 3D,} \end{cases} \tag{3.26}$$

$$\delta\epsilon^{(4,6)} = \left(-\frac{1}{\epsilon_0}\right)^2\delta\epsilon^{(2)}\delta\epsilon^{(1)}\delta\epsilon^{(1)} = -\frac{1}{d}\frac{\langle(\delta\epsilon)^2\rangle_c\langle\delta\epsilon\rangle_c^2}{\epsilon_0^3}. \tag{3.27}$$

Diagram equations expressing $\delta\epsilon^{(4,4)}$ and $\delta\epsilon^{(4,6)}$ are depicted in Fig. 3. In general, nested diagrams are factorizable into lower-order one-vertex diagrams, whereas crossed diagrams are not reducible to one-vertex diagrams.

Contributions from single-site diagrams including one-vertex and nested diagrams are counted by the following simple recipes:

(i) Assign the sth-order cumulant $\langle(\delta\epsilon)^s\rangle_c$ to s dashed interaction lines originating from the same cross vertex;

(ii) allot $-1/\epsilon_0$ to each propagator;

(iii) allot $1/d$ to each independent propagator;

(iv) take the product of all factors thus calculated.

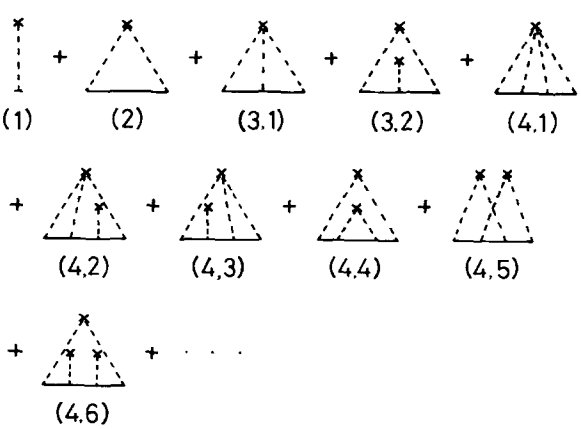


FIG. 2. Diagrammatic representation of Eq. (3.5).

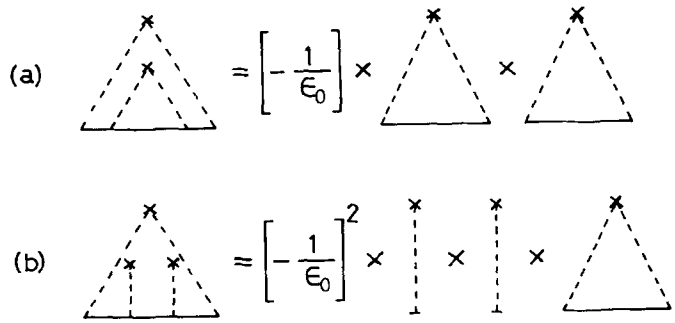


FIG. 3. Process for factorizing nested diagram into lower-order one-vertex diagrams. The diagram equations (a) and (b) correspond to Eqs. (3.25) and (3.27), respectively.

From a slightly different point of view, single-site diagrams can be summed by means of renormalized propagators as shown in Fig. 4. As for crossed diagrams, on the other hand, we are not successful in evaluating all contributions rigorously, but some of them are determined in conformity with the following additional rule:

(v) For a 3D material, associate zero with crossed diagrams which are contained as addends in Fig. 5(a). Also in the 1D and 2D cases, contributions from these diagrams can easily be computed, although they do not vanish as in a 3D case. Vanishing crossed diagrams in 1D or 2D are illustrated in Fig. 6 where closed-circle vertices indicate moments of $\delta\epsilon$ rather than its cumulant.

Let us derive several approximate solutions for $\delta\epsilon^*$. First we pick up only one-vertex cumulant diagrams; namely,

$$\delta\epsilon^* = \langle\delta\epsilon\rangle_c - \frac{\langle(\delta\epsilon)^2\rangle_c}{d\epsilon_0} + \frac{\langle(\delta\epsilon)^3\rangle_c}{(d\epsilon_0)^2} - \frac{\langle(\delta\epsilon)^4\rangle_c}{(d\epsilon_0)^3} + \cdots, \tag{3.28}$$

whence

$$\epsilon^* = \langle\epsilon\rangle_c - \frac{\langle\epsilon^2\rangle_c}{d\epsilon_0} + \frac{\langle\epsilon^3\rangle_c}{(d\epsilon_0)^2} - \frac{\langle\epsilon^4\rangle_c}{(d\epsilon_0)^3} + \cdots$$

$$= \int_0^1 \langle\epsilon z^{\epsilon/d\epsilon_0}\rangle/\langle z^{\epsilon/d\epsilon_0}\rangle\, dz \tag{3.29}$$

(see IV). The non-self-consistent cumulant solution ϵ_0^* and the self-consistent cumulant solution ϵ_c^* are given

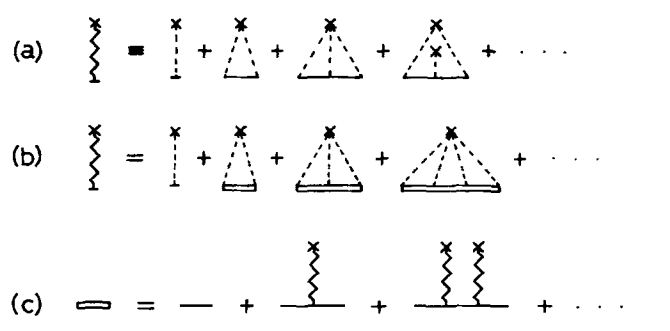


FIG. 4. Sum of all possible single-site diagrams in the wide sense. (a) Direct expression up to third order. (b) Indirect expression in terms of renormalized propagators. (c) Definition of a renormalized propagator.

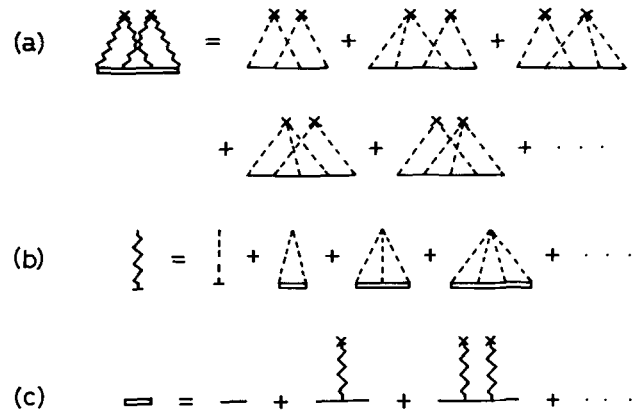


FIG. 5. Crossed diagrams whose contribution is zero for the 3D case. (a) Sum of vanishing crossed diagrams. (b) Definition of a wavy interaction line. (c) Definition of a double line propagator.

by

$$\epsilon_0^* = \int_0^1 \langle \epsilon z^{\epsilon/d\langle\epsilon\rangle} \rangle / \langle z^{\epsilon/d\langle\epsilon\rangle} \rangle \, dz, \tag{3.30}$$

$$\epsilon_C^* = \int_0^1 \langle \epsilon z^{\epsilon/d\epsilon_C^*} \rangle / \langle z^{\epsilon/d\epsilon_C^*} \rangle \, dz, \tag{3.31}$$

where we have put $\epsilon_0 = \langle \epsilon \rangle$ and $\epsilon_0 = \epsilon_C^*$, respectively. Diagrammatic representation of $\delta\epsilon_0^*$ and $\delta\epsilon_C^*$ is shown in Fig. 7. We should notice that the self-consistent cumulant approximation takes care of all single-site cumulant diagrams.

Next we use moment averages instead of cumulant averages in Eq. (3.28); then

$$\delta\epsilon^* = \langle \delta\epsilon \rangle - \frac{\langle (\delta\epsilon)^2 \rangle}{d\epsilon_0} + \frac{\langle (\delta\epsilon)^3 \rangle}{(d\epsilon_0)^2} - \frac{\langle (\delta\epsilon)^4 \rangle}{(d\epsilon_0)^3} + \cdots$$

$$= d\epsilon_0 \left\langle \frac{\delta\epsilon}{\delta\epsilon + d\epsilon_0} \right\rangle = d\epsilon_0 \left\langle \frac{\epsilon - \epsilon_0}{\epsilon + (d-1)\epsilon_0} \right\rangle. \tag{3.32}$$

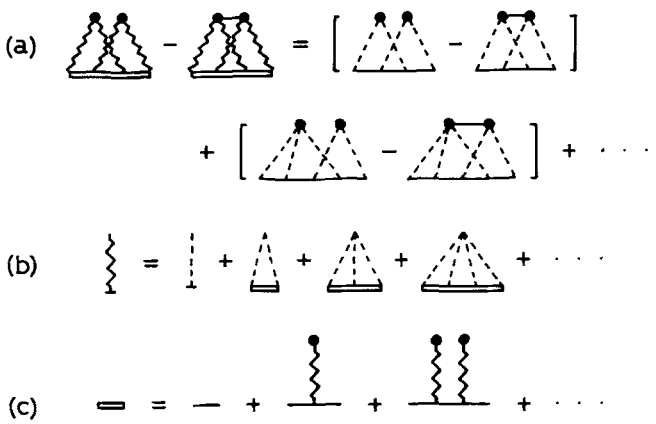


FIG. 6. Crossed diagrams whose contribution is zero for the 1D or 2D case. (a) Sum of vanishing crossed diagrams. Each difference in square brackets contributes zero. The solid line connecting the two closed-circle vertices indicates that the two points should be taken as the same (see IV for details). (b) Definition of a wavy interaction line. (c) Definition of a double line propagator.

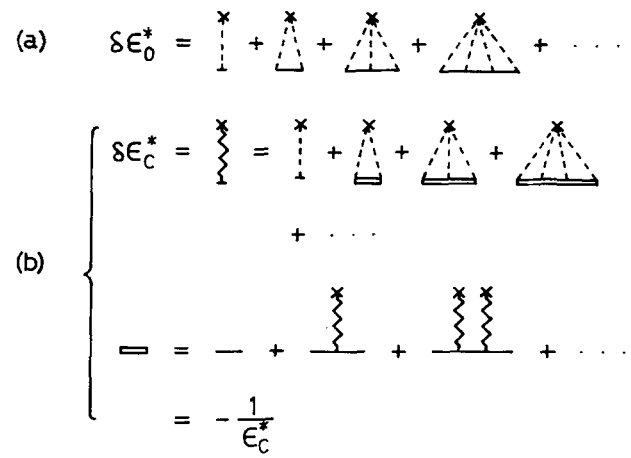


FIG. 7. Diagram equations expressing $\delta\epsilon_0^*$ and $\delta\epsilon_C^*$. (a) Non-self-consistent cumulant solution $\delta\epsilon_0^*$. (b) Self-consistent cumulant solution $\delta\epsilon_C$.

When $\epsilon_0 = \langle \epsilon \rangle$ this series reduces to

$$\epsilon_K^* = \langle \epsilon \rangle \left\langle \frac{(d+1)\epsilon - \langle \epsilon \rangle}{\epsilon + (d-1)\langle \epsilon \rangle} \right\rangle, \tag{3.33}$$

which coincides with Kröner's approximation.[12,31] Furthermore, substitution of $\epsilon_0 = \epsilon_{EM}^*$ into Eq. (3.32) yields

$$\left\langle \frac{\epsilon - \epsilon_{EM}^*}{\epsilon + (d-1)\epsilon_{EM}^*} \right\rangle = 0. \tag{3.34}$$

As argued in IV, V, and Ref. 19, the solution of Eq. (3.34) is equivalent to the self-contained single-site approximation[32] or the EM approximation.[33–38] Actually, it is readily seen that the self-contained single-site

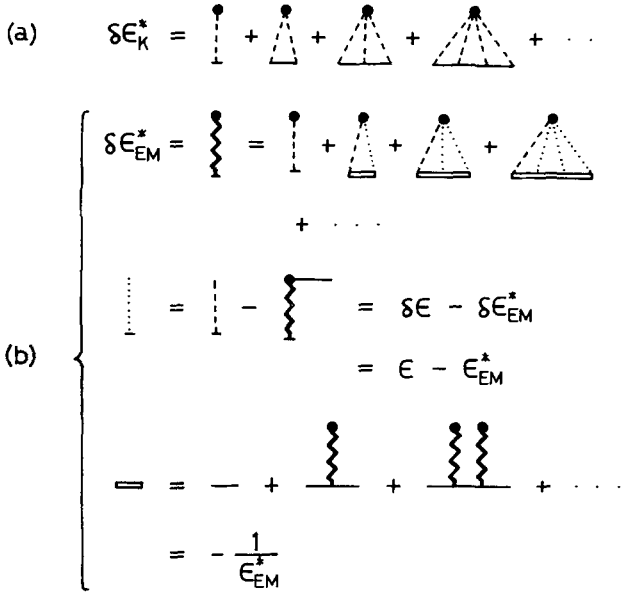


FIG. 8. Diagram equations expressing $\delta\epsilon_K^*$ and $\delta\epsilon_{EM}^*$. (a) Kröner's approximation $\delta\epsilon_K$. (b) Self-contained single-site approximation $\delta\epsilon_{EM}^*$.

solution

$$\delta\epsilon^*_{EM} = \langle\delta\epsilon\rangle - \frac{\langle\delta\epsilon(\delta\epsilon - \delta\epsilon^*_{EM})\rangle}{d\epsilon^*_{EM}} + \frac{\langle\delta\epsilon(\delta\epsilon - \delta\epsilon^*_{EM})^2\rangle}{(d\epsilon^*_{EM})^2}$$

$$- \frac{\langle\delta\epsilon(\delta\epsilon - \delta\epsilon^*_{EM})^3\rangle}{(d\epsilon^*_{EM})^3} + \cdots \tag{3.35}$$

satisfies Eq. (3.34) irrespective of $\epsilon_0 = \epsilon - \delta\epsilon$. Diagram equations for ϵ^*_K and ϵ^*_{EM} are given in Fig. 8.

4. PERTURBATION EXPANSIONS IN TERMS OF THE ELECTRIC DISPLACEMENT

In this section we expand the electric displacement **D** instead of **E**. Starting from Eq. (2.28) we have

$$\begin{aligned}\mathbf{D} &= [1 + \Gamma(\delta\gamma - \langle\delta\gamma\rangle) + \Gamma(\delta\gamma\Gamma\delta\gamma - \delta\gamma\Gamma\langle\delta\gamma\rangle \\ &\quad - \langle\delta\gamma\Gamma\delta\gamma\rangle + \langle\delta\gamma\rangle\Gamma\langle\delta\gamma\rangle) + \cdots]\langle\mathbf{D}\rangle \\ &\equiv \sum_{n=0}^{\infty}\mathbf{D}^{(n)},\end{aligned} \tag{4.1}$$

$$\begin{aligned}\delta\gamma^* &= \langle\delta\gamma\rangle + [\langle\delta\gamma\Gamma\delta\gamma\rangle - \langle\delta\gamma\rangle\Gamma\langle\delta\gamma\rangle] + [\langle\delta\gamma\Gamma\delta\gamma\Gamma\delta\gamma\rangle \\ &\quad - \langle\delta\gamma\Gamma\delta\gamma\rangle\Gamma\langle\delta\gamma\rangle - \langle\delta\gamma\rangle\Gamma\langle\delta\gamma\Gamma\delta\gamma\rangle \\ &\quad + \langle\delta\gamma\rangle\Gamma\langle\delta\gamma\rangle\Gamma\langle\delta\gamma\rangle] + \cdots \\ &= \langle\delta\gamma\rangle_c + \langle\delta\gamma\Gamma\delta\gamma\rangle_c + [\langle\delta\gamma\Gamma\delta\gamma\Gamma\delta\gamma\rangle_c \\ &\quad + \langle\delta\gamma\Gamma\langle\delta\gamma\rangle_c\Gamma\delta\gamma\rangle_c] + \cdots \\ &\equiv \sum_{n=1}^{\infty}\delta\gamma^{(n)}.\end{aligned} \tag{4.2}$$

The prescriptions (i)—(iv) for constructing diagrams described in Sec. 3 apply equally well to the present case.

The second-order term $\delta\gamma^{(2)}$ is calculated as

$$\begin{aligned}\delta\gamma^{(2)}_{ij} &= \int d\mathbf{r}_2\,\Gamma_{ij}(\mathbf{r}_{12})\langle\delta\gamma(\mathbf{r}_1)\delta\gamma(\mathbf{r}_2)\rangle_c \\ &= \langle(\delta\gamma)^2\rangle_c\left[-\frac{\delta_{ij}}{\gamma_0} + \int d\mathbf{r}_{12}\frac{\partial}{\partial x_{12,i}}\right. \\ &\quad \left.\times\left(\frac{g(\mathbf{r}_{12})}{\gamma_0^2}\right)\frac{\partial f(\mathbf{r}_{12})}{\partial x_{12,j}}\right].\end{aligned} \tag{4.3}$$

Accordingly, for a statistically isotropic system

$$\delta\gamma^{(2)} = -\left(1 - \frac{1}{d}\right)\frac{\langle(\delta\gamma)^2\rangle_c}{\gamma_0}, \tag{4.4}$$

and for an ellipsoidal (elliptic) cell material

$$\delta\gamma^{(2)}_{ij} = -(\delta_{ij} - L_{ij})\frac{\langle(\delta\gamma)^2\rangle_c}{\gamma_0}. \tag{4.5}$$

Likewise, formulas corresponding to Eqs. (3.21)—(3.28) are:

$$\delta\gamma^{(3,1)} = \left(1 - \frac{1}{d}\right)^2\frac{\langle(\delta\gamma)^3\rangle_c}{\gamma_0^2}, \tag{4.6}$$

$$\delta\gamma^{(3,2)} = \left(1 - \frac{1}{d}\right)\frac{\langle(\delta\gamma)^2\rangle_c\langle\delta\gamma\rangle_c}{\gamma_0^3}, \tag{4.7}$$

$$\delta\gamma^{(4,1)} = -\left(1 - \frac{1}{d}\right)^3\frac{\langle(\delta\gamma)^4\rangle_c}{\gamma_0^3}, \tag{4.8}$$

$$\delta\gamma^{(4,2)} = \delta\gamma^{(4,3)} = -\left(1 - \frac{1}{d}\right)^2\frac{\langle(\delta\gamma)^3\rangle_c\langle\delta\gamma\rangle_c}{\gamma_0^3}, \tag{4.9}$$

$$\delta\gamma^{(4,4)} = -\left(1 - \frac{1}{d}\right)^2\frac{\langle(\delta\gamma)^2\rangle_c^2}{\gamma_0^3}, \tag{4.10}$$

$$\delta\gamma^{(4,5)} = \begin{cases} -\left(1 - \frac{1}{d}\right)^3\frac{\langle(\delta\gamma)^2\rangle_c^2}{\gamma_0^3} & \text{in 1D or 2D,} \\ -\frac{\langle(\delta\gamma)^2\rangle_c^2}{3\gamma_0^3} & \text{in 3D,} \end{cases} \tag{4.11}$$

$$\delta\gamma^{(4,6)} = -\left(1 - \frac{1}{d}\right)\frac{\langle(\delta\gamma)^2\rangle_c\langle\delta\gamma\rangle_c^2}{\gamma_0^3}; \tag{4.12}$$

the proof of these equations is shown in Appendix A.

We note that the process of factorizing single-site diagrams as in Fig. 3 also holds for $\delta\gamma$. More generally, the expedients (i)—(iv) for counting contributions from single-site diagrams and the procedure for renormalizing propagators stated in Fig. 4 are available with appropriate changes such as $\delta\epsilon \to \delta\gamma$, $\epsilon_0 \to \gamma_0$, and $1/d \to 1 - 1/d$. However, crossed diagrams as expressed by Fig. 5 do not vanish even in a 3D case, so that recipe (v) is invalid. In a 1D or 2D case, crossed moment diagrams comprised in Fig. 6 are found to contribute zero to $\delta\gamma^*$ as well as to $\delta\epsilon^*$.

As the non-self-consistent and self-consistent cumulant solutions we get

$$\gamma_0^* = \int_0^1 \langle\gamma z^{(d-1)\gamma/d\langle\gamma\rangle}\rangle/\langle z^{(d-1)\gamma/d\langle\gamma\rangle}\rangle\,dz, \tag{4.13}$$

$$\gamma_C^* = \int_0^1 \langle\gamma z^{(d-1)\gamma/d\gamma_C^*}\rangle/\langle z^{(d-1)\gamma/d\gamma_C^*}\rangle\,dz. \tag{4.14}$$

Besides, Kröner's approximation provides

$$\gamma_K^* = \langle\gamma\rangle\left\langle\frac{(2d-1)\gamma - (d-1)\langle\gamma\rangle}{(d-1)\gamma + \langle\gamma\rangle}\right\rangle, \tag{4.15}$$

while the self-contained single-site approximation leads us to

$$\left\langle\frac{\gamma - \gamma^*_{EM}}{(d-1)\gamma + \gamma^*_{EM}}\right\rangle = 0. \tag{4.16}$$

Out of the four approximations γ_0^*, γ_C^*, γ_K^*, γ_{EM}^* mentioned above, γ_{EM}^* alone satisfies the relationship

$$\epsilon^*_{EM}\gamma^*_{EM} = 1, \tag{4.17}$$

which should be true for the exact solution. Equation (4.17) means that ϵ_{EM}^* and γ_{EM}^* are of the same degree of approximation, owing to the self-containedness condition that correction terms due to exclusion effects are self-contained within the approximation under consideration.[32]

Finally we compare merits and demerits of the approximate effective permittivities ϵ_0^*, $1/\gamma_0^*$, ϵ_C^*, $1/\gamma_C^*$, ϵ_K^*, $1/\gamma_K^*$, $\epsilon_{EM}^* = 1/\gamma_{EM}^*$. For a 1D material, $1/\gamma_0^*$, $1/\gamma_C^*$, $1/\gamma_K^*$, and $\epsilon_{EM}^* = 1/\gamma_{EM}^*$ produce an exact result,

$$\begin{aligned}\epsilon^* &= 1/\langle 1/\epsilon\rangle = 1/\langle\gamma\rangle \\ &= 1/\gamma_0^* = 1/\gamma_C^* = 1/\gamma_K^* = 1/\gamma_{EM}^* = \epsilon_{EM}^*,\end{aligned} \tag{4.18}$$

whereas ϵ_0^*, ϵ_K^*, and ϵ_C^* are not correct in fourth or higher order. In the 2D problem, it was pointed out in V that $\epsilon_{EM}^* = 1/\gamma_{EM}^*$ is not only the best possible single-site approximation to the effective permittivity $\epsilon^* = 1/\gamma^*$, but also it includes contributions from important crossed diagrams. In fact, both ϵ_{EM}^* and γ_{EM}^* prove to be strictly valid up to the fourth order. In a 3D case, crossed cumulant diagrams comprised in Fig. 5 vanish for $\delta\epsilon$, so that ϵ_C^* is considered to be the best approximation at the present stage. As an estimate of γ^*, however, we need to employ $1/\epsilon_C^*$ instead of γ_C^*, because the counting rule (v) does not apply to $\delta\gamma^*$. Recall that ϵ_C^* is exact up to the fourth order but γ_C^* is not true in the fourth order.

Another merit of $\epsilon_{EM}^* = 1/\gamma_{EM}^*$ for a 2D material is that it is compatible with the so-called "phase interchange" theorem.[39-41] Let us specify by $\epsilon^*(\epsilon_1, \epsilon_2)$ the effective permittivity of a 2D binary mixture with phase permittivities ϵ_1 and ϵ_2. The phase interchange theorem asserts that

$$\epsilon^*(\epsilon_1, \epsilon_2)\epsilon^*(\epsilon_2, \epsilon_1) = \epsilon_1\epsilon_2, \tag{4.19}$$

where $\epsilon^*(\epsilon_2, \epsilon_1)$ is the effective permittivity when the constituting phases are exchanged. We have not changed interphase geometry, but only changed phase properties. In the 3D case, this theorem can no longer be applied, though an inequality

$$\epsilon^*(\epsilon_1, \epsilon_2)\epsilon^*(\epsilon_2, \epsilon_1) \geq \epsilon_1\epsilon_2 \tag{4.20}$$

holds instead of the above equality.[17] When the constituents with ϵ_1 and ϵ_2 occupy fractions of the total volume v_1 and v_2, respectively, the 2D EM solution takes on the form

$$\epsilon_{EM}^* = \frac{(\epsilon_1 - \epsilon_2)(v_1 - v_2) + [(\epsilon_1 - \epsilon_2)^2(v_1 - v_2)^2 + 4\epsilon_1\epsilon_2]^{1/2}}{2}. \tag{4.21}$$

Therefore, it is easy to demonstrate that Eq. (4.19) is actually satisfied by $\epsilon_{EM}^* = 1/\gamma_{EM}^*$ but not by ϵ_0^*, $1/\gamma_0^*$, ϵ_C^*, $1/\gamma_C^*$, ϵ_K^*, $1/\gamma_K^*$. This also ensures the accuracy of the EM approximation in 2D.

5. PERTURBATION EXPANSIONS IN TERMS OF THE LORENTZ FIELD

Next we shall develop another type of perturbation formulation taking account of the fact that the Green's function G has a singular point $\mathbf{r}_{12} = 0$. Let us introduce a kind of Lorentz field such that

$$\mathbf{F}(\mathbf{r}) = \frac{\delta\epsilon(\mathbf{r}) + d\epsilon_0}{d\epsilon_0}\mathbf{E}(\mathbf{r}) = \frac{\epsilon(\mathbf{r}) + (d-1)\epsilon_0}{d\epsilon_0}\mathbf{E}(\mathbf{r}) \tag{5.1}$$

and define

$$\kappa(\mathbf{r}) = \frac{d\delta\epsilon(\mathbf{r})}{\delta\epsilon(\mathbf{r}) + d\epsilon_0} = \frac{d[\epsilon(\mathbf{r}) - \epsilon_0]}{\epsilon(\mathbf{r}) + (d-1)\epsilon_0}. \tag{5.2}$$

By virtue of Eq. (2.14) or (2.15) the basic integral equation (2.12) is converted into

$$F_i(\mathbf{r}) = E_i^0(\mathbf{r}) + \int d\mathbf{r}_2 \Lambda_{ij}(\mathbf{r}_{12})\kappa(\mathbf{r}_2)F_j(\mathbf{r}_2), \tag{5.3}$$

or

$$\mathbf{F} = \mathbf{E}^0 + \Lambda\kappa\mathbf{F}, \tag{5.4}$$

with

$$\Lambda_{ij}(\mathbf{r}_{12}) = \frac{1}{d}\delta_{ij}\delta(\mathbf{r}_{12}) + \epsilon_0 G_{ij}(\mathbf{r}_{12}) = \begin{cases} \rho\,\dfrac{1}{2\pi r_{12}^2}\left(2\,\dfrac{x_{12,i}}{r_{12}}\dfrac{x_{12,j}}{r_{12}} - \delta_{ij}\right) & \text{in 2D,} \\ \rho\,\dfrac{1}{4\pi r_{12}^3}\left(3\,\dfrac{x_{12,i}}{r_{12}}\dfrac{x_{12,j}}{r_{12}} - \delta_{ij}\right) & \text{in 3D.} \end{cases} \tag{5.5}$$

Equation (5.3) or (5.4) is essentially equivalent to that employed by Brown,[7,8] Finkel'berg,[42,43] or Davies.[44]

Analogously to Eq. (2.4), the effective constant κ_{ij}^* is defined by

$$\langle\kappa(\mathbf{r})F_i(\mathbf{r})\rangle = \kappa_{ij}^*\langle F_i(\mathbf{r})\rangle \tag{5.6}$$

which, for a statistically isotropic medium, gives

$$\kappa^* = \frac{d\delta\epsilon^*}{\delta\epsilon^* + d\epsilon_0} = \frac{d(\epsilon^* - \epsilon_0)}{\epsilon^* + (d-1)\epsilon_0}. \tag{5.7}$$

Formal interation of Eq. (5.4) shows

$$\begin{aligned}\kappa^* &= \langle\kappa\rangle + [\langle\kappa\Lambda\kappa\rangle - \langle\kappa\rangle\Lambda\langle\kappa\rangle] + [\langle\kappa\Lambda\kappa\Lambda\kappa\rangle - \langle\kappa\Lambda\kappa\rangle\Lambda\langle\kappa\rangle \\ &\quad - \langle\kappa\rangle\Lambda\langle\kappa\Lambda\kappa\rangle + \langle\kappa\rangle\Lambda\langle\kappa\rangle\Lambda\langle\kappa\rangle] + \cdots \\ &= \langle\kappa\rangle_c + \langle\kappa\Lambda\kappa\rangle_c + [\langle\kappa\Lambda\kappa\Lambda\kappa\rangle_c + \langle\kappa\Lambda\langle\kappa\rangle_c\Lambda\kappa\rangle_c + \cdots \\ &\equiv \sum_{n=1}^{\infty}\kappa^{(n)}.\end{aligned} \tag{5.8}$$

In the completely random case, leading terms of κ^* up to fourth order are expressed as follows:

$$\kappa^{(2)} = \int d\mathbf{r}_2\left(\frac{1}{d}\delta_{i(i)}\delta(\mathbf{r}_{12}) + \epsilon_0 G_{i(i)}(\mathbf{r}_{12})\right)\langle\kappa(\mathbf{r}_1)\kappa(\mathbf{r}_2)\rangle_c = \langle\kappa^2\rangle_c\left(\frac{1}{d} - \int d\mathbf{r}_{12}\,\frac{\partial[\epsilon_0 g(\mathbf{r}_{12})]}{\partial x_{12,i}}\,\frac{\partial f(\mathbf{r}_{12})}{\partial x_{12,(i)}}\right) = 0, \tag{5.9}$$

$$\kappa^{(3,1)} = 0, \tag{5.10}$$

$$\kappa^{(3,2)} = \frac{1}{d}\left(1 - \frac{1}{d}\right)\langle\kappa^2\rangle_c\langle\kappa\rangle_c, \tag{5.11}$$

$$\kappa^{(4,1)} = 0, \tag{5.12}$$

$$\kappa^{(4,2)} = \kappa^{(4,3)} = 0, \tag{5.13}$$

$$\kappa^{(4,4)} = 0, \tag{5.14}$$

$$\kappa^{(4,5)} = \begin{cases} 0 & \text{in 1D or 2D,} \\ \dfrac{\langle\kappa^2\rangle_c^2}{27} & \text{in 3D,} \end{cases} \tag{5.15}$$

$$\kappa^{(4,6)} = -\frac{1}{d}\left(1 - \frac{1}{d}\right)\left(1 - \frac{2}{d}\right)\langle\kappa^2\rangle_c\langle\kappa\rangle_c^2, \tag{5.16}$$

whose detailed derivation is presented in Appendix C.

Equations (5.11) and (5.16) suggest that the decomposition law as shown in Fig. 3(b) is not available to κ. Unfortunately, therefore, contributions from nested diagrams cannot be evaluated by using the rules (i)—(iv) elucidated in Sec. 3. However, it is easily seen that every single-site diagram vanishes when $\langle\kappa\rangle_c=\langle\kappa\rangle=0$. Choosing $\epsilon_0=\epsilon^*$ we obtain

$$\kappa_C^*=\langle\kappa\rangle_c=0, \tag{5.17}$$

$$\kappa_{EM}^*=\langle\kappa\rangle=0; \tag{5.18}$$

in other words, the self-consistent cumulant solution κ_C^* proves to agree with the EM solution κ_{EM}^*. Because of the self-containedness of the approximation involved, κ_{EM}^* might be expected to provide the same result as that of ϵ_{EM}^* or γ_{EM}^*. Actually, the combination of Eq. (5.2) reproduces Eq. (3.34) or (4.16). Moreover, it follows from Eq. (5.15) that crossed diagrams illustrated in Fig. 5 contribute zero as far as the value of κ^* for a 2D material is concerned. This implies that $\kappa_C^*=\kappa_{EM}^*$ and so ϵ_{EM}^* and γ_{EM}^* give very accurate estimates, thus confirming that the self-contained single-site theory for ϵ^* and γ^* serves as a better approximation than the cumulant expansion method in 2D systems.

Unless we set $\epsilon_0=\epsilon^*$, on the other hand, we arrive at worse approximations than κ_C^* or κ_{EM}^*. In this case, Eqs. (5.17) and (5.18) must be replaced by

$$\kappa^*=\langle\kappa\rangle_c=\langle\kappa\rangle\neq 0, \tag{5.19}$$

which corresponds to Eqs. (3.28) and (3.32). The expression (5.19), together with Eq. (5.7), gives

$$\frac{\epsilon^*-\epsilon_0}{\epsilon^*+(d-1)\epsilon_0}=\left\langle\frac{\epsilon-\epsilon_0}{\epsilon+(d-1)\epsilon_0}\right\rangle_c=\left\langle\frac{\epsilon-\epsilon_0}{\epsilon+(d-1)\epsilon_0}\right\rangle, \tag{5.20}$$

or

$$\epsilon^*=\left\langle\frac{\epsilon}{\epsilon+(d-1)\epsilon_0}\right\rangle\Big/\left\langle\frac{1}{\epsilon+(d-1)\epsilon_0}\right\rangle. \tag{5.21}$$

If we interpret ϵ_0 as the permittivity of a particular phase of the material under investigation, say the permittivity of a matrix of a suspension, Eq. (5.21) reduces to the famous Maxwell—Wagner formula[45,46] in the classical theory of dilute suspensions. Of course, this formula was originally proposed as an approximate solution in the dilute limit, but Eq. (5.21) was found by Hashin and Shtrikman[47] to give an upper or lower bound for the effective permittivity of a multiphase medium, when ϵ_0 is the maximum or minimum of the phase permittivities. Meanwhile, the result for $\epsilon_0=\langle\epsilon\rangle$, that is,

$$\epsilon_{OP}^*=\left\langle\frac{\epsilon}{\epsilon+(d-1)\langle\epsilon\rangle}\right\rangle\Big/\left\langle\frac{1}{\epsilon+(d-1)\langle\epsilon\rangle}\right\rangle \tag{5.22}$$

agrees with the optical potential approximation discussed in the subsequent sections.

6. T MATRIX EXPANSIONS

Although T matrices have been widely used in the quantum mechanics of disordered systems, it seems that there have appeared only two attempts to handle classical mixtures by means of T matrix expansions; one is an approach to the continuum problem made by Dederichs and Zeller[9–11] and the other is an analysis of the network problem due to Kirkpatrick[21] (see also Blackman[26]). The aim of this section is to clarify the meaning of the T-matrix expansion in our formulation and to demonstrate the relationship to the expansions derived in previous sections. By repeated application of Eqs. (2.20) and (2.21) we have

$$\begin{aligned}\mathbf{T}&=\delta\epsilon+\delta\epsilon\mathbf{G}\delta\epsilon+\delta\epsilon\mathbf{G}\delta\epsilon\mathbf{G}\delta\epsilon+\cdots\\&=\delta\epsilon(1-\mathbf{G}\delta\epsilon)^{-1},\end{aligned} \tag{6.1}$$

$$\begin{aligned}\langle\mathbf{T}\rangle&=\langle\delta\epsilon\rangle+\langle\delta\epsilon\mathbf{G}\delta\epsilon\rangle+\langle\delta\epsilon\mathbf{G}\delta\epsilon\mathbf{G}\delta\epsilon\rangle+\cdots\\&=\delta\epsilon^*+\delta\epsilon^*\mathbf{G}\delta\epsilon^*+\delta\epsilon^*\mathbf{G}\delta\epsilon^*\mathbf{G}\delta\epsilon^*+\cdots\\&=\langle\delta\epsilon(1-\mathbf{G}\delta\epsilon)^{-1}\rangle=\delta\epsilon^*(1-\mathbf{G}\delta\epsilon^*)^{-1},\end{aligned} \tag{6.2}$$

or inversely,

$$\begin{aligned}\delta\epsilon^*&=\langle\mathbf{T}\rangle-\langle\mathbf{T}\rangle\mathbf{G}\langle\mathbf{T}\rangle+\langle\mathbf{T}\rangle\mathbf{G}\langle\mathbf{T}\rangle\mathbf{G}\langle\mathbf{T}\rangle-\cdots\\&=\langle\mathbf{T}\rangle(1+\mathbf{G}\langle\mathbf{T}\rangle)^{-1}.\end{aligned} \tag{6.3}$$

It is interesting to note that the self-consistency condition $\delta\epsilon^*=0$ is equivalent to the statement $\langle\mathbf{T}\rangle=0$.

In analogy with the T matrix for the whole system, we define a T matrix t_α for a single cell α as

$$\mathbf{t}_\alpha=\delta\epsilon_\alpha+\delta\epsilon_\alpha\mathbf{G}\mathbf{t}_\alpha=\delta\epsilon_\alpha(1-\mathbf{G}\delta\epsilon_\alpha)^{-1}, \tag{6.4}$$

where $\delta\epsilon=\sum_\alpha\delta\epsilon_\alpha$. Then, the familiar T matrix expansion in the multiple scattering theory becomes

$$\mathbf{T}=\sum_\alpha\mathbf{t}_\alpha+\sum_{\alpha\neq\beta}\sum\mathbf{t}_\alpha\mathbf{G}\mathbf{t}_\beta+\sum_{\alpha\neq\beta\neq\gamma}\sum\sum\mathbf{t}_\alpha\mathbf{G}\mathbf{t}_\beta\mathbf{G}\mathbf{t}_\gamma+\cdots, \tag{6.5}$$

each sum being taken so that no two successive subscripts are equal. From Eqs. (6.3) and (6.5) we find

$$\langle\mathbf{T}\rangle=\sum_\alpha\langle\mathbf{t}_\alpha\rangle+\sum_{\alpha\neq\beta}\sum\langle\mathbf{t}_\alpha\mathbf{G}\mathbf{t}_\beta\rangle+\sum_{\alpha\neq\beta\neq\gamma}\sum\sum\langle\mathbf{t}_\alpha\mathbf{G}\mathbf{t}_\beta\mathbf{G}\mathbf{t}_\gamma\rangle+\cdots, \tag{6.6}$$

and

$$\begin{aligned}\delta\epsilon^*=&\sum_\alpha\langle\mathbf{t}_\alpha\rangle-\left(\sum_{\alpha\beta}\sum\langle\mathbf{t}_\alpha\rangle\mathbf{G}\langle\mathbf{t}_\beta\rangle-\sum_{\alpha\neq\beta}\sum\langle\mathbf{t}_\alpha\mathbf{G}\mathbf{t}_\beta\rangle\right)\\&+\left(\sum_{\alpha\beta\gamma}\sum\sum\langle\mathbf{t}_\alpha\rangle\mathbf{G}\langle\mathbf{t}_\beta\rangle\mathbf{G}\langle\mathbf{t}_\gamma\rangle-\sum_{\alpha\beta\neq\gamma}\sum\sum\langle\mathbf{t}_\alpha\rangle\mathbf{G}\langle\mathbf{t}_\beta\mathbf{G}\mathbf{t}_\gamma\rangle\right.\\&\left.-\sum_{\alpha\neq\beta\gamma}\sum\sum\langle\mathbf{t}_\alpha\mathbf{G}\mathbf{t}_\beta\rangle\mathbf{G}\langle\mathbf{t}_\gamma\rangle+\sum_{\alpha\neq\beta\neq\gamma}\sum\sum\langle\mathbf{t}_\alpha\mathbf{G}\mathbf{t}_\beta\mathbf{G}\mathbf{t}_\gamma\rangle\right)-\cdots.\end{aligned} \tag{6.7}$$

In full notation $\mathbf{t}_\alpha$ is expressed as

$$\begin{aligned}\mathbf{t}_{\alpha,ij}(\mathbf{r}_1,\mathbf{r}_2)=&\,\delta\epsilon^\alpha\delta_{ij}\xi_\alpha(\mathbf{r}_1)\delta(\mathbf{r}_{12})+\delta\epsilon^\alpha\xi_\alpha(\mathbf{r}_1)\\&\times\int d\mathbf{r}_3\,G_{ik}(\mathbf{r}_{13})\mathbf{t}_{\alpha,kj}(\mathbf{r}_3,\mathbf{r}_2).\end{aligned} \tag{6.8}$$

Integration of Eq. (6.8) with respect to $\mathbf{r}_2$ gives

$$\begin{aligned}t_{\alpha,ij}(\mathbf{r}_1)\equiv&\int d\mathbf{r}_2\,\mathbf{t}_{\alpha,ij}(\mathbf{r}_1,\mathbf{r}_2)\\=&\,\delta\epsilon^\alpha\xi_\alpha(\mathbf{r}_1)\delta_{ij}+\delta\epsilon^\alpha\xi_\alpha(\mathbf{r}_1)\\&\times\int d\mathbf{r}_2\,G_{ik}(\mathbf{r}_{12})\mathbf{t}_{\alpha,kj}(\mathbf{r}_2).\end{aligned} \tag{6.9}$$

In Appendix D we establish that for ellipsoidal or elliptic cells

$$t_{\alpha,ij}(\mathbf{r}_1)=\tilde{t}_{\alpha,ij}\xi_\alpha(\mathbf{r}_1), \tag{6.10}$$

$$\tilde{t}_{\alpha,ij}=\delta\epsilon^{\alpha}\left(\delta_{ij}+L_{ij}\frac{\delta\epsilon^{\alpha}}{\epsilon_0}\right)^{-1}. \tag{6.11}$$

If the medium is completely random, the result is simplified to

$$\mathbf{t}_{\alpha,ij}(\mathbf{r}_1,\mathbf{r}_2)=t_{\alpha,ij}(\mathbf{r}_1)\delta(\mathbf{r}_{12})$$
$$=\tilde{t}_{\alpha}\delta_{ij}\xi_{\alpha}(\mathbf{r}_1)\delta(\mathbf{r}_{12}), \tag{6.12}$$

with

$$\tilde{t}_{\alpha}=\delta\epsilon^{\alpha}\Big/\left(1+\frac{\delta\epsilon^{\alpha}}{d\epsilon_0}\right). \tag{6.13}$$

Comparing Eq. (6.13) with Eq. (5.2) we conclude that the factor $\tilde{t}_{\alpha}$ in the T matrix substantially represents the quantity κ in the cell α.

In the completely random case the expansion series (6.7) is transformed into

$$\delta\epsilon^*=\sum_{\alpha}\langle\bar{\mathbf{t}}_{\alpha}\rangle-\sum_{\alpha}\langle\bar{\mathbf{t}}_{\alpha}\mathbf{G}\bar{\mathbf{t}}_{\alpha}\rangle+\Big(\sum_{\alpha}\langle\bar{\mathbf{t}}_{\alpha}\mathbf{G}\bar{\mathbf{t}}_{\alpha}\mathbf{G}\bar{\mathbf{t}}_{\alpha}\rangle$$
$$-\sum_{\alpha\neq\beta}\sum\langle\bar{\mathbf{t}}_{\alpha}\mathbf{G}\bar{\mathbf{t}}_{\beta}\mathbf{G}\bar{\mathbf{t}}_{\alpha}\rangle+\sum_{\alpha\neq\beta}\sum\langle\overline{\mathbf{t}_{\alpha}\mathbf{G}\bar{\mathbf{t}}_{\beta}\mathbf{G}\mathbf{t}_{\alpha}}\rangle\Big)-\cdots, \tag{6.14}$$

which is readily verified by means of Eq. (6.12). Here the bar indicates an average over the phases of a single cell according to their frequency. This average is taken only over the material properties of cells; for instance,

$$\bar{\mathbf{t}}_{\alpha,ij}(\mathbf{r}_1,\mathbf{r}_2)=\langle\tilde{t}_{\alpha}\rangle\delta_{ij}\xi_{\alpha}(\mathbf{r}_1)\delta(\mathbf{r}_{12}). \tag{6.15}$$

As an explicit form of Eq. (6.14) we can write

$$\delta\epsilon^*=\langle\tilde{t}\rangle-\langle\tilde{t}\rangle^2\int d\mathbf{r}_2\,G_{i(i)}(\mathbf{r}_{12})\delta_{\mathbf{r}_{12}}$$
$$+\Big(\langle\tilde{t}\rangle^3\int\int d\mathbf{r}_2d\mathbf{r}_3\,G_{ik}(\mathbf{r}_{12})G_{k(i)}(\mathbf{r}_{23})\delta_{\mathbf{r}_{12}}\delta_{\mathbf{r}_{23}}$$
$$-\langle\tilde{t}\rangle^3\int\int d\mathbf{r}_2d\mathbf{r}_3\,G_{ik}(\mathbf{r}_{12})G_{k(i)}(\mathbf{r}_{23})\delta_{\mathbf{r}_{13}}$$
$$\times(1-\delta_{\mathbf{r}_{12}})+\langle\tilde{t}^2\rangle\langle\tilde{t}\rangle\int\int d\mathbf{r}_2d\mathbf{r}_3\,G_{ik}(\mathbf{r}_{12})$$
$$\times G_{k(i)}(\mathbf{r}_{23})\delta_{\mathbf{r}_{13}}(1-\delta_{\mathbf{r}_{12}})\Big)-\cdots, \tag{6.16}$$

making use of such relations as Eqs. (2.32)–(2.43). We remark that the T matrix expansion (6.14) is valid even for ellipsoidal or elliptic cell materials if every term is integrated over $\mathbf{r}_2$ (refer also to Ref. 9).

The first approximation to $\delta\epsilon^*$ in Eq. (6.14) will be

$$\delta\epsilon^*=\sum_{\alpha}\langle\bar{\mathbf{t}}_{\alpha}\rangle,\quad \delta\epsilon^*=\langle\tilde{t}\rangle, \tag{6.17}$$

yielding

$$\epsilon^*=\epsilon_0\left\langle\frac{(d+1)\epsilon-\epsilon_0}{\epsilon+(d-1)\epsilon_0}\right\rangle. \tag{6.18}$$

This is just Kröner's approximation (3.33) in the case of $\epsilon_0=\langle\epsilon\rangle$. As the next step we make the approximation

$$\delta\epsilon^*=\sum_{\alpha}\langle\bar{\mathbf{t}}_{\alpha}\rangle-\sum_{\alpha}\langle\bar{\mathbf{t}}_{\alpha}\mathbf{G}\bar{\mathbf{t}}_{\alpha}\rangle+\sum_{\alpha}\langle\bar{\mathbf{t}}_{\alpha}\mathbf{G}\bar{\mathbf{t}}_{\alpha}\mathbf{G}\bar{\mathbf{t}}_{\alpha}\rangle-\cdots$$
$$=\sum_{\alpha}\langle\bar{\mathbf{t}}_{\alpha}(1+\mathbf{G}\bar{\mathbf{t}}_{\alpha})^{-1}\rangle. \tag{6.19}$$

The above assumption is equivalent to

$$\delta\epsilon^*=\langle\tilde{t}\rangle+\frac{\langle\tilde{t}\rangle^2}{d\epsilon_0}+\frac{\langle\tilde{t}\rangle^3}{(d\epsilon_0)^2}+\cdots=\langle\tilde{t}\rangle\Big/\left(1-\frac{\langle\tilde{t}\rangle}{d\epsilon_0}\right), \tag{6.20}$$

which implies Eq. (5.21). The approximation (6.19) when $\epsilon_0=\langle\epsilon\rangle$ is called the optical potential approximation[9–11] and gives an identical result with Eq. (5.22).

Lastly let us choose $\epsilon_0=\epsilon^*$ so that $\delta\epsilon^*=0$. Then Eqs. (6.17) and (6.19) become

$$\sum_{\alpha}\langle\bar{\mathbf{t}}_{\alpha}\rangle=0, \tag{6.21}$$

$$\sum_{\alpha}\langle\bar{\mathbf{t}}_{\alpha}(1+\mathbf{G}\bar{\mathbf{t}}_{\alpha})^{-1}\rangle=0, \tag{6.22}$$

both of which lead to

$$\bar{\mathbf{t}}_{\alpha}=0,\quad \langle\tilde{t}\rangle=d\epsilon^*\left\langle\frac{\epsilon-\epsilon^*}{\epsilon+(d-1)\epsilon^*}\right\rangle=0; \tag{6.23}$$

we have thus obtained Eq. (3.34) again. As a matter of fact, many terms entering in Eq. (6.14) vanish under the condition $\bar{\mathbf{t}}_{\alpha}=0$. In particular, all terms up to the third order contribute nothing, while there is only one possible nonzero term in the fourth order. That is to say,

$$\delta\epsilon^*=\langle\tilde{t}^2\rangle^2\int\int\int d\mathbf{r}_2d\mathbf{r}_3d\mathbf{r}_4\,G_{ik}(\mathbf{r}_{12})G_{kh}(\mathbf{r}_{23})$$
$$\times G_{h(i)}(\mathbf{r}_{34})\delta_{\mathbf{r}_{13}}\delta_{\mathbf{r}_{24}}(1-\delta_{\mathbf{r}_{12}})+\cdots. \tag{6.24}$$

Analogously to the calculations in earlier sections, it is shown that this leading term amounts to zero in 1D or 2D and to $\langle\tilde{t}^2\rangle^2/27\epsilon_0^3$ in 3D. For a cell material consisting of ellipsoids or ellipses, nevertheless, it must be noticed that third-order terms $\sum\sum_{\alpha\neq\beta}\langle\overline{\mathbf{t}_{\alpha}\mathbf{G}\bar{\mathbf{t}}_{\beta}\mathbf{G}\mathbf{t}_{\alpha}}\rangle$ cannot be discarded, because the integral of $\bar{\mathbf{t}}_{\alpha}$ rather than $\bar{\mathbf{t}}_{\alpha}$ itself is equal to zero. In addition, we may point out that the arguments based on T matrix expansions are applicable to the reciprocal permittivity $\delta\gamma^*$ as well as the permittivity $\delta\epsilon^*$.

7. VARIATIONAL TREATMENTS

We now turn to a discussion of variational methods. Variational principles have been repeatedly employed to obtain upper and lower bounds for the effective properties of inhomogeneous materials. Since different approaches including our work in II have recently been made,[9,11,14,16,18,28] we shall mention here the main results and add a few comments.

The statistical variational principles usually adopted are as follows[2,9,11]:

$$\int d\mathbf{r}\langle E_i(\mathbf{r})\epsilon^*_{ij}\langle E_j(\mathbf{r})\rangle\leq\int d\mathbf{r}\,\langle E^{\mathrm{A}}_i(\mathbf{r})\epsilon(\mathbf{r})E^{\mathrm{A}}_i(\mathbf{r})\rangle, \tag{7.1}$$

$$\int d\mathbf{r}\,\langle D_i(\mathbf{r})\gamma^*_{ij}\langle E_j(\mathbf{r})\rangle\leq\int d\mathbf{r}\,\langle D^{\mathrm{A}}_i(\mathbf{r})\gamma(\mathbf{r})D^{\mathrm{A}}_i(\mathbf{r})\rangle, \tag{7.2}$$

where the superscript A refers to a trial function. Instead of a boundary condition we impose the requirement that any trial function must have the same average as the true function. Denote by $\mathbf{X}\cdot\mathbf{Y}$ the spatial average of the inner product of the two vectors $\mathbf{X}$ and $\mathbf{Y}$; in a 3D case, for example,

$$\mathbf{X}\cdot\mathbf{Y}=\frac{1}{V}\int d\mathbf{r}X_i(\mathbf{r})Y_i(\mathbf{r}). \tag{7.3}$$

Then, the inequalities (7.1) and (7.2) are expressed as

$$\langle\mathbf{E}\rangle\cdot\boldsymbol{\epsilon}^*\langle\mathbf{E}\rangle\leq\langle\mathbf{E}^{\mathrm{A}}\cdot\boldsymbol{\epsilon}\mathbf{E}^{\mathrm{A}}\rangle, \tag{7.4}$$

$$\langle \mathbf{D} \rangle \cdot \gamma^* \langle \mathbf{D} \rangle \leq \langle \mathbf{D}^A \cdot \gamma \mathbf{D}^A \rangle. \tag{7.5}$$

Dederichs and Zeller[9,11] have introduced random operators $\mathbf{P}$ and $\mathbf{Q}$ such that

$$\mathbf{E}^A = \mathbf{P}\langle \mathbf{E} \rangle, \quad \mathbf{P} = 1 + G\delta\mathbf{P}, \tag{7.6}$$

$$\mathbf{D}^A = \mathbf{Q}\langle \mathbf{D} \rangle, \quad \mathbf{Q} = 1 + \Gamma\delta\mathbf{Q}. \tag{7.7}$$

In order that these equations actually give trial functions, it is sufficient that

$$\langle \delta\mathbf{P} \rangle = 0, \quad \langle \delta\mathbf{Q} \rangle = 0. \tag{7.8}$$

Substituting Eq. (7.6) into Eq. (7.4) we find

$$\langle \mathbf{E} \rangle \cdot \epsilon^* \langle \mathbf{E} \rangle \leq \langle \mathbf{E} \rangle \cdot \langle \mathbf{P}^\dagger \epsilon \mathbf{P} \rangle \langle \mathbf{E} \rangle. \tag{7.9}$$

Here the dagger indicates an adjoint operator; namely,

$$P^\dagger_{ij}(\mathbf{r}_1, \mathbf{r}_2) = P_{ji}(\mathbf{r}_2, \mathbf{r}_1). \tag{7.10}$$

Since the inequality (7.9) holds for any $\langle \mathbf{E} \rangle$, we may write symbolically

$$\epsilon^* \leq \langle \mathbf{P}^\dagger \epsilon \mathbf{P} \rangle, \tag{7.11}$$

which means that every eigenvalue of ϵ^* is smaller than or equal to the corresponding eigenvalue of $\langle \mathbf{P}^\dagger \epsilon \mathbf{P} \rangle$. Similarly,

$$\gamma^* \leq \langle \mathbf{Q}^\dagger \gamma \mathbf{Q} \rangle. \tag{7.12}$$

The above two inequalities enable us to derive upper and lower bound on ϵ^* $(=\gamma^{*-1})$; the breadth of gap between the bounds depends on the choice of $\delta\mathbf{P}$ and $\delta\mathbf{Q}$.

(i) The simplest assumption we can make is

$$\delta\mathbf{P} = 0, \quad \delta\mathbf{Q} = 0. \tag{7.13}$$

As shown in II, this leads to Wiener's elementary bounds

$$\langle \epsilon^{-1} \rangle^{-1} \leq \epsilon^* \leq \langle \epsilon \rangle, \quad \langle \gamma^{-1} \rangle^{-1} \leq \gamma^* \leq \langle \gamma \rangle. \tag{7.14}$$

(ii) Next we choose

$$\delta\mathbf{P} = \epsilon' = \epsilon - \langle \epsilon \rangle, \quad \delta\mathbf{Q} = \gamma' = \gamma - \langle \gamma \rangle; \tag{7.15}$$

or equivalently, we suppose

$$\mathbf{E}^A = \langle \mathbf{E} \rangle + \mathbf{E}^{(1)} = \langle \mathbf{E} \rangle + G\epsilon' \langle \mathbf{E} \rangle, \tag{7.16}$$

$$\mathbf{D}^A = \langle \mathbf{D} \rangle + \mathbf{D}^{(1)} = \langle \mathbf{D} \rangle + \Gamma\gamma' \langle \mathbf{D} \rangle, \tag{7.17}$$

where $\mathbf{E}^{(1)}$ and $\mathbf{D}^{(1)}$ have been defined in Eqs. (3.4) and (4.1). In Appendix E it is proved that

$$-\mathrm{G}^\dagger \epsilon_0 \mathrm{G} = \mathrm{G}^\dagger = \mathrm{G}, \quad -\Gamma^\dagger \gamma_0 \Gamma = \Gamma^\dagger = \Gamma. \tag{7.18}$$

Therefore, insertion of Eqs. (7.15) into Eq. (7.11) and (7.12) yields

$$\epsilon^* \leq \langle \epsilon \rangle + \langle \epsilon' G \epsilon' \rangle + \langle \epsilon' G \epsilon' G \epsilon' \rangle, \tag{7.19}$$

$$\gamma^* \leq \langle \gamma \rangle + \langle \gamma' \Gamma \gamma' \rangle + \langle \gamma' \Gamma \gamma' \Gamma \gamma' \rangle. \tag{7.20}$$

The right-hand sides are just the perturbation expansions of ϵ^* and γ^* up to the third order.

(iii) In place of Eqs. (7.16) and (7.17) we utilize the trial functions

$$\mathbf{E}^A = \langle \mathbf{E} \rangle + \mathbf{E}^{(1)} + \cdots + \mathbf{E}^{(N)} = \sum_{n=0}^{N} \mathbf{E}^{(n)}, \tag{7.21}$$

$$\mathbf{D}^A = \langle \mathbf{D} \rangle + \mathbf{D}^{(1)} + \cdots + \mathbf{D}^{(N)} = \sum_{n=0}^{N} \mathbf{D}^{(n)}, \tag{7.22}$$

which, for $N=2$, imply

$$\delta\mathbf{P} = \epsilon' + \epsilon' G \epsilon' - \langle \epsilon' G \epsilon' \rangle, \tag{7.23}$$

$$\delta\mathbf{Q} = \gamma' + \gamma' \Gamma \gamma' - \langle \gamma' \Gamma \gamma' \rangle. \tag{7.24}$$

In the same way as we did in (ii), we obtain

$$\epsilon^* \leq \langle \epsilon \rangle + \epsilon'^{(1)} + \cdots + \epsilon'^{(2N+1)}, \tag{7.25}$$

$$\gamma^* \leq \langle \gamma \rangle + \gamma'^{(1)} + \cdots + \gamma'^{(2N+1)}; \tag{7.26}$$

here, for instance, $\epsilon'^{(n)}$ is taken to be the value of $\delta\epsilon^{(n)}$ when $\epsilon_0 = \langle \epsilon \rangle$. Consequently, it is concluded that a finite perturbation series of ϵ^* or γ^* terminated at odd order provides an upper bound (for details see II and Ref. 9). Needless to say, a lower bound for ϵ^* is derived from the expansion series of $\gamma^* = \epsilon^{*-1}$ up to the $(2N+1)$th order.

(iv) The previous bounds can be improved by the inclusion of a set of adjustable constants $\{\lambda_n\}$ and $\{\mu_n\}$ such that

$$\mathbf{E}^A = \langle \mathbf{E} \rangle + \lambda_1 \mathbf{E}^{(1)} + \cdots + \lambda_n \mathbf{E}^{(N)}, \tag{7.27}$$

$$\mathbf{D}^A = \langle \mathbf{D} \rangle + \mu_1 \mathbf{D}^{(1)} + \cdots + \mu_N \mathbf{D}^{(N)}. \tag{7.28}$$

The multiplicative constants are to be chosen to minimize the upper bound when Eq. (7.27) or (7.28) are used as an admissible solution. The cases of $N=1$ and 2 have already been investigated by several authors.[2,9,11,14,16,18,28–30,48–50]

For simplicity consider the case of $N=1$, which corresponds to assuming that

$$\delta\mathbf{P} = \lambda_1 \epsilon', \quad \delta\mathbf{Q} = \mu_1 \gamma'. \tag{7.29}$$

Analogously to Eqs. (3.17) and (4.17) of II, and Eqs. (2.28) and (2.29) of V, then, we have for the eigenvalues ϵ_i^* and γ_i^* $(=1/\epsilon_i^*)$

$$\epsilon_i^* \leq \langle \epsilon \rangle \left(1 - \frac{(A_i^{(2)} \langle \epsilon'^2 \rangle / \langle \epsilon \rangle^2)^2}{A_i^{(2)} \langle \epsilon'^2 \rangle / \langle \epsilon \rangle^2 + A_i^{(3)} \langle \epsilon'^3 \rangle / \langle \epsilon \rangle^3} \right), \tag{7.30}$$

$$\gamma_i^* \leq \langle \gamma \rangle \left(1 - \frac{(B_i^{(2)} \langle \gamma'^2 \rangle / \langle \gamma \rangle^2)^2}{B_i^{(2)} \langle \gamma'^2 \rangle / \langle \gamma \rangle^2 + B_i^{(3)} \langle \gamma'^3 \rangle / \langle \gamma \rangle^3} \right). \tag{7.31}$$

The coefficients $A_i^{(2)}$, $A_i^{(3)}$, $B_i^{(2)}$, $B_i^{(3)}$ are the eigenvalues of

$$A_{ij}^{(2)} = -\epsilon'^{(2)}_{ij} / (\langle \epsilon'^2 \rangle / \langle \epsilon \rangle)$$

$$= -\frac{\langle \epsilon \rangle}{\langle \epsilon'^2 \rangle} \int d\mathbf{r}_2 \, G_{ij}(\mathbf{r}_{12}) \langle \epsilon'(\mathbf{r}_1) \epsilon'(\mathbf{r}_2) \rangle, \tag{7.32}$$

$$A_{ij}^{(3)} = \epsilon'^{(3)}_{ij} / (\langle \epsilon'^3 \rangle / \langle \epsilon \rangle^2)$$

$$= \frac{\langle \epsilon \rangle^2}{\langle \epsilon'^3 \rangle} \int d\mathbf{r}_2 \int d\mathbf{r}_3 \, G_{ik}(\mathbf{r}_{12}) G_{kj}(\mathbf{r}_{23})$$

$$\times \langle \epsilon'(\mathbf{r}_1) \epsilon'(\mathbf{r}_2) \epsilon'(\mathbf{r}_3) \rangle, \tag{7.33}$$

$$B_{ij}^{(2)} = -\gamma'^{(2)}_{ij} / (\langle \gamma'^2 \rangle / \langle \gamma \rangle)$$

$$= -\frac{\langle \gamma \rangle}{\langle \gamma'^2 \rangle} \int d\mathbf{r}_2 \, \Gamma_{ij}(\mathbf{r}_{12}) \langle \gamma'(\mathbf{r}_1) \gamma'(\mathbf{r}_2) \rangle, \tag{7.34}$$

$$B_{ij}^{(3)} = -\gamma_{ij}'^{(3)}/(\langle\gamma'^3\rangle/\langle\gamma\rangle^2)$$

$$= \frac{\langle\gamma\rangle^2}{\langle\gamma'^3\rangle}\int d\mathbf{r}_2\int d\mathbf{r}_3\,\Gamma_{ik}(\mathbf{r}_{12})\Gamma_{kj}(\mathbf{r}_{23})$$

$$\times\langle\gamma'(\mathbf{r}_1)\gamma'(\mathbf{r}_2)\gamma'(\mathbf{r}_3)\rangle. \tag{7.35}$$

We remark that Eq. (7.31) gives a lower bound on ϵ_i^*,

$$\epsilon_i^* \geq \frac{1}{\langle 1/\epsilon\rangle}\left[1 - \left(B_i^{(2)}\frac{\langle(1/\epsilon)'^2\rangle}{\langle 1/\epsilon\rangle^2}\right)^2\right.$$

$$\left.\times\left(B_i^{(2)}\frac{\langle(1/\epsilon)'^2\rangle}{\langle 1/\epsilon\rangle^2} + B_i^{(3)}\frac{\langle(1/\epsilon)'^3\rangle}{\langle 1/\epsilon\rangle^3}\right)^{-1}\right]^{-1}, \tag{7.36}$$

somewhat different in form from Eqs. (4.17) in II and (2.29) in V.

If the medium is statistically isotropic, Eqs. (3.16) and (4.4) show that $A^{(2)} = 1/d$ and $B^{(2)} = 1 - 1/d$. For a symmetric cell material or for a two-phase material, we have $B_{ij}^{(2)} = \delta_{ij} - A_{ij}^{(2)}$ and $B_{ij}^{(3)} = \delta_{ij} - 2A_{ij}^{(2)} + A_{ij}^{(3)}$ [use Eqs. (2.25), (7.34), and (7.35)]. For a spherical-cell material with $A^{(3)} = 1/d^2$ and $B^{(3)} = (1 - 1/d)^2$, furthermore, the bounds are expressed by

$$\epsilon^* \leq \langle\epsilon\rangle\left(1 - \frac{\langle\epsilon'^2\rangle^2/\langle\epsilon\rangle^4}{d\langle\epsilon'^2\rangle/\langle\epsilon\rangle^2 + \langle\epsilon\rangle^3}\right), \tag{7.37}$$

$$\gamma^* \leq \langle\gamma\rangle\left(1 - \frac{\langle\gamma'^2\rangle^2/\langle\gamma\rangle^4}{d\langle\gamma'^2\rangle/(d-1)\langle\gamma\rangle^2 + \langle\gamma'^3\rangle/\langle\gamma\rangle^3}\right). \tag{7.38}$$

Especially when the system consists of two phases with permittivities ϵ_1 $(=1/\gamma_1)$ and ϵ_2 $(=1/\gamma_2)$, these inequalities reduce to

$$\epsilon^* \leq \epsilon_1 v_1 + \epsilon_2 v_2 - \frac{(\epsilon_1-\epsilon_2)^2 v_1 v_2}{[(d-1)\epsilon_1+\epsilon_2]v_1 + [\epsilon_1+(d-1)\epsilon_2]v_2} \tag{7.39}$$

and

$$\gamma^* \leq \gamma_1 v_1 + \gamma_2 v_2 - \frac{(d-1)(\gamma_1-\gamma_2)^2 v_1 v_2}{[\gamma_1+(d-1)\gamma_2]v_1 + [(d-1)\gamma_1+\gamma_2]v_2}, \tag{7.40}$$

the latter being transformed into

$$\epsilon^* \geq \epsilon_1\epsilon_2\left(\epsilon_2 v_1 + \epsilon_1 v_2\right.$$

$$\left. - \frac{(d-1)(\epsilon_1-\epsilon_2)^2 v_1 v_2}{[(d-1)\epsilon_1+\epsilon_2]v_1 + [\epsilon_1+(d-1)\epsilon_2]v_2}\right)^{-1}. \tag{7.41}$$

(v) Another type of trial function applicable to symmetric cell materials is

$$\delta\mathbf{P} = \sum_\alpha (\mathbf{t}_\alpha - \bar{\mathbf{t}}_\alpha)(1 + \mathbf{G}\bar{\mathbf{t}}_\alpha)^{-1}. \tag{7.42}$$

According to the arguments advanced by Dederichs and Zeller,[9,11] the use of this trial function results in

$$\epsilon^* \leq \langle\epsilon\rangle + \sum_\alpha \langle\bar{\mathbf{t}}_\alpha(1 + \mathbf{G}\bar{\mathbf{t}}_\alpha)^{-1}\rangle, \tag{7.43}$$

whence the bound thus presented agrees with the optical potential approximation (6.19). Defining the T matrix $\mathbf{t}_\alpha$ for γ as well, we get a similar expression

$$\gamma^* \leq \langle\gamma\rangle + \sum_\alpha \langle\bar{\mathbf{t}}_\alpha(1 + \Gamma\bar{\mathbf{t}}_\alpha)^{-1}\rangle, \tag{7.44}$$

which provides us with a lower bound on ϵ^*. The inequalities (7.43) and (7.44) reduce for a spherical-cell material to

$$\epsilon^* \leq \left\langle\frac{\epsilon}{\epsilon + (d-1)\langle\epsilon\rangle}\right\rangle \Big/ \left\langle\frac{1}{\epsilon + (d-1)\langle\epsilon\rangle}\right\rangle, \tag{7.45}$$

$$\gamma^* \leq \left\langle\frac{\gamma}{(d-1)\gamma + \langle\gamma\rangle}\right\rangle \Big/ \left\langle\frac{1}{(d-1)\gamma + \langle\gamma\rangle}\right\rangle. \tag{7.46}$$

In the case of binary mixtures, Eqs. (7.45) and (7.46) coincide with Eqs. (7.37) and (7.38), respectively; for Eqs. (7.45) and (7.46) in fact yield Eqs. (7.39) and (7.40). Notice that this coincidence is only accidental and occurs neither for more-than-two phase materials nor for nonspherical cells.

(vi) Upper and lower bounds for the effective permittivity ϵ^* of a statistical isotropic material were derived by Hashin and Shtrikman[47,51] and Prager[52] without knowledge about correlation functions. Hashin and Shtrikman[47] found that their bounds for a two-phase mixture are the best possible in terms of phase permittivities and volume fractions alone. In order to rederive the Hashin–Shtrikman bounds, Dederichs and Zeller[9,11] proposed an admissible solution

$$\delta\mathbf{P} = (\mathcal{T} - \langle\mathcal{T}\rangle)(\mathbf{1} + \mathcal{G}\langle\mathcal{T}\rangle)^{-1}, \tag{7.47}$$

where

$$\mathcal{G} = -\mathbf{1}/d\epsilon_0, \tag{7.48}$$

$$\mathcal{T} = \delta\epsilon + \delta\epsilon\mathcal{G}\mathcal{T} = \delta\epsilon\left(1 + \frac{\delta\epsilon}{d\epsilon_0}\right)^{-1}. \tag{7.49}$$

As a matter of fact, it is not difficult to show that Eqs. (7.47)–(7.49) lead to

$$\epsilon^* \leq \left\langle\frac{\epsilon}{\epsilon + (d-1)\epsilon_0}\right\rangle \Big/ \left\langle\frac{1}{\epsilon + (d-1)\epsilon_0}\right\rangle, \tag{7.50}$$

whenever $\epsilon(\mathbf{r}) \leq \epsilon_0$. Similarly, for the reciprocal permittivity $\gamma^* = 1/\epsilon^*$,

$$\gamma^* \leq \left\langle\frac{1}{(d-1)\gamma + \gamma_0}\right\rangle \Big/ \left\langle\frac{1}{(d-1)\gamma + \gamma_0}\right\rangle, \tag{7.51}$$

if $\gamma(\mathbf{r}) \leq \gamma_0$. Hence the lower bound on ϵ^* is

$$\epsilon^* \geq \left\langle\frac{\epsilon}{\epsilon + (d-1)\epsilon_0}\right\rangle \Big/ \left\langle\frac{1}{\epsilon + (d-1)\epsilon_0}\right\rangle, \tag{7.52}$$

provided that $\epsilon(\mathbf{r}) \geq \epsilon_0$. If we take ϵ_0 to be the maximum or the minimum of the phase permittivities, it is confirmed that the Maxwell–Wager formula gives an upper- or lower bound on ϵ^*.

ACKNOWLEDGMENTS

The author has benefited from stimulating conversations with Dr. F. Yonezawa. Thanks are also due to the referee of Paper V for suggesting some aspects of the problem.

APPENDIX A

Consider a two-phase material with phase permittivities ϵ_1 and ϵ_2, and introduce an indicator $\eta(\mathbf{r})$ such that $\eta(\mathbf{r}) = 1$ or 0 according to whether the point $\mathbf{r}$ belongs to the first or the second phase. Then we get

$$\delta\epsilon(\mathbf{r}) = \epsilon(\mathbf{r}) - \epsilon_0 = (\epsilon_2 - \epsilon_0) + (\epsilon_1 - \epsilon_2)\eta(\mathbf{r}), \tag{A1}$$

so that

$$\langle\delta\epsilon(\mathbf{r}_1)\delta\epsilon(\mathbf{r}_2)\cdots\delta\epsilon(\mathbf{r}_n)\rangle_c = (\epsilon_1-\epsilon_2)^n\langle\eta(\mathbf{r}_1)\eta(\mathbf{r}_2)\cdots\eta(\mathbf{r}_n)\rangle_c, \tag{A2}$$

and

$$\langle(\delta\epsilon)^n\rangle_c = (\epsilon_1-\epsilon_2)^n\langle\eta^n\rangle_c. \tag{A3}$$

Thus from Eq. (2.30), it follows that the correlation function $f(\mathbf{r}_{12},\mathbf{r}_{23},\ldots,\mathbf{r}_{n-1,n})$ is determined by the phase concentrations and geometries independently of the permittivities ϵ_1 and ϵ_2. This conclusion is a generalization of the statement given by Schulgasser,[18] who argues the three-point correlation alone.

APPENDIX B

It is a simple matter to establish Eqs. (4.6)–(4.9) and (4.12); as an illustration

$$\begin{aligned}\delta\gamma^{(3,1)} &= \int d\mathbf{r}_2\int d\mathbf{r}_3\,\Gamma_{ik}(\mathbf{r}_{12})\Gamma_{k(i)}(\mathbf{r}_{23})\\ &\quad\times\langle\delta\gamma(\mathbf{r}_1)\delta\gamma(\mathbf{r}_2)\delta\gamma(\mathbf{r}_3)\rangle_c\\ &= \langle(\delta\gamma)^3\rangle_c\int d\mathbf{r}_{12}\,\Gamma_{ik}(\mathbf{r}_{12})\delta_{\mathbf{r}_{12}}\\ &\quad\times\int d\mathbf{r}_{23}\,\Gamma_{k(i)}(\mathbf{r}_{23})\delta_{\mathbf{r}_{23}}\\ &= \left(1-\frac{1}{d}\right)^2\langle(\delta\gamma)^3\rangle_c/\gamma_0^2,\end{aligned} \tag{B1}$$

$$\begin{aligned}\delta\gamma^{(3,2)} &= \langle(\delta\gamma)^2\rangle_c\langle\delta\gamma\rangle_c\int d\mathbf{r}_{12}\int d\mathbf{r}_{23}\,\Gamma_{ik}(\mathbf{r}_{12})\\ &\quad\times\Gamma_{k(i)}(\mathbf{r}_{23})\delta_{\mathbf{r}_{13}}\\ &= \frac{\langle(\delta\gamma)^2\rangle_c\langle\delta\gamma\rangle_c}{\gamma_0^2}\left(1+\frac{2}{\gamma_0}\int d\mathbf{r}_{12}G_{i(i)}(\mathbf{r}_{12})\delta_{\mathbf{r}_{12}}\right.\\ &\quad\left.+\frac{1}{\gamma_0^2}\int d\mathbf{r}_{12}\int d\mathbf{r}_{23}\,G_{ik}(\mathbf{r}_{12})G_{k(i)}(\mathbf{r}_{23})\delta_{\mathbf{r}_{12}+\mathbf{r}_{23}}\right)\\ &= \left(1-\frac{1}{d}\right)\frac{\langle(\delta\gamma)^2\rangle_c\langle\delta\gamma\rangle_c}{\gamma_0^2}.\end{aligned} \tag{B2}$$

We now turn to a calculation of $\delta\gamma^{(4,4)}$. By definition,

$$\begin{aligned}\delta\gamma^{(4,4)} &= \langle(\delta\gamma)^2\rangle_c^2\int d\mathbf{r}_{12}\int d\mathbf{r}_{23}\int d\mathbf{r}_{34}\,\Gamma_{ik}(\mathbf{r}_{12})\Gamma_{kh}(\mathbf{r}_{23})\\ &\quad\times\Gamma_{h(i)}(\mathbf{r}_{34})\delta_{\mathbf{r}_{14}}\delta_{\mathbf{r}_{23}}.\end{aligned} \tag{B3}$$

As in III–V, the null function $\delta_{\mathbf{r}}$ will be taken to be the limit of a characteristic function $I_\rho(r)$ that assumes the value 1 or 0 according as $r\lesseqgtr\rho$. By reference to Eqs. (4.3) in III and (4.19) in V we introduce

$$\begin{aligned}J_{\rho,ij}(\mathbf{r}_{23}) &= -\int d\mathbf{r}_{12}G_{ij}(\mathbf{r}_{12})I_\rho(r_{13})\\ &= C_\rho(r_{23})\frac{x_{23,i}}{r_{23}}\frac{x_{23,j}}{r_{23}}+D_\rho(r_{23})\delta_{ij},\end{aligned} \tag{B4}$$

where

$$C_\rho(r_{23}) = \begin{cases}0 & \text{for } r<\rho,\\ -\rho^d/r^d & \text{for } r>\rho,\end{cases} \tag{B5}$$

$$D_\rho(r_{23}) = \begin{cases}1/d & \text{for } r<\rho,\\ \rho^d/dr_{23}^d & \text{for } r>\rho.\end{cases} \tag{B6}$$

The formula corresponding to Eq. (C4) in III is

$$J_{\rho,ij}(\mathbf{r}_{23}) = -\int d\mathbf{r}_{12}\,G_{ik}(\mathbf{r}_{12})J_{\rho,kj}(\mathbf{r}_{13}). \tag{B7}$$

Making use of Eqs. (2.25), (B4), and (B7), we have

$$\begin{aligned}\delta\gamma^{(4,4)} &= -\lim_{\rho\downarrow0}\left(1-\frac{1}{d}\right)^2\frac{\langle(\delta\gamma)^2\rangle_c^2}{\gamma_0^3}\\ &\quad\times\left(1-\frac{1}{1-1/d}\int_0^\infty\frac{I_\rho(r_{23})C_\rho(r_{23})}{r_{23}}dr_{23}\right).\end{aligned} \tag{B8}$$

Substitution of Eqs. (B5) and (B6) into Eq. (B8) yields

$$\delta\gamma^{(4,4)} = -\left(1-\frac{1}{d}\right)^2\frac{\langle(\delta\gamma)^2\rangle_c^2}{\gamma_0^3}. \tag{B9}$$

The procedure to evaluate $\delta\gamma^{(4,5)}$ goes in like manner. The term $\delta\gamma^{(4,5)}$ is given by

$$\begin{aligned}\delta\gamma^{(4,5)} &= \langle(\delta\gamma)^2\rangle_c^2\int d\mathbf{r}_{12}\int d\mathbf{r}_{23}\int d\mathbf{r}_{34}\,\Gamma_{ik}(\mathbf{r}_{12})\\ &\quad\times\Gamma_{kh}(\mathbf{r}_{23})\Gamma_{h(i)}(\mathbf{r}_{34})\delta_{\mathbf{r}_{13}}\delta_{\mathbf{r}_{24}}.\end{aligned} \tag{B10}$$

Converting Γ in Eq. (B10) to G, we obtain for the 3D case

$$\begin{aligned}\delta\gamma^{(4,5)} &= -\lim_{\rho\downarrow0}\frac{\langle(\delta\gamma)^2\rangle_c^2}{3\gamma_0^3}\left(1-2\int_0^\infty\frac{I_\rho(r_{23})C_\rho(r_{23})}{r_{23}}dr_{23}\right)\\ &= -\frac{\langle(\delta\gamma)^2\rangle_c^2}{3\gamma_0^3}.\end{aligned} \tag{B11}$$

In the 2D case, the result is

$$\begin{aligned}\delta\gamma^{(4,5)} &= -\lim_{\rho\downarrow0}\frac{\langle(\delta\gamma)^2\rangle_c}{8\gamma_0^3}\left(1-4\int_0^\infty\frac{I_\rho(r_{23})C_\rho(r_{23})}{r_{23}}dr_{23}\right)\\ &= -\frac{\langle(\delta\gamma)^2\rangle_c^2}{8\gamma_0^3}.\end{aligned} \tag{B12}$$

The validity of Eq. (4.11) when $d=1$ is almost evident, because $\delta\gamma^*=\langle\delta\gamma\rangle$ in 1D.

APPENDIX C

Analogously to the proof of Eqs. (4.6)–(4.12), we can prove Eqs. (5.10)–(5.16); it is only necessary to use Eq. (5.5) instead of Eq. (2.25). By way of example let us show the results for $\kappa^{(3,1)}$, $\kappa^{(3,2)}$, $\kappa^{(4,4)}$, $\kappa^{(4,5)}$:

$$\begin{aligned}\kappa^{(3,1)} &= \langle\kappa^3\rangle_c\int d\mathbf{r}_{12}\int d\mathbf{r}_{23}\,\Lambda_{ik}(\mathbf{r}_{12})\\ &\quad\times\Lambda_{k(i)}(\mathbf{r}_{23})\delta_{\mathbf{r}_{12}}\delta_{\mathbf{r}_{23}}\\ &= \langle\kappa^3\rangle_c\int d\mathbf{r}_{12}\,\Lambda_{ik}(\mathbf{r}_{12})\delta_{\mathbf{r}_{12}}\int d\mathbf{r}_{23}\,\Lambda_{k(i)}(\mathbf{r}_{23})\delta_{\mathbf{r}_{23}}\\ &= 0,\end{aligned} \tag{C1}$$

$$\kappa^{(3,2)} = \langle\kappa^2\rangle_c\langle\kappa\rangle_c \int d\mathbf{r}_{12}\int d\mathbf{r}_{23}\,\Lambda_{ik}(\mathbf{r}_{12})\Lambda_{k(i)}(\mathbf{r}_{23})\delta_{\mathbf{r}_{13}}$$

$$= \langle\kappa^2\rangle_c\langle\kappa\rangle_c\left(\frac{1}{d^2} + \frac{2}{d}\epsilon_0\int d\mathbf{r}_{12}\,G_{i(i)}(\mathbf{r}_{12})\delta_{\mathbf{r}_{12}}\right.$$

$$\left. + \epsilon_0^2\int d\mathbf{r}_{12}\int d\mathbf{r}_{23}\,G_{ik}(\mathbf{r}_{12})G_{k(i)}(\mathbf{r}_{23})\delta_{\mathbf{r}_{13}}\right)$$

$$= \frac{1}{d}\left(1 - \frac{1}{d}\right)\langle\kappa^2\rangle_c\langle\kappa\rangle_c, \tag{C2}$$

$$\kappa^{(4,4)} = \langle\kappa^2\rangle_c^2\int d\mathbf{r}_{12}\int d\mathbf{r}_{23}\int d\mathbf{r}_{34}\,\Lambda_{ik}(\mathbf{r}_{12})\Lambda_{kh}(\mathbf{r}_{23})$$

$$\times\Lambda_{h(i)}(\mathbf{r}_{34})\delta_{\mathbf{r}_{14}}\delta_{\mathbf{r}_{23}}$$

$$= \lim_{\rho\downarrow 0}\tfrac{2}{9}\langle\kappa^2\rangle_c\int_0^\infty \frac{I_\rho(r_{23})C_\rho(r_{23})}{r_{23}}dr_{23}$$

$$= 0 \text{ in 3D}, \tag{C3}$$

$$\kappa^{(4,4)} = 0 \text{ in 2D}, \tag{C4}$$

$$\kappa^{(4,5)} = \langle\kappa^2\rangle_c^2\int d\mathbf{r}_{12}\int d\mathbf{r}_{23}\int d\mathbf{r}_{34}\,\Lambda_{ik}(\mathbf{r}_{12})\Lambda_{kh}(\mathbf{r}_{23})$$

$$\times\Lambda_{h(i)}(\mathbf{r}_{34})\delta_{\mathbf{r}_{13}}\delta_{\mathbf{r}_{24}}$$

$$= \lim_{\rho\downarrow 0}\frac{\langle\kappa^2\rangle_c^2}{27}\left(1 + 6\int_0^\infty \frac{I_\rho(r_{23})C_\rho(r_{23})}{r_{23}}dr_{23}\right)$$

$$= \frac{\langle\kappa^2\rangle_c^2}{27} \text{ in 3D}, \tag{C5}$$

$$\kappa^{(4,5)} = 0 \text{ in 2D}. \tag{C6}$$

In contrast to the 3D case, Eqs. (C4) and (C6) hold regardless of the form of $I_\rho(r_{23})$ or $C_\rho(r_{23})$.

APPENDIX D

Substituting Eq. (6.10) into Eq. (6.9) we obtain

$$\tilde{t}_{\alpha,ij}\xi_\alpha(\mathbf{r}_1) = \delta\epsilon^\alpha\xi_\alpha(\mathbf{r}_1)\delta_{ij} + \delta\epsilon^\alpha\tilde{t}_{\alpha,kj}\xi_\alpha(\mathbf{r}_1)$$

$$\times\int d\mathbf{r}_2 G_{ik}(\mathbf{r}_{12})\xi_\alpha(\mathbf{r}_2). \tag{D1}$$

According to Eqs. (3.13) of I and (3.6) of V, it holds for a cell material composed of uniformly oriented ellipsoids or ellipses that

$$\xi_\alpha(\mathbf{r}_1)\int d\mathbf{r}_2\,G_{ik}(\mathbf{r}_{12})\xi_\alpha(\mathbf{r}_2) = -\frac{L_{ik}}{\epsilon_0}\xi_\alpha(\mathbf{r}_1). \tag{D2}$$

This leads to

$$\tilde{t}_{\alpha,ij} = \delta\epsilon^\alpha - \frac{\delta\epsilon^\alpha}{\epsilon_0}L_{ik}\tilde{t}_{\alpha,kj}, \tag{D3}$$

from which Eq. (6.11) follows. Since the solution of Eq. (6.8) is unique, the proof of Eqs. (6.10) and (6.11) is completed. Especially for a completely random material, the depolarization tensor becomes $L_{ij} = \delta_{ij}/d$ and $\xi_\alpha(\mathbf{r}_1)\xi_\alpha(\mathbf{r}_2)$ vanishes unless $\mathbf{r}_1 = \mathbf{r}_2$. From Eq. (D2), thereupon, we find

$$\xi_\alpha(\mathbf{r}_1)G_{ik}(\mathbf{r}_{12})\xi_\alpha(\mathbf{r}_2) = -\frac{\delta_{ik}}{d\epsilon_0}\xi_\alpha(\mathbf{r}_1)\delta(\mathbf{r}_{12}). \tag{D4}$$

In this case, insertion of Eq. (6.12) into Eq. (6.8) gives

$$\tilde{t}_\alpha = \delta\epsilon^\alpha - \frac{\delta\epsilon^\alpha}{d\epsilon_0}\tilde{t}_\alpha, \tag{D5}$$

which implies Eq. (6.13).

APPENDIX E

Equation (7.18) is a dielectric analog of Eq. (31) in Ref. 11 and has already been used implicitly in II. In full notation we can write $\mathbf{G}^\dagger\epsilon_0\mathbf{G}$ as

$$\int d\mathbf{r}_3 G^\dagger_{ik}(\mathbf{r}_{13})\epsilon_0 G_{kj}(\mathbf{r}_{32})$$

$$= \epsilon_0\frac{\partial^2}{\partial x_{1,i}\partial x_{2,j}}\int d\mathbf{r}_3\,\frac{\partial g(\mathbf{r}_{31})}{\partial x_{3,k}}\frac{\partial g(\mathbf{r}_{32})}{\partial x_{3,k}}. \tag{E1}$$

By integration by parts the right-hand side is transformed into

$$-\epsilon_0\frac{\partial^2}{\partial x_{1,i}\partial x_{2,j}}\int d\mathbf{r}_3\,g(\mathbf{r}_{13})\frac{\partial^2 g(\mathbf{r}_{32})}{\partial x_{3,k}^2}$$

$$= -\epsilon_0\frac{\partial^2}{\partial x_{1,i}\partial x_{2,j}}\int d\mathbf{r}_3\,g(\mathbf{r}_{13})\delta(\mathbf{r}_{32})$$

$$= -G_{ij}(\mathbf{r}_{12}) = -G^\dagger_{ij}(\mathbf{r}_{12}). \tag{E2}$$

Similarly,

$$\int d\mathbf{r}_3\,\Gamma^\dagger_{ik}(\mathbf{r}_{13})\gamma_0\Gamma_{kj}(\mathbf{r}_{32}) = -\Gamma_{ij}(\mathbf{r}_{12}) = \Gamma^\dagger_{ij}(\mathbf{r}_{12}). \tag{E3}$$

[1] M. Hori, J. Math. Phys. 14, 514 (1973).
[2] M. Hori, J. Math. Phys. 14, 1942 (1973).
[3] M. Hori and F. Yonezawa, J. Math. Phys. 15, 2177 (1974).
[4] M. Hori and F. Yonezawa, J. Math. Phys. 16, 352 (1975).
[5] M. Hori and F. Yonezawa, J. Math. Phys. 16, 365 (1975).
[6] M. Hori, J. Math. Phys. 16, 1772 (1975); 17, 598(E) (1976).
[7] W. F. Brown Jr., J. Chem. Phys. 23, 1514 (1955).
[8] W. F. Brown Jr., Trans. Soc. Rheol. 9, 1, 357 (1965).
[9] P. H. Dederichs and R. Zeller, Kernforschungsanlage Jülich Berichte JÜL-877-FF (1972).
[10] R. Zeller and P. H. Dederichs, Phys. Status Solidi B 55, 831 (1973).
[11] P. H. Dederichs and R. Zeller, Z. Phys. 259, 103 (1973).
[12] E. Kröner, Int. J. Eng. Sci. 11, 171 (1973).
[13] J. Korringa, J. Math. Phys. 14, 509 (1973).
[14] M. A. Elsayed and J. J. McCoy, J. Compos. Mater. 7, 466 (1973).
[15] W. F. Brown Jr., J. Math. Phys. 15, 1516 (1974).
[16] M. A. Elsayed, J. Math. Phys. 15, 2001 (1974).
[17] K. Schulgasser, J. Math. Phys. 17, 378 (1976).
[18] K. Schulgasser, J. Math. Phys. 17, 382 (1976).
[19] M. Hori and F. Yonezawa, J. Phys. C 10, 229 (1977).
[20] S. Kirkpatrick, Phys. Rev. Lett. 27, 1722 (1971).
[21] S. Kirkpatrick, Rev. Mod. Phys. 45 574 (1973).
[22] B. P. Watson and P. L. Leath, Phys. Rev. B 9, 4893 (1974).
[23] R. B. Stinchcombe, J. Phys. C 6, L1 (1973).
[24] R. B. Stinchcombe, J. Phys. C 7, 179 (1974).
[25] J. W. Essam, C. M. Place, and E. H. Sondheimer, J. Phys. C 7, L258 (1974).
[26] J. A. Blackman, J. Phys. C 9, 2049 (1976).
[27] J. E. Molyneux, J. Math. Phys. 10, 912 (1969).

[28]J.E. Molyneux, J. Math. Phys. 11, 1172 (1970).
[29]M.N. Miller, J. Math. Phys. 10, 1988 (1969).
[30]M.N. Miller, J. Math. Phys. 10, 2005 (1969); 12, 1057 (1971).
[31]E. Kröner, J. Mech. Phys. Solids 15, 319 (1967).
[32]F. Yonezawa, Prog. Theor. Phys. 40, 734 (1968).
[33]D.A.G. Bruggeman, Ann. Phys. (Leipzig) 24, 636 (1935).
[34]C.J.F. Böttcher, Recl. Trav. Chim. Pays Bas 64, 47 (1945).
[35]V.I. Odelevskii, Zh. Tekh. Fiz. 21, 678 (1951).
[36]R. Landauer, J. Appl. Phys. 23, 779 (1952).
[37]E.H. Kerner, Proc. Phys. Soc. (London) B 69, 802 (1956).
[38]M.H. Cohen and J. Jortner, Phys. Rev. Lett. 30, 696 (1973).
[39]J.B. Keller, J. Math. Phys. 5, 548 (1964).
[40]J.E. Flaherty and J.B. Keller, Commun. Pure Appl. Math. 26, 565 (1973).
[41]K.S. Mendelson, J. Appl. Phys. 46, 917 (1975).
[42]V.M. Finkel'berg, Zh. Eksp. Teor. Fiz. 46, 725 (1964) [Sov. Phys. JETP 19, 494 (1964)].
[43]V.M. Finekl'berg, Zh. Tekh. Fiz. 34, 509 (1964) [Sov. Phys. Tech. Phys. 9, 396 (1964)].
[44]W.E.A. Davies, J. Phys. D 4, 318 (1971).
[45]J.C. Maxwell, *Treatise on Electricity and Magnetism* (Oxford U.P., London. 1904), 3rd ed., Vol. 1, p. 440.
[46]K.W. Wagner, Arch. Elektrotech. Berlin 2, 371 (1914).
[47]Z. Hasin and S. Shtrikman, J. Appl. Phys. 33, 3125 (1962).
[48]M. Beran, Nuovo Cimento 38, 771 (1965).
[49]M. Beran and J. Molyneux, Q. Appl. Math. 24, 107 (1966).
[50]M.J. Beran and N.R. Silnutzer, J. Compos. Mater. 5, 246 (1971).
[51]Z. Hashin and S. Shtrikman, Phys. Rev. 130, 129 (1963).
[52]S. Prager, J. Chem. Phys. 50, 4305 (1969).

On invariance groups and Lagrangian theories

J. C. Houard

Laboratoire de Physique Théorique et Mathématique, Université de Paris 7, Paris-Cedex 05, France 75221
(Received 11 February 1976)

The problem of determining the Lagrangian theories whose equations of motion are invariant under a given transformation group is formulated and studied.

1. INTRODUCTION

It is known that, for a mechanical system defined from a Lagrange function, the invariance of the equations of motion under a transformation group does not necessarily imply the invariance of the Lagrangian.[1-4] The reason is found in that the equations of motion remain unchanged if the Lagrangian is modified by the addition of a total derivative with respect to time. So, invariance is preserved if, under the action of any element of the group, the transformed Lagrangian only differs from the original one by such a derivative. By writing the conditions expressing the group law we find an interesting structure in which the basic equations were established in Ref. 2. The invariance of the Lagrangian up to a derivative at first implies the existence of a "gauge function" defined on the product $G\times M$, where G is the group and M is the configuration space including the time. The group property then furnishes a functional equation to be satisfied by any gauge function and depending on a second function defined on $G\times G$. This last function must, in turn, obey another functional equation which is identified as the functional equation of what is called "exponents" (or 2-cocycles) of G.[5] These exponents are related to the central extensions of the group G.

It follows that the problem of determining the Lagrangian systems invariant under a given transformation group requires the successive solutions of the above functional equations, starting from the exponents to end with the Lagrange function. In these questions, which in fact belong to the domain of cohomology theory, some partial answers and examples are already known.[2-5] Here we intend to examine the whole of the problem as directly as possible. We limit ourselves to the case where the configuration space is a C^∞ manifold and G a Lie group. The equations will be solved only locally both on the group and on the configuration space. This is not necessarily a restriction because a local Lagrangian may lead to equations of motion which have a sense everywhere in the configuration space except for singularities, and an example of such a situation will be given. Moreover, the construction of conserved quantities by Noether's theorem only needs the invariance under a local group.

In Sec. 2 the equations defining the problem are established. Section 3 is devoted to the equation for the exponents; after introducing some differential forms from any exponent, we reciprocally deduce a general formula giving the local exponents in terms of the closed left invariant 2-forms on the group. The local gauge functions are examined in Sec. 4; for transitive groups, by an alternative reasoning as that of Ref. 2, a complete solution is given, while for intransitive groups the problem is reduced to the solution of a system of partial differential equations. Section 5 deals with the determination of the Lagrange function. Finally we show in Sec. 6 that the conserved quantities given by Noether's theorem generate an algebra which is identical to the central extension defined by the associated exponent, and in Sec. 7 we give some new examples.

2. BASIC EQUATIONS

The space of the "events" (t,q), where q belongs to the configuration space of the system, is denoted by M and assumed to be a C^∞ manifold. A group G is also given which is a Lie transformation group of M; the transformed element of $x\in M$ by $\omega\in G$ will be denoted by x^ω.

Let $L(x,\dot{x})$ be a Lagrangian describing the system, where $x\in M$ and $\dot{x}$ corresponds to the derivative dx/ds with s an arbitrary parameter; the function L is a homogeneous function of the first degree with respect to $\dot{x}$. One knows[1] that the equations of motion are unchanged if (and only if) one replaces L by a function of the form

$$L_\Lambda(x,\dot{x}) = L(x,\dot{x}) + \dot{\widehat{\Lambda(x)}}, \tag{2.1}$$

in which Λ is an arbitrary real function and $\dot{\Lambda}$ is defined by

$$\dot{\widehat{\Lambda(x)}} = \dot{x}^k \frac{\partial}{\partial x^k}\Lambda(x), \tag{2.2}$$

the quantities x^k representing the coordinates of x in any chart on M. For these equations being derived by the action principle one immediately sees that equivalent equations are obtained for the variables x^ω, with ω fixed, if one uses the transformed Lagrangian L_ω which is defined by

$$L_\omega(x^\omega,\dot{x}^\omega) = L(x,\dot{x}), \tag{2.3}$$

where the variables $\dot{x}^\omega$ are given by the formula

$$(\dot{x}^\omega)^k = \dot{x}^i \frac{\partial}{\partial x^i}(x^\omega)^k. \tag{2.4}$$

Invariance of the equations of motion under the action of the group G means that the variables x^ω obey the same equations as the variables x; thus the equations of motion will be invariant under the group G if a function $\Lambda^\omega(x)$ exists such that

$$(L_\omega)_{\Lambda^\omega} \equiv L \quad \forall\ \omega\in G \tag{2.5}$$

or, explicitly,

$$L(x,\dot{x}) + \dot{\widehat{\Lambda^\omega(x^\omega)}} = L(x^\omega,\dot{x}^\omega). \tag{2.6}$$

By writing this condition for $\omega=\omega_1\omega_2$ and using the group law, one easily finds

$$d[\Lambda^{\omega_1\omega_2}(x) - \Lambda^{\omega_2}(x^{\omega_1^{-1}}) - \Lambda^{\omega_1}(x)] = 0,$$

where the differentiation operates on the x variable. An integration then gives

$$\Lambda^{\omega_1\omega_2}(x) = \Lambda^{\omega_2}(x^{\omega_1^{-1}}) + \Lambda^{\omega_1}(x) - \lambda(\omega_1, \omega_2). \tag{2.7}$$

This last equation imposes some restriction on the function λ: Expressing the equality $\Lambda^{(\omega_1\omega_2)\omega_3} = \Lambda^{\omega_1(\omega_2\omega_3)}$, one obtains

$$\lambda(\omega_1, \omega_2\omega_3) - \lambda(\omega_1\omega_2, \omega_3) = \lambda(\omega_1, \omega_2) - \lambda(\omega_2, \omega_3). \tag{2.8}$$

From (2.7) and (2.8) one derives the equalities

$$\Lambda^e(x) = \lambda(e, \omega) = \lambda(\omega, e) = \lambda(e, e),$$

where e stands for the neutral element of G. Next, one observes that the preceding equations are left unchanged if one replaces Λ by $\Lambda - \lambda(e, e)$, and λ by $\lambda - \lambda(e, e)$; so one may impose the conditions

$$\Lambda^e(x) = 0, \tag{2.9}$$

$$\lambda(e, e) = 0. \tag{2.10}$$

A real function satisfying (2.8) and (2.10) will be called an exponent[5] (or a 2-cocycle) on the group G, and a real function satisfying (2.7) and (2.9) will be called a gauge function on M associated with λ.[2] We are interested in the problem of determining the Lagrangians leading to equations of motion invariant under the group G. For that purpose we have to solve successively the Eqs. (2.8), (2.7), and (2.6). In what follows, the solutions will be given only locally as well on the group G as on the space M, and we shall refer to them as local exponents or local gauge functions. In fact we are only interested in determining the classes of Lagrangians modulo a transformation such that (2.1) is satisfied. Two functions $\Lambda^\omega(x^\omega)$ differing by an expression of the form

$$k(\omega, x) = \phi(x) - \phi(x^\omega) - \mu(\omega), \tag{2.11}$$

lead to equivalent Lagrangians and will be called equivalent gauge functions.[2] The associated exponents then differ by the function

$$\mu(\omega_1\omega_2) - \mu(\omega_1) - \mu(\omega_2) \tag{2.12}$$

and will be called equivalent exponents (modμ).[5] An exponent or a gauge function equivalent to zero will be called trivial.

3. CONSTRUCTION OF LOCAL EXPONENTS

We restrict ourselves to the case of continuous exponents. In fact this amounts to considering only C^∞ exponents in view of the theorem of Bargmann[5] which asserts that, on a Lie group, any continuous local exponent is locally equivalent to a C^∞ one; moreover, if $\lambda \equiv 0$ (modμ) is a C^∞ local exponent, and if μ is continuous in the neighborhood of e it is also C^∞ in some neighborhood of e. We begin by associating some differential forms to any C^∞ exponent λ; if λ is only a local exponent, subsequent results remain valid in a neighborhood of the neutral element. The left and right translations by $\omega \in G$ will be respectively denoted by γ_ω and δ_ω and the symmetry $\omega \to \omega^{-1}$ by S; if X is a left invariant vector field on G, the transformed[6] $\bar{X} = -S_* X$ is right invariant and coincides with X at e, $\bar{X}_e = X_e$, and the brackets are connected by the relation

$$\overline{[X, Y]} = -[\bar{X}, \bar{Y}]. \tag{3.1}$$

Let us now write the Eq. (2.8) for $\omega_1 = \omega$, $\omega_2 = \exp(tX)$, and $\omega_3 = \omega'$. Taking the derivative with respect to t, at $t = 0$, we obtain

$$(\bar{X}' - X)\lambda(\omega, \omega') = \Phi(X)(\omega) - \bar{\Phi}(\bar{X})(\omega'), \tag{3.2}$$

where $\bar{X}'$ represents the vector field $\bar{X}$ acting on the variable ω' and where we have put

$$\Phi(X)(\omega) = \frac{d}{dt}\lambda(\omega, \exp tX)\Big|_{t=0} = X'\lambda(\omega, \omega')\Big|_{\omega'=e}, \tag{3.3}$$

$$\bar{\Phi}(\bar{X})(\omega') = \frac{d}{dt}\lambda(\exp tX, \omega')\Big|_{t=0} = \bar{X}\lambda(\omega, \omega')\Big|_{\omega=e}. \tag{3.4}$$

If ω (resp. ω') is held fixed, the last member of (3.3) [resp. (3.4)] is a linear form on X_ω (resp. $\bar{X}_{\omega'}$) and it then defines a differential 1-form Φ (resp. $\bar{\Phi}$). We easily see that Φ and $\bar{\Phi}$ are C^∞. From (3.2) we now deduce

Proposition 3.1: The following 2-forms:

$$F = -2d\Phi, \quad \bar{F} = 2d\bar{\Phi}, \tag{3.5}$$

are respectively left invariant and right invariant and connected by the symmetry $\bar{F} = S^* F$.

Proof: Let Y be a left invariant vector field. By applying $\bar{Y}' - Y$ to Eq. (3.2), antisymmetrizing with respect to X and Y, and using (3.1) and (3.2), we find

$$-(X\Phi(Y) - Y\Phi(X) - \Phi([X, Y]))_{(\omega)}$$

$$= (\bar{X}\bar{\Phi}(\bar{Y}) - \bar{Y}\bar{\Phi}(\bar{X}) - \Phi([\bar{X}, \bar{Y}]))_{(\omega')};$$

that is

$$F(X, Y)_{(\omega)} = \bar{F}(\bar{X}, \bar{Y})_{(\omega')}.$$

Since ω and ω' are independent variables, the two members of this last equation are some constants that evidently express the left (resp. right) invariance of F (resp. $\bar{F}$). The same equation then means that each of F and $\bar{F}$ is transformed into the other by the symmetry S.

The 2-forms F and $\bar{F}$ are therefore closed and left or right invariant. They may be characterized as follows:

Proposition 3.2: Let $\{X_\alpha\}$ be a basis of the Lie algebra of G, let $\{\zeta^\alpha\}$ (resp. $\{\bar{\zeta}^\alpha\}$) be the dual basis of $\{X_\alpha\}$ (resp. $\{\bar{X}_\alpha\}$), and let $C^\gamma_{\alpha\beta}$ be the associated structure constants.

(1) The 2-forms F and $\bar{F}$ are of the general form

$$F = F_{\alpha\beta}\zeta^\alpha \wedge \zeta^\beta, \quad \bar{F} = F_{\alpha\beta}\bar{\zeta}^\alpha \wedge \bar{\zeta}^\beta, \tag{3.6}$$

where the constant coefficients $F_{\alpha\beta}$ are antisymmetrical and satisfy

$$C^\sigma_{\alpha\beta}F_{\sigma\gamma} + C^\sigma_{\beta\gamma}F_{\sigma\alpha} + C^\sigma_{\gamma\alpha}F_{\sigma\beta} = 0. \tag{3.7}$$

(2) For a trivial exponent the coefficients $F_{\alpha\beta}$ are given by a formula of the type

$$F_{\alpha\beta} = C^\gamma_{\alpha\beta}f_\gamma. \tag{3.8}$$

Proof: (1) Since the forms ζ^α (resp. $\bar{\zeta}^\alpha$) are left

(resp. right) invariant, formulas (3.6) represent the general form of left or right invariant 2-forms, the coefficients $F_{\alpha\beta}$ being some constants, antisymmetrical with respect to the interchange of the indices, and equal in the two formulas in view of the symmetry relation between F and $\bar{F}$. Relation (3.7) immediately follows by expressing F and $\bar{F}$ as closed and by using the formula

$$d\xi^\gamma = -\tfrac{1}{2} C^\gamma_{\alpha\beta}\, \xi^\alpha \wedge \xi^\beta. \tag{3.9}$$

(2) If λ is given by

$$\lambda(\omega, \omega') = \mu(\omega\omega') - \mu(\omega) - \mu(\omega'), \tag{3.10}$$

where μ is C^∞ in the neighborhood of e, definition (3.3) locally gives

$$\Phi(X)_{(\omega)} = X\mu(\omega) - X\mu(e) = d\mu(X)_{(\omega)} - d\mu(X)_{(e)}.$$

Then, we have

$$\begin{aligned} F(X, Y) &= -2d\Phi(X, Y) = (-X\Phi(Y) + Y\Phi(X) + \Phi([X, Y]))_{(\omega)} \\ &= -X\, d\mu(Y)_{(\omega)} + Y\, d\mu(X)_{(\omega)} + d\mu([X, Y])_{(\omega)} \\ &\quad - d\mu([X, Y])_{(e)} \end{aligned}$$

or simply, since $d\mu$ is closed,

$$F(X, Y) = -d\mu([X, Y])_{(e)}. \tag{3.11}$$

By choosing X and Y among the X_α we find (3.8) with

$$f_\gamma = -X_\gamma \mu(e). \tag{3.12}$$

Remarks: (1) The coefficients $F_{\alpha\beta}$ may be defined directly from the exponent λ by the following formula, easily deduced from (3.3), (3.5), and (3.6),

$$F_{\alpha\beta} = (X_\beta X'_\alpha - X_\alpha X'_\beta)\lambda(\omega, \omega')\big|_{\omega=\omega'=e}. \tag{3.13}$$

(2) The particular expression (3.8) automatically satisfies condition (3.7) in virtue of the Jacobi identity for structure constants.

(3) Formula (3.11) shows that for a trivial exponent $\lambda \equiv 0 \pmod{\mu}$, the condition $d\mu_e = 0$ leads to $F = 0$. The converse is true in the sense that if $F = 0$ then λ is trivial and a function μ may be found such that $d\mu_e = 0$ and $\lambda \equiv 0 \pmod{\mu}$. That result will appear as a consequence of the general expression (3.24) which we shall demonstrate below.

Up to now we have associated a closed left invariant 2-form F to any C^∞ exponent λ. Conversely, we will prove that any such 2-form determines a family of C^∞ local exponents by giving an explicit construction of the latter. Let V be an open neighborhood of the neutral element and let ω^α be a coordinate system in V, for which the coordinates of e vanish, and such that if $\omega = (\omega^\alpha) \in V$ and $0 \leq t \leq 1$ one also has $t\omega = (t\omega^\alpha) \in V$. Our main tool will be Poincaré's lemma.[7] We shall use it under the following form:

Poincaré's Lemma: Let B be a closed m-form ($m \geq 1$) on V. The general solution of the equation $dA = B$ on V is given by the formula

$$\begin{aligned} A(Z_2, \ldots, Z_m) &= da(Z_2, \ldots, Z_m) \\ &\quad + m \int_0^1 \frac{dt}{t} h_t^* B([\omega], Z_2, \ldots, Z_m), \end{aligned} \tag{3.14}$$

in which a denotes an arbitrary $(m-2)$-form (or $da = c^{te}$ if $m = 1$) and $Z_2, \ldots, Z_m$ arbitrary vector fields, $[\omega]$ being the vector field whose the components at the point (ω^α) are equal to ω^α, and h_t being the "homothety" $(\omega^\alpha) \to (t\omega^\alpha)$.

Given a closed left invariant 2-form F we have to solve successively the equations (3.5) and (3.2) with the following conditions, derived from (2.10),

$$\Phi_e = \bar{\Phi}_e = 0. \tag{3.15}$$

Poincaré's lemma immediately gives the solutions of (3.5) (note that $h_{t*}[\omega] = [\omega]$),

$$\begin{aligned} \Phi(Z)_{(\omega)} &= d\varphi(Z)_{(\omega)} - \int_0^1 \frac{dt}{t} F([\omega], h_{t*} Z)_{(t\omega)}, \\ \bar{\Phi}(Z)_{(\omega)} &= d\bar{\varphi}(Z)_{(\omega)} + \int_0^1 \frac{dt}{t} \bar{F}([\omega], h_{t*} Z)_{(t\omega)}, \end{aligned} \tag{3.16}$$

in which we have, due to (3.15),

$$d\varphi_e = d\bar{\varphi}_e = 0. \tag{3.17}$$

In order to solve (3.2), let us introduce the function

$$\mu_{\omega_0}(\omega) = \lambda(\omega, \omega^{-1}\omega_0). \tag{3.18}$$

Owing to the formula

$$(\bar{X}f)_{(\omega^{-1}\omega_0)} = -X \cdot f(\omega^{-1}\omega_0)$$

which results from the definition of $\bar{X}$, by putting $\omega' = \omega^{-1}\omega_0$ in (3.2), we find

$$X\mu_{\omega_0} = -\Phi(X) + \bar{\Phi}(\bar{X}) \circ \delta_{\omega_0} \circ S. \tag{3.19}$$

The last term transforms as follows:

$$\begin{aligned} \bar{\Phi}(\bar{X}) \circ \delta_{\omega_0} \circ S &= -\bar{\Phi}(S_* X) \circ S \circ \gamma_{\omega_0^{-1}} = -S^* \bar{\Phi}(X) \circ \gamma_{\omega_0^{-1}} \\ &= -S^* \bar{\Phi}(\gamma_{\omega_0^{-1}*} X) \circ \gamma_{\omega_0^{-1}} = -\gamma^*_{\omega_0^{-1}} S^* \bar{\Phi}(X) \\ &= -S^* \delta^*_{\omega_0} \bar{\Phi}(X), \end{aligned}$$

so that (3.19) gives

$$d\mu_{\omega_0} = -\Phi - S^* \delta^*_{\omega_0} \bar{\Phi}.$$

The differential form in the right member is indeed closed as shown by (3.5) and the assumed properties of F and $\bar{F}$. Poincaré's lemma then gives

$$\mu_{\omega_0}(\omega) = -\int_0^1 \frac{du}{u} (\Phi + S^* \delta^*_{\omega_0} \bar{\Phi})([\omega])_{(u\omega)} \tag{3.20}$$

in which we have incorporated the initial condition $\lambda(e, \omega_0) = 0$ which for any exponent follows from (2.10).

This last expression will be cast into a simpler form by specializing the coordinate system, namely by choosing the canonical coordinate system associated with a basis $\{X_\alpha\}$ of the Lie algebra of G. In such a system we have the following relations:

$$\omega^{-1} = -\omega, \quad (u\omega)^{-1} = u\omega^{-1}, \quad S_*[\omega] = [\omega],$$

and (3.20) becomes

$$\mu_{\omega_0}(\omega) = -\int_e^\omega \Phi - \int_e^{\omega^{-1}} \delta^*_{\omega_0} \bar{\Phi}, \tag{3.21}$$

the integration paths being straight lines in canonical coordinates that are segments of one-parameter subgroups. Returning to the definition of μ_{ω_0}, putting

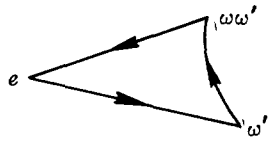


FIG. 1. The contour $\bar{C}(\omega,\omega')$. The straight lines represent segments of one-parameter subgroups and the curved line represents the right translated of such a subgroup.

$\omega_0=\omega\omega'$, and transforming the last term of (3.21) by the translation δ_ω, we find

$$\lambda(\omega,\omega')=\int_e^{\omega}(\delta_{\omega'}^{*}\bar{\Phi}-\Phi).$$

The condition $\lambda(\omega,e)=0$ also gives the constraint

$$\int_e^{\omega}\bar{\Phi}=\int_e^{\omega}\Phi$$

which, by (3.16), is equivalent to

$$\bar{\varphi}(\omega)-\bar{\varphi}(e)=\varphi(\omega)-\varphi(e). \tag{3.22}$$

We then have

$$\lambda(\omega,\omega')=\int_e^{\omega}(\delta_{\omega'}^{*}\bar{\Phi}-\bar{\Phi}). \tag{3.23}$$

In this last formula, the differential form appearing under the integral sign being closed, the integral does not depend on the integration path joining e to ω. Finally, (3.23) may be directly expressed in terms of the 2-form $\bar{F}$; adding and subtracting to (3.23) the quantity

$$\int_e^{\omega'}\bar{\Phi}-\int_e^{\omega\omega'}\bar{\Phi}=\bar{\varphi}(\omega')-\bar{\varphi}(\omega\omega')=\varphi(\omega')-\varphi(\omega\omega'),$$

we find

$$\lambda\equiv\xi(\mathrm{mod}(\varphi-\varphi(e))) \tag{3.24}$$

with

$$\xi(\omega,\omega')=\int_e^{\omega}\delta_{\omega'}^{*}\bar{\Phi}+\int_e^{\omega'}\bar{\Phi}-\int_e^{\omega\omega'}\bar{\Phi}.$$

The first term in the right-hand side may be transformed by the translation $\delta_{\omega'}$, thereby giving

$$\xi(\omega,\omega')=\int_{\bar{C}(\omega,\omega')}\bar{\Phi}=\int_{\bar{S}(\omega,\omega')}\tfrac{1}{2}\bar{F}, \tag{3.25}$$

where the contour $\bar{C}$ is represented in Fig. 1 and where $\bar{S}$ is any two-dimensional integration domain with $\bar{C}$ as boundary. With the help of some change of variables, the preceding formulas may be expressed in terms of the differential forms Φ and F, we have

$$\xi(\omega,\omega')=\int_{C(\omega,\omega')}\Phi=\int_{S(\omega,\omega')}-\tfrac{1}{2}F \tag{3.26}$$

and

$$\lambda(\omega,\omega')=\int_e^{\omega'}(\gamma_{\omega}^{*}\Phi-\Phi). \tag{3.27}$$

In the last formula the integral does not depend on the integration path, while the contour C is described in Fig. 2.

Since Φ and $\bar{\Phi}$ are defined in the canonical neighborhood V by (3.16), all the preceding relations are valid in a sufficiently restricted neighborhood of e; in particular, formulas (3.23) and (3.27) make sense if ω and ω' belong to any neighborhood v such that $v^2\subset V$, and the same is true for (3.25) and (3.26). What remains to prove is that the function ξ just obtained is indeed a local exponent of G. In fact, formula (2.8) is easily obtained, in the case where $\omega_i\in v$, $i=1,2,3$, by using the expression of ξ in terms of F and by a simple application of Stokes' formula. So, formulas (3.24) and (3.25) [or (3.26)] completed by (3.17) define the family of local exponents associated with a given 2-form F. In particular, that result contains the reciprocal of Proposition 3.2 concerning trivial exponents: If the coefficients $F_{\alpha\beta}$ are of the form (3.8), we have at first, with the help of (3.9),

$$F=-2d(f_\gamma\zeta^\gamma),$$

and then

$$\xi(\omega,\omega')=\int_{S(\omega,\omega')}d(f_\gamma\zeta^\gamma)=\int_{C(\omega,\omega')}f_\gamma\zeta^\gamma.$$

Since ζ^γ is left invariant we easily find $\xi\equiv 0\ (\mathrm{mod}\mu)$ with

$$\mu(\omega)=-\int_e^{\omega}f_\gamma\zeta^\gamma.$$

To sum up, we have obtained:

Theorem 3.3: For a Lie group, any C^∞ local exponent λ determines a closed left invariant 2-form F by the formulas

$$F=F_{\alpha\beta}\zeta^\alpha\wedge\zeta^\beta,$$

$$F_{\alpha\beta}=(X_\beta X'_\alpha-X_\alpha X'_\beta)\lambda(\omega,\omega')\big|_{\omega=\omega'=e}.$$

Reciprocally, to any closed left invariant 2-form F corresponds a family of C^∞ local exponents which is given by

$$\lambda\equiv\xi(\mathrm{mod}\mu)$$

with $d\mu_e=0$ and

$$\xi(\omega,\omega')=\int_{S(\omega,\omega')}-\tfrac{1}{2}F.$$

The trivial local exponents correspond to the forms F which are differential of a left invariant 1-form.

Corollary 3.4: The classes of local exponents on G correspond biunivocally to the classes of closed left invariant 2-forms, two such forms being equivalent if they differ by the differential of a left invariant 1-form.

4. LOCAL GAUGE FUNCTIONS

We turn now to the determination of gauge functions which is the solution of Eq. (2.7). It will be convenient to replace the function Λ by the function K defined by

$$K(\omega,x)=\Lambda^{\omega}(x^{\omega}). \tag{4.1}$$

The Eqs. (2.7) and (2.9) then become

$$K(\omega_1,x^{\omega_2})+K(\omega_2,x)-K(\omega_1\omega_2,x)=\lambda(\omega_1,\omega_2), \tag{4.2}$$

$$K(e,x)=0. \tag{4.3}$$

We first examine the case of a transitive transformation group for which a direct solution may be given. Our procedure here will be fairly similar to that of Ref. 2.

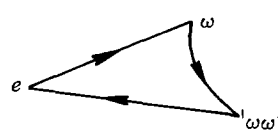


FIG. 2. The contour $C(\omega,\omega')$. The straight lines represent segments of one-parameter subgroups and the curved line represents the left translated of such a subgroup.

A. Transitive group

Suppose G is a connected and transitive Lie transformation group of M. For any fixed element x_0 of M, it may then be given an open neighborhood $u \times v$ of 0 in the Lie algebra of G so that the following conditions hold[8]:

(1) v is a neighborhood of 0 in the Lie algebra of the little group G_{x_0}, and u is a neighborhood of 0 in a given supplementary subspace of this subalgebra.

(2) The mapping $u \times v \ni (X, Y) \to \exp X \cdot \exp Y$ is a diffeomorphism of $u \times v$ on an open neighborhood of e in G.

(3) The mapping $u \ni X \to x_0^{\exp X}$ is a diffeomorphism of u on an open neighborhood of x_0 in M.

In particular, these conditions mean that the set $U = \exp u$, endowed with the chart $\exp X \to X$, is a submanifold of G, while $V = x_0^{\exp u}$ is an open submanifold of M with the chart $x_0^{\exp X} \to X$. For any element x of V we denote by $\alpha(x)$ the unique element of U such that $x = x_0^{\alpha(x)}$; the mapping $x \to \alpha(x)$ is evidently C^∞. These properties imply at first the following theorem whose proof is given in Appendix A:

Theorem 4.1: For a connected Lie group acting transitively on M:

(1) Any continuous local gauge function is locally equivalent to a C^∞ local gauge function.

(2) If the local gauge function k is given by (2.11), and is C^∞ in a neighborhood of (e, x_0), and if μ and ϕ are continuous in some respective neighborhoods of e and x_0, they also are C^∞ in some neighborhoods.

As for the exponents, that result reduces the study of continuous gauge functions to that of C^∞ ones. Now let λ be a C^∞ local exponent on $\exp u \cdot \exp v$ and let K be a C^∞ local gauge function on $(\exp u \cdot \exp v) \times V$ associated with λ. Let us consider the following quantity, which is defined for (ω, x) sufficiently close to (e, x_0),

$$K(\alpha(x^\omega), x_0) + K(\alpha(x^\omega)^{-1}\omega\alpha(x), x_0) + K(\alpha(x)^{-1}, x). \quad (4.4)$$

[All subsequent calculations are valid in a sufficiently restricted neighborhood of (e, x_0). To simplify, we shall not explicitly indicate the precise conditions which validate each step.] The sum of the last two terms transforms by applying (4.2) with the substitutions $\omega_1 \to \alpha(x^\omega)^{-1}\omega\alpha(x)$, $\omega_2 \to \alpha(x)^{-1}$ so that (4.4) reads

$$K(\alpha(x^\omega), x_0) + K(\alpha(x^\omega)^{-1}\omega, x) + \lambda(\alpha(x^\omega)^{-1}\omega\alpha(x), \alpha(x)^{-1}).$$

A second application of (4.2) with $\omega_1 \to \alpha(x^\omega)$, $\omega_2 \to \alpha(x^\omega)^{-1}\omega$ gives

$$K(\omega, x) + \lambda(\alpha(x^\omega), \alpha(x^\omega)^{-1}\omega) + \lambda(\alpha(x^\omega)^{-1}\omega\alpha(x), \alpha(x)^{-1}). \quad (4.5)$$

By using the relation [take $\omega_1 = \alpha(x)$, $\omega_2 = \alpha(x)^{-1}$ in (4.2)]

$$K(\alpha(x)^{-1}, x) = -K(\alpha(x), x_0) + \lambda(\alpha(x), \alpha(x)^{-1}),$$

the comparison of (4.4) and (4.5) leads to

$$K(\omega, x) = K(\alpha(x^\omega), x_0) - K(\alpha(x), x_0) + K(\alpha(x^\omega)^{-1}\omega\alpha(x), x_0) + \lambda(\alpha(x), \alpha(x)^{-1}) - \lambda(\alpha(x^\omega), \alpha(x^\omega)^{-1}\omega) - \lambda(\alpha(x^\omega)^{-1}\omega\alpha(x), \alpha(x)^{-1})$$

and then, by using the functional equation for λ,

$$K(\omega, x) = \lambda(\alpha(x), \alpha(x)^{-1}) - K(\alpha(x), x_0) - \lambda(\alpha(x^\omega), \alpha(x^\omega)^{-1}) + K(\alpha(x^\omega), x_0) + \lambda(\alpha(x^\omega)^{-1}, \omega\alpha(x)) - \lambda(\omega\alpha(x), \alpha(x)^{-1}) + K(\alpha(x^\omega)^{-1}\omega\alpha(x), x_0).$$

The sum of the first four terms represents a trivial gauge function corresponding to a vanishing exponent. The last term only depends on the values of K on $G_{x_0} \times x_0$. By putting

$$\chi(\Omega) = K(\Omega, x_0), \quad \Omega \in G_{x_0} \quad (4.6)$$

we have then, up to an equivalence leaving λ unchanged [that is with $\mu = 0$ in formula (2.11)],

$$K(\omega, x) = \lambda(\alpha(x^\omega)^{-1}, \omega\alpha(x)) - \lambda(\omega\alpha(x), \alpha(x)^{-1}) + \chi(\alpha(x^\omega)^{-1}\omega\alpha(x)). \quad (4.7)$$

From (4.2) one derives the following functional equation for χ on G_{x_0}:

$$\chi(\Omega_1) + \chi(\Omega_2) - \chi(\Omega_1\Omega_2) = \lambda(\Omega_1, \Omega_2). \quad (4.8)$$

Reciprocally one checks that the formulas (4.7) and (4.8) actually define a local gauge function in the neighborhood of (e, x_0). It then remains to determine the solutions of (4.8).

We at first observe that (4.8) may have solutions only if λ is trivial on G_{x_0}. That condition is easily expressed with the help of the 2-form F associated with λ; if $\{X_\alpha\}$, $1 \le \alpha \le n$, denotes a basis of the Lie algebra of G such that $\{X_{\alpha'}\}$, $1 \le \alpha' \le m$, constitutes a basis of the Lie algebra of G_{x_0}, Theorem 3.3 shows that the condition on λ is equivalent to

$$F_{\alpha'\beta'} = C^{\gamma'}_{\alpha'\beta'} f_{\gamma'}, \quad (4.9)$$

where the primed indices vary from 1 to m. Formula (3.24) then gives the following general expression for λ on G_{x_0}:

$$\lambda(\Omega_1, \Omega_2) = \mu(\Omega_1) + \mu(\Omega_2) - \mu(\Omega_1\Omega_2) \quad (4.10)$$

with

$$\mu(\Omega) = \int_e^\Omega f_{\gamma'}\xi^{\gamma'} + \nu(\Omega), \quad \nu(e) = 0, \quad d\nu(e) = 0. \quad (4.11)$$

Now putting $\chi = \mu + \theta$, formula (4.8) shows that θ is a C^∞ local homomorphism of G_{x_0} into $\mathbb{R}$. Such an homomorphism is determined by a formula of the type (see Appendix B)

$$\theta(\Omega) = \int_e^\Omega \theta_{\alpha'}\xi^{\alpha'}, \quad \Omega \in G_{x_0}, \quad (4.12)$$

where the coefficients $\theta_{\alpha'}$ have to satisfy the relations

$$C^{\gamma'}_{\alpha'\beta'}\theta_{\gamma'} = 0. \quad (4.13)$$

In fact, the division of χ into $\mu + \theta$ is arbitrary as we see, for instance, by adding (4.11) and (4.12). The combination $\chi_{\alpha'} = f_{\alpha'} + \theta_{\alpha'}$ appears, which is, from (4.9) and (4.13), the general solution of the equation

$F_{\alpha'\beta'} = C^{\gamma'}_{\alpha'\beta'}\chi_{\gamma'}$, the coefficients $F_{\alpha'\beta'}$ being given. We then have

Theorem 4.2: Let λ be a C^∞ local exponent on G. Local gauge functions may be associated with λ in the neighborhood of x_0 if and only if λ is locally trivial on the little group G_{x_0}. The C^∞ representatives of these gauge functions are given [modulo a function of the type $\phi(x) - \phi(x^\omega)$] by the formula (4.7) in which χ is a C^∞ solution of (4.8).

In terms of the 2-form F associated with λ, the condition on λ is expressed by (4.9), while χ is given by

$$\chi(\Omega) = \int_e^\Omega \chi_{\alpha'}\xi^{\alpha'} + \nu(\Omega), \quad \nu(e) = 0, \quad d\nu(e) = 0, \qquad (4.14)$$

where ν is determined by λ and where the coefficients $\chi_{\alpha'}$ represent a solution of

$$F_{\alpha'\beta'} = C^{\gamma'}_{\alpha'\beta'}\chi_{\gamma'}, \quad 1 \leq \alpha', \beta', \gamma' \leq m. \qquad (4.15)$$

Let us now determine the equivalence classes of local gauge functions. According to (4.7) a gauge function is associated with a pair (λ, χ). The basic equation (4.2) firstly shows that if K_i, $i = 1, 2$, is associated with λ_i, the relation $K_1 \sim K_2$ implies $\lambda_1 \sim \lambda_2$. Reciprocally, let λ_1 and λ_2 be such that

$$\lambda_1(\omega_1, \omega_2) = \lambda_2(\omega_1, \omega_2) + \mu(\omega_1\omega_2) - \mu(\omega_1) - \mu(\omega_2) \qquad (4.16)$$

and let χ_1 and χ_2 be corresponding solutions of (4.8). We observe that, from (4.8) and (4.16), the function $\rho = \chi_1 - \chi_2 + \mu$ is necessarily a local homomorphism of G_{x_0} into $\mathbb{R}$. By using (4.7) and (4.16), the equivalence condition of the gauge functions respectively associated with (λ_1, χ_1) and (λ_2, χ_2) reads

$$\rho(\alpha(x^\omega)^{-1}\omega\alpha(x)) = \phi(x) - \phi(x^\omega) + \psi(\omega). \qquad (4.17)$$

Let us first remark that, for $x = x_0$ and $\Omega \in G_{x_0}$, relation (4.17) gives

$$\rho(\Omega) = \psi(\Omega). \qquad (4.18)$$

The following quantity

$$\rho(\alpha(x^{\omega_1\omega_2})^{-1}\omega_1\omega_2\alpha(x)) - \rho(\alpha(x^{\omega_2})^{-1}\omega_2\alpha(x)) - \rho(\alpha(x^{\omega_1\omega_2})^{-1}\omega_1\alpha(x^{\omega_2}))$$

vanishes since ρ is a local homomorphism of G_{x0}. On the other hand, due to (4.17), it is equal to $\psi(\omega_1\omega_2) - \psi(\omega_1) - \psi(\omega_2)$. This shows that ψ must be a local homomorphism of G into $\mathbb{R}$. Comparing it with (4.18) this means that the local homomorphism ρ of G_{x_0} may be extended in a local homomorphism of G. Reciprocally this last condition immediately leads to (4.17).

Let us finally give the expression of the preceding conditions in terms of the quantities $\chi_{\alpha'}$. Let F_1 and F_2 be the 2-forms respectively associated with λ_1 and λ_2; formula (4.16) gives, with the help of (3.8) and (3.12),

$$F_{1\,\alpha\beta} = F_{2\,\alpha\beta} + C^\gamma_{\alpha\beta}f_\gamma, \quad f_\gamma = -X_\gamma\mu(e). \qquad (4.19)$$

The property of ρ to be a local homomorphism of G_{x_0} is expressed by (see Appendix B)

$$C^{\gamma'}_{\alpha'\beta'}\chi_{\gamma'}\rho(e) = 0,$$

which is indeed verified with the help of (4.15) and (4.19). The condition that ρ may be extended locally on G now implies the existence of quantities R_γ, $1 \leq \gamma \leq n$, such that $R_{\gamma'} = X_{\gamma'}\rho(e)$ for $1 \leq \gamma' \leq m$ and $C^\gamma_{\alpha\beta}R_\gamma = 0$. By putting $\bar{f}_\gamma = f_\gamma + R_\gamma$, taking into account the definition of ρ, we finally have

Proposition 4.3: The local gauge functions $K_{(\lambda_1,\chi_1)}$ and $K_{(\lambda_2,\chi_2)}$ defined by (4.7) are equivalent under the following necessary and sufficient conditions:

(a) $\lambda_1 \equiv \lambda_2 (\mathrm{mod}\mu)$,

(b) The function $\chi_1 - \chi_2 + \mu$ may be extended in a local homomorphism of G into $\mathbb{R}$.

In terms of the 2-forms F_1 and F_2 and of the quantities $\chi_{i\,\alpha'} = X_{\alpha'}\chi_i(e)$, $i = 1, 2$, these conditions are equivalent to

(a′) $F_{1\,\alpha\beta} - F_{2\,\alpha\beta} = C^\gamma_{\alpha\beta}\bar{f}_\gamma$,

(b′) $\chi_{1\gamma'} - \chi_{2\gamma'} = \bar{f}_{\gamma'}$, $1 \leq \gamma' \leq m$,

for some values of the coefficients $\bar{f}_\gamma$.

B. General case

In cases where the transformation group G is not transitive, a general study along the line used in the preceding subsection would be much more difficult in view of a possible complicated structure of the orbits, and would require some additional hypotheses. Instead of that we shall limit ourselves to reducing the problem for C^∞ gauge functions to the solution of a system of partial differential equations, which eventually may be handled in practical cases. It will be convenient to introduce the notation $x^\omega = U_x(\omega)$ where U_x, $x \in M$, is a C^∞ mapping from G into M. The velocity field is the set of the vector fields $\mathcal{D}_X$ on M, X being any element in the Lie algebra of G, which are defined by

$$\mathcal{D}_X(x) = -dU_x(X_e). \qquad (4.20)$$

Definition (4.20) is equivalent to the following one, f denoting a C^∞ function on M:

$$\mathcal{D}_X f(x) = -\frac{d}{dt} f(x^{\exp tX})\Big|_{t=0}. \qquad (4.21)$$

From this formula we get

$$\mathcal{D}_X f(x^\omega) = -\frac{d}{dt} f(x^{\exp tX\cdot\omega})\Big|_{t=0} = -\bar{X}\cdot f(x^\omega)$$

or equivalently (Lie equations)

$$\mathcal{D}_X f \circ U_x = -\bar{X}(f \circ U_x). \qquad (4.22)$$

This leads to the following property of the brackets:

$$[\mathcal{D}_X, \mathcal{D}_Y] = \mathcal{D}_{[X,Y]}. \qquad (4.23)$$

Now let K be a C^∞ (local) gauge function on M associated with the C^∞ (local) exponent λ on G. By writing the formula (4.2) for $\omega_1 = \omega$, $\omega_2 = \exp tX$, taking the derivative with respect to t, at $t = 0$, we find, with the help of (3.3) and (4.21),

$$-\mathcal{D}_X K(\omega, x) + XK(e, x) - XK(\omega, x) = \Phi(X)_{(\omega)}.$$

By putting

$$K_X(x) = \frac{d}{dt} K(\exp tX, x)\Big|_{t=0} = XK(e, x) \qquad (4.24)$$

the last equation becomes

$$(X+\partial_X)K(\omega,x)=K_X(x)-\Phi(X)_{(\omega)}. \tag{4.25}$$

By applying the operator $Y+\partial_Y$ to (4.25) and antisymmetrizing with respect to X and Y we get, with the help of (4.23), (4.25), and (3.5),

$$\partial_X K_Y-\partial_Y K_X-K_{[X,Y]}=-F(X,Y), \tag{4.26}$$

where F is the differential 2-form introduced in Sec. 3. A second equation analogous to (4.25) may be obtained by putting $\omega_1=\exp tX$, $\omega_2=\omega$ in (4.2); by differentiation at $t=0$ we obtain

$$\bar{X}K(\omega,x)=K_X(x^\omega)-\bar{\Phi}(\bar{X})_{(\omega)}. \tag{4.27}$$

That equation allows to construct the function $K(\omega,x)$ in terms of the functions $K_X(x)$. Let us first remark that K_X depends linearly on X. For any fixed x in M let us introduce the differential 1-form ψ_x on G which is defined by

$$K_X(x^\omega)=(K_X\circ U_x)_{(\omega)}=\psi_x(\bar{X})_{(\omega)}. \tag{4.28}$$

From the formula $U_{x\omega}=U_x\circ\delta_\omega$ we get

$$\psi_{x\omega}=\delta_\omega^*\psi_x. \tag{4.29}$$

We also have, due to (4.22),

$$\begin{aligned}2d\psi_x(\bar{X},\bar{Y})&=\bar{X}\psi_x(\bar{Y})-\bar{Y}\psi_x(\bar{X})-\psi_x([\bar{X},\bar{Y}])\\&=\bar{X}(K_Y\circ U_x)-\bar{Y}(K_X\circ U_x)+K_{[X,Y]}\circ U_x\\&=(-\partial_X K_Y+\partial_Y K_X+K_{[X,Y]})\circ U_x\end{aligned}$$

and then, from (4.26) and the symmetry relation between F and $\bar{F}$,

$$d\psi_x=\tfrac{1}{2}\bar{F}. \tag{4.30}$$

The formula (4.27) becomes

$$\bar{X}K(\omega,x)=(\psi_x-\bar{\Phi})(\bar{X})_{(\omega)}.$$

Since the differential form on the right-hand side is closed according to (4.30) and (3.5), that equation may be integrated on a neighborhood of the neutral element of G in the following form, which takes into account condition (4.3):

$$K(\omega,x)=\int_e^\omega(\psi_x-\bar{\Phi}). \tag{4.31}$$

Conversely, let K_X be any C^∞ solution of (4.26), and let us define ψ_x and K respectively by (4.28) and (4.31). The last function is a C^∞ gauge function associated with λ: Indeed, we successively have

$$\begin{aligned}&K(\omega_1,x^{\omega_2})+K(\omega_2,x)-K(\omega_1\omega_2,x)\\&=\int_e^{\omega_1}(\psi_{x\omega_2}-\bar{\Phi})-\int_{\omega_2}^{\omega_1\omega_2}(\psi_x-\bar{\Phi})\\&=\int_e^{\omega_1}(\delta_{\omega_2}^*\psi_x-\bar{\Phi})-\int_e^{\omega_1}\delta_{\omega_2}^*(\psi_x-\bar{\Phi})\\&=\int_e^{\omega_1}(\delta_{\omega_2}^*\bar{\Phi}-\bar{\Phi})=\lambda(\omega_1,\omega_2).\end{aligned}$$

We have shown

Proposition 4.4: With any C^∞ solution $K_X(x)$, linear on X, of the equations

$$\partial_X K_Y-\partial_Y K_X-K_{[X,Y]}=-F(X,Y), \tag{4.32}$$

where F denotes the left invariant 2-form associated with the C^∞ exponent λ, corresponds a C^∞ gauge function associated with λ in the neighborhood of e and defined by the formula

$$K(\omega,x)=\int_e^\omega(\psi_x-\bar{\Phi}), \tag{4.33}$$

the differential form ψ_x being given by

$$\psi_x(\bar{X})=K_X\circ U_x. \tag{4.34}$$

This result remains obviously valid if the functions K_X are defined only on a neighborhood of an element x_0 of M; the preceding formulas then define a local gauge function in the neighborhood of (e,x_0).

Finally, the equivalence conditions of gauge functions may be formulated as follows:

Proposition 4.5: The local gauge function K which is defined from solutions K_X of (4.32) by (4.33) and (4.34) is trivial if and only if the functions K_X are of the form

$$K_X(x)=\partial_X\phi(x)-a(X), \tag{4.35}$$

where a denotes a linear form on the Lie algebra of G.

Proof: If K is given by (2.11) and λ by (2.12), definition (4.24) immediately gives, with the help of (4.22),

$$K_X(x)=\partial_X\phi(x)-X\mu(e).$$

Conversely, let us assume that K_X is of the form (4.35). From (4.32) and (4.23) we at first obtain

$$F(X,Y)=-a([X,Y]).$$

The exponent λ is therefore trivial according to Proposition 3.2 and Theorem 3.3, and we have

$$F_{\alpha\beta}=-C^\gamma_{\alpha\beta}a(X_\gamma)$$

and accordingly $\lambda\equiv 0\ (\mathrm{mod}\mu)$ with

$$\mu(\omega)=\nu(\omega)+\int_e^\omega a(X_\gamma)\zeta^\gamma,\quad \nu(e)=0,\quad d\nu(e)=0,$$

where the integration path is a straight line in canonical coordinates. From (3.4) we next deduce

$$\bar{\Phi}(\bar{X})_{(\omega)}=\bar{X}\mu(\omega)-\bar{X}\mu(e)=\bar{X}\mu(\omega)-a(X).$$

Now, (4.34) and (4.35) give, with the help of (4.22),

$$\psi_x(\bar{X})=\partial_X\phi\circ U_x-a(X)=-\bar{X}(\phi\circ U_x)-a(X),$$

and then

$$(\psi_x-\bar{\Phi})(\bar{X})=-\bar{X}(\phi\circ U_x)-\bar{X}\mu$$

or

$$\psi_x-\bar{\Phi}=-d(\phi\circ U_x)-d\mu.$$

Thus, the formula (4.33) finally gives

$$K(\omega,x)=\phi(x)-\phi(x^\omega)-\mu(\omega).$$

5. LAGRANGIANS

The equation we have to solve is Eq. (2.6) or, by using the gauge function K defined by (4.1),

$$L(x^\omega,\dot{x}^\omega)=L(x,\dot{x})+K\widehat{(\omega,x)}. \tag{5.1}$$

Geometrically, the pair $(x,\dot{x})$ is the set constituted by a point $x\in M$ and an element $\dot{x}$ of M_x, the tangent vector space to M at the point x [in other words, $(x,\dot{x})$ belongs to the tangent fiber space of M]. We denote by $J_\omega(x)$ the differential at the point x of the application $x\to x^\omega$,

ω being fixed, and by $dK(\omega, x)$ the differential of $x \to K(\omega, x)$. The formulas (2.4) and (2.2) then read

$$\dot{x}^{\omega} = J_{\omega}(x)(\dot{x}),$$
$$\dot{K}(\omega, x) = dK(\omega, x)(\dot{x}). \tag{5.2}$$

We note the relation

$$J_{\omega}(x)^{-1} = J_{\omega^{-1}}(x^{\omega}). \tag{5.3}$$

Let $l \in M_{x\omega}$; the invariance condition (1) takes the form

$$L(x^{\omega}, l) = L(x, J_{\omega}(x)^{-1}l) + dK(\omega, x)(J_{\omega}(x)^{-1}l). \tag{5.4}$$

We only consider the case of a connected transitive group G and take again the notations of Sec. 4A.

Proposition 5.1: For a connected transitive transformation group G and a given local gauge function K, the Lagrangians fulfilling the invariance condition are locally given by the formula

$$L(x, l) = f(J_{\alpha(x)^{-1}}(x)l) + dK(\alpha(x), x_0)(J_{\alpha(x)^{-1}}(x)l), \tag{5.5}$$

in which the function f is a homogeneous function of the first degree on M_{x_0}, a solution of the equation

$$f(j(\Omega)l_0) - f(l_0) = dK(\Omega, x_0)(l_0) \tag{5.6}$$

with $\Omega \in G_{x_0}$, $l_0 \in M_{x_0}$ and

$$j(\Omega) = J_{\Omega}(x_0), \quad \Omega \in G_{x_0}. \tag{5.7}$$

Proof: By writing (5.4) for $x = x_0$ and $\omega = \alpha(x)$ we obtain the following necessary condition for invariance, valid in some neighborhood of x_0,

$$L(x, l) = L(x_0, J_{\alpha(x)}(x_0)^{-1}l) + dK(\alpha(x), x_0)(J_{\alpha(x)}(x_0)^{-1}l). \tag{5.8}$$

Let f be the homogeneous function of the first degree on M_{x_0} which is defined by

$$f(l_0) = L(x_0, l_0). \tag{5.9}$$

With the help of (5.3), the formula (5.8) is immediately rewritten into the form (5.5). That relation defines L in terms of K and f. Conversely, L being given by (5.5), we have to express the full invariance condition (5.4). Introducing (5.5) into (5.4) we have at once

$$\begin{aligned} &L(x^{\omega}, l) - L(x, J_{\omega}(x)^{-1}l) - dK(\omega, x)(J_{\omega}(x)^{-1}l) \\ &= f(J_{\alpha(x^{\omega})^{-1}}(x^{\omega})l) - f(J_{(\omega\alpha(x))^{-1}}(x^{\omega})l) \\ &\quad + dK(\alpha(x^{\omega}), x_0)(J_{\alpha(x^{\omega})^{-1}}(x^{\omega})l) - dK(\alpha(x), x_0) \\ &\quad \times (J_{(\omega\alpha(x))^{-1}}(x^{\omega})l) - dK(\omega, x)(J_{\omega^{-1}}(x^{\omega})l). \end{aligned} \tag{5.10}$$

On the other hand, by differentiating the two members of (4.2) with respect to x and by applying the result just obtained to the vector $J_{\omega_2^{-1}}(x^{\omega_2})l$, $l \in M_{x\omega_2}$, we obtain

$$\begin{aligned} &dK(\omega_1, x^{\omega_2})(l) + dK(\omega_2, x)(J_{\omega_2^{-1}}(x^{\omega_2})l) \\ &= dK(\omega_1\omega_2, x)(J_{\omega_2^{-1}}(x^{\omega_2})l). \end{aligned} \tag{5.11}$$

By making the substitutions $x \to x_0$, $\omega_1 \to \omega$, $\omega_2 \to \alpha(x)$, $l \to J_{\omega^{-1}}(x^{\omega})l$, we find that the last two terms of (5.10) may be replaced by

$$-dK(\omega\alpha(x), x_0)(J_{(\omega\alpha(x))^{-1}}(x^{\omega})l).$$

The sum of that result and of the term in dK remaining in (5.10) may be transformed by a second application of (5.11) with the substitutions $x \to x_0$, $\omega_1 \to \alpha(x^{\omega})$, $\omega_2 \to \alpha(x^{\omega})^{-1}\omega\alpha(x)$, $l \to J_{\alpha(x\omega)^{-1}}(x^{\omega})l$, and we find for this sum the expression

$$-dK(\alpha(x^{\omega})^{-1}\omega\alpha(x), x_0)(J_{(\omega\alpha(x))^{-1}}(x^{\omega})l).$$

Now putting $l_0 = J_{(\omega\alpha(x))^{-1}}(x^{\omega})l \in M_{x_0}$, the right member of (5.10) reads

$$f(J_{\alpha(x\omega)^{-1}\omega\alpha(x)}(x_0)l_0) - f(l_0) - dK(\alpha(x^{\omega})^{-1}\omega\alpha(x), x_0)(l_0).$$

The invariance condition will be fulfilled as soon as this last expression identically vanishes. Since the product $\alpha(x^{\omega})^{-1}\omega\alpha(x)$ belongs to the little group G_{x_0}, by putting (5.7), we indeed obtain (5.6). Proposition 5.1 reduces the problem of solving (5.4) to that of solving (5.6). This last equation may be directly treated in any particular case. Here we restrict ourselves to transforming (5.6) into a system of partial differential equations. The correspondence $\Omega \to j(\Omega)$ defines a linear representation of G_{x_0} into M_{x_0}. The corresponding generators are given by

$$j_X = Xj(e) \tag{5.12}$$

in which X belongs to the Lie algebra of the little group G_{x_0}. The brackets are

$$[j_X, j_Y] = j_{[X,Y]}. \tag{5.13}$$

The velocity field D_X associated with this representation is then given by a formula similar to (4.20) and is

$$D_X(l_0) = -dj_{l_0}(X_e) = -j_X l_0, \tag{5.14}$$

where we have put $j_{l_0}(\Omega) = j(\Omega)l_0$ and where the last member represents the action of the linear operator j_X on the vector l_0. We also have, as in (4.22),

$$D_X f \circ j_{l_0} = -\bar{X}(f \circ j_{l_0}) \tag{5.15}$$

for any C^{∞} function f on M_{x_0}. We now have

Proposition 5.2: The solution of Eq. (5.6) is equivalent to the solution of

$$D_X f(l_0) = -dK_X(x_0)(l_0) \tag{5.16}$$

for any element X of the Lie algebra of G_{x_0}.

Proof: The equation (5.6) is of the general form

$$f(j(\Omega)l_0) = f(l_0) + k(\Omega) \cdot l_0 \tag{5.17}$$

in which $k(\Omega)$ is a linear form on M_{x_0} given here by

$$k(\Omega) = dK(\Omega, x_0). \tag{5.18}$$

We easily see that (5.17) can have a solution only if the following necessary condition holds:

$$k(\Omega_1\Omega_2) = k(\Omega_2) + \widetilde{j(\Omega_2)}k(\Omega_1), \tag{5.19}$$

where $\tilde{j}$ denotes the transpose of j. That condition is automatically satisfied in the present case as is shown by (5.11) written for $x = x_0$ and $\omega_1, \omega_2 \in G_{x_0}$, and by using (5.7). By applying the operator $\bar{X}_{\Omega_1}$ at the point $\Omega_1 = e$ to (5.19) we find

$$\bar{X}k(\Omega) = \widetilde{j(\Omega)}\bar{X}k(e). \tag{5.20}$$

Now, the application of $\bar{X}_e$ to Eq. (5.17) gives, with the help of (5.15),

$$-D_X f(l_0) = \bar{X}k(e)\cdot l_0 \tag{5.21}$$

Equation (5.21) is equivalent to (5.17). Indeed, with (5.15) and (5.20), we successively deduce from (5.21),

$$\bar{X}\cdot f(j(\Omega)l_0) = -D_X f(j(\Omega)l_0) = \bar{X}k(e)\cdot j(\Omega)l_0$$

$$= \widetilde{j(\Omega)}\bar{X}k(e)\cdot l_0 = \bar{X}k(\Omega)\cdot l_0,$$

which gives, by integration,

$$f(j(\Omega)l_0) = k(\Omega)\cdot l_0 + \varphi(l_0),$$

and then $\varphi(l_0) = f(l_0)$ by taking $\Omega = e$. It remains to be shown that (5.21) is identical to (5.16); in fact, we have

$$\bar{X}k(e) = \bar{X}_e(dK(\Omega, x_0)) = d(XK(e,x))\big|_{x=x_0} = dK_X(x_0).$$

Finally, let us give the form of (5.16) in a coordinate system. Denoting a basis of the Lie algebra of G_{x_0} by $\{X_\alpha\}$ we find

$$(j_\alpha)^m_n l_0^n \frac{\partial}{\partial l_0^m} f(l_0) = k_{\alpha n} l_0^n$$

with

$$k_{\alpha n} = \frac{\partial}{\partial x^n} X_\alpha K(\omega, x)\Big|_{\omega=e,\ x=x_0}.$$

6. THE ALGEBRA OF CONSERVED QUANTITIES

As we know, with any Lagrangian theory possessing an invariance group of the type just described, corresponds, by Noether's theorem, a set of conserved quantities.[1] We will determine the Poisson brackets of these quantities, and show that these brackets define an algebra which is an extension of the Lie algebra of the group, the one corresponding to the exponent λ.

The manifold M is the product $\mathbb{R}\times V$ where V is the configuration space of the system. For a given local chart on V we denote the coordinates of $x = (t, q) \in \mathbb{R}\times V$ by $(x^\mu) = (x^0, (x^k))$. From the homogeneity property of L it is easy to see that the Lagrange equation for x^0 is not independent of the others. The same property holds for the conjugate momenta

$$\pi_\mu = \frac{\partial L}{\partial \dot{x}^\mu}(x, \dot{x}). \tag{6.1}$$

Indeed, the function L is connected to the Lagrange function usually considered (from which the action integral is $\int dt\,\hat{L}$) by the formula

$$L(x, \dot{x}) = \dot{t}\hat{L}(t, q, \dot{q}/\dot{t}). \tag{6.2}$$

This implies the equalities

$$\pi_k = p_k(t, q, \dot{q}/\dot{t}), \tag{6.3}$$

where p_k are the conjugate momenta associated with $\hat{L}$, and the constraint

$$\pi_0 = -H(t, q, \pi), \pi = (\pi_k), \tag{6.4}$$

where H denotes the Hamiltonian. On the other hand, the value $\dot{t} = 0$ must be excluded since L generally has a singularity at this point. With this restriction, the correspondence $(x, \dot{x}) \to (x, \pi_0, \pi)$ maps the tangent fiber space of M (velocity phase space) into the submanifold S of the cotangent fiber space of M (momentum phase space) which is defined by Eq. (6.4). According to the usual hypotheses,[7] this mapping is, for any $\dot{t} \neq 0$ fixed, a local diffeomorphism into S. The momentum phase space is naturally endowed with the symplectic form $\Omega = dx^\mu \wedge d\pi_\mu$ for which the Poisson bracket of two functions of (x^μ) and (π_μ) is[7]

$$\{f, g\}_\Omega = \frac{\partial f}{\partial x^\mu}\frac{\partial g}{\partial \pi_\mu} - \frac{\partial f}{\partial \pi_\mu}\frac{\partial g}{\partial x^\mu}. \tag{6.5}$$

The form $\bar{\Omega}$ induced by Ω on the section S_t of the submanifold S by the hyperplane $x^0 = t$ is $\bar{\Omega} = dx^k \wedge d\pi_k$ and the associated Poisson bracket is the usual one,

$$\{f, g\}_{\bar{\Omega}} = \frac{\partial f}{\partial x^k}\frac{\partial g}{\partial \pi_k} - \frac{\partial f}{\partial \pi_k}\frac{\partial g}{\partial x^k}, \tag{6.6}$$

where f and g are functions on S expressed with the help of the coordinates (t, x^k, π_k). Now if $\bar{f}$ denotes the restriction on S of a function f defined on the momentum phase space, an easy calculation gives the following relation between the two kinds of brackets:

$$\{\bar{f}, \bar{g}\}_{\bar{\Omega}} = \overline{\{f, g\}_\Omega} + \frac{\overline{\partial f}}{\partial \pi_0}\left(\frac{\partial \bar{g}}{\partial t} + \{\bar{g}, H\}_{\bar{\Omega}}\right) - \frac{\overline{\partial g}}{\partial \pi_0}\left(\frac{\partial \bar{f}}{\partial t} + \{\bar{f}, H\}_{\bar{\Omega}}\right). \tag{6.7}$$

The action integral is defined by

$$A^{s_2}_{s_1}[x] = \int_{s_1}^{s_2} ds\, L(x(s), x'(s)). \tag{6.8}$$

The invariance condition (5.1) then reads

$$A^{s_2}_{s_1}[x^\omega] - A^{s_2}_{s_1}[x] = K(\omega, x(s_2)) - K(\omega, x(s_1)). \tag{6.9}$$

Taking $\omega = \delta\omega$ to be infinitesimal, a classical calculation gives

$$\delta A^{s_2}_{s_1}[x] = \big|K(\delta\omega, x(s))\big|^{s_2}_{s_1} \tag{6.10}$$

with

$$\delta A^{s_2}_{s_1}[x] = \big|\pi_\mu \delta x^\mu\big|^{s_2}_{s_1} + \int_{s_1}^{s_2} ds\left[\frac{\partial L}{\partial x^\mu} - \frac{d}{ds}\left(\frac{\partial L}{\partial \dot{x}^\mu}\right)\right]\delta x^\mu(s) \tag{6.11}$$

in which π_μ is defined by (6.1). When $x(s)$ is a solution of the equations of motion, (6.10) reduces to

$$\big|\pi_\mu \delta x^\mu - K(\delta\omega, x)\big|^{s_2}_{s_1} = 0.$$

The quantity

$$\delta Q = \pi_\mu \delta x^\mu - K(\delta\omega, x) \tag{6.12}$$

is therefore a constant of motion. Alternatively, if we put $\delta\omega = \exp(tX)$, where X belongs to the Lie algebra of the group G, we obtain the finite quantities [see (4.21) and (4.24)]

$$Q_X = -\pi_\mu D_X x^\mu - K_X(x). \tag{6.13}$$

As it stands, this last formula defines Q_X as a function on the whole momentum phase space, the corresponding conserved quantity $\bar{Q}_X$ being obtained by restricting Q_X on S. Due to the form of the equations of motion written with Poisson brackets, the conservation property implies that the function $(\partial/\partial t)\bar{Q}_X + \{Q_X, H\}_{\bar{\Omega}}$ identically vanishes. Thus, the Poisson brackets of the $\bar{Q}_X$'s are, by (6.7), directly given by those of the Q_X's. From (6.13) we have

$$\{Q_X, Q_Y\}_\Omega = \pi_\mu\{D_X x^\mu, \pi_\nu\}_\Omega D_Y x^\nu + \pi_\nu\{\pi_\mu, D_Y x^\nu\}_\Omega D_X x^\mu + \{\pi_\mu, K_Y(x)_\Omega D_X x^\mu + \{K_X(x), \pi_\nu\}_\Omega D_Y x^\nu.$$

Definition (6.5) implies the following relation:

$$\{\pi_\mu, f(x)\}_\Omega \partial_X x^\mu = -(\partial_X x^\mu)\frac{\partial f(x)}{\partial x^\mu} = -\partial_X f(x),$$

and then

$$\{Q_X, Q_Y\}_\Omega = \pi_\mu \partial_Y(\partial_X x^\mu) - \pi_\nu \partial_X(\partial_Y x^\nu) - \partial_X K_Y(x) + \partial_Y K_X(x).$$

According to (4.23) and (4.26), this finally leads to the formula

$$\{Q_X, Q_Y\}_\Omega = Q_{[X,Y]} + F(X, Y). \tag{6.14}$$

If X and Y are chosen among the basis fields X_α, setting $Q_\alpha = Q_{X_\alpha}$, we have

$$\{Q_\alpha, Q_\beta\} = C^\gamma_{\alpha\beta} Q_\gamma + F_{\alpha\beta}. \tag{6.15}$$

This shows that the Lie algebra generated by the conserved quantities is identical to the extension of the Lie algebra of the invariance group which is defined by the coefficients $F_{\alpha\beta}$. On the other hand, this Lie algebra is that of the group extension defined by the exponent χ.[5]

Remark: From the action of the group G on M, two different canonical transformations may be constructed on the momentum phase space, corresponding to different transformations of the Lagrangian.

1. Transformation $(x, L) \to (x^\omega, L_\omega)$

The transformation of the conjugate momenta is defined by [see (2.3) and (5.2)]

$$\pi^\omega_\mu = \frac{\partial}{\partial(\dot{x}^\omega)^\mu} L_\omega(x^\omega, \dot{x}^\omega) = (J_\omega(x)^{-1})^\nu_\mu \pi_\nu. \tag{6.16}$$

For an infinitesimal transformation we have

$$\delta\pi_\mu(x, \pi_0, \pi) = -\pi_\nu \frac{\partial}{\partial x^\mu}\delta x^\nu(x). \tag{6.17}$$

The Hamiltonian is not invariant; its transformed is

$$H_{\delta\omega}(t, q, \pi) = \frac{1}{\dot{t}}[\pi_k \dot{x}^k - L_{\delta\omega}(x, \dot{x})] \tag{6.18}$$

in which the variables $\dot{x}^k/\dot{t}$ must be replaced by the momenta through the relations [see (2.3)],

$$\pi_k = \frac{\partial}{\partial \dot{x}^k} L_{\delta\omega}(x, \dot{x}) = \frac{\partial}{\partial \dot{x}^k}\left[L(x, \dot{x}) - \frac{\partial L}{\partial x^\mu}\delta x^\mu - \frac{\partial L}{\partial \dot{x}^\mu}\delta \dot{x}^\mu\right].$$

Taking into account the invariance condition (5.1) written in infinitesimal form,

$$\frac{\partial L}{\partial x^\mu}\delta x^\mu + \frac{\partial L}{\partial \dot{x}^\mu}\delta \dot{x}^\mu = \dot{\overline{\delta K}}, \quad \delta K(x) = K(\delta\omega, x),$$

we get

$$\pi_k = \frac{\partial}{\partial \dot{x}^k}\left[L(x, \dot{x}) - \dot{\overline{\delta K}}\right]$$

or

$$\pi_k + \frac{\partial}{\partial \dot{x}^k}\delta K = \frac{\partial}{\partial \dot{x}^k} L(x, \dot{x}). \tag{6.19}$$

On the other hand, (6.18) may be written

$$H_{\delta\omega}(x, \pi) = \frac{1}{\dot{t}}\left[\left(\pi_k + \frac{\partial}{\partial \dot{x}^k}\delta K\right)\dot{x}^k - L(x, \dot{x})\right] + \frac{\partial}{\partial t}\delta K,$$

from which we deduce, with the help of (6.19),

$$H_{\delta\omega}(x, \pi) = H(x, \pi + \partial(\delta K)) + \frac{\partial}{\partial t}\delta K. \tag{6.20}$$

The infinitesimal generating function is

$$\delta\mathcal{G} = \pi_\mu \delta x^\mu(x) \tag{6.21}$$

and we have

$$\{x^\mu, \delta\mathcal{G}\}_\Omega = \delta x^\mu(x), \quad \{\pi_\mu, \delta\mathcal{G}\}_\Omega = \delta\pi_\mu(x, \pi_0, \pi). \tag{6.22}$$

For the motions of the system, the restriction $\overline{\delta\mathcal{G}} = \pi_k \delta x^k - H\delta t$ generates the "local variations" on the shell S (Schwinger principle). From (6.7) we deduce

$$\{x^k, \overline{\delta\mathcal{G}}\}_\Omega = \delta x^k - \delta t\frac{dx^k}{dt} \equiv \bar{\delta} x^k, \tag{6.23}$$

$$\{\pi_k, \overline{\delta\mathcal{G}}\}_\Omega = \overline{\delta\pi_k} - \delta t\frac{d\pi_k}{dt} \equiv \bar{\delta}\pi_k.$$

2. Transformation $(x,L) \to (x^\omega, (L_\omega)_{\Lambda\omega}) = (x^\omega, L)$

Formula (6.16) is replaced by[4]

$$\pi^\omega_\mu = \frac{\partial}{\partial(\dot{x}^\omega)^\mu} L(x^\omega, \dot{x}^\omega) = \pi_\mu(x^\omega, \dot{x}^\omega) = (J_\omega(x)^{-1})^\nu_\mu\left[\pi_\nu + \frac{\partial}{\partial x^\nu}K(\omega, x)\right]. \tag{6.24}$$

Infinitesimally we have

$$\delta\pi_\mu = \delta\left(\frac{\partial L}{\partial \dot{x}^\mu}\right) = -\pi_\nu \frac{\partial}{\partial x^\mu}\delta x^\nu + \frac{\partial}{\partial x^\mu}\delta K. \tag{6.25}$$

The Hamiltonian is now invariant so that the shell S is conserved. Finally the formulas (6.22) and (6.23) remain valid for the new variations with the following expression of the generating function:

$$\delta\mathcal{G} = \pi_\mu \delta x^\mu(x) - \delta K(x) = \delta Q. \tag{6.26}$$

7. EXAMPLES

A. Translation group

Let us consider $M = G = \mathbb{R}^n$ and the group action $x^\omega = x + \omega$. The coefficients $F_{\alpha\beta}$ are only restricted by the antisymmetry condition $F_{\alpha\beta} = -F_{\beta\alpha}$. The corresponding exponents are directly defined by the formula (3.26),

$$\xi(\omega_1, \omega_2) = -\tfrac{1}{2}\int_{S(\omega_1, \omega_2)} F_{\alpha\beta}\, d\omega^\alpha \wedge d\omega^\beta = -\tfrac{1}{2}F_{\alpha\beta}\omega_1^\alpha \omega_2^\beta. \tag{7.1}$$

The gauge functions are calculated by (4.7) in which we choose $x_0 = 0$ and $\alpha(x) = x$, and we find

$$K(\omega, x) = -\tfrac{1}{2}F_{\alpha\beta}\omega^\alpha x^\beta. \tag{7.2}$$

(Since $G_{x_0} = \{0\}$ we have $\chi = 0$.) Then, (5.5) and (5.6) furnish the general form of the Lagrangian

$$L(x, \dot{x}) = f(\dot{x}) - \tfrac{1}{2}F_{\alpha\beta}x^\alpha \dot{x}^\beta, \tag{7.3}$$

where f is any homogeneous function of the first degree. The conserved quantities are

$$Q_\alpha = -\pi_\beta \partial_\alpha x^\beta - K_\alpha(x) = \pi_\alpha + \tfrac{1}{2}F_{\alpha\beta}x^\beta \tag{7.4}$$

and the relations (6.15) are indeed verified. If $n = 4$, by putting $e\mathbf{E} = (F_{0i})$ and $e\mathbf{B} = (-\frac{1}{2}\epsilon^{ijk}F_{jk})$ for $1 \le i, j, k \le 3$, we find [see (6.2)]

$$\hat{L}(t, q, \dot{q}) = \hat{f}(\dot{q}) + \tfrac{1}{2}e(\mathbf{B}\wedge\mathbf{q})\cdot\dot{\mathbf{q}} + \tfrac{1}{2}e[\mathbf{E}\cdot\mathbf{q} - t\mathbf{E}\cdot\dot{\mathbf{q}}]. \tag{7.5}$$

The coupling terms describe the interaction of a charged particle with a constant and uniform electromagnetic field. The usual form of the interaction with the electric field is recovered by adding to the Lagrangian the "divergence term"

$$\tfrac{1}{2}et\mathbf{E}\cdot\mathbf{q}=\tfrac{1}{2}e\mathbf{E}\cdot\mathbf{q}+\tfrac{1}{2}et\mathbf{E}\cdot\dot{\mathbf{q}}.$$

The invariance under translation of this interaction is evidently well known,[3,4] but the above calculation characterizes it as being the only one having that property.

B. Forced harmonic oscillator

Let us consider a system with one degree of freedom ($M=\mathbb{R}^2$), and with G the two-dimensional translation group with elements $a=(a^1,a^2)$ acting by

$$\begin{aligned} &t\rightharpoonup t,\\ &q\rightharpoonup q+a^1\cos\omega t+a^2\sin\omega t, \end{aligned}\tag{7.6}$$

where ω is given. The exponents on G are given by (7.1), in which we have set $F_{12}=f$,

$$\xi(a,a')=-\tfrac{1}{2}f(a^1a'^2-a^2a'^1)=-\tfrac{1}{2}f\epsilon_{\alpha\beta}a^\alpha a'^\beta.\tag{7.7}$$

From (3.4) and the expression of the vector fields $X_\alpha=\bar{X}_\alpha=\partial/\partial a^\alpha$, we deduce

$$\bar{\Phi}(\bar{X}_\alpha)_{(a)}=-\tfrac{1}{2}f\epsilon_{\alpha\beta}a^\beta$$

and then

$$\bar{\Phi}(a)=\tfrac{1}{2}f\epsilon_{\alpha\beta}a^\alpha da^\beta.\tag{7.8}$$

On the other hand, we have

$$D_1=-\cos\omega t\frac{\partial}{\partial q},\quad D_2=-\sin\omega t\frac{\partial}{\partial q}.$$

The Eq. (4.32), obeyed by the functions $K_\alpha(x)=X_\alpha K(e,x)$, where K is a gauge function associated with ξ, is here

$$D_1K_2-D_2K_1=-f.$$

Its solution is

$$K_1\sin\omega t-K_2\cos\omega t=-fq+\varphi(t).$$

According to Proposition 4.5 we easily see that, up to an equivalence leaving the corresponding exponent unchanged, we may choose

$$\begin{aligned} K_1(t,q)&=(-fq+\varphi(t))\sin\omega t,\\ K_2(t,q)&=-(-fq+\varphi(t))\cos\omega t. \end{aligned}\tag{7.9}$$

The differential 1-form ψ_x introduced in Proposition 4.4 is then given by

$$\psi_x(a)=\psi_x(X_\alpha)_{(a)}\,da^\alpha=K_\alpha(x^a)\,da^\alpha$$

or

$$\begin{aligned} \psi_x(a)=&(-fq+\varphi(t))(\sin\omega t\,da^1-\cos\omega t\,da^2)\\ &-f(\sin^2\omega t\,a^2da^1-\cos^2\omega t\,a^1da^2\\ &+\sin\omega t\cos\omega t(a^1da^1-a^2da^2)). \end{aligned}$$

Formula (4.33) gives the associated gauge function

$$\begin{aligned} K(a,x)=&(-fq+\varphi(t))(a^1\sin\omega t-a^2\cos\omega t)\\ &-\tfrac{1}{4}f[((a^1)^2-(a^2)^2)\sin2\omega t-2a^1a^2\cos2\omega t]. \end{aligned}\tag{7.10}$$

The Lagrangian must satisfy the relation (5.4) which may be written here as

$$\begin{aligned} &L(t,q+a^1\cos\omega t+a^2\sin\omega t,\dot{t},\dot{q}+\omega(-a^1\sin\omega t+a^2\cos\omega t)\dot{t})\\ &=L(t,q,\dot{t},\dot{q})+\left(-f\dot{q}+\dot{t}\frac{d\varphi}{dt}\right)(a^1\sin\omega t-a^2\cos\omega t)\\ &\quad+(-fq+\varphi)(a^1\cos\omega t+a^2\sin\omega t)\omega\dot{t}\\ &\quad-\tfrac{1}{2}f((a^1)^2-(a^2)^2)\omega\dot{t}\cos2\omega t-fa^1a^2\omega\dot{t}\sin2\omega t. \end{aligned}$$

The solution is easily found by calculating the derivatives with respect to a^1 and a^2 at $a=0$, that furnishes the derivatives $\partial L/\partial q$ and $\partial L/\partial\dot{q}$, and we obtain

$$\begin{aligned} &L(t,q,\dot{t},\dot{q})\\ &=\dot{t}l(t)+\frac{1}{2}\frac{f}{\omega}\left(\frac{\dot{q}^2}{\dot{t}}-\omega^2q^2\dot{t}\right)\\ &\quad+\frac{1}{\omega}\left(-\dot{q}\frac{d\varphi}{dt}+\omega^2q\varphi\dot{t}\right). \end{aligned}\tag{7.11}$$

Adding the divergence term

$$\frac{1}{\omega}q\frac{d\varphi}{dt}-\int^t\dot{l}(\tau)\,d\tau=\frac{1}{\omega}\left[\dot{q}\frac{d\varphi}{dt}+q\frac{d^2\varphi}{dt^2}\dot{t}\right]-\dot{t}\,l(t)$$

and putting

$$F(t)=\frac{1}{\omega}\left[\frac{d^2\varphi(t)}{dt^2}+\omega^2\varphi(t)\right]\tag{7.12}$$

and

$$f=m\omega\tag{7.13}$$

we find for the usual Lagrange function, up to an equivalence,

$$\hat{L}(t,q,\dot{q})=\tfrac{1}{2}m\dot{q}^2-\tfrac{1}{2}m\omega^2q^2+qF(t).\tag{7.14}$$

This is the Lagrangian of a harmonic oscillator submitted to the driving force $F(t)$. With this choice of the Lagrangian the gauge function becomes

$$\begin{aligned} K(a,x)=&-m\omega q(a^1\sin\omega t-a^2\cos\omega t)\\ &-\tfrac{1}{4}m\omega[((a^1)^2-(a^2)^2)\sin2\omega t-2a^1a^2\cos2\omega t]\\ &+\int^t d\tau F(\tau)[a^1\cos\omega\tau+a^2\sin\omega\tau], \end{aligned}\tag{7.15}$$

the associated exponent being unchanged. The conserved quantities are then

$$\begin{aligned} Q_1&=p\cos\omega t+m\omega q\sin\omega t-\int^t d\tau F(\tau)\cos\omega\tau,\\ Q_2&=p\sin\omega t-m\omega q\cos\omega t-\int^t d\tau F(\tau)\sin\omega\tau, \end{aligned}\tag{7.16}$$

and, according to the general result (6.15), their bracket is

$$\{Q_1,Q_2\}=f.\tag{7.17}$$

By eliminating p or q between the two Eqs. (7.16) we immediately obtain the general solution of the Hamilton equations of motion

$$\begin{aligned} fq(t)&=Q_1\sin\omega t-Q_2\cos\omega t+\int^t d\tau F(\tau)\sin\omega(t-\tau),\\ p(t)&=Q_1\cos\omega t+Q_2\sin\omega t+\int^t d\tau F(\tau)\cos\omega(t-\tau). \end{aligned}\tag{7.18}$$

If $F(t)$ acts during a finite time interval, and if the lower bound of the integrals in (7.18) is chosen equal to

$\mp\infty$, we see that the corresponding constants of motion Q_1 and Q_2 are the coefficients characterizing the asymptotic behavior of the canonical variables, so that (7.17) simply expresses the fundamental Poisson brackets between these variables. More generally, there is a connection between the gauge functions and the Green functions associated with the equation of motion.

These results may be generalized by replacing the functions cos and sin in (7.6) by h_1 and h_2, two independent real solutions of the second order equation

$$h''(t)+u(t)h'(t)+v(t)h(t)=0, \quad u \text{ and } v \text{ real.} \tag{7.19}$$

The Lagrange function replacing (7.14) is

$$\hat{L}(t,q,\dot{q})=\exp\Big[\int_0^t d\tau\, u(\tau)\Big]\Big[\tfrac{1}{2}m\dot{q}^2-\tfrac{1}{2}mv(t)q^2+qF(t)\Big] \tag{7.20}$$

and the associated gauge function is

$$\begin{aligned}K(a,x)=m\exp\Big[\int_0^t d\tau\, u(\tau)\Big]\Big[a^1h_1'(t)+a^2h_2'(t)\Big]\\ \times\Big[q+\tfrac{1}{2}(a^1h_1(t)+a^2h_2(t))\Big]\\ +\int^t d\tau\exp\Big(\int_0^\tau d\tau'\,u(\tau')\Big)F(\tau)\\ \times\Big[a^1h_1(\tau)+a^2h_2(\tau)\Big].\end{aligned} \tag{7.21}$$

The constant f is here $mW(0)$ where W is the Wronskian of h_1 and h_2,

$$W(t)=h_1(t)h_2'(t)-h_2(t)h_1'(t)=W(0)\exp\Big(-\int_0^t d\tau\,u(\tau)\Big) \tag{7.22}$$

and the conserved quantities $Q_i(i=1,2)$ are given by

$$Q_i=ph_i(t)-mqU(t)h_i'(t)-\int^t d\tau F(\tau)\ U(\tau)h_i(\tau), \tag{7.23}$$

where we have put

$$U(t)=\exp\int_0^t d\tau\,u(\tau). \tag{7.24}$$

As for (7.16), the Eqs. (7.23) allow us to obtain the general solution of the equation of motion. Finally, let us note that a simple manipulation of that equation directly furnishes the expression (7.23): By multiplying the equation of motion

$$m[q''+uq'+vq]-F=0$$

by any solution h of (7.19) we obtain

$$m\left[\frac{d}{dt}(q'h-qh')+u(q'h-qh')\right]-Fh=0$$

or

$$mU^{-1}\frac{d}{dt}\left[(q'h-qh')U\right]-Fh=0.$$

Multiplying by U we find $(d/dt)Q_h=0$ with

$$Q_h=m(q'h-qh')U-\int^t d\tau F(\tau)U(\tau)h(\tau),$$

which is identical with (7.23) if we introduce in that formula the expression of p derived from (7.20).

C. Rotation group

As a last example we take $G=SO_3$ acting on $M=\mathbb{R}\times\mathbb{R}^3$. The rotations will be parametrized by the components of their rotation vector ω. The form of the structure constants, $C^\gamma_{\alpha\beta}=\epsilon_{\alpha\beta\gamma}$, immediately implies that the exponents on G are trivial, so that it will be sufficient to consider the case of a vanishing exponent. The first equation we have to solve is Eq. (4.32) which here reads

$$\partial_\alpha K_\beta-\partial_\beta K_\alpha-\epsilon_{\alpha\beta\gamma}K_\gamma=0, \quad 1\leq\alpha,\beta,\gamma\leq 3 \tag{7.25}$$

with

$$\partial_\alpha=-\epsilon_{\alpha\beta\gamma}x^\beta\frac{\partial}{\partial x^\gamma}. \tag{7.26}$$

In vector notation this equation takes the form

$$\mathrm{grad}(\mathbf{x}\cdot\mathbf{K})=\mathbf{x}\,\mathrm{div}\mathbf{K}. \tag{7.27}$$

The integrability condition gives $\partial_\alpha(\mathrm{div}\mathbf{K})=0$, which implies that $\mathrm{div}\mathbf{K}$ is a function of $\mathbf{x}^2$ only (except for an arbitrary dependence on the time t). Equation (7.27) is then integrated in the form

$$\begin{aligned}&\mathrm{div}\mathbf{K}=k(\mathbf{x}^2,t),\\ &\mathbf{x}\cdot\mathbf{K}=C+\tfrac{1}{2}\int_{\mathbf{x}_0^2}^{\mathbf{x}^2}ds\,k(s,t)\equiv K(\mathbf{x}^2,t).\end{aligned} \tag{7.28}$$

Introducing the longitudinal and transverse parts in the $\mathbf{x}$ space of K, the second equation in (7.28) gives

$$\mathbf{K}_{\parallel}=\frac{\mathbf{x}}{\mathbf{x}^2}K(\mathbf{x}^2,t), \tag{7.29}$$

while the first becomes

$$\mathrm{div}\mathbf{K}_\perp=-K(\mathbf{x}^2,t)/\mathbf{x}^2. \tag{7.30}$$

To solve that equation in the neighborhood of the given point $\mathbf{x}_0\neq 0$ let us introduce a direct orthonormal basis $(\mathbf{e}_1,\mathbf{e}_2,\mathbf{e}_3)$ of $\mathbb{R}^3$ such that $\mathbf{e}_1=\mathbf{x}_0/|\mathbf{x}_0|$. The transverse part of $\mathbf{K}$ may be written

$$\mathbf{K}_\perp(\mathbf{x},t)=(\mathbf{e}_2\wedge\mathbf{x})f(\mathbf{x},t)+(\mathbf{e}_3\wedge\mathbf{x})g(\mathbf{x},t). \tag{7.31}$$

Expressing (7.30) with the help of the coordinates $r=|\mathbf{x}|$, $y=\mathbf{x}\cdot\mathbf{e}_2$, $z=\mathbf{x}\cdot\mathbf{e}_3$, we easily find

$$\frac{\partial g}{\partial y}-\frac{\partial f}{\partial z}=\frac{-K(r^2,t)}{r^2(r^2-y^2-z^2)^{1/2}}.$$

This equation allows us to express g in terms of f and K. We find that the only nontrivial term is the one coming from K, thereby giving, up to trivial terms (that is terms of the form $\mathbf{x}\wedge\mathrm{grad}\Phi$),

$$\mathbf{K}_\perp=-\frac{K(\mathbf{x}^2,t)}{\mathbf{x}^2}\varphi(\mathbf{x})\,\mathbf{e}_3\wedge\mathbf{x},$$

where $\varphi(\mathbf{x})$ is the usual polar angle of $\mathbf{x}$ in the plane $(\mathbf{e}_1,\mathbf{e}_2)$. A simpler expression may be obtained by adding the trivial function

$$\begin{aligned}&-\mathbf{x}\wedge\mathrm{grad}\left[\frac{K(\mathbf{x}^2,t)}{\mathbf{x}^2}z\varphi(\mathbf{x})\right]\\ &\quad=\frac{K(\mathbf{x}^2,t)}{\mathbf{x}^2}\left[\varphi(\mathbf{x})\mathbf{e}_3\wedge\mathbf{x}+\frac{z^2\mathbf{x}-r^2z\mathbf{e}_3}{x^2+y^2}\right],\end{aligned}$$

which finally gives, by combining it with (7.29),

$$\begin{aligned}\mathbf{K}(\mathbf{x},t)&=K(r^2,t)\frac{x\mathbf{e}_1+y\mathbf{e}_2}{x^2+y^2}\\ &=K(r^2,t)\frac{\mathbf{x}-(\mathbf{e}_3\cdot\mathbf{x})\mathbf{e}_3}{\mathbf{x}^2-(\mathbf{e}_3\cdot\mathbf{x})^2}.\end{aligned} \tag{7.32}$$

The corresponding gauge function is now given by the formula (4.33), namely

$$K(\omega, \mathbf{x}, t) = \int_0^{\omega} K_\alpha(R_\omega, (\mathbf{x}), t)\bar{\xi}^\alpha(\omega') \tag{7.33}$$

in which R_ω is the rotation defined by the vector ω, and $\bar{\xi}^\alpha$ are the right invariant 1-forms on G. These forms are given by the formulas

$$\bar{\xi}^\alpha(\omega) = \bar{\lambda}^\alpha_\beta(\omega)\, d\omega^\beta,$$

$$\bar{\lambda}^\alpha_\beta(\omega) = \frac{\omega^\alpha\omega^\beta}{\omega^2} + \frac{\sin\omega}{\omega}\left(\delta^{\alpha\beta} - \frac{\omega^\alpha\omega^\beta}{\omega^2}\right) - \frac{1-\cos\omega}{\omega^2}\epsilon_{\alpha\beta\gamma}\omega^\gamma. \tag{7.34}$$

The integral (7.33) calculated along the integration path $s \to R_{s\omega}(\mathbf{x}), 0 \leq s \leq 1$, becomes

$$K(\omega, \mathbf{x}, t) = K(r^2, t)\int_0^1 ds\, \frac{\omega\cdot\mathbf{x} - (\omega\cdot\mathbf{e}_3)(\mathbf{e}_3\cdot R_{s\omega}(\mathbf{x}))}{\mathbf{x}^2 - (\mathbf{e}_3\cdot R_{s\omega}(\mathbf{x}))^2}. \tag{7.35}$$

The integral may be explicitly calculated[9] by introducing the expression $R_{s\omega}(\mathbf{x})$ in terms of ω and $\mathbf{x}$, but the final result which is a sum of four functions arctg with rather complicated arguments will be not given here. The formula (7.35) then defines a family of gauge functions depending on an arbitrary function $K(r^2, t)$ and on a given direction $\mathbf{e}_3$. As we shall see later, this last dependence will disappear at the level of the equations of motion.

Let us now consider the determination of the Lagrange function. Although the action of SO_3 on $\mathbb{R}^3$ is not transitive, the results of Proposition 5.1 are readily adapted and furnishes the expression

$$L(\mathbf{x}, t, \dot{\mathbf{x}}, \dot{t}) = f(|\mathbf{x}|, t, R^{-1}_{\alpha(\mathbf{x})}(\dot{\mathbf{x}}), \dot{t}) + dK(\alpha(\mathbf{x}), |\mathbf{x}|\mathbf{e}_1, t)(R^{-1}_{\alpha(\mathbf{x})}(\dot{\mathbf{x}}), \dot{t}) \tag{7.36}$$

in which f is a solution of the equation

$$f(|\mathbf{x}|, t, R_\Omega(\dot{\mathbf{x}}), \dot{t}) - f(|\mathbf{x}|, t, \dot{\mathbf{x}}, \dot{t}) = dK(\Omega, |\mathbf{x}|\mathbf{e}_1, t)(\dot{\mathbf{x}}, \dot{t}). \tag{7.37}$$

In these equations the reference point x_0 of Proposition 5.1 is replaced by a set of vectors playing the same role in each orbit, namely the vectors $|\mathbf{x}|\mathbf{e}_1$, the rotation vector $\alpha(\mathbf{x})$ is taken equal to

$$\alpha(\mathbf{x}) = \frac{\mathbf{e}_1\wedge\mathbf{x}}{|\mathbf{e}_1\wedge\mathbf{x}|}\arccos x/r, \tag{7.38}$$

and Ω is of the form $\Omega = \Omega\mathbf{e}_1$. The right member of (7.37) is easily calculated from (7.35) and we find

$$dK(\Omega\mathbf{e}_1, |\mathbf{x}|\mathbf{e}_1, t) = \Omega\left[\frac{\partial}{\partial t}\left(\frac{K}{r}\right)dt + \frac{\partial}{\partial r}\left(\frac{K}{r}\right)dx_1\right].$$

Equation (7.37) is then reduced to an equation of the form

$$g(l_1, ze^{i\Omega}) - g(l_1, z) = \Omega[A + Bl_1],$$

with $\mathbf{l} = \dot{\mathbf{x}}/\dot{t}$ and $z = l_2 + il_3$. The solution is

$$g(\mathbf{l}) = g_0(l_1, l_2^2 + l_3^2) + (A + Bl_1)\,\text{arc}tg\,\frac{l_3}{l_2}.$$

From that we deduce the solution of (7.37),

$$f(|\mathbf{x}|, t, \dot{\mathbf{x}}, \dot{t}) = \dot{t}\left[f_0(|\mathbf{x}|, t, l_1, l_2^2 + l_3^2) + \left(\frac{\partial}{\partial t}\left(\frac{K}{r}\right) + \frac{\partial}{\partial r}\left(\frac{K}{r}\right)l_1\right)\text{arc}tg\,\frac{l_3}{l_2}\right]. \tag{7.39}$$

The contribution to the Lagrange function (7.36) of the first term in the right-hand side of (7.39) is the most general form-invariant Lagrangian; this is due to the relations

$$(R^{-1}_{\alpha(\mathbf{x})}(\mathbf{l}))_1 = \frac{\mathbf{x}\cdot\dot{\mathbf{x}}}{r\dot{t}},$$

$$\sum_{i=2}^{3}(R^{-1}_{\alpha(\mathbf{x})}(\mathbf{l}))_i^2 = \frac{\dot{\mathbf{x}}^2}{\dot{t}^2} - \frac{(\mathbf{x}\cdot\dot{\mathbf{x}})^2}{r^2\dot{t}^2}.$$

Collecting the other terms we obtain

$$L = L_{\text{inv}} + \frac{\hat{K}}{r}\text{arc}tg\,\frac{R_{\alpha(\mathbf{x})}(\mathbf{e}_3)\cdot\dot{\mathbf{x}}}{R_{\alpha(\mathbf{x})}(\mathbf{e}_2)\cdot\dot{\mathbf{x}}} + dK(\alpha(\mathbf{x}), |\mathbf{x}|\mathbf{e}_1, t)(R^{-1}_{\alpha(\mathbf{x})}(\dot{\mathbf{x}}), \dot{t}). \tag{7.40}$$

To cast this expression into an explicit formula we have to introduce expression (7.38) for $\alpha(\mathbf{x})$ and explicitly calculate the differential dK. The calculations are tedious and finally give

$$L = L_{\text{inv}} - \frac{K}{r}\cdot\frac{z}{r}\frac{x\dot{y} - y\dot{x}}{x^2 + y^2} - \frac{\hat{K}}{r}\text{arc}tg\,\frac{x(z\dot{x} - x\dot{z}) - y(y\dot{z} - z\dot{y})}{r(x\dot{y} - y\dot{x})}. \tag{7.41}$$

This Lagrangian is in general a multiform function and depends on a particular direction of space. These two features are absent from the equations of motion deduced from (7.41). If, for example, we choose for L_{inv} the usual nonrelativistic kinetic energy term $\frac{1}{2}m\dot{\mathbf{x}}^2/\dot{t}$, the equations are in fact the following:

$$\frac{d}{dt}(m\dot{\mathbf{x}}) + \frac{K}{r}\frac{\mathcal{L}}{r^3} + \frac{d}{dt}\left(\frac{K}{r}\right)\frac{\mathbf{x}\cdot\dot{\mathbf{x}}}{r}\frac{\mathcal{L}}{\mathcal{L}^2} + \frac{\partial}{\partial r}\left(\frac{K}{r}\right)\frac{\ddot{\mathbf{x}}\cdot\mathcal{L}}{\mathcal{L}^2}\mathbf{x} + \frac{d}{dt}\left[\frac{d}{dt}\left(\frac{K}{r}\right)\frac{r\mathcal{L}}{\mathcal{L}^2}\right] = 0, \tag{7.42}$$

where we have put

$$\mathcal{L} = \mathbf{x}\wedge\dot{\mathbf{x}}. \tag{7.43}$$

The rotational invariance of these equations is evident. The fact that they do not depend on the direction of space defining the Lagrangian proves that the various Lagrangians associated with the various directions are all equivalent. The same is then true for the corresponding gauge functions (7.35). The expression of the energy is easily derived and reads

$$E = \tfrac{1}{2}m\dot{\mathbf{x}}^2 + \frac{\partial}{\partial t}\left(\frac{K}{r}\right)A, \tag{7.44}$$

in which A denotes the arctg term contained in (7.41). The function is uniform as soon as the function K is time independent. The conserved quantities, as given by the general theory of Sec. 6, are

$$\mathbf{Q} = \mathbf{x}\wedge\mathbf{p} - \mathbf{K}. \tag{7.45}$$

These quantities are different from the usual generators of the rotation group in phase space. In terms of the velocities, the expression for $\mathbf{Q}$ recovers a symmetrical aspect, namely

$$\mathbf{Q} = m\mathcal{L} - \frac{K}{r}\frac{\mathbf{x}}{r} + r\frac{\hat{K}}{r}\frac{\mathbf{x}\wedge\mathcal{L}}{\mathcal{L}^2}. \tag{7.46}$$

No more than $\mathbf{x}\wedge\mathbf{p}$ the components of the angular momentum $m\mathbf{x}\wedge\dot{\mathbf{x}}$ are conserved.

A particularly simple case is obtained when $K=\lambda r$, with λ a constant. The Lagrangian then becomes uniform, and is reduced to that of a charged particle in a static magnetic field. This field may be identified in the Eqs. (7.42) as being that of a magnetic monopole situated at the origin, $eB=\lambda\mathbf{x}/r^3$. The Hamiltonian formalism is, in this particular case, easily defined. The conjugate momenta and the Hamiltonian are respectively

$$\begin{aligned} p_x &= m\dot{x} + \lambda\,\frac{yz}{r(x^2+y^2)} \\ p_y &= m\dot{y} - \lambda\,\frac{xz}{r(x^2+y^2)}\,, \\ p_z &= m\dot{z}, \end{aligned} \tag{7.47}$$

$$H=\frac{1}{2m}\,(p_x^2+p_y^2+p_z^2)+\frac{\lambda}{m}\,\frac{z}{r(x^2+y^2)}\,(xp_y-yp_x) + \frac{\lambda^2}{2m}\,\frac{z^2}{r^2(x^2+y^2)}\,. \tag{7.48}$$

Thus, although the equations of motion in velocity phase space are perfectly invariant, we see that any Lagrangian or Hamiltonian formulation of the theory must necessarily introduce a preferential direction in space. At the same time neither the generators $\mathbf{x}\wedge\mathbf{p}$ of the rotations nor the components of the angular momentum $m\mathbf{x}\wedge\dot{\mathbf{x}}$ are conserved. However, the equal time brackets of the components of $\mathbf{x}\wedge\mathbf{p}$ are evidently the usual ones. These features look similar to that which are postulated for the so called broken symmetries in particle physics. This aspect is presently under investigation.

APPENDIX A

Proof of Theorem 4.1: (1) Let K be a continuous local gauge function which is defined and bounded on the neighborhood $v\times w$ of (e,x_0), and associated with the continuous local exponent λ. Owing to the theorem of Bargmann, the neighborhood v may be assumed so chosen that $\lambda\equiv\lambda_1(\mathrm{mod}\mu)$, where λ_1 is C^∞ and μ is continuous and bounded on v. The local gauge function K_1 defined by $K_1(\omega,x)=K(\omega,x)+\mu(\omega)$ is locally equivalent to K and associated with λ_1. Let us first show that K_1 is locally equivalent to a local gauge function C^∞ with respect to ω. Let v_1 and w_1 be open neighborhoods of e and x_0 respectively such that $v_1^2\subset v$ and $w_1^{v_1}\subset w$. Following Bargmann we introduce a function $\nu(\omega)$ with the following properties:

(a) ν is C^∞ on G,

(b) ν vanishes outside an open neighborhood v_2 concontained in v_1,

(c) $\int_G \nu\, d\omega = 1$, where $d\omega$ denotes a right invariant measure on G.

The following function:

$$\phi(x)=\int d\omega'\,\nu(\omega')K_1(\omega',x)$$

is defined on w and satisfies, for $\omega\in v_1$ and $x\in w_1$,

$$\phi(x)-\phi(x^\omega)=\int d\omega'\,\nu(\omega')[K_1(\omega',x)-K_1(\omega',x^\omega)].$$

That is, by using (4.2),

$$\phi(x)-\phi(x^\omega)=K_1(\omega,x)-\int d\omega'\,\nu(\omega')\lambda_1(\omega',\omega) + \int d\omega'\,\nu(\omega')[K_1(\omega',x)-K_1(\omega'\omega,x)].$$

This shows that K_1 is locally equivalent to the function

$$K_2(\omega,x)=\int d\omega'[\nu(\omega'\omega^{-1})-\nu(\omega')]K_1(\omega',x).$$

For any $x\in w_1$ that function is C^∞ with respect to $\omega\in v_1$. The corresponding functional equation is

$$K_2(\omega_1,x^{\omega_2})=K_2(\omega_1\omega_2,x)-K_2(\omega_2,x) + \int d\omega'[\nu(\omega'\omega_1^{-1})-\nu(\omega')]\lambda_1(\omega',\omega_2),$$

where we have used the functional equation for λ_1. By putting $x=x_0$ and choosing the neighborhood v_3 in G such that $v_3^2\subset v_1$ and $x_0^{v_3}\subset w_1$, one sees that the mapping

$$v_3\times v_3\ni(\omega_1,\omega_2)\to K_2(\omega_1,x_0^{\omega_2}) \tag{A1}$$

is C^∞. If we now assume that G is connected and transitive, the conditions we have recalled at the beginning of subsection A of Sec. 4 imply the existence of a submanifold $U\subset v_3$ of G which is mapped diffeomorphically onto an open neighborhood of x_0 by the mapping $\omega\to x_0^\omega$. The restriction of the mapping (A1) to $v_3\times U$ is therefore C^∞ and the same property holds for the mapping

$$v_3\times x_0^U\ni(\omega,x)\to K_2(\omega,x).$$

(2) Let k be a C^∞ local gauge function defined on the neighborhood $v\times w$ of (e,x_0) by the formula (2.11) with continuous functions ϕ and μ. From (4.2) the local exponent $\lambda\equiv 0\ (\mathrm{mod}\mu)$ associated with k is C^∞ in the neighborhood of e. Bargmann's theorem then implies that μ is C^∞ in the neighborhood of e. Since k is defined on $v\times w$, the function ϕ must be defined on w^v. For any fixed x in w the mapping $\omega\to\phi(x)-\phi(x^\omega)$ is C^∞ on some neighborhood $v'\subset v$. As in the preceding paragraph, let $U_x\subset v'$ be a submanifold of G diffeomorphically mapped onto an open neighborhood of x by $\omega\to x^\omega$; the function $U_x\ni\omega\to\phi(x^\omega)$ being C^∞, the same is true for the function $x^{U_x}\ni y\to\phi(y)$. It follows that ϕ is C^∞ on w.

APPENDIX B: LOCAL HOMOMORPHISMS FROM G INTO $\mathbb{R}$

We will show that C^∞ local homomorphisms $\omega\to\theta(\omega)$ from G into $\mathbb{R}$ are given by the formulas

$$\theta(\omega)=\int_0^\omega \theta_\alpha\xi^\alpha=\int_0^\omega \theta_\alpha\bar{\xi}^\alpha, \tag{B1}$$

where the coefficients $\theta_\alpha=X_\alpha\theta(e)=\bar{X}_\alpha\theta(e)$ only have to satisfy the conditions

$$C^\gamma_{\alpha\beta}\theta_\gamma=0. \tag{B2}$$

Proof: Let θ be a C^∞ local homomorphism. By writing the relation $\theta(\omega_1\omega_2)=\theta(\omega_1)+\theta(\omega_2)$ for $\omega_1=\omega$, $\omega_2=\exp tX$, taking the derivative with respect to t, at $t=0$, we find

$$d\theta(X)(\omega)=X\theta(\omega)=X\theta(e).$$

The differential $d\theta$ is then locally left invariant and may be written

$$d\theta = \theta_\alpha \zeta^\alpha, \tag{B3}$$

the coefficients θ_α being some constants. With the help of (3.9), the closure condition $d(d\theta)=0$ immediately gives (B2), while (B3) integrates to give the first formula (B1). The second of these formulas is easily derived by permuting the roles played above by ω_1 and ω_2. In particular, $d\theta$ is simultaneously left and right invariant.

Conversely, the condition (B2) implies that the 1-form $\theta_\alpha \zeta^\alpha$ is closed and therefore that the first integral in (B1) does not depend (locally) on the integration path. Let $\theta(\omega)$ be the C^∞ function so obtained. We have $\theta(e)=0$ and, due to the left invariance of the form under the integral sign,

$$\theta(\omega_1\omega_2) - \theta(\omega_1) - \theta(\omega_2) = \int_e^{\omega_1\omega_2} \theta_\alpha \zeta^\alpha - \int_e^{\omega_1} \theta_\alpha \zeta^\alpha - \int_{\omega_1}^{\omega_1\omega_2} \theta_\alpha \zeta^\alpha.$$

This last expression locally vanishes as the integral of a closed form along a closed path.

[1] E. L. Hill, Rev. of Mod. Phys. **23**, 253 (1951).
[2] J. M. Levy-Leblond, Commun. Math. Phys. **12**, 64 (1969).
[3] J. M. Levy-Leblond, Am. J. Phys. **39/5**, 502 (1971).
[4] L. J. Tassie and H. A. Buchdahl, Aust. J. Phys. **17**, 431 (1964); H. A. Buchdahl and L. J. Tassie, Aust. J. Phys. **18**, 109 (1965).
[5] V. Bargmann, Ann. Math. (N. Y.) **59**, 1 (1954).
[6] Let us recall that if ϕ is a diffeomorphism and X a vector field, the transformed $\phi_* X$ is defined by $(\phi_* X)_\omega = d\phi(X)_{\phi^{-1}(\omega)}$ or equivalently, for any C^∞ function f, $\phi_* X(f) \circ \phi = X(f \circ \phi)$. For a differential m-form ζ the transformed $\phi^* \zeta$ obtained by duality is given by $\phi^* \zeta(X_1, \ldots, X_m) = \zeta(\phi_* X_1, \ldots, \phi_* X_m) \circ \phi$. For a function this last formula reduces to $\phi^* f = f \circ \phi$.
[7] See, for example, R. Abraham, *Foundations of Mechanics* (Benjamin, New York, 1967).
[8] S. Helgason, *Differential Geometry and Symmetric Spaces* (Academic, New York, 1962), Chap. 2.
[9] I. M. Ryshik and I. S. Grandstein, *Tables* (Deutscher Verlag, Berlin, 1963), p. 105.

On Bogoliubov transformations for systems of relativistic charged particles

S. N. M. Ruijsenaars

Instituut-Lorentz, University of Leiden, Leiden, The Netherlands
(Received 11 May 1976)

A simple expression for the normal form of the unitary operator implementing a Bogoliubov transformation on a system of relativistic charged particles is obtained. Necessary and sufficient conditions for the transformation to be unitarily implementable are rederived.

1. INTRODUCTION

It is well known that the interaction of relativistic particles with external fields should be considered as a many-particle problem. The classical (i.e., single-particle) theory leads to difficulties which are connected with the unphysical negative energy solutions of relativistic wave equations. In the many-particle framework the wave equation is looked upon as an equation for a quantized field, which is an operator-valued distribution acting on a Fock space. (For a more algebraic viewpoint see Refs. 1 and 2.) If the particle has a distinct antiparticle (which will be assumed in this paper) this space is the symmetric or antisymmetric Fock space over the direct sum of a one-particle and a one-antiparticle space, depending on whether the particle is a boson or a fermion. When the classical theory can be formulated in a Hilbert space it is convenient to smear field operators with vectors from this space instead of with test functions from a Schwartz space, since one can then easily use various operators from the classical theory, for instance the time-evolution operator. If these operators are pseudo-unitary, resp. unitary (in the boson, resp. the fermion case) they generate transformations of the field operators which amount to Bogoliubov transformations of the annihilation and creation operators, i.e., linear transformations which leave the canonical commutation relations (CCR), resp. canonical anticommutation relations (CAR) invariant. When these transformations are unitarily implementable the resulting unitary Fock space operator is assumed to be the physical operator corresponding to the unphysical operator from the classical theory.

More information on the connection between this type of Bogoliubov transformation and the external field problem can be found in Refs. 3–5. General Bogoliubov transformations are treated in the books by Friedrichs[6] and Berezin[7] and, for bosons, in Ref. 8.

The main result of this paper is a simple expression for the normal form of the unitary operator $\mathcal{U}$ which implements the field operator transformation generated by a (pseudo-) unitary operator U acting on the classical Hilbert space. We prove that on the dense subspace of "physical vectors," to be defined below, $\mathcal{U}$ equals a strongly convergent infinite series, the terms of which contain creation and annihilation operators in the normal order. The coefficients of the terms are determined by an operator Λ which is closely related to U. In forthcoming papers on the interaction of relativistic charged spin-0 and spin-$\frac{1}{2}$ particles with external fields[9] we will use this result to establish the connection between the formal Feynman–Dyson series for the Fock space S operator and the unitary operator implementing the transformation generated by the classical S operator. Our results might also be useful for higher spin theories.

Section 2 contains definitions and a summary of various equivalent requirements for the transformation to be unitarily implementable. In Sec. 3 we introduce operators which are used in Sec. 4 to obtain the normal form of $\mathcal{U}$. In the fermion case there is a restriction on U that is dropped in Sec. 5, in which an expression for the normal form of $\mathcal{U}$ is obtained for the general fermion case. Section 6 contains a new proof that a certain well-known condition is necessary for our kind of Bogoliubov transformation to be unitarily implementable, and remarks about unbounded pseudo-unitary operators.

2. PRELIMINARIES

The classical Hilbert space will be denoted by $\mathcal{H}$. It is the direct sum of two subspaces $\mathcal{H}_+$ and $\mathcal{H}_-$, with corresponding projections P_+ and P_-. $\mathcal{H}_+$ will be the one-particle space, $\mathcal{H}_-$ the one-antiparticle space. This decomposition is closely connected with the occurrence of unphysical negative energies in the classical theory. For more details we refer the reader to Refs. 10, 1, 4, and 9. It is convenient to assume

$$\mathcal{H}_+, \mathcal{H}_- = L^2(R^3, dp)^M, \quad M < \infty. \tag{2.1}$$

This assumption has definite notational advantages and corresponds to physical applications.[9] We will indicate at various points how one could proceed in a coordinate-free way. It will also become clear that our results hold true as well if $\mathcal{H}_+$ or $\mathcal{H}_-$ are finite dimensional.

We shall now summarize some results on second quantization, most of which are well-known. The elements of the (anti) symmetric Fock space $\mathcal{F}_\epsilon$ over $\mathcal{H}$ ($\epsilon = a, s$) can be written as

$$\{\psi^{n,r}(p_1, \alpha_1, \ldots, p_n, \alpha_n; q_1, \beta_1, \ldots, q_r, \beta_r)\},$$

where $n, r \in N$ and $\alpha_i, \beta_j = 1, \ldots, M$; $\psi^{n,r}$ is (anti) symmetric in particle and antiparticle variables separately. The inner product in $\mathcal{F}_\epsilon$ is given by

$$(\psi_1, \psi_2) = \sum_{n,r=0}^{\infty} \sum_{\alpha_1,\ldots,\alpha_n,\beta_1,\ldots,\beta_r=1}^{M} \int dp_1 \cdots dp_n\, dq_1 \cdots dq_r \times \overline{\psi_1^{n,r}}(p_1, \alpha_1, \ldots, p_n, \alpha_n; q_1, \beta_1, \ldots, q_r, \beta_r) \times \psi_2^{n,r}(p_1, \alpha_1, \ldots, p_n, \alpha_n; q_1, \beta_1, \ldots, q_r, \beta_r). \tag{2.2}$$

An element $\psi \in \mathcal{J}_\epsilon$ will be called a finite vector if there are $N, R < \infty$ such that $\psi^{n,r} = 0$ if $n > N$, $r > R$. The dense subspace of finite vectors will be denoted by D_f. In $\mathcal{J}_\epsilon$ one has particle and antiparticle creation and annihilation operators $a^{(*)}(f)$, resp. $b^{(*)}(g)$, where $f \in \mathcal{H}_+$, $g \in \mathcal{H}_-$. On a finite vector ψ they are defined by

$$
\begin{aligned}
&(a(f)\psi)^{n,r}(p_1, \alpha_1, \ldots, p_n, \alpha_n; q_1, \beta_1, \ldots, q_r, \beta_r) \\
&\quad = (n+1)^{1/2} \sum_{\alpha=1}^{M} \int dp \bar{f}_\alpha(p) \\
&\qquad \cdot \psi^{n+1,r}(p, \alpha, p_1, \alpha_1, \ldots, p_n, \alpha_n; q_1, \beta_1, \ldots, q_r, \beta_r), \\
&(b(g)\psi)^{n,r}(p_1, \alpha_1, \ldots, p_n, \alpha_n; q_1, \beta_1, \ldots, q_r, \beta_r) \\
&\quad = (r+1)^{1/2}(\mp)^n \sum_{\beta=1}^{M} \int dq \bar{g}_\beta(q) \\
&\qquad \cdot \psi^{n,r+1}(p_1, \alpha_1, \ldots, p_n, \alpha_n; q, \beta, q_1, \beta_1, \ldots, q_r, \beta_r), \\
&(a^*(f)\psi)^{n,r}(p_1, \alpha_1, \ldots, p_n, \alpha_n; q_1, \beta_1, \ldots, q_r, \beta_r) \\
&\quad = n^{-1/2} \sum_{i=1}^{n} (\mp)^{i+1} f_{\alpha_i}(p_i) \\
&\qquad \cdot \psi^{n-1,r}(p_1, \alpha_1, \ldots, \hat{p}_i, \hat{\alpha}_i, \ldots, p_n, \alpha_n; q_1, \beta_1, \ldots, q_r, \beta_r), \\
&(b^*(g)\psi)^{n,r}(p_1, \alpha_1, \ldots, p_n, \alpha_n; q_1, \beta_1, \ldots, q_r, \beta_r) \\
&\quad = r^{-1/2} \sum_{j=1}^{r} (\mp)^{n+j+1} g_{\beta_j}(q_j) \\
&\qquad \cdot \psi^{n,r-1}(p_1, \alpha_1, \ldots, p_n, \alpha_n; q_1, \beta_1, \ldots, \hat{q}_j, \hat{\beta}_j, \ldots, q_r, \beta_r).
\end{aligned}
\tag{2.3}
$$

We will suppress the indices from now on. In (2.3) the upper sign refers to fermions, the lower to bosons. This convention will be used in the whole paper. One can easily show that these operators are bounded in the fermion case and unbounded, but closable, in the boson case. It is straightforward to verify that on D_f the well-known CAR (CCR) hold,

$$
\begin{aligned}
&[a(f_1), a(f_2)]_\pm = [b(g_1), b(g_2)]_\pm = [a^{(*)}(f), b(g)]_\pm = 0, \\
&[a(f_1), a^*(f_2)]_\pm = (f_1, f_2), \quad [b(g_1), b^*(g_2)]_\pm = (g_1, g_2).
\end{aligned}
\tag{2.4}
$$

We will denote the spectral projection of the number operator $N \equiv \Omega(1)$ (for this notation and additional information see Ref. 11) on the interval $[0, M]$ by P_M. One easily sees that the domain of the closure of $a^{(*)}(f)$ (which will be denoted by the same symbol) can be characterized as the set of vectors ψ for which $s\text{-}\lim_{M\to\infty} a^{(*)}(f) P_M \psi$ exists, and that $D(a(f)) = D(a^*(f))$; this is also true for $b^{(*)}(g)$. Hence,

$$
\begin{aligned}
&a^{(*)}(f)\psi = s\text{-}\lim_{M\to\infty} a^{(*)}(f) P_M \psi \quad \forall\ \psi \in D(a(f)) = D(a^*(f)), \\
&b^{(*)}(g)\psi = s\text{-}\lim_{M\to\infty} b^{(*)}(g) P_M \psi \quad \forall\ \psi \in D(b(g)) = D(b^*(g)).
\end{aligned}
\tag{2.5}
$$

This implies that (2.3) holds true for any ψ in the domain of the respective operators. One also concludes, using relations like

$$
a(f)\, a^*(f) P_M \leq \| f \|^2 (N+1) P_M, \tag{2.6}
$$

that the domain of $N^{1/2}$ belongs to the intersection of the domains of all creation and annihilation operators. The latter subspace will be denoted by $\tilde{D}$. (For fermions $\tilde{D} = \mathcal{J}_a$ of course.)

We will also have occasion to use the dense subspace D_∞ on which all powers of the number operator are defined,

$$
D_\infty = \bigcap_{k=1}^{\infty} D(N^k). \tag{2.7}
$$

From well-known (and easily proved) relations like (2.6) and

$$
a(f) N^k P_M = (N+1)^k a(f) P_M, \tag{2.8}
$$

one concludes that, for $\psi \in D_\infty$,

$$
s\text{-}\lim_{M\to\infty} \prod_{i=1}^{n} a^{(*)}(f_i) \prod_{j=1}^{r} b^{(*)}(g_j) P_M \psi
$$

exists and belongs to D_∞, i.e., the closure of any finite product of creation and/or annihilation operators (w.r.t. the subspace of finite vectors) is defined on D_∞ and leaves D_∞ invariant. One also verifies that on D_∞ the closure of the product equals the product of the closures. The relations (2.4) clearly hold true on D_∞.

We need one more subspace. Let Ω be the vacuum; then we will call "physical vectors" the finite linear combinations of vectors of the form $\Pi_{i=1}^{n} a^*(f_i) \Pi_{j=1}^{r} b^*(g_j)\Omega$, where $n, r \geq 0$. (From a physical point of view these vectors are the relevant ones in describing initial states in a scattering theory.) The physical vectors form a dense subspace, denoted by D.

We define field operators on $\tilde{D}$ by

$$
\Phi(v) = a(P_+ v)^+ b^*(\overline{P_- v}) \quad \forall\ v \in \mathcal{H}, \tag{2.9}
$$

where the bar denotes complex conjugation on $\mathcal{H}_-$; in a coordinate free approach one could take any conjugation K which maps $\mathcal{H}_-$ onto itself. [The connection between $\Phi(v)$ and the usual field operators from the Klein–Gordon and Dirac theories can be found in Ref. 9.] In the fermion case we consider transformations $\Phi(v) \to \hat{\Phi}(v)$, generated by unitary operators on $\mathcal{H}$ as follows:

$$
\hat{\Phi}(v) = \Phi(U^* v) \quad \forall\ v \in \mathcal{H}. \tag{2.10}
$$

In the boson case we also have (2.10) but now U is pseudo-unitary, i.e.,

$$
UqU^* = U^* q U = q, \tag{2.11}
$$

where

$$
q = P_+ - P_-. \tag{2.12}
$$

(We will assume that U is bounded. At the end of the paper we shall comment on the case that U is unbounded.) Defining

$$
U_{\epsilon\epsilon'} = P_\epsilon U P_{\epsilon'}, \quad \epsilon, \epsilon' = +, -, \tag{2.13}
$$

we observe that

$$
U_{\epsilon\epsilon'}{}^* = U^*{}_{\epsilon'\epsilon} \tag{2.14}
$$

and that the (pseudo-) unitarity of U is equivalent to the relations

$$
\begin{aligned}
&U^*_{++}U_{++} = 1_{++} \mp U^*_{+-}U_{-+}, \quad U_{++}U^*_{++} = 1_{++} \mp U_{+-}U^*_{-+},\\
&U^*_{--}U_{--} = 1_{--} \mp U^*_{-+}U_{+-}, \quad U_{--}U^*_{--} = 1_{--} \mp U_{-+}U^*_{+-},\\
&U^*_{++}U_{+-} = \mp U^*_{+-}U_{--}, \quad U_{++}U^*_{+-} = \mp U_{+-}U^*_{--},\\
&U^*_{--}U_{-+} = \mp U^*_{-+}U_{++}, \quad U_{--}U^*_{-+} = \mp U_{-+}U^*_{++}.
\end{aligned}
\tag{2.15}
$$

Decomposing $\hat{\Phi}(v)$ in new annihilation and creation operators as in (2.9), i.e., setting

$$
\hat{\Phi}(v) \equiv \hat{a}(P_+ v) + \hat{b}^*(\overline{P_- v}), \tag{2.16}
$$

one easily sees that (2.10) is equivalent to the transformation

$$
a(f) \to \hat{a}(f), b(\bar{g}) \to \hat{b}(\bar{g}),
$$

where

$$
\begin{aligned}
&\hat{a}(f) = a(U^*_{++}f) + b^*(\overline{U^*_{-+}f}) \quad \forall\ f \in \mathcal{H}_+\\
&\hat{b}(\bar{g}) = b(\overline{U^*_{--}g}) + a^*(U^*_{+-}g) \quad \forall\ g \in \mathcal{H}_-.
\end{aligned}
\tag{2.17}
$$

Using (2.15) it is straightforward to verify that these operators also fulfil the CAR (CCR). The reader will have no difficulty in writing our transformation in terms of "one-body" annihilation and creation operators $c^{(*)}(v) \equiv a^{(*)}(P_+ v) + b^{(*)}(P_- v)$, and establishing the special character of the resulting Bogoliubov transformation.

The transformation (2.10) by definition is unitarily implementable if there exists a unitary operator U, mapping $\tilde{D}$ onto $\tilde{D}$, such that

$$
\hat{\Phi}(v) = U^* \Phi(v) U \quad \forall\ v \in \mathcal{H} \tag{2.18}
$$

or, equivalently, such that

$$
\begin{aligned}
&\hat{a}(f) = U^* a(f) U \quad \forall\ f \in \mathcal{H}_+\\
&\hat{b}(\bar{g}) = U^* b(\bar{g}) U \quad \forall\ g \in \mathcal{H}_-.
\end{aligned}
\tag{2.19}
$$

It is well known that this is equivalent to the existence of a nonzero vector $\hat{\Omega} \in \tilde{D}$ such that

$$
\hat{a}(f)\hat{\Omega} = \hat{b}(\bar{g})\hat{\Omega} = 0 \quad \forall\ f \in \mathcal{H}_+ \ \forall\ g \in \mathcal{H}_-; \tag{2.20}
$$

if such a vector exists it is a scalar multiple of $U^*\Omega$.

Because we want to obtain an expression for $U\Omega$ it is convenient to consider as well the transformation generated by the inverse of U, i.e., the transformation $\Phi(v) \to \Phi'(v)$ where

$$
\Phi'(v) \equiv \Phi(Uv) \quad \forall\ v \in \mathcal{H} \tag{2.21}
$$

in the fermion case, and

$$
\Phi'(v) \equiv \Phi(qUqv) \quad \forall\ v \in \mathcal{H} \tag{2.22}
$$

in the boson case. Existence of a unitary operator U satisfying (2.18) is obviously equivalent to existence of a unitary operator U satisfying

$$
\Phi'(v) = U\Phi(v)U^* \quad \forall\ v \in \mathcal{H}, \tag{2.23}
$$

or, equivalently, satisfying

$$
\begin{aligned}
&a'(f) = U a(f) U^* \quad \forall\ f \in \mathcal{H}_+,\\
&b'(\bar{g}) = U b(\bar{g}) U^* \quad \forall\ g \in \mathcal{H}_-,
\end{aligned}
\tag{2.24}
$$

where

$$
\begin{aligned}
&a'(f) \equiv a(U_{++}f) \pm b^*(\overline{U_{-+}f}),\\
&b'(\bar{g}) \equiv b(\overline{U_{--}g}) \pm a^*(U_{+-}g).
\end{aligned}
\tag{2.25}
$$

From (2.15) one again concludes that these operators fulfil the CAR (CCR), so implementability is also equivalent to existence of a nonzero vector $\Omega' \in \tilde{D}$ satisfying

$$
a'(f)\Omega' = b'(\bar{g})\Omega' = 0 \quad \forall\ f \in \mathcal{H}_+ \ \forall\ g \in \mathcal{H}_-; \tag{2.26}
$$

if such a vector exists it is a scalar multiple of $U\Omega$.

From the work of several authors[12–14] it follows that the transformation (2.10) or, equivalently, (2.21)–(2.22) is unitarily implementable if and only if

$$
U_{+-},\ U_{-+} \in \mathrm{HS}, \tag{2.27}
$$

where HS is the set of all Hilbert–Schmidt (H.S.) operators on $\mathcal{H}$. In Secs. 3, 4, and 5 it will be assumed that (2.27) holds true. We will denote the corresponding unitary operator on Fock space by U. From the fact that the Fock–Cook representation of the Clifford algebra, resp. the Weyl algebra over $\mathcal{H}$, is irreducible it follows that U is up to a phase factor uniquely determined.

We remark that the sufficiency of (2.27) will be a consequence of our results, while we will give a new proof of the necessity in Sec. 6, so in this respect the paper is self-contained.

3. THE OPERATORS Λ AND EXP $(\Lambda_{+-} a^* b^*)$

In the boson case one easily concludes from (2.15) that U_{++} and U_{--} have bounded inverses (as operators on $\mathcal{H}_+$ resp., $\mathcal{H}_-$). In the fermion case this also follows from (2.15) and (2.27) if we make the additional assumption

$$
\mathrm{Ker}\,U_{++} = \mathrm{Ker}\,U_{--} = 0. \tag{3.1}
$$

We will assume (3.1) in this section and the next one, and deal with the general case in Sec. 5.

We now introduce a bounded operator Λ on $\mathcal{H}$ which will enable us to obtain a simple expression for the normal form of U. Λ is defined by

$$
\begin{aligned}
&\Lambda_{--} = \pm(1_{--} - U_{--}^{-1}),\\
&\Lambda_{-+} = \pm U_{--}^{-1}U_{-+},\\
&\Lambda_{+-} = U_{+-}U_{--}^{-1},\\
&\Lambda_{++} = U_{++} - 1_{++} - U_{+-}U_{--}^{-1}U_{-+}.
\end{aligned}
\tag{3.2}
$$

From (2.15) it follows that this is equivalent to

$$
\begin{aligned}
&\Lambda_{--} = \pm(1_{--} - U^*_{--} + U^*_{-+}U^{*-1}_{++}U^*_{+-}),\\
&\Lambda_{-+} = -U^*_{-+}U^{*-1}_{++},\\
&\Lambda_{+-} = \mp U^{*-1}_{++}U^*_{+-},\\
&\Lambda_{++} = -1_{++} + U^{*-1}_{++}.
\end{aligned}
\tag{3.3}
$$

One easily verifies that (3.2) is also equivalent to

$$
\text{(fermions)}\quad (U-1) - \Lambda - (U-1)P_-\Lambda = (U-1) - \Lambda - \Lambda P_-(U-1) = 0, \tag{3.4}
$$

$$
\text{(bosons)}\quad (U-1) - q\Lambda + (U-1)P_-\Lambda = (U-1) - q\Lambda - q\Lambda P_-(U-1) = 0. \tag{3.5}
$$

We will need the following relations, which follow from (3.2) and (3.3):

$$U_{++}-\Lambda_{+-}U_{-+}=1_{++}+\Lambda_{++},\quad \Lambda_{-+}\mp U_{-+}+\Lambda_{--}U_{-+}=0, \tag{3.6}$$

$$U_{--}\pm\Lambda^*_{-+}U_{+-}=1_{--}\mp\Lambda^*_{--},\quad U_{+-}+\Lambda^*_{+-}+\Lambda^*_{++}U_{+-}=0. \tag{3.7}$$

From (2.27) and (3.2) we infer that Λ_{+-} and Λ_{-+} are H.S.; moreover, one obtains from (2.15), in the boson case,

$$\|\Lambda_{+-}\|^2=\sup_{v\in H}\frac{(U_{+-}U_{--}^{-1}v,U_{+-}U_{--}^{-1}v)}{(v,v)}=\sup_{w\in H}\frac{(U_{+-}w,U_{+-}w)}{(U_{--}w,U_{--}w)}$$

$$=\sup_{W\in H}\frac{(U^*_{-+}U_{+-}w,w)}{(U^*_{-+}U_{+-}w,w)+(w,w)}<1. \tag{3.8}$$

Consequently,

$$(\text{bosons})\Lambda_{+-}=\sum_{i=1}^{M_0}\lambda_iF_i(G_i,\cdot)\quad M_0\leq\infty,\quad 0<\lambda_i\leq\theta<1, \tag{3.9}$$

where $\{F_i\}$, $\{G_i\}$ are orthonormal sets in H_+, H_- and where $\lambda_i\leq\lambda_j$ if $i>j$ (see Ref. 15); furthermore,

$$\sum_{i=1}^{M_0}\lambda_i^2<\infty. \tag{3.10}$$

We set

$$\Lambda_{+-}a^*b^*\equiv\int dp\,dp'\Lambda_{+-}(p,p')a^*(p)b^*(p'), \tag{3.11}$$

where $\Lambda_{+-}(p,p')$ is the kernel of Λ_{+-}. The operator $\Lambda_{+-}a^*b^*$ and its powers are clearly defined on D_f. The next lemma shows that the operator $\exp(\Lambda_{+-}a^*b^*)$ is defined on D.

Lemma 3.1: Let $\phi\in D$ and let

$$\phi_n=\frac{(\Lambda_{+-}a^*b^*)^n}{n!}\phi. \tag{3.12}$$

Then $s\text{-}\lim_{N\to\infty}\sum_{n=0}^{N}\phi_n$ exists and belongs to D_∞.

Proof (*A. bosons*): We assume first that $\phi=\Omega$. Existence of the limit is then obviously equivalent to existence of $\lim_{N\to\infty}\sum_{n=0}^{N}a_n$, where

$$a_n=\left\|\frac{(\Lambda_{+-}a^*b^*)^n}{n!}\Omega\right\|^2. \tag{3.13}$$

One easily obtains ($n\geq 1$)

$$a_n=\frac{1}{n!}\int dp_1\cdots dp_n\,dq_1\cdots dq_n$$

$$\times\sum_{\sigma\in S_n}\prod_{i=1}^{n}\overline{\Lambda_{+-}}(q_i,p_i)\Lambda_{+-}(q_i,p_{\sigma(i)}), \tag{3.14}$$

where S_n is the symmetric group. We now define

$$a_{n,N}=\left\|\frac{(\Lambda_Na^*b^*)^n}{n!}\Omega\right\|^2, \tag{3.15}$$

where

$$\Lambda_N\equiv\sum_{i=1}^{N}\lambda_iF_i(G_i,\cdot). \tag{3.16}$$

The analog of (3.14) for $a_{n,N}$ implies

$$\lim_{N\to\infty}a_{n,N}=a_n. \tag{3.17}$$

On the other hand,

$$a_{n,N}=\left\|\frac{[\sum_{i=1}^{N}\lambda_ia^*(F_i)b^*(\overline{G_i})]^n}{n!}\Omega\right\|^2$$

$$=\left\|\sum_{\substack{k_1,\ldots,k_N=0\\ k_1+\cdots+k_N=n}}^{n}\prod_{i=1}^{N}\frac{[\lambda_ia^*(F_i)b^*(\overline{G_i})]^{k_i}}{k_i!}\Omega\right\|^2$$

$$=\sum_{\substack{k_1,\ldots,k_N=0\\ k_1+\cdots+k_N=n}}^{n}\prod_{i=1}^{N}\lambda_i^{2k_i}. \tag{3.18}$$

We now introduce a function

$$F_N(\alpha)=\prod_{i=1}^{N}(1-\alpha\lambda_i^2)^{-1}, \tag{3.19}$$

which is clearly analytic in the disc 0, defined by

$$0=\{\alpha\in C\mid|\alpha|<\theta^{-2}\}. \tag{3.20}$$

Using (3.18) one easily verifies

$$F_N(\alpha)=\sum_{n=0}^{\infty}a_{n,N}\alpha^n. \tag{3.21}$$

We shall prove that $\lim_{N\to\infty}F_N(\alpha)$ exists on 0 and that the limit function is analytic in 0. Let r be such that

$$0<r<\theta^{-2}. \tag{3.22}$$

If $|\alpha|\leq r$,

$$|F_n(\alpha)|\leq\prod_{i=1}^{N}|1-\alpha\lambda_i^2|^{-1}\leq\prod_{i=1}^{N}(1-r\lambda_i^2)^{-1}$$

$$=\exp\left[-\sum_{i=1}^{N}\ln(1-r\lambda_i^2)\right]$$

$$\leq\exp\left(-N'\ln(1-r\theta^2)+2r\sum_{i=N'+1}^{\infty}\lambda_i^2\right)\equiv C_r<\infty, \tag{3.23}$$

where N' is such that

$$r\lambda_{N'+1}^2\leq\tfrac{1}{2}. \tag{3.24}$$

From (3.23), ($|\alpha|\leq r, M>N$),

$$|F_N(\alpha)-F_M(\alpha)|\leq C_r\left|1-\prod_{i=N+1}^{M}(1-\alpha\lambda_i^2)^{-1}\right|$$

$$=C_r\left|1-\exp\left[-\sum_{i=N+1}^{M}\ln(1-\alpha\lambda_i^2)\right]\right|. \tag{3.25}$$

We now observe, ($N>N'$),

$$\left|\sum_{i=N+1}^{M}\ln(1-\alpha\lambda_i^2)\right|$$

$$\leq\sum_{i=N+1}^{M}\left|\ln|1-\alpha\lambda_i^2|\,\right|+\sum_{i=N+1}^{M}|\arg(1-\alpha\lambda_i^2)|$$

$$\leq-\sum_{i=N+1}^{\infty}\ln(1-r\lambda_i^2)+\frac{\pi}{2}\sum_{i=N+1}^{\infty}|\sin[\arg(1-\alpha\lambda_i^2)]|$$

$$\leq\left(2r+\frac{\pi r}{2}\right)\sum_{i=N+1}^{\infty}\lambda_i^2. \tag{3.26}$$

It evidently follows from (3.25) and (3.26) that

$$\lim_{N\to\infty}F_N(\alpha)=F(\alpha)\quad\alpha\in 0, \tag{3.27}$$

where $F(\alpha)$ is analytic in 0.

Using (3.17) and (3.21) we infer

$$F(\alpha)=\sum_{n=0}^{\infty}a_n\alpha^n. \tag{3.28}$$

Thus, since $\theta^{-2}>1$,

$$\sum_{n=0}^{\infty}a_n=F(1)<\infty, \tag{3.29}$$

which proves that $s\text{-}\lim_{N\to\infty}\sum_{n=0}^{N}\phi_n$ exists if $\phi=\Omega$. We notice that

$$F(1)\equiv\lim_{N\to\infty}F_N(1)=\prod_{i=1}^{M_0}(1-\lambda_i^2)^{-1}. \tag{3.30}$$

Hence (see Ref. 16),

$$\text{(bosons)}\,\|\exp(\Lambda_{+-}a^*b^*)\Omega\|^2$$

$$=\det(1_{--}-\Lambda_{+-}{}^*\Lambda_{+-})^{-1}. \tag{3.31}$$

We now introduce the functions

$$G_k(\alpha)=\sum_{n=0}^{\infty}(2n)^k a_n\alpha^n \quad k\in N. \tag{3.32}$$

These functions are analytic in 0 because the power series on the rhs of (3.32) has the same convergence radius as the rhs of (3.28). Therefore,

$$\sum_{n=0}^{\infty}(2n)^k a_n=G_k(1)<\infty. \tag{3.33}$$

From (3.13) and (3.33) it then follows that

$$\exp(\Lambda_{+-}a^*b^*)\Omega\in D_\infty. \tag{3.34}$$

One obviously has

$$\sum_{n=0}^{N}\frac{(\Lambda_{+-}a^*b^*)^n}{n!}\prod_{i=1}^{n}a^*(f_i)\prod_{j=1}^{r}b^*(g_j)\Omega$$

$$=\prod_{i=1}^{n}a^*(f_i)\prod_{j=1}^{r}b^*(g_j)P_{2N}\exp(\Lambda_{+-}a^*b^*)\Omega. \tag{3.35}$$

From (3.34) and (3.35) we finally conclude that the limit of the lhs of (3.35) exists and belongs to D_∞, which proves the lemma for bosons.

(B. *fermions*): Proceeding in the same way as for bosons, one obtains instead of (3.14),

$$a_n=\frac{1}{n!}\int dp_1\cdots dp_n\,dq_1\cdots dq_n$$

$$\times\sum_{\sigma\in S_n}\operatorname{sgn}\sigma\prod_{i=1}^{n}\overline{\Lambda_{+-}}(q_i,p_i)\Lambda_{+-}(q_i,p_{\sigma(i)}). \tag{3.36}$$

Defining

$$T=\Lambda_{+-}{}^*\Lambda_{+-}, \tag{3.37}$$

we have

$$a_n=(1/n!)T_n, \tag{3.38}$$

where

$$T_n\equiv\int dp_1\cdots dp_n T(p_1,\ldots,p_n) \tag{3.39}$$

and $T(p_1,\ldots,p_n)$ is the determinant the elements of which are $T(p_i,p_j)$ $(i,j=1,\ldots,n)$. Introducing the entire function

$$\tilde{d}(\lambda)=\det(1_{--}+\lambda T), \tag{3.40}$$

one has (see Ref. 16)

$$\tilde{d}(\lambda)=\sum_{n=0}^{\infty}\frac{1}{n!}P_n\lambda^n, \tag{3.41}$$

where

$$P_0=1,$$

$$P_n=\begin{vmatrix}\sigma_1 & 1 & 0 & \cdots & 0 & 0\\ \sigma_2 & \sigma_1 & 2 & \cdots & 0 & 0\\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots\\ \sigma_{n-1} & \sigma_{n-2} & \sigma_{n-3} & \cdots & \sigma_1 & n-1\\ \sigma_n & \sigma_{n-1} & \sigma_{n-2} & \cdots & \sigma_2 & \sigma_1\end{vmatrix}, \tag{3.42}$$

and

$$\sigma_i=\operatorname{Tr}(T^i). \tag{3.43}$$

Expanding the determinant, we obtain the recurrence relation

$$P_n=\sum_{k=1}^{n}(-1)^{k-1}\frac{(n-1)!}{(n-k)!}\sigma_k P_{n-k} \quad (n\geq 1). \tag{3.44}$$

Expanding $T(p_1,\ldots,p_n)$ in (3.39), one easily sees that T_n obeys the same recurrence relation. Thus, since $T_1=P_1$,

$$T_n=P_n \quad \forall\, n\in N^*. \tag{3.45}$$

Therefore,

$$\tilde{d}(\lambda)=\sum_{n=0}^{\infty}a_n\lambda^n \tag{3.46}$$

so the limit exists if $\phi=\Omega$ and

$$\text{(fermions)}\quad\|\exp(\Lambda_{+-}a^*b^*)\Omega\|^2=\det(1_{--}+\Lambda_{+-}{}^*\Lambda_{+-}). \tag{3.47}$$

Arguing in the same way as for bosons, one concludes

$$\exp(\Lambda_{+-}a^*b^*)\Omega\in D_\infty. \tag{3.48}$$

The lemma now follows from (3.48) and (3.35). ∎

We point out that the proof of the lemma could be shortened in the boson case by using more results on infinite determinants.[17–19]

If one does not assume (2.1), one should define

$$\text{(bosons)}\quad\Lambda_{+-}a^*b^*=\sum_{i=1}^{M_0}\lambda_i a^*(F_i)b^*(KG_i), \tag{3.49}$$

where K is the conjugation chosen in (2.9). One could then map $\mathcal{H}$ onto L^2 spaces as in (2.1) in such a way that K becomes complex conjugation on $\mathcal{H}_-$, use the lemma, and transform back. Using the analog of (3.49) for fermions one could prove the lemma in a similar way for fermions.

4. THE NORMAL FORM OF U

The result of the next lemma was obtained in different forms by several authors.[7,8,3] It essentially dates back to the work of Friedrichs.[6]

Lemma 4.1: The following relation holds true:

$$U\Omega = \exp(i\theta)\det(1_{--} \pm \Lambda_{+-}{}^{*}\Lambda_{+-})^{\mp 1/2}$$

$$\times \exp(\Lambda_{+-} a^{*} b^{*})\Omega \qquad 0 \leq \theta < 2\pi. \tag{4.1}$$

Proof: By (3.31) and (3.47) the norm of the rhs equals 1, so by (2.26) and (2.25) it suffices to prove

$$[a(U_{++} f) \pm b^{*}(\overline{U_{-+} f})]\exp(\Lambda_{+-} a^{*} b^{*})\Omega = 0 \quad \forall\, f \in \mathcal{H}_{+},$$

$$[b(\overline{U_{--} g}) \pm a^{*}(U_{+-} g)]\exp(\Lambda_{+-} a^{*} b^{*})\Omega = 0 \quad \forall\, g \in \mathcal{H}_{-}. \tag{4.2}$$

Notice that the lhs of these equations are well defined in virtue of (3.34). [In the fermion case this of course already follows from the relation $\exp(\Lambda_{+-} a^{*} b^{*})\Omega \in \mathcal{F}_{a}$.] Using (2.5) and the CAR (CCR), we conclude that (4.2) is equivalent to

$$[b^{*}(\overline{\Lambda_{+-}{}^{*} U_{++} f}) \pm b^{*}(\overline{U_{-+} f})]\exp(\Lambda_{+-} a^{*} b^{*})\Omega = 0 \quad \forall\, f \in \mathcal{H}_{+}$$

$$[\mp a^{*}(\Lambda_{+-} U_{--} g) \pm a^{*}(U_{+-} g)]\exp(\Lambda_{+-} a^{*} b^{*})\Omega = 0 \quad \forall\, g \in \mathcal{H}_{-}. \tag{4.3}$$

However, (4.3) follows immediately from (3.2) and (3.3). ∎

We normalize U by setting $\theta = 0$ in (4.1):

$$U\Omega = \det(1_{--} \pm \Lambda_{+-}{}^{*}\Lambda_{+-})^{\mp 1/2}\exp(\Lambda_{+-} a^{*} b^{*})\,\Omega. \tag{4.4}$$

We now introduce the operators which are needed for the normal form of U. Let K, L, M be bounded operators on $\mathcal{H}$. We set

$$K_{++}^{j} L_{--}^{k} M_{-+}^{l}\, a^{*j} b^{*k} b^{k} b^{l} a^{l} a^{j}$$

$$\equiv \int dk_1 \circ\circ\circ dk_j' dp_1 \cdots dp_k' dq_1 \cdots dq_l' \prod_{\rho=1}^{j} (K_{++})(k_\rho, k_\rho') \prod_{\sigma=1}^{k} (L_{--})$$

$$(p_\sigma, p_\sigma') \prod_{\tau=1}^{l} (M_{-+})(q_\tau, q_\tau')\, a^{*}(k_1) \cdots a^{*}(k_j)\, b^{*}(p_1) \circ\circ\circ b^{*}(p_k)$$

$$\times b(p_k') \circ\circ\circ b(p_1')\, b(q_1) \cdots b(q_l)(q_l') \cdots a(q_1')\, a(k_j') \cdots a(k_1'), \tag{4.5}$$

where, e.g., $(K_{++})(k, k')$ is the tempered distribution which corresponds to K_{++} by the nuclear theorem. The formal expression at the rhs of (4.5) is defined on D by writing

$$a^{*}(f) = \int dp\, a^{*}(p) f(p), \quad b^{*}(g) = \int dp\, b^{*}(p) g(p) \tag{4.6}$$

and then using the formal CAR (CCR)

$$[a(p), a(p')]_{\pm} = [b(p), b(p')]_{\pm} = [a^{(*)}(p), b(p')]_{\pm} = 0,$$
$$[a(p), a^{*}(p')]_{\pm} = [b(p), b^{*}(p')]_{\pm} = \delta(p - p') \tag{4.7}$$

and the relation

$$a(p)\,\Omega = b(p)\,\Omega = 0, \tag{4.8}$$

to get rid of all annihilation operators in (4.5). One should then set, e.g.,

$$\int dk\, dk'\, (K_{++})(k, k')\, a^{*}(k) f(k') \equiv a^{*}(K_{++} f),$$
$$\int dq\, dq'\, (M_{-+})(q, q')\, g(q)\, f(q') \equiv (\bar{g}, M_{-+} f). \tag{4.9}$$

One easily convinces oneself that this gives rise to a well-defined linear operator mapping D into D. It is clear that one could define this operator on D without using (2.1) but this would obviously give rise to very unwieldy formulas. Denoting the operator by $O_{j,k,l}$, it is straightforward to verify relations like

$$O_{j,k,l}\, a^{*}(f) = a^{*}(f)\, O_{j,k,l} + j a^{*}(K_{++} f)\, O_{j-1,k,l}$$
$$+ l O_{j,k,l-1}\, b(\overline{M_{-+} f}) \tag{4.10}$$

which hold on D.

Defining the transpose N^{T} of a bounded operator N on $\mathcal{H}$ by

$$(N^{T} f)(p) = \overline{(N^{*}\bar{f})}(p) \tag{4.11}$$

we can now proceed to the first theorem. We define an operator $\tilde{\Gamma}(U)$ by

$$\tilde{\Gamma}(U) = \det(1_{--} \pm \Lambda_{+-}{}^{*}\Lambda_{+-})^{\pm 1/2} \sum_{L=0}^{\infty} U_L, \tag{4.12}$$

where

$$U_L = \sum_{\substack{i,j,k,l=0 \\ i+j+k+l=L}}^{L} \frac{1}{i!\,j!\,k!\,l!}\, \Lambda_{+-}^{i} \Lambda_{++}^{j} (\mp \Lambda_{--}^{T})^{k} \Lambda_{-+}^{l}$$

$$\times a^{*i} b^{*i} a^{*j} b^{*k} b^{k} b^{l} a^{l} a^{j}. \tag{4.13}$$

Notice that U_L is well defined on D. We can abbreviate (4.12) and (4.13) as follows:

$$\tilde{\Gamma}(U) = \det(1_{--} \pm \Lambda_{+-}{}^{*}\Lambda_{+-})^{\mp 1/2}$$
$$\times :\exp(\Lambda_{+-} a^{*} b^{*} + \Lambda_{++} a^{*} a + \Lambda_{--} b b^{*} + \Lambda_{-+} b a):. \tag{4.14}$$

Defining the domain of $\tilde{\Gamma}(U)$ by

$$D(\tilde{\Gamma}(U)) = \{\varphi \in D \mid \operatorname*{s-lim}_{N\to\infty} \sum_{L=0}^{N} U_L \varphi \text{ exists}\} \tag{4.15}$$

we have the following theorem.

Theorem 4.1: The domain of $\tilde{\Gamma}(U)$ equals D,

$$D(\tilde{\Gamma}(U)) = D, \tag{4.16}$$

and

$$\tilde{\Gamma}(U) D \subset D_{\infty}. \tag{4.17}$$

The operator U is equal to $\tilde{\Gamma}(U)$ on D,

$$U\phi = \det(1_{--} \pm \Lambda_{+-}{}^{*}\Lambda_{+-})^{\mp 1/2}$$
$$\times :\exp(\Lambda_{+-} a^{*} b^{*} + \Lambda_{++} a^{*} a + \Lambda_{--} b b^{*} + \Lambda_{-+} b a):\phi \quad \forall\, \phi \in D. \tag{4.18}$$

Proof: From Lemma 3.1 and (4.4) it evidently follows that $\Omega \in D(\tilde{\Gamma}(U))$, $\tilde{\Gamma}(U)\,\Omega = U\Omega$, and $U\Omega \in D_{\infty}$. To prove the existence of the limit in (4.15) for a vector ϕ of the form

$$\phi = \prod_{i=1}^{n} a^{*}(f_i) \prod_{j=1}^{r} b^{*}(g_j)\,\Omega, \tag{4.19}$$

we observe that the individual term in (4.13) only contributes if

$$l + j \leq n, \quad l + k \leq r. \tag{4.20}$$

Since there is only a finite number of (j, k, l) which fulfill (4.20) we conclude from Lemma 3.1 that the limit in (4.15) exists and belongs to D_{∞}. It remains to prove (4.18).

In view of (2.24) and the relation $\tilde{\Gamma}(U)\,\Omega = U\Omega$, it suffices to show that on D

$$\tilde{\Gamma}(U)\,a^*(f)=a'^*(f)\,\tilde{\Gamma}(U) \quad \forall\, f\in H_+, \tag{4.21}$$

$$\tilde{\Gamma}(U)\,b^*(\bar{g})=b'^*(\bar{g})\,\tilde{\Gamma}(U) \quad \forall\, g\in H_-. \tag{4.22}$$

Using relations like (4.10) and the relations (3.6) and (2.25) we now have on D [observe that, e.g., $\lim_{N\to\infty} a^*(f)\sum_{L=0}^{N}\cdots = a^*(f)\lim_{N\to\infty}\sum_{L=0}^{\infty}\cdots$ on D according to Lemma 3.1, (2.5), and the argument after (4.20)]

$$\begin{aligned}\tilde{\Gamma}(U)\,a^*(f)&=[a^*(f)+a^*(\Lambda_{++}f)]\,\tilde{\Gamma}(U)+\tilde{\Gamma}(U)\,b(\overline{\Lambda_{-+}f})\\&=[a^*(U_{++}f)-a^*(\Lambda_{+-}U_{-+}f)]\,\tilde{\Gamma}(U)+\tilde{\Gamma}(U)\,b(\overline{\Lambda_{-+}f})\\&=[a^*(U_{++}f)\pm b(\overline{U_{-+}f})]\,\tilde{\Gamma}(U)+\tilde{\Gamma}(U)\,b(\overline{\Lambda_{-+}f})\\&\quad+\det{}^{\mp}\cdots\sum\sum\cdots\Lambda^{i}_{+-}\,a^{*i}b^{*i}(\mp b(\overline{U_{-+}f}))\\&\quad\times\Lambda^{j}_{++}\cdots a^{*j}\cdots a^{j}\\&=a'^*(f)\,\tilde{\Gamma}(U)+\tilde{\Gamma}(U)\,b(\overline{(\Lambda_{-+}\mp U_{-+}+\Lambda_{--}U_{-+})f})\\&=a'^*(f)\,\tilde{\Gamma}(U),\end{aligned}$$

which proves (4.21). Similarly, using (3.7), we obtain

$$\begin{aligned}\tilde{\Gamma}(U)\,b^*(\bar{g})&=[b^*(\bar{g})\mp b^*(\overline{\Lambda^*_{--}g})]\,\tilde{\Gamma}(U)+\tilde{\Gamma}(U)\,a(\mp\Lambda_{+-}{}^*g)\\&=[b^*(\overline{U_{--}g})\pm b^*(\overline{\Lambda^*_{-+}U_{+-}g})]\,\tilde{\Gamma}(U)+\tilde{\Gamma}(U)\,a(\mp\Lambda_{+-}{}^*g)\\&=[b^*(\overline{U_{--}g})\pm a(U_{+-}g)]\,\tilde{\Gamma}(U)+\tilde{\Gamma}(U)\,a(\mp\Lambda_{+-}{}^*g)\\&\quad+\det{}^{\mp}\cdots b^{*i}(\mp a(U_{+-}g))\cdots a^{j}\\&=b'^*(\bar{g})\,\tilde{\Gamma}(U)\mp\tilde{\Gamma}(U)\,a((\Lambda_{+-}{}^*+U_{+-}+\Lambda^*_{++}U_{+-})g)\\&=b'^*(g)\,\tilde{\Gamma}(U),\end{aligned}$$

which proves (4.22). ∎

It should be noticed that as a consequence of this theorem one can write the "matrix element" $(\phi, \mathcal{U}\psi)$ for "physical vectors" ϕ, ψ as a finite sum of terms each of which is a finite product of the "matrix elements" of the operator Λ on H and the scalar $\det(\cdots)^{\mp 1/2}$.

We further observe that a pseudo-unitary U is unitary if and only if

$$U_{+-}=U_{-+}=0. \tag{4.23}$$

Assuming (4.23) for bosons and fermions one can define a unitary operator $\tilde{U}$ by

$$\tilde{U}_{+-}=\tilde{U}_{-+}=0, \quad \tilde{U}_{++}=U_{++}, \quad \tilde{U}_{--}=\overline{U}_{--}, \tag{4.24}$$

where

$$(\overline{U}_{--}v)(p)\equiv(\overline{U_{--}\bar{v}})(p) \quad \forall\, v\in H. \tag{4.25}$$

Then

$$\tilde{\Gamma}(U)\subset\Gamma(\tilde{U}) \tag{4.26}$$

which motivates our notation [for a definition of $\Gamma(\tilde{U})$ see, e.g., Ref. 15].

Using Stone's theorem, one can conclude from the Weyl algebra formulation of the CCR that $\mathcal{U}$ maps $\tilde{D}$ onto $\tilde{D}$. This also follows from our theorem. To show this, let $\psi\in D_f$. Then $P_N\psi=\psi$ if N is big enough. Now let $\psi_n\in D$ be such that $P_N\psi_n=\psi_n$ and $\psi_n\to\psi$. Then, e.g.,

$$a(f)\,\mathcal{U}P_N\psi_n=\mathcal{U}a'(f)\,P_N\psi_n. \tag{4.27}$$

The lhs is well defined since by the theorem

$$\mathcal{U}D\subset D_\infty. \tag{4.28}$$

Now the limit $n\to\infty$ in (4.27) exists since $a'(f)$ is bounded on $P_N\mathcal{F}_\epsilon$. Hence

$$\mathcal{U}\psi\in D(a(f)) \quad a(f)\mathcal{U}\psi=\mathcal{U}a'(f)\psi. \tag{4.29}$$

Thus,

$$\mathcal{U}D_f\subset\tilde{D}. \tag{4.30}$$

If $\phi\in\tilde{D}$, then by (4.29),

$$a(f)\,\mathcal{U}P_M\phi=\mathcal{U}a'(f)\,P_M\phi. \tag{4.31}$$

The limit $M\to\infty$ in (4.31) exists in view of (2.5). Therefore,

$$\mathcal{U}\phi\in D(a(f)), \quad a(f)\,\mathcal{U}\phi=\mathcal{U}a'(f)\,\phi, \tag{4.32}$$

so

$$\mathcal{U}\tilde{D}\subset\tilde{D}. \tag{4.33}$$

Repeating the argument for $\mathcal{U}^*$ we infer

$$\mathcal{U}\tilde{D}=\tilde{D} \tag{4.34}$$

as asserted.

In a coordinate-free approach one should define the operator in (4.5) directly on D, replacing complex conjugation by the conjugation K. One could then proceed as indicated at the end of Sec. 3.

5. THE GENERAL FERMION CASE

We shall now treat the general fermion case, i.e., we drop the assumption (3.1). Let $\{g'_j\}_{j=1}^{M}$ and $\{f'_i\}_{i=1}^{L}$ be orthonormal bases for $\mathrm{Ker}U_{++}$, resp. $\mathrm{Ker}U_{--}$. In view of our standing assumption (2.27), one has $M, L<\infty$. Defining

$$f_i=Uf'_i, \quad g_j=Ug'_j, \tag{5.1}$$

one easily verifies that $\{f_i\}_{i=1}^{L}$ and $\{g_j\}_{j=1}^{M}$ are orthonormal bases for $\mathrm{Ker}U^*_{++}$, resp. $\mathrm{Ker}\,U^*_{--}$. From (2.15) and (2.27) we now infer that U_{--}, as an operator from $(\mathrm{Ker}U_{--})^\perp$ to $(\mathrm{Ker}U^*_{--})^\perp$, has a bounded inverse mapping $(\mathrm{Ker}U^*_{--})^\perp$ onto $(\mathrm{Ker}U_{--})^\perp$. We extend this inverse to H_- by setting it equal to zero on $\mathrm{Ker}U^*_{--}$ and denote the resulting operator on H_- by U_{--}^{-1}. In an analogous fashion we define the bounded operator $U^{*\,-1}_{++}$.

Defining a bounded operator Λ by (3.2) it is straightforward to verify, using the unitarity relations (2.15), that (3.3), (3.6), and (3.7) again hold true. However, it should be noticed that (3.4) only holds if $L=M=0$, since it implies that U_{--}, as an operator from H_- to H_-, has the inverse $1_{--}-\Lambda_{--}$.

The next lemma is the generalization of Lemma 4.1. An analogous result has been obtained in Ref. 5.

Lemma 5.1: The following relation holds true:

$$\begin{aligned}\mathcal{U}\Omega=\exp(i\theta)\det(1_{--}+\Lambda_{+-}{}^*\Lambda_{+-})^{-1/2}\prod_{i=1}^{L}a^*(f_i)\\\times\prod_{j=1}^{M}b^*(\bar{g}_j)\exp(\Lambda_{+-}a^*b^*)\,\Omega, \quad 0\le\theta<2\pi.\end{aligned} \tag{5.2}$$

Proof: From

$$\Lambda_{+-}{}^*f_i=\Lambda_{+-}g_j=0, \quad i=1,\ldots,L, \quad j=1,\ldots,M, \tag{5.3}$$

it follows that $a(f_i)$, $b(\bar{g}_j)$ commute with $\exp(\Lambda_{+-}a^*b^*)$.

Consequently the norm of the rhs of (5.2) equals 1. It remains to prove

$$[a(U_{++}f)+b^*(\overline{U_{-+}f})]\prod_{i=1}^{L}a^*(f_i)$$
$$\times\prod_{j=1}^{M}b^*(\overline{g_j})\exp(\Lambda_{+-}a^*b^*)\Omega=0,\quad \forall f\in\mathcal{H}_+, \tag{5.4}$$

$$[b(\overline{U_{--}g})+a^*(U_{+-}g)]\prod_{i=1}^{L}a^*(f_i)$$
$$\times\prod_{j=1}^{M}b^*(\overline{g_j})\exp(\Lambda_{+-}a^*b^*)\Omega=0,\quad \forall g\in\mathcal{H}_-. \tag{5.5}$$

It follows from (5.1) that these relations hold if $f\in\mathrm{Ker}U_{++}$, resp. $g\in\mathrm{Ker}U_{--}$. If $f\in(\mathrm{Ker}U_{++})^\perp$ then $a(U_{++}f)$ in (5.4) (anti)-commutes with $\Pi\cdots\Pi\cdots$. Since

$$\Lambda_{+-}{}^*U_{++}f+U_{-+}f=0\quad \forall f\in(\mathrm{Ker}U_{++})^\perp \tag{5.6}$$

we conclude as in Lemma 4.1 that (5.4) holds. Similarly, (5.5) follows from

$$-\Lambda_{+-}U_{--}g+U_{+-}g=0\quad \forall g\in(\mathrm{Ker}U_{--})^\perp. \tag{5.7}$$

■

We normalize $\mathcal{U}$ by setting

$$\mathcal{U}\Omega=\det(1_{--}+\Lambda_{+-}{}^*\Lambda_{+-})^{-1/2}\prod_{i=1}^{L}a^*(f_i)$$
$$\times\prod_{j=1}^{M}b^*(\overline{g_j})\exp(\Lambda_{+-}a^*b^*)\Omega, \tag{5.8}$$

where the products are in the natural order of the indices. This convention will also be used in the sequel.

Defining the operator $\tilde{\Gamma}(U):D\to D_\infty$ by (4.14), one concludes in the same way as in the proof of Theorem 4.1 that on D,

$$\tilde{\Gamma}(U)a^*(f)=a'^*(f)\tilde{\Gamma}(U)\quad \forall f\in\mathcal{H}_+, \tag{5.9}$$
$$\tilde{\Gamma}(U)b^*(\bar{g})=b'^*(\bar{g})\tilde{\Gamma}(U)\quad \forall g\in\mathcal{H}_-. \tag{5.10}$$

From this proof one also infers that on D

$$\tilde{\Gamma}(U)a^*(f)=\tilde{\Gamma}(U)b^*(\bar{g})=0\quad \forall f\in\mathrm{Ker}U_{++}\quad \forall g\in\mathrm{Ker}U_{--}. \tag{5.11}$$

Hence, by (5.9) and (5.10),

$$b(\overline{f'})\tilde{\Gamma}(U)=a(g')\tilde{\Gamma}(U)=0\quad \forall f'\in\mathrm{Ker}U_{--}^*\quad \forall g'\in\mathrm{Ker}U_{++}^*. \tag{5.12}$$

We note that $\tilde{\Gamma}(-U)$ also satisfies (5.9)—(5.12), apart from a minus sign at the rhs of (5.9) and (5.10).

Now let P be the set of all partitions of the index set $\{1,\ldots,L\}\cup\{1,\ldots,M\}$ into two subsets. P clearly contains 2^{L+M} elements. An element $(\rho,\tau)\in P$ is specified by two subsets $\{\rho_1,\ldots,\rho_l\}\cup\{\tau_1,\ldots,\tau_m\}$ and $\{\rho_{l+1},\ldots,\rho_L\}\cup\{\tau_{m+1},\ldots,\tau_M\}$ in which we take by convention the indices in the natural order. We now define a function on P by

$$\mathrm{sgn}(\rho,\tau)=\mathrm{sgn}(\rho_{l+1},\ldots,\rho_L,\tau_{m+1}+L,\ldots,\tau_M+L,\rho_1,\ldots,\rho_l,\tau_1+L,\ldots,\tau_m+L), \tag{5.13}$$

i.e., $\mathrm{sgn}(\rho,\tau)$ is the sign of the permutation of the indices $\{1,\ldots,L+M\}$ which occur in the rhs of (5.13).

Defining the operator $\mathcal{U}':D\to D_\infty$ by

$$\mathcal{U}'=\sum_{(\rho,\tau)\in P}\mathrm{sgn}(\rho,\tau)\prod_{i=1}^{l}a^*(f_{\rho_i})\prod_{j=1}^{m}b^*(\overline{g_{\tau_j}})\,\tilde{\Gamma}((-)^{L+M}U)$$
$$\times\prod_{i=l+1}^{L}b(\overline{f'_{\rho_i}})\prod_{j=m+1}^{M}a(g'_{\tau_j}) \tag{5.14}$$

we are in a position to state the following theorem.

Theorem 5.1: The operator $\mathcal{U}$ is equal to $\mathcal{U}'$ on D,

$$\mathcal{U}\phi=\sum_{(\rho,\tau)\in P}\mathrm{sgn}(\rho,\tau)\prod_{i=1}^{l}a^*(f_{\rho_i})\prod_{j=1}^{m}b^*(\overline{g_{\tau_j}})\,\tilde{\Gamma}((-)^{L+M}U)$$
$$\cdot\prod_{i=l+1}^{L}b(\overline{f'_{\rho_i}})\prod_{j=m+1}^{M}a(g'_{\tau_j})\,\phi\quad \forall\phi\in D. \tag{5.15}$$

Proof: It follows from (5.8) that $\mathcal{U}'\Omega=\mathcal{U}\Omega$, so it suffices to prove

$$\mathcal{U}'a^*(f)=a'^*(f)\mathcal{U}'\quad \forall f\in\mathcal{H}_+, \tag{5.16}$$
$$\mathcal{U}'b^*(\bar{g})=b'^*(\bar{g})\mathcal{U}'\quad \forall g\in\mathcal{H}_-, \tag{5.17}$$

which should hold on D. To show this, first take $f\in(\mathrm{Ker}U_{++})^\perp$. Then $a^*(f)$ anticommutes with the a and b in (5.14) so we can use (5.9). Both $a^*(U_{++}f)$ and $b(\overline{U_{-+}f})$ now anticommute with all a^* and b^* in (5.14) since $U_{++}f\in(\mathrm{Ker}U_{++}^*)^\perp$ and $U_{-+}f\in(\mathrm{Ker}U_{--}^*)^\perp$. If $L+M$ is odd the resulting minus sign is compensated by the extra minus sign from (5.9). We conclude that (5.16) and, similarly, (5.17) hold true if $f\in(\mathrm{Ker}U_{++})^\perp$, resp. $g\in(\mathrm{Ker}U_{--})^\perp$. It therefore suffices to show

$$\mathcal{U}'a^*(g'_{j_0})=b(\overline{g_{j_0}})\mathcal{U}'\quad \forall j_0\in\{1,\ldots,M\}, \tag{5.18}$$
$$\mathcal{U}'b^*(\overline{f'_{i_0}})=a(f_{i_0})\mathcal{U}'\quad \forall i_0\in\{1,\ldots,L\}. \tag{5.19}$$

To prove (5.18) we observe that from (5.11) it follows that $\mathcal{U}'a^*(g'_{j_0})$ equals the sum of all terms in (5.14) in which the index j_0 is at the right of $\tilde{\Gamma}$, with the factor $a(g'_{j_0})$ suppressed, while from (5.12) it follows that $b(\overline{g_{j_0}})\mathcal{U}'$ is equal to the sum of all terms in which it is at the left, with the factor $b^*(\overline{g_{j_0}})$ suppressed; the terms get an extra minus sign if the number of transpositions required to pull the suppressed factor to the right, resp. the left, is odd. It is easily seen that the same terms occur in the lhs and the rhs of (5.18). To show that they have the same sign, let (ρ,τ) be a partition such that j_0 is at the right of $\tilde{\Gamma}$ and let (ρ,τ') be the corresponding partition, i.e., it equals (ρ,τ) except that j_0 is at the left. We should then prove that

$$\mathrm{sgn}(\rho,\tau')(-)^{l+j'_1-1}=\mathrm{sgn}(\rho,\tau)(-)^{M-j_1}, \tag{5.20}$$

where j'_1,j_1 are such that

$$j_0=\tau_{j_1}=\tau'_{j'_1}. \tag{5.21}$$

However, this follows immediately from (5.13), so (5.18) is proved. The proof of (5.19) is similar. ■

6. THE NECESSITY OF (2.27)

We observe that the sufficiency of (2.27) for implementability, i.e., for the existence of a nonzero vector $\Omega'\in\tilde{D}$ satisfying (2.26), follows from Lemmas 3.1, 4.1, and 5.1. We now give a proof of the necessity of these conditions. For notational convenience (in the boson

case) we again assume (2.1). One easily sees that the result does not depend on this choice.

Theorem 6.1: Let U be a (pseudo-) unitary operator on $\mathcal{H}$. If there exists a nonzero vector $\Omega' \in \tilde{D}$ such that

$$a'(f)\Omega' = b'(\bar{g})\Omega' = 0 \quad \forall f \in \mathcal{H}_+ \quad \forall g \in \mathcal{H}_-, \tag{6.1}$$

where

$$\begin{aligned} a'(f) &= a(U_{++}f) \pm b^*(\overline{U_{-+}f}), \\ b'(\bar{g}) &= b(\overline{U_{--}g}) \pm a^*(U_{+-}g), \end{aligned} \tag{6.2}$$

then U_{+-} and U_{-+} are H.S.

Proof (*A. bosons*): We define projections $P^{n,r}$ $(n, r \in Z)$ by [see (2.2)]

$$(P^{n,r}\psi)^{n',r'} = \delta_{nn'}\,\delta_{rr'}\,\psi^{n,r}, \tag{6.3}$$

where the notation should be clear. One easily verifies relations like

$$P^{n,r}a(f) = a(f)P^{n+1,r}, \tag{6.4}$$

which of course holds on $D(a(f))$. Since $P^{n,r}a'(f)\Omega' = 0$ we have $(n, r \geq 0)$

$$a(U_{++}f)P^{n+1,r}\Omega' = b^*(\overline{U_{-+}f})\,P^{n,r-1}\Omega'. \tag{6.5}$$

It evidently follows from (2.15) that $\mathrm{Ran}U_{++}$ equals $\mathcal{H}_+$, so from (6.5) we conclude that if $P^{n,r-1}\Omega' = 0$, then also $P^{n+1,r}\Omega' = 0$. This implies

$$P^{n+k,k}\Omega' = 0 \quad \forall\, n > 0 \quad \forall\, k \geq 0. \tag{6.6}$$

From $P^{n,r}b'(\bar{g})\Omega' = 0$ we infer analogously

$$P^{l,r+l}\Omega' = 0 \quad \forall\, r > 0 \quad \forall\, l \geq 0. \tag{6.7}$$

Since $\|\Omega'\| \neq 0$ it follows from (6.6), (6.7), and the argument given above that we must have

$$P^{0,0}\Omega' \equiv \alpha \neq 0. \tag{6.8}$$

Defining

$$\psi(p,q) = (1/\alpha)(P^{1,1}\Omega')^{1,1}(p,q), \tag{6.9}$$

one has, using $P^{0,1}a'(f)\Omega' = 0$, resp. $P^{1,0}b'(\bar{g})\Omega' = 0$, and (2.3),

$$\int dp\,(\overline{U_{++}f})(p)\,\psi(p,q) = (\overline{U_{-+}f})(q) \quad \forall f \in \mathcal{H}_+, \tag{6.10}$$

$$\int dq\,(U_{--}g)(q)\,\psi(p,q) = (U_{+-}g)(p) \quad \forall\, g \in \mathcal{H}_-. \tag{6.11}$$

Introducing a H.S. operator $H_B : \mathcal{H}_- \to \mathcal{H}_+$ by

$$(H_B g)(p) = \int dq\,\psi(p,q)\,g(q), \tag{6.12}$$

we can write (6.10), resp. (6.11), as

$$H_B^* U_{++} = U_{-+} \tag{6.13}$$

$$H_B U_{--} = U_{+-}. \tag{6.14}$$

We conclude that U_{-+} and U_{+-} are H.S.

(*B. fermions*): Let $\{f_i\}_{i=1}^{L}$ and $\{g_j\}_{j=1}^{M}$ be orthonormal (o.n.) bases for $\mathrm{Ker}U_{++}^*$, resp. $\mathrm{Ker}U_{--}^*$. Let $\{f_i\}_{i=L+1}^{\infty}$ and $\{g_j\}_{j=M+1}^{\infty}$ be o.n. bases for $\overline{\mathrm{Ran}U_{++}}$, resp. $\overline{\mathrm{Ran}U_{--}}$. Then $\{f_i\}_{i=1}^{\infty}$ and $\{g_j\}_{j=1}^{\infty}$ obviously are o.n. bases for $\mathcal{H}_+$ resp. $\mathcal{H}_-$. We now introduce an o.n. basis for $\mathcal{F}_a$ by setting

$$\phi_{\rho_1,\rho_2,\ldots;\tau_1,\tau_2,\ldots} \equiv \prod_{i=1}^{\infty} a^*(f_i)^{\rho_i} \prod_{j=1}^{\infty} b^*(\bar{g}_j)^{\tau_j}\Omega, \tag{6.15}$$

where

$$\rho_i, \tau_j = 0, 1, \quad \sum_{i=1}^{\infty}\rho_i + \sum_{j=1}^{\infty}\tau_j < \infty. \tag{6.16}$$

Then

$$\Omega' = \sum_{\rho_i,\tau_j} \alpha_{\rho_1,\ldots;\tau_1,\ldots}\,\phi_{\rho_1,\ldots;\tau_1,\ldots}. \tag{6.17}$$

From (6.1) and (6.2) [cf. Sec. 5, esp. (5.1)],

$$a^*(f_i)\Omega' = b^*(\bar{g}_j)\Omega' = 0, \quad i = 1,\ldots,L, \quad j = 1,\ldots,M. \tag{6.18}$$

Using (6.17) one now concludes that $L, M < \infty$ (this was anticipated above for notational convenience) and that

$$\sum_{i=1}^{L}\rho_i < L \quad \text{or} \quad \sum_{j=1}^{M}\tau_j < M \Rightarrow \alpha_{\rho_1,\ldots;\tau_1,\ldots} = 0. \tag{6.19}$$

Thus,

$$P^{n,r}\Omega' = 0 \quad \forall\, n < L \quad \forall\, r < M, \tag{6.20}$$

$$P^{L,M}\Omega' = \beta \prod_{i=1}^{L} a^*(f_i) \prod_{j=1}^{M} b^*(\bar{g}_j)\Omega, \tag{6.21}$$

$$P^{L+1,M+1}\Omega' = \sum_{k=L+1}^{\infty}\sum_{l=M+1}^{\infty}\gamma_{kl}\,a^*(f_k)\,b^*(\bar{g}_l)\,P^{L,M}\Omega'. \tag{6.22}$$

From $P^{L+n,M+r}a'(f)\Omega' = 0$ $(n, r \geq 0)$ it follows that

$$a(U_{++}f)P^{L+n+1,M+r}\Omega' = -\,b^*(\overline{U_{-+}f})\,P^{L+n,M+r-1}\Omega'. \tag{6.23}$$

Using (6.19) one easily concludes that (6.23) implies: If $P^{L+n,M+r-1}\Omega' = 0$ then $P^{L+n+1,M+r}\Omega' = 0$. Hence, from (6.20),

$$P^{L+n+k,M+k}\Omega' = 0 \quad \forall\, n > 0 \quad \forall\, k \geq 0. \tag{6.24}$$

Analogously,

$$P^{L+l,M+r+l}\Omega' = 0 \quad \forall\, r > 0 \quad \forall\, l \geq 0. \tag{6.25}$$

We therefore must have $\beta \neq 0$ in (6.21).

Defining a H.S. operator H_F: $\overline{\mathrm{Ran}U_{--}} \to \overline{\mathrm{Ran}U_{++}}$ by

$$H_F g = \sum_{k=L+1}^{\infty}\sum_{l=M+1}^{\infty} f_k\,\gamma_{kl}\,(g_l, g), \tag{6.26}$$

one infers from (6.21) and (6.22), using $P^{L,M+1}a'(f)\Omega' = 0$, resp. $P^{L+1,M}b'(\bar{g})\Omega' = 0$,

$$H_F^* U_{++} f = -\,U_{-+}f \quad \forall\, f \in (\mathrm{Ker}U_{++})^{\perp}, \tag{6.27}$$

$$H_F U_{--} g = U_{+-}g \quad \forall\, g \in (\mathrm{Ker}U_{--})^{\perp}. \tag{6.28}$$

Thus, U_{-+} and U_{+-} are direct sums of a H.S. operator and a finite-rank operator. Therefore, U_{-+} and U_{+-} are H.S. ∎

We finally make some remarks about unbounded pseudo-unitary operators. It seems reasonable to require that (2.11) hold on a dense subspace M belonging to the domains of U and U^* and invariant under $P_\pm, U$ and U^*. If U_{+-} and U_{-+} are H.S. one concludes from (2.15), which holds on M, and from the relation

$$(U_{\epsilon\epsilon'} \restriction M)^* \supset U^*_{\epsilon'\epsilon} \restriction M, \tag{6.29}$$

which follows from our assumptions, that U must be bounded.

On the other hand, if the conditions of Theorem 6.1 are met (for any $f, g \in M$), one is again led to (6.13) and

(6.14) which now hold on M [use (2.15) and (6.29) to establish that $\mathrm{Ran}(U_{++}\restriction M)$ and $\mathrm{Ran}(U_{--}\restriction M)$ are dense in $\mathcal{H}_+$, resp. $\mathcal{H}_-$]. Now from (6.29) and (2.15) it follows that $A \equiv U_{--}\restriction M$ has a bounded inverse A^{-1} and that

$$A^{-1*}A^{-1} = 1_{--} - H_B^* H_B \tag{6.30}$$

on $U_{--}M$, where (6.14) has been used. In virtue of the relation $\overline{A^{-1}} = \overline{A^{-1}}$ it follows from (6.30) that

$$\overline{A}^{*-1}\overline{A}^{-1} = 1_{--} - H_B^* H_B \tag{6.31}$$

on $\mathcal{H}_-$. Since H_B is compact,

$$\|A^{-1}g\| \geq \epsilon\|g\|, \quad \epsilon > 0 \quad \forall\, g \in U_{--}M. \tag{6.32}$$

Thus, $U_-\restriction M$ is bounded. Similarly, $U_{++}\restriction M$ is bounded, so U must be bounded. We conclude that unbounded pseudo-unitary operators (as defined above) cannot give rise to implementable Bogoliubov transformations.

ACKNOWLEDGMENT

This investigation is part of the research program of the "Stichting voor Fundamenteel Onderzoek der Materie (F.O.M.)," which is financially supported by the "Organisatie voor Zuiver Wetenschappelijk Onderzoek (Z.W.O)."

[1] P. J. M. Bongaarts, Ann. Phys. (N.Y.) **56**, 108—39 (1970).
[2] P. J. M. Bongaarts and S. N. M. Ruijsenaars, "The Klein paradox as a many particle problem," Ann. Phys. (N.Y.) (to be published).
[3] R. Seiler, Commun. Math. Phys. **25**, 127—51 (1972).
[4] A. S. Wightman, Symp. Pure Math. **23**, 441—77 (1973).
[5] G. Labonté, Commun. Math. Phys. **36**, 59—72 (1974).
[6] K. O. Friedrichs, *Mathematical aspects of the quantum theory of fields* (Interscience, New York, 1953).
[7] F. A. Berezin, *The method of second quantization* (Academic, New York, 1966).
[8] P. Kristensen, L. Mejlbo, and E. Thue Poulsen, Commun. Math. Phys. **6**, 29—48 (1967).
[9] S. N. M. Ruijsenaars, "Charged particles in external fields. I. Classical theory," to be published in J. Math. Phys.; "Charged particles in external fields. II. The quantized Dirac and Klein—Gordon theories," Leiden University preprint (1976).
[10] M. Weinless, J. Funct. An. 4, 350—79 (1969).
[11] J. M. Cook, Trans. Am. Math. Soc. **74**, 222—45 (1953).
[12] D. Shale, Trans. Am. Math. Soc. **103**, 149—67 (1962).
[13] J. M. Chadam, J. Math. Phys. **9**, 386—96 (1968).
[14] D. Shale and W. F. Stinespring, J. Math. Mech. **14**, 315—22 (1965).
[15] M. Reed and B. Simon, *Functional Analysis* (Academic, New York, 1972).
[16] J. R. Ringrose, *Compact non-self-adjoint operators* (Van Nostrand Reinhold, London, 1971).
[17] N. Dunford and J. T. Schwartz, *Linear operators, Part II.* (Interscience, New York, 1958).
[18] I. C. Gohberg and M. G. Krein, *Introduction to the theory of linear nonselfadjoint operators* (American Mathematical Society, Providence, R.I., 1969).
[19] B. Simon, "Notes on infinite determinants of Hilbert space operators," Princeton University preprint (1975).

The singular coupling and weak coupling limits

P. F. Palmer*

Mathematical Institute, University of Oxford, Oxford, OX1 3LB England
(Received 17 May 1976)

We show that the analysis of the motion of a system coupled to an infinite reservoir in the singular coupling limit and the weak coupling limit can be performed in the same mathematical framework. This enables us to clarify the way in which the motion of the system in the singular coupling limit depends on the temperature of the reservoir.

1. INTRODUCTION

Open quantum mechanical systems have recently been studied to provide some understanding of irreversible behavior in quantum statistical mechanics and in certain limiting situations it has been possible to prove rigorously that the time evolution of an open system obeys a semigroup law.[1-4] It seems that two distinct limiting procedures are possible, the weak coupling limit,[1,4] and the singular coupling limit.[2,3]

In this paper we show that the two limits are mathematically very similar and that the distinction depends on which of two possible time scales is regarded as natural or "physical". We describe a simple singular coupling model and show how by suitable rescaling it may be considered as a weak coupling model. By applying the general theory of Ref. 1 to this case, we consider how the evolution of a system coupled to a heat bath depends on the temperature in the singular coupling limit.

2. THE SINGULAR COUPLING MODEL

We describe a model of a system coupled to a heat bath of fermions, which is essentially that used by Hepp and Lieb,[2] and Gorini and Kossakowski.[3]

The system is described by a Hilbert space $\mathcal{H}_S$ with free Hamiltonian H_S. The heat bath is described by a quasifree representation of the CAR with an infinite number of degrees of freedom. To be specific we take the single particle space $\mathcal{V}^\omega$ to be either $L^2(\mathbb{R}, d\omega)$ or $L^2((0,\infty), d\omega)$, where $d\omega$ is Lebesgue measure. The single particle free evolution is given by

$$f_\tau(\omega) = \exp(i\omega\tau) f(\omega), \quad f \in \mathcal{V}^\omega. \tag{2.1}$$

For each $f \in \mathcal{V}^\omega$ we have a bounded operator $\phi^\omega(f)$ on a space $\mathcal{H}_B$ satisfying the CAR

$$\phi^\omega(f)\phi^\omega(g) + \phi^\omega(g)\phi^\omega(f) = 2\mathrm{Re}\langle f, g\rangle \tag{2.2}$$

There is a cyclic vector Ω in $\mathcal{H}_B$ and a Hamiltonian H_B^ω on $\mathcal{H}_B$ such that

$$H_B^\omega \Omega = 0 \tag{2.3}$$

$$\exp(iH_B^\omega \tau)\phi^\omega(f)\exp(-iH_B^\omega \tau) = \phi^\omega(f_\tau). \tag{2.4}$$

The representation is determined by its two point correlation functions. For a bath at inverse temperature β and chemical potential μ we know[5,6]

$$\langle \Omega, \phi^\omega(f)\phi^\omega(g)\Omega\rangle$$

$$= \langle f, g\rangle - 2i\,\mathrm{Im}\langle f, \{1 + \exp[\beta(Q - \mu)]\}^{-1} g\rangle, \tag{2.5}$$

where

$$Qf(\omega) = \omega f(\omega). \tag{2.6}$$

The Hilbert space of the composite system is

$$\mathcal{H} = \mathcal{H}_S \otimes \mathcal{H}_B \tag{2.7}$$

and the full Hamiltonian is

$$H_\lambda = H_S \otimes 1 + 1 \otimes H_B^\omega + H_I, \tag{2.8}$$

where

$$H_I = Q \otimes \phi^\omega(f^\lambda). \tag{2.9}$$

Here Q is an arbitrary bounded self-adjoint operator on $\mathcal{H}_S$, $f \in \mathcal{V}^\omega$, and

$$f^\lambda(\omega) = f(\lambda^2\omega). \tag{2.10}$$

With the correct choice of f, β, μ the time correlation function

$$\langle \phi^\omega(f_\tau^\lambda)\phi^\omega(f^\lambda)\rangle_{\beta,\mu} \tag{2.11}$$

can be shown to become a δ function in the limit $\lambda \to 0$. This is the "singular coupling limit," and in this limit we may expect the states of the system to obey a Markovian time evolution.[2,3]

3. THE EQUIVALENT WEAK COUPLING PROBLEM

We call τ the "slow" time and define a rescaled "fast" time t by

$$\lambda^2 t = \tau. \tag{3.1}$$

We also rescale the underlying variable ω by

$$w = \lambda^2\omega \tag{3.2}$$

and let $\mathcal{V}^w$ be $L^2(\mathbb{R}, dw)$ or $L^2((0,\infty), dw)$ according to the choice of $\mathcal{V}^\omega$ made in Sec. 2.

The unitary dilation

$$f \to \lambda f^\lambda \tag{3.3}$$

gives a unitary map

$$U_\lambda : \mathcal{V}^w \to \mathcal{V}^\omega. \tag{3.4}$$

Using U_λ and ϕ^ω we can construct a representation of the CAR over $\mathcal{V}^w$ on the same space $\mathcal{H}_B$, and then construct a free bath Hamiltonian H_B^w from the single particle evolution

$$f_t(w) = \exp(iwt) f(w). \tag{3.5}$$

The correspondence between the singular coupling problem and the weak coupling problem is contained in the following rather easy theorem.

Theorem 3.1: With the above notation let H_λ^ω be the singular coupling Hamiltonian

$$H_\lambda^\omega = H_S \otimes 1 + 1 \otimes H_B^\omega + Q \otimes \phi^\omega(f_\lambda) \tag{3.6}$$

and H_λ^w be the weak coupling Hamiltonian

$$H_\lambda^w = \lambda^2 H_S \otimes 1 + 1 \otimes H_B^w + \lambda Q \otimes \phi^w(f). \tag{3.7}$$

Then

$$\exp(iH_\lambda^\omega \tau) = \exp(iH_\lambda^w t). \tag{3.8}$$

Proof: Since $\exp(i\omega\tau) = \exp(iwt)$ it follows that

$$\exp(iH_B^\omega \tau) = \exp(iH_B^w t). \tag{3.9}$$

ϕ^ω, ϕ^w satisfy their respective forms of the CAR so that

$$\lambda \phi^\omega(f_\lambda) = \phi^w(f) \tag{3.10}$$

and Eq. (3.8) follows from the fact that $\tau = \lambda^2 t$.

If β^ω is the temperature of the bath in the initial problem, when we rescale we must ensure that

$$\beta^\omega(H_B^\omega - \mu^\omega) = \beta^w(H_B^w - \mu^w), \tag{3.11}$$

where β^w is the temperature of the rescaled problem and μ^ω, μ^w are the initial and rescaled chemical potentials, respectively.

Since

$$H_B^w = \lambda^2 H_B^\omega \tag{3.12}$$

then we must have

$$\beta^w = \lambda^{-2}\beta^\omega, \quad \mu^w = \lambda^2 \mu^\omega. \tag{3.13}$$

Thus we must in general allow a temperature rescaling, except when $\beta^\omega = 0$ (infinite temperature limit) or $1/\beta^\omega = 0$ (zero temperature).

In order to compare these two limits we extend the results of Ref. 1. First, for $\rho \in \mathcal{T}(\mathcal{H}_S)$, the space of trace class operators on $\mathcal{H}_S$, let

$$\rho_\lambda(t) = \mathrm{tr}_{\mathcal{H}_B}\{\exp(-iH_\lambda^w t)(\rho \otimes |\Omega\rangle\langle\Omega|)\exp(iH_\lambda^w t)\} \tag{3.14}$$

and let

$$h_\lambda(t) = \langle \Omega, \phi^w(f_t)\phi^w(f)\Omega\rangle_{\beta,\mu}, \tag{3.15}$$

where we allow β, μ and hence h_λ to depend on λ.

Theorem 3.2: Suppose that H_S is bounded. Suppose also that for some $g(t)$, $t \geq 0$, $\epsilon > 0$,

$$|h_\lambda(t)| \leq g(t) \quad t \geq 0 \tag{3.16}$$

and

$$\int_0^\infty g(t)(1+t)^\epsilon \, dt < \infty. \tag{3.17}$$

If $h_\lambda(t) \to h(t)$ pointwise for $t \geq 0$ then for all $\tau_0 \geq 0$

$$\lim_{\lambda \to 0} \sup_{0 \leq t \leq \lambda^{-2}\tau_0} \{||\rho_\lambda(t) - U_t^\lambda(\rho)||_{\mathrm{tr}}\} = 0, \tag{3.18}$$

where

$$U_t = \exp[(Z+K)\lambda^2 t]. \tag{3.19}$$

Here

$$Z(\rho) = -i[H_S, \rho] \tag{3.20}$$

and

$$K(\rho) = -cQQ\rho + \bar{c}Q\rho Q + cQ\rho Q - \bar{c}\rho QQ \tag{3.21}$$

where

$$c = \int_0^\infty h(t)\, dt. \tag{3.22}$$

Proof: This theorem may be proved by the method of Ref. 1, Theorems (3.1)–(3.5) with the following modifications[7]:

(i) We require H_S bounded so that

$$\lim_{\lambda \to 0} ||\exp(i\lambda^2 H_S t) - 1|| = 0. \tag{3.23}$$

(ii) The interaction is linear rather than quadratic in the field operators. This only makes the proof easier.

(iii) The system evolution is λ dependent. We note that in Ref. 1 all terms arising are eventually written as a sum of tensor products and simple norm estimates made on the system variables. This λ dependence disappears at this stage.

(iv) The error estimates eventually involve multiple integrals of h_λ. We may compute these estimates uniformly in λ using $g(t)$

(v) $$c_\lambda = \int_0^\infty h_\lambda(t)\, dt \to c \tag{3.24}$$

by dominated convergence theorem.

Remarks: (i) This theorem holds for a general H_B^w, not just the specific one we have discussed.

(ii) This result holds for systems linearly coupled to Boson heat baths. The proof may be on the lines above, though the usual difficulties with unbounded field operators arise. See also Ref. 4.

(iii) In the limit $\lambda \to 0$, the spectrum of $\lambda^2 H_S$ is $\{0\}$ and K does not depend on H_S. We contrast this with the result of Ref. 1 for the weak coupling problem with

$$H_\lambda^w = H_S \otimes 1 + 1 \otimes H_B^w + \lambda Q \otimes \phi^w(f). \tag{3.25}$$

For this H_λ^w the properties of K depend strongly on H_S.

4. THE SINGULAR COUPLING LIMIT

We use the general theory of Sec. 3 to study the dynamics of the system of Sec. 2 in the singular coupling limit.

In the case of the vacuum $h_\lambda(t)$ is independent of λ,

$$h_\lambda(t) = \int_{-\infty}^\infty \exp(-iwt)|f(w)|^2\, dw. \tag{4.1}$$

If

$$c = \int_0^\infty h(t)\, dt \tag{4.2}$$

then

$$\begin{aligned} 2\mathrm{Re}\, c &= \int_{t=0}^\infty \int_{w=-\infty}^\infty \exp(-iwt)|f(w)|^2\, dw \\ &\quad + \int_{t=0}^\infty \int_{w=-\infty}^\infty \exp(iwt)|f(w)|^2\, dw \\ &= \int_{t=-\infty}^\infty \int_{-\infty}^\infty \exp(-iwt)|f(w)|^2\, dw \\ &= 2\pi |f(0)|^2, \end{aligned} \tag{4.3}$$

so that if $f(0) = 0$, c is imaginary

$$c = id, \quad d \in \mathbb{R}, \tag{4.4}$$

and

$$K(\rho) = -id QQ\rho - id Q\rho Q + id Q\rho Q + id\rho QQ$$

$$= -id[Q^2, \rho]. \tag{4.5}$$

The condition (3.20) is satisfied if $|f(w)|^2$ has a continuous $(1+\epsilon)$th derivative which is integrable. Certainly it is enough that f is twice continuously differentiable and f', $f'' \in L^2$. However in the semibounded case we also require that at least

$$f(0) = f'(0) = 0. \tag{4.6}$$

Thus we obtain dissipative behavior only in the unphysical case where the bath spectrum is unbounded below. In the semibounded case K only gives a shift of energy levels.

In the infinite temperature limit $\beta^\omega = 0$ and

$$h_\lambda(t) = \tfrac{1}{2}\int_{-\infty}^{\infty} \exp(-iwt) g(w)\, dw, \tag{4.7}$$

where

$$g(w) = |f(w)|^2 + |f(-w)|^2. \tag{4.8}$$

Clearly the nonsemibounded and semibounded cases are essentially the same since if $\mathrm{supp} f \subseteq (0, \infty)$ we may use $(1/\sqrt{2}) f(|w|)$ instead of f.

$\mathrm{Im} c = 0$, since g is even and

$$c = \pi |f(0)|^2 \tag{4.9}$$

so

$$K(\rho) = -cQQ\rho + 2cQ\rho Q - c\rho QQ. \tag{4.10}$$

This generates a Gaussian semigroup and the limiting evolution is dissipative.

When $\beta^\omega = \beta$, $\mu^\omega = \mu$ with β, μ constants and $\beta > 0$ then we must rescale the temperatures β^w, and chemical potential μ^w by

$$\beta^w = \beta/\lambda^2, \quad \mu^w = \lambda^2 \mu. \tag{4.11}$$

Now

$$h_\lambda(t) = \int_{-\infty}^{\infty} \frac{\exp(iwt)|f(w)|^2}{\{\exp[\lambda^{-2}\beta(w-\lambda^2\mu)]+1\}} + \frac{\exp(-iwt)|f(w)|^2}{\{\exp[-\lambda^{-2}\beta(w-\lambda^2\mu)]+1\}}\, dw. \tag{4.12}$$

Since $f \in L^2$ the dominated convergence theorem shows that

$$h_\lambda(t) \to h(t) = \int_{-\infty}^{0} \exp(iwt)|f(w)|^2\, dw + \int_0^{\infty} \exp(-iwt)|f(w)|^2\, dw. \tag{4.13}$$

The condition (3.16) is satisfied if the integrands of Eq. (4.12) have $(1+\epsilon)$th derivatives with uniform L^1 bound. By inspection we require at least that

$$f(0) = f'(0) = 0. \tag{4.14}$$

As was the case with the vacuum, K therefore only gives a shift of energy levels. Note also that K does not depend on β.

Davies[8] has shown that where it is not true that $f(0) = f'(0) = 0$ we may expect non-Markovian behavior. This provides some illumination on the analysis of Frigerio and Gorini[9] who discovered non-Markovian behavior for the particular test function $f(w) = \exp(-w^2)$ at finite nonzero temperatures.

ACKNOWLEDGMENTS

The author would like to thank his supervisor, Dr. E.B. Davies, for his encouragement and support.

*Financial support given by the Science Research Council and International Computers Limited.

[1] E.B. Davies, Commun. Math. Phys. 39, 91–110 (1974).
[2] K. Hepp and E.H. Lieb, Helv. Phys. Acta 46, 573–603 (1973).
[3] V. Gorini and A. Kossakowski, J. Math. Phys. (to be published).
[4] J.V. Pulè, Commun. Math. Phys. 38, 241–56. (1974).
[5] H. Araki and W. Wyss, Helv. Phys. Acta 37, 137–59 (1964).
[6] E. Balslev and A. Verbeure, Commun. Math. Phys. 7, 55–76 (1968).
[7] P.F. Palmer, Ph.D. thesis, Oxford University (1976).
[8] E.B. Davies, J. Funct. Anal. 17, 2123 (1976).
[9] A. Frigerio and V. Gorini, J. Math. Phys. (to be published).

Soliton solutions and the higher order Korteweg–de Vries equations

H. C. Morris

School of Mathematics, Trinity College, Dublin, Ireland
(Received 15 March 1976)

A prolongation structure is determined for a single equation from which the Korteweg-de Vries and all of its higher order forms can be derived. As result, inverse scattering problems for all of the equations are determined simultaneously.

1. INTRODUCTION

It is well known that the Korteweg—deVries equation

$$u_t = u_{xxx} + 6uu_x \tag{1.1}$$

owes many of its remarkable properties to the fact that it admits a Lax representation[1] in the form

$$\dot{L} = [B_2, L], \tag{1.2}$$

where the differential operators L and B_2 are given by

$$L = \frac{d^2}{dx^2} + u \quad \text{and} \quad B_2 = 4\frac{d^3}{dx^3} + 3\left(u\frac{d}{dx} + \frac{d}{dx}u\right). \tag{1.3}$$

One of the most interesting features of the Korteweg—de Vries equation (1.1) is that is posses an infinite number of conservation laws[2] of the form

$$I_n(u) = \int_{-\infty}^{\infty} P_n(u, \ldots, u^{(n)})\, dx, \tag{1.4}$$

where $P_n(u_1, \ldots, u^{(n)})$ is a polynomial in the first $n+1$ space derivatives of u, $u^1, u^2, \ldots, u^{(n)}$.

Lax and Gardner[1,2] have shown that each of the known polynomial integrals $I_n(u)$ determines an equation

$$\dot{u} = \frac{d}{dx}\left(\frac{\delta I_n}{\delta u(x)}\right) \tag{1.5}$$

admitting a Lax representation

$$\dot{L} = [B_n, L], \tag{1.6}$$

where the B_n are skew symmetric differential operators of order $(2n+1)$. The first three of the conserved quantities are given by

$$I_1 = \int_{-\infty}^{\infty} 2u^2\, dx, \tag{1.7}$$

$$I_2 = \int_{0}^{\infty} [(u_x)^2 + u^3]\, dx, \tag{1.8}$$

$$I_3 = \int_{-\infty}^{\infty} \tfrac{1}{8}[(u_{xx})^2 + 5u^2u_{xx} + 5u^4]\, dx, \tag{1.9}$$

and by Eq. (1.5) they give rise to the equations

$$u_t = 4u_x, \tag{1.10}$$

$$u_t = u_{xxx} + 6uu_x, \tag{1.11}$$

$$u_t = \tfrac{1}{4}u_{xxxxx} + 5u_xu_{xx} + \tfrac{5}{2}u_{xxx}u + \tfrac{15}{2}u^2u_x. \tag{1.12}$$

Equation (1.12) and the equations which result from I_n, $n \geq 3$ are known as the higher order Korteweg—de Vries equations. In the following section we will show how all of these equations can be determined from the single equation

$$u_t = \tfrac{1}{2}C_{xxx} + 2C_x(u-\lambda) + u_xC \tag{1.13}$$

which we will refer to as the generating equation. Equation (1.13) has previously been discussed by Gardner, Green, and Miura.[3]

In Sec. 3 we will determine a prolongation structure[4] for Eq. (1.13) and by so doing obtain simultaneously prolongation structures for all of the Korteweg—de Vries equations. In Sec. 4 we construct one and two dimensional representations of the prolongation structure and examine the resulting inverse scattering problems that they lead to.

2. THE GENERATING EQUATION

The equation

$$\tfrac{1}{2}C_{xxx} + 2C_x(u-\lambda) + Cu_x = u_t \tag{2.1}$$

yields the Koretweg— de Vries equations in the following way. If we suppose that $u(x,t)$ is independent of λ and that $C(u,\lambda)$ can be expressed as a polynomial of degree n in λ given by

$$C(u,\lambda) = \sum_{i=0}^{n} C_i(u)\lambda^{n-i}, \tag{2.2}$$

then Eq. (2.1) yields the following relationships:

$$\tfrac{1}{2}C_{ixxx} + 2C_{ix}u + C_iu_x = 2C_{i+1x} \quad i = 1, \ldots, n-1, \tag{2.3}$$

$$\tfrac{1}{2}C_{nxxx} + 2C_{nx}u + C_nu_x = u_t. \tag{2.4}$$

Thus if we define the sequence of functions $\{C_n\}_{n=0}^{\infty}$ by the recurrence relation

$$\tfrac{1}{2}C_{nxxx} + 2C_{nx}u + C_nu_x = 2C_{n+1\,x} \tag{2.5}$$

it is clear that we are considering the evolution equations

$$u_t = (2C_n)_x \quad n = 1, 2, \ldots, \infty. \tag{2.6}$$

If we choose

$$C_0 = 4, \tag{2.7}$$

then the early members of the sequence $\{C_n(u)\}_{n=0}^{\infty}$ are given by

$$C_0 = 4, \quad C_1 = 2u, \quad C_2 = \tfrac{1}{2}(u_{xx} + 3u^2)$$

$$C_3 = \tfrac{1}{8}u_{xxxx} + \tfrac{5}{4}u_{xx}u + \tfrac{5}{8}(u_x)^2 + \tfrac{5}{4}u^3, \tag{2.8}$$

which by (2.6) gives rise to the Eqs. (1.10)–(1.12).

Clearly we must have

$$C_n(u) = \frac{1}{2}\frac{\delta I_n}{\delta u} \tag{2.9}$$

and so the recurrence relation (2.5) is capable of generating the polynomial conserved quantities $\{I_n\}_{n=1}^{\infty}$.

3. A PROLONGATION STRUCTURE FOR THE GENERATING EQUATION

If we define $U=u-\lambda$ then a set of forms equivalent to the equation

$$\tfrac{1}{2}C_{xxx}+2C_xU+U_xC=U_t \tag{3.1}$$

is given by

$$\alpha_1=dC\wedge dt-Z\,dx\wedge dt, \tag{3.2}$$

$$\alpha_2=dZ\wedge dt-P\,dx\wedge dt, \tag{3.3}$$

$$\alpha_3=dU\wedge dx+\tfrac{1}{2}dP\wedge dt+U\,dC\wedge dt+C\,dU\wedge dt + ZU\,dx\wedge dt. \tag{3.4}$$

It may be easily checked that

$$d\alpha_1=\alpha_2\wedge dx, \tag{3.5}$$

$$d\alpha_2=2\alpha_3\wedge(dx+C\,dt)-2\alpha_1\wedge(U\,dx), \tag{3.6}$$

$$d\alpha_3=\alpha_3\wedge(Z\,dt)-\alpha_2\wedge(U\,dx). \tag{3.7}$$

We seek a prolongation structure in the form

$$\Omega=d\xi+F(C,Z,P,U,\xi)\,dx+G(C,Z,P,U,\xi)\,dt \tag{3.8}$$

and by the standard procedure[4] we arrive at the equations

$$[F,G]=Z\left(\frac{\partial G}{\partial C}-U\frac{\partial F}{\partial U}\right)+P\frac{\partial G}{\partial Z}-ZU\frac{\partial F}{\partial U},$$

$$2\frac{\partial G}{\partial P}=C^{-1}\frac{\partial G}{\partial U}=\frac{\partial F}{\partial U}, \tag{3.9}$$

$$\frac{\partial F}{\partial C}=\frac{\partial F}{\partial Z}=\frac{\partial F}{\partial P}=0,$$

where $[F,G]$ is the Lie bracket defined by

$$[F,G]=(FG_\xi-GF_\xi). \tag{3.10}$$

We can determine a solution to (3.9) with F and G having the forms

$$F=X_1+UX_2 \tag{3.11}$$

$$G=\tfrac{1}{2}PX_2+CUX_2+ZX_3+CX_4+X_5. \tag{3.12}$$

The substitution of (3.11) and (3.12) into (3.9) yields the Lie bracket relations

$$\begin{aligned}&[X_1,X_2]=2X_3,\quad [X_1,X_4]=0,\quad [X_1,X_3]=X_4,\\ &[X_2,X_3]=-X_2,\quad [X_2,X_4]=-2X_3,\quad [X_5,X_1]=0,\\ &[X_5,X_2]=0.\end{aligned} \tag{3.13}$$

The Jacobi identity supplies the bracket relations

$$[X_3,X_4]=-X_4,\quad [X_5,X_3]=0,\quad [X_5,X_4]=0, \tag{3.14}$$

which completes the algebraic relations (3.13) into the Lie algebra

$$\begin{aligned}&[X_1,X_2]=2X_3,\quad [X_1,X_3]=X_4,\quad [X_1,X_4]=0,\quad [X_1,X_5]=0,\\ &[X_2,X_3]=-X_2,\quad [X_2,X_4]=-2X_3,\quad [X_2,X_5]=0,\\ &[X_3,X_4]=-X_4,\quad [X_3,X_5]=0,\\ &[X_4,X_5]=0.\end{aligned} \tag{3.15}$$

Examining this Lie algebra we see that we can consistently make the identification,

$$X_1=X_4\,, \tag{3.16}$$

which simplifies the algebra to

$$[X_1,X_2]=2X_3,\quad [X_1,X_3]=X_1,\quad [X_2,X_3]=-X_2\,, \tag{3.17}$$

$$[X_i,X_5]=0\,,\quad i=1,2,3. \tag{3.18}$$

There are two simple representations of the algebra (3.17) and (3.18) which prove useful and we construct these in the following section.

4. REPRESENTATIONS OF THE PROLONGATION STRUCTURE

(a) A one-dimensional representation of the Lie algebra (3.17)–(3.18) is given by

$$X_1=\phi^2\frac{\partial}{\partial\phi}\,,\quad X_2=\frac{\partial}{\partial\phi}\,,\quad X_3=-\phi\frac{\partial}{\partial\phi}\,, \tag{4.1}$$

$$X_5=0.$$

The resulting one-dimensional prolongation structure is given by

$$\Omega=d\phi+(U+\phi^2)\,dx+[(\tfrac{1}{2}P+Cu)+C\phi^2-Z\phi]\,dt. \tag{4.2}$$

Sectioning onto a solution manifold of Eq. (3.1) gives the inverse scattering problem

$$\phi_x=-U-\phi^2 \tag{4.3}$$

$$\phi_t=-(\tfrac{1}{2}P+CU)+Z\phi-C\phi^2. \tag{4.4}$$

(b) A two-dimensional representation is given by

$$\begin{aligned}&X_1=-\xi^2b_1,\quad X_2=\xi^2b_2,\\ &X_3=\tfrac{1}{2}(\xi^1b_1-\xi^2b_2),\quad X_5=\mu(\xi^1b_1+\xi^2b_2),\end{aligned} \tag{4.5}$$

where $b_i=\partial/\partial\xi^i$ and μ is completely arbitrary and may be a function of λ. This representation gives rise to the prolongation structure

$$\begin{aligned}&\Omega^1=d\xi^1-\xi^2\,dx+[(\tfrac{1}{2}Z+\mu)\xi^1-C\xi^2]\,dt\\ &\Omega^2=d\xi^2+\xi^1U\,dx+[(\tfrac{1}{2}P+CU)\xi^1+(\mu-\tfrac{1}{2}Z)\xi^2]\,dt.\end{aligned} \tag{4.6}$$

The related inverse scattering problem is given by

$$\begin{aligned}&\xi_x=\begin{bmatrix}0 & 1\\ -U & 0\end{bmatrix}\xi,\\ &\xi_t=\begin{bmatrix}-(\tfrac{1}{2}Z+\mu), & C\\ -\tfrac{1}{2}P-CU, & (\tfrac{1}{2}Z-\mu)\end{bmatrix}\xi.\end{aligned} \tag{4.7}$$

In terms of $u=U+\lambda$ and ξ^1 Eqs. (4.7) become

$$\xi^1_{xx} + u\xi^1 = \lambda\xi^1,$$
$$\xi^1_t - C\xi^1_x + \tfrac{1}{2}Z\xi^1 = \mu\xi^1. \tag{4.8}$$

It is clear that the inverse scattering problems for the various higher order Korteweg–de Vries equations are given by

$$\xi^1_{xx} + u\xi^1 = \lambda\xi^1,$$
$$\xi^1_t - C^{(n)}(u)\xi^1_x + \tfrac{1}{2}C^{(n)}_x(u)\xi^1 = \mu\xi^1, \tag{4.9}$$

where

$$C^{(n)}(u) = \sum_{i=0}^{n} C_i \lambda^{n-i}. \tag{4.10}$$

For example if $n=2$, we have shown that the equation

$$u_t = \tfrac{1}{4}u_{xxxxx} + 5u_x u_{xx} + \tfrac{5}{2}u_{xxx}u + \tfrac{15}{2}u^2 u_x \tag{4.11}$$

has the inverse scattering formulation

$$\xi^1_{xx} + u\xi^1 = \lambda\xi^1$$
$$\xi^1_t - [4\lambda^2 + 2\mu\lambda + \tfrac{1}{2}(p + 3u^2)]\xi^1_x + [(\tfrac{3}{2}u + \lambda)z + \tfrac{1}{4}r - \mu]\xi^1 = 0, \tag{4.12}$$

where $z = u_x$, $p = u_{xx}$, $r = u_{xxx}$.

If we note that as well as $u \to 0$ we have $C^{(n)} \to 42^n$ when $x \to \pm\infty$ we can solve the relevant Gelfand–Levitan equations for the general inverse scattering problem (4.9) and easily obtain the general result

$$u(x,t) = 2\alpha^2 \operatorname{sech}^2(\alpha x + 4\alpha^{2n+1}t + \delta_0) \tag{4.13}$$

for the single soliton solution of the nth order Korteweg–de Vries equation. The multisoliton solutions may also be constructed in general for any order equation.

[1] P. D. Lax, Commun. Pure Appl. Math. 21, 467–490 (1968).
[2] C. Gardner, H. Kruskal, and R. Muira, J. Math. Phys. 9, 1202–09 (1968).
[3] C. Gardner, J. Green, M. Kruskal, and R. Miura, Commun. Pure Appl. Math. 27, 97 (1974). The author wishes to thank the referee for bringing this important reference to his attention.
[4] H. D. Wahlquist and F. B. Estabrook, J. Math. Phys. 16, 1–7 (1974).

A prolongation structure for the AKNS system and its generalization

H. C. Morris

School of Mathematics, Trinity College, Dublin 2, Ireland
(Received 17 May 1976)

A prolongation structure is determined for the AKNS system of equations. Using an interpretation in terms of Cartan–Ehresmann connections, a generalization is then constructed which leads to an inverse scattering problem for the multicomponent nonlinear Schrodinger equation $iq_{\alpha t} = (1/2)q_{\alpha xx} + q_\alpha(\Sigma_{\beta=1}^{n}|q_\beta|^2)$ $(\alpha = 1,\ldots,n)$.

1. INTRODUCTION

Ablowitz, Kaup, Segur, and Newell[1] have shown that the equations

$$A_x = qC - rB, \tag{1.1}$$

$$D_x = rB - qC, \tag{1.2}$$

$$q_t = B_x + (A - D)q + 2\lambda Bi, \tag{1.3}$$

$$r_t = C_x + (D - A)r - 2\lambda Ci, \tag{1.4}$$

can yield most of the special equations such as the Korteweg–de Vries equation, the nonlinear Schrödinger equation, and the sine Gordon equation which have been solved by the inverse scattering method. We will refer to this system as the AKNS system. In this paper we will determine the prologation structure of these equations and also the prolongation structure of a multicomponent form of these equations. We do not make the conventional identification of D with $-A$ until it is convenient to do so as the identification fails to hold in the generalized form of the equations that we will later determine. The plan of the work is as follows: In the following section we will establish a closed set of forms equivalent to (1.1)–(1.4) and determine a prolongation structure for that ideal of forms. We then look at a recent interpretation[2,3] of such a structure in terms of Cartan–Ehresmann connections and in Sec. 3 determine from such a connection a generalized form of the AKNS system. The most important physical equation associated with this new system is the two-component nonlinear Schrödinger equation previously considered by Manakov.[4] Finally in Sec. 4 we generalize to the n-dimensional case and determine an inverse scattering problem for the multicomponent nonlinear Schrödinger equation.

2. THE AKNS SYSTEM

The forms

$$\alpha_1 = dA \wedge dt + (rB - qC)\,dx \wedge dt, \tag{2.1}$$

$$\alpha_2 = dB \wedge dt + dq \wedge dx + [2iB + (A - D)q]\,dx \wedge dt, \tag{2.2}$$

$$\alpha_3 = dC \wedge dt + dr \wedge dx - [2iC + (A - D)r]\,dx \wedge dt, \tag{2.3}$$

$$\alpha_4 = dD \wedge dt + (qC - rB)\,dx \wedge dt, \tag{2.4}$$

form a closed ideal which is equivalent to the AKNS system (1.1)–(1.4). If, following the normal procedure,[5] we seek a prolongation structure in the form

$$\Omega = d\xi + F(A, B, C, D, r, q, \xi)\,dx + G(A, B, C, D, r, q, \xi)\,dt \tag{2.5}$$

we find that we can choose F and G in the forms

$$F = x_1 + x_2 q + x_3 r \tag{2.6}$$

$$G = x_5 + Bx_2 + Cx_3 + Ax_4 + Dx_6. \tag{2.7}$$

These objects must satisfy the equation

$$[F, G] = -(rB - qC)\frac{\partial G}{\partial A} - [2\lambda iB + (A - D)q]\frac{\partial G}{\partial B} + [2\lambda Ci + (A - D)r]\frac{\partial G}{\partial C} - (qC - rB)\frac{\partial G}{\partial D} \tag{2.8}$$

and their substitution into (2.8) gives rise to the bracket relations

$$\begin{aligned}
&[x_3, x_2] = x_6 - x_4, \quad [x_3, x_4] = x_3,\\
&[x_1, x_2] = -2\lambda x_2 i, \quad [x_3, x_6] = -x_3,\\
&[x_2, x_4] = -x_2, \quad [x_1, x_4] = 0,\\
&[x_2, x_6] = x_2, \quad [x_1, x_6] = 0,\\
&[x_1, x_3] = 2\lambda i x_3, \quad [x_i, x_5] = 0 \text{ for } i = 1, 2, 3.
\end{aligned} \tag{2.9}$$

The Jacobi identity yields all the additional relations to complete the algebraic structure (2.9) into the Lie algebra

$$\begin{aligned}
&[x_1, x_2] = -2i\lambda x_2, \quad [x_1, x_3] = 2i\lambda x_3, \quad [x_1, x_4] = 0,\\
&[x_1, x_6] = 0, \quad [x_2, x_3] = (x_4 - x_6), \quad [x_2, x_4] = -x_2,\\
&[x_2, x_6] = x_2, \quad [x_3, x_4] = x_3, \quad [x_3, x_6] = -x_3,\\
&[x_4, x_6] = 0, \quad [x_5, x_i] = 0 \ \forall i
\end{aligned} \tag{2.10}$$

If we define

$$\begin{aligned}
&Y_0 = \tfrac{1}{2}(x_4 - x_6), \quad Y_2 = \tfrac{1}{2}(x_4 + x_6),\\
&Y_1 = x_2, \quad Y_3 = x_1 - i\lambda(x_6 - x_4),\\
&Y_{-1} = x_3, \quad Y_4 = x_5,
\end{aligned} \tag{2.11}$$

the Lie algebra (2.10) takes the form

$$[Y_0, Y_1] = Y_1, \quad [Y_0, Y_{-1}] = -Y_{-1}, \quad [Y_1, Y_{-1}] = 2Y_0, \tag{2.12}$$

$$[Y_\alpha, Y_i] = 0 \quad (\alpha = 2, 3, 4) \quad (i = \pm 1, 0). \tag{2.13}$$

The semisimple component (2.12) is the Lie algebra $sl(2, R)$. A one-dimensional representation of $sl(2, R)$

is given by

$$Y_0 = y\frac{\partial}{\partial y}, \quad Y_1 = y^2\frac{\partial}{\partial y}, \quad Y_{-1} = -\frac{\partial}{\partial y}. \tag{2.14}$$

This representation together with the trivial one-dimensional representation of (2.13),

$$Y_2 = Y_3 = Y_5 = 0, \tag{2.15}$$

gives the representation of the Lie algebra (2.10) defined by

$$x_1 = -2i\lambda y\,\partial/\partial y, \quad x_2 = y^2\,\partial/\partial y, \quad x_3 = -\partial/\partial y,$$
$$x_4 = -x_6 = y\,\partial/\partial y, \quad x_5 = 0. \tag{2.16}$$

The resulting prolongation structure is given by

$$\Omega = dy - [r + 2i\lambda y - qy^2]dx - [C - 2Ay - By^2]dt, \tag{2.17}$$

where we have now made the usual identification of D with $-A$.

The variable y is known as a pseudopential.[5] Hermann[2] has recently suggested an interesting geometric interpretation of the analagous result for the Korteweg–de Vries equation in terms of a Cartan–Ehresmann connection[3] with structure group SL(2, R). As the Korteweg–de Vries equation is only one of the equations obtainable from the AKNS equations it is of interest to see how his interpretation is easily extended to that system.

A connection ω is said to be associated with a differential equation if its connection forms θ generate the ideal I which define the differential equation. For the AKNS system (2.1)–(2.3) with $D = -A$ the quadratic connection

$$\omega = \omega_0 + \omega_1 y + \omega_2 y^2, \tag{2.18}$$

where

$$\omega_0 = r\,dx + C\,dt, \tag{2.19}$$
$$\omega_1 = 2(i\lambda\,dx - A\,dt), \tag{2.20}$$
$$\omega_2 = -(q\,dx + B\,dt), \tag{2.21}$$

gives rise to the curvature forms

$$\theta_1 = d\omega_1 + 2\omega_0 \wedge \omega_2 = -\alpha_1, \tag{2.22}$$
$$\theta_2 = d\omega_2 - \omega_1 \wedge \omega_2 = -\alpha_2, \tag{2.23}$$
$$\theta_0 = d\omega_0 + \omega_0 \wedge \omega_1 = \alpha_3, \tag{2.24}$$

and so are clearly associated with the ideal (2.1)–(2.3) with $D = -A$.

The two-dimensional representation of the algebra given by

$$Y_0 = \tfrac{1}{2}(\xi^2 b_2 - \xi^1 b_1), \quad Y_1 = -\xi^2 b_1, \quad Y_{-1} = -\xi^1 b_2, \tag{2.25}$$

gives rise to the following representation of the algebra (2.10):

$$x_1 = i\lambda(\xi^1 b_1 - \xi^2 b_2), \quad x_2 = -\xi^2 b_1, \quad x_3 = -\xi^1 b_2,$$
$$x_4 = -\xi^1 b_1, \quad x_5 = \mu(\xi^1 b_1 + \xi^2 b_2), \quad \text{where } \mu \text{ is arbitrary},$$
$$x_6 = -\xi^2 b_2.$$

This representation gives rise to the normal inverse scattering problem of AKNS[1] and Zakharov and Shabat.[6]

We note that for any given equation such as the Korteweg–de Vries equation a much larger prolongation structure may exist. This is well demonstrated by the results of Wahlquist and Estabrook,[5] and will also be apparent from the following section.

3. GENERALIZING THE AKNS SYSTEM BY AN SL (3, R) CONNECTION

The two 1-forms

$$\omega^1 = \omega^1_0 + \omega^1_{11}y^1 + \omega^1_{12}y^2 + \omega^1_{211}(y^1)^2 + 2\omega^1_{212}y^1y^2, \tag{3.1}$$
$$\omega^2 = \omega^2_0 + \omega^2_{11}y^1 + \omega^2_{12}y^2 + \omega^2_{222}(y^2)^2 + 2\omega^2_{212}y^1y^2, \tag{3.2}$$

where

$$\omega^1_0 = (r\,dx + C\,dt), \quad \omega^2_0 = (s\,dx + F\,dt), \tag{3.3}$$
$$\omega^1_{11} = 2i\lambda\,dx - (A - D)\,dt, \quad \omega^1_{12} = G\,dt, \tag{3.4}$$
$$\omega^2_{11} = H\,dt, \quad \omega^2_{12} = 2i\lambda\,dx - (A - I)\,dt, \tag{3.5}$$
$$\omega^1_{211} = +2\omega^2_{212} = -(q\,dx + B\,dt), \tag{3.6}$$
$$+2\omega^1_{212} = \omega^2_{222} = -(p\,dx + E\,dt), \tag{3.7}$$

define a Cartan–Ehresmann connection[3] with structure group SL(3, R) which clearly generalizes that defined in equations (2.18)–(2.21).

The curvature 2-forms of this connection generate an ideal spanned by the nine 2-forms

$$\alpha_1 = dA \wedge dT + [(rB - qc) + (pF - Es)]dx \wedge dt, \tag{3.8}$$
$$\alpha_2 = dB \wedge dq \wedge dx + [2\lambda iB + (A - D)q - pH]dx \wedge dt, \tag{3.9}$$
$$\alpha_3 = dC \wedge dt + dr \wedge dx + [-2\lambda iC + (D - A)r + sG]dx \wedge dt, \tag{3.10}$$
$$\alpha_4 = dD \wedge dt + (qC - rB)\,dx \wedge dt, \tag{3.11}$$
$$\alpha_5 = dE \wedge dt + dp \wedge dx + [2\lambda iE + (A - I)p - qG]dx \wedge dt, \tag{3.12}$$
$$\alpha_6 = dF \wedge dt + ds \wedge dx + [-2\lambda iF + (I - A)s + rH]dx \wedge dt, \tag{3.13}$$
$$\alpha_7 = dG \wedge dt + (pC - rE)\,dx \wedge dt, \tag{3.14}$$
$$\alpha_8 = dH \wedge dt + (qF - sB)\,dx \wedge dt, \tag{3.15}$$
$$\alpha_9 = dI \wedge dt + (Fp - sE)\,dx \wedge dt. \tag{3.16}$$

The most interesting new equation of this extended system is the two-component nonlinear Schrödinger equation

$$iq_t = \tfrac{1}{2}q_{xx} + q(|q|^2 + |p|^2),$$
$$ip_t = \tfrac{1}{2}p_{xx} + p(|q|^2 + |p|^2). \tag{3.17}$$

This is obtained from expanding the $A, B, C, D, E, F, G, H, I$ as quadratic series in λ. The exact expressions which give rise to (3.17) are

$$A = \frac{1}{2i}[|q|^2 + |p|^2] + 2i\lambda^2, \quad E = \frac{1}{2i}p_x - \lambda p,$$
$$B = \frac{1}{2i}q_x - \lambda q, \qquad F = \frac{1}{2i}p^*_x + \lambda p^*,$$
$$C = \frac{1}{2i}q^*_x + \lambda q^*, \qquad G = -\frac{1}{2i}(q^*p), \tag{3.18}$$

$$D=-\frac{1}{2i}|q|^2, \qquad H=-\frac{1}{2i}(p^*q),$$

$$I=-\frac{1}{2i}|p|^2.$$

The prolongation structure for the system (3.8)–(3.16) has the form

$$\Omega=d\zeta+L(p,q,r,s,\zeta)\,dx+M(A,B,C,D,\zeta)\,dt, \tag{3.19}$$

with L and M having the forms

$$L=x_1+x_2q+x_3r+x_7p+x_8s \tag{3.20}$$

$$M=x_5+Bx_2+Cx_3+Ax_4+Dx_6+Ex_7+Fx_8+Gx_9 +Hx_{10}+Ix_{11}. \tag{3.21}$$

The x_i satisfy the Lie algebra SL(3, R) and the representation is given by

$$\begin{aligned} &x_1=-2i\lambda(y^1b_1+y^2b_2), \quad x_2=((y^1)^2b_1+y^1y^2b_2)\\ &x_3=-b_1, \quad x_4=(y^1b_1+y^2b_2), \quad x_5=0,\\ &x_6=-y^1b_1, \quad x_7=(y^1y^2b_1+(y^2)^2b_2), \quad x_8=-b_2,\\ &x_9=-y^2b_1, \quad x_{10}=-y^1b_2, \quad x_{11}=-y^2b_2, \end{aligned} \tag{3.22}$$

where $b_i=\partial/\partial y^i$ and $\zeta=(y^1,y^2)$.

Changing to projective coordinates $\zeta^2/\zeta^1=y^1$, $\zeta^3/\zeta^1=y^2$ yields the linear three-dimensional representation

$$\begin{aligned} &x_1=i\lambda(\zeta^1b_1-\zeta^2b_2-\zeta^3b_3), \quad x_2=-\zeta^2b_2,\\ &x_3=-\zeta^1b_2, \quad x_4=-\zeta^1b_1, \quad x_5=0, \quad x_6=-\zeta^2b_2,\\ &x_7=-\zeta^3b_1, \quad x_8=-\zeta^1b_3, \quad x_9=-\zeta^3b_2,\\ &x_{10}=-\zeta^2b_3, \quad x_{11}=-\zeta^3b_3. \end{aligned} \tag{3.23}$$

The prolongation structure which results for the two-component nonlinear Schrödinger equation is given by

$$\begin{aligned} \Omega^1=&d\zeta^1-(-i\lambda\zeta^1+q\zeta^2+p\zeta^3)\,dx\\ &-[(|q|^2+|p|^2/2i+2i\lambda^2)\zeta^1+(q_x/2i-\lambda q)\zeta^2\\ &+(q_x/2i-\lambda p)\zeta^3]\,dt, \end{aligned} \tag{3.24}$$

$$\Omega^2=d\zeta^2-(i\lambda\zeta^2+r\zeta^1)\,dx-\left[(q_x^*/2i+\lambda q^*)\zeta^1 -\frac{1}{2i}|q|^2\,\zeta^2+\left(-\frac{1}{2i}(q^*p)\right)\zeta^3\right]dt, \tag{3.25}$$

$$\Omega^3=d\zeta^3-(i\lambda\zeta^3+s\zeta^1)\,dx-\left[(p_x^*/2i+\lambda p^*)\zeta^1 -\frac{1}{2i}(p^*q)\zeta^2+\left(-\frac{1}{2i}|p|^2\right)\zeta^3\right]dt. \tag{3.26}$$

The equations

$$\tilde{\Omega}^1=\tilde{\Omega}^2=\tilde{\Omega}^3=0$$

yield the inverse scattering problem previously determined by Manakov.[4] The linear operator involved is

$$L\overset{\text{def}}{=}i\begin{bmatrix}\partial/\partial x & p & q\\ r & -\partial/\partial x & 0\\ s & 0 & -\partial/\partial x\end{bmatrix} \tag{3.27}$$

and the scattering theory of this operator has been analysed by both Manakov[4] and Date.[7]

4. THE GENERAL CASE OF SL(n, R) AND THE MULTICOMPONENT NONLINEAR SCHRÖDINGER EQUATION

The extension to n dimensions is easily established. The quadratic connection forms

$$\omega^\alpha=\omega_0^\alpha+\omega_{1\ \beta}^{\alpha}y^\beta+\omega_{2\beta\gamma}^{\alpha}y^\beta y^\gamma, \tag{4.1}$$

where

$$\omega_0^\alpha=(r^\alpha\,dx+C^\alpha\,dt), \tag{4.2}$$

$$\omega_{1\ \beta}^{\alpha}=(2i\lambda\delta_\beta^\alpha\,dx+(A\delta_\beta^\alpha-D_\beta^\alpha)\,dt), \tag{4.3}$$

$$\omega_{2\ \beta\gamma}^{\alpha}=-\tfrac{1}{2}[(q_\beta\delta_\gamma^\alpha+q_\gamma\delta_\beta^\alpha)\,dx+(B_\beta\delta_\gamma^\alpha+B_\gamma\delta_\beta^\alpha)\,dt], \tag{4.4}$$

define a Cartan–Ehresmann connection with structure group SL(n, R) which generalizes those previously considered.

The curvature forms determine an ideal spanned by the n^2 2-forms

$$\alpha_1=dA\wedge dt+(r^\alpha B_\alpha-q_\alpha C^\alpha)\,dx\wedge dt, \tag{4.5}$$

$$\alpha_{2\alpha}=dB_\alpha\wedge dt+dq_\alpha\wedge dx+[2i\lambda B_\alpha+(A\delta_\alpha^\beta-D_\alpha^\beta)q_\beta] \times dx\wedge dt, \tag{4.6}$$

$$\alpha_3^\alpha=dC^\alpha\wedge dt+dr^\alpha\wedge dx-[+2i\lambda C^\alpha +(A\delta_\beta^\alpha-D_\beta^\alpha)r^\beta]dx\wedge dt, \tag{4.7}$$

$$\alpha_{4\ \beta}^{\alpha}=dD^\alpha{}_\beta\wedge dt+(C^\alpha q_\beta-r^\alpha B_\beta)\,dx\wedge dt, \tag{4.8}$$

and the corresponding prolongation structure is defined by

$$\Omega^\alpha=dy^\alpha-\omega^\alpha. \tag{4.9}$$

The choice

$$A=2i\lambda^2-\frac{1}{2i}(q_\alpha r^\alpha), \tag{4.10}$$

$$B=\left(\frac{1}{2i}q_{\alpha x}-\lambda q_\alpha\right), \tag{4.11}$$

$$C^\alpha=-\left(\frac{1}{2i}r_x^\alpha+\lambda r^\alpha\right), \tag{4.12}$$

$$D^\alpha{}_\beta=(r^\alpha q_\beta)\frac{1}{2i}, \tag{4.13}$$

gives rise to the equations

$$iq_{\alpha t}=\tfrac{1}{2}q_{\alpha xx}-(q_\alpha q_\beta r^\beta), \tag{4.14}$$

$$+ir^\alpha{}_t=-\tfrac{1}{2}r_{xx}^\alpha+(r^\alpha r^\beta q_\beta). \tag{4.15}$$

If $r^\alpha=-q_\alpha^*$ these equations reduce to the single equation

$$iq_{\alpha t}=\tfrac{1}{2}q_{\alpha xx}+q_\alpha\left(\sum_{\gamma=1}^{n}|q_\gamma|^2\right) \tag{4.16}$$

which we refer to as the multicomponent nonlinear Schrödinger equation. The inverse scattering problem which results from the prolongation structure (4.9) by changing to projective coordinates is given by

$$\zeta_x=\left[\begin{array}{c|c}-i\lambda & q_\alpha\\ \hline -q_\alpha^* & i\lambda\end{array}\right]\zeta \tag{4.17}$$

$$\zeta_t = \begin{bmatrix} 2i\lambda^2 + \frac{1}{2i}\sum_\alpha (q_\alpha q_\alpha^*) & \frac{1}{2i} q_{\alpha x} - \lambda q_\alpha \\ \frac{1}{2i} q_{\alpha x}^* \mp \lambda q_\alpha^* & -\frac{1}{2i} q_\alpha^* q_\beta \end{bmatrix} \zeta. \qquad (4.18)$$

Equation (4.16) has the internal symmetry group U(n).

Other multiple component analogs of the standard equations of the AKNS system may also be obtained. For example, one can easily determine from a cubic expansion in λ an inverse scattering problem for the equation

$$q_{\alpha t} + \left(q_{\alpha xx} \pm 2\left(\sum_\gamma q_\gamma^2\right) q_\alpha\right)_x = 0 \quad (\alpha = 1, \ldots, n), \qquad (4.19)$$

which is a multicomponent form of the modified Korteweg–de Vries equation. A more detailed analysis of these equations will be presented elsewhere.

[1]M.J. Ablowitz, D.J. Kaup, A.C. Newell, and H. Segur, Phys. Rev. Lett. **31**, 125 (1973).
[2]R. Hermann, Phys. Rev. Lett. **36**, 835 (1976).
[3]R. Hermann, *Interdisciplinary Mathematics* (Math. Sci. Press, Brookline, Mass., 1975), Vol. X.
[4]S.V. Manakov, Zh. Eksp. Teor. Fiz. **65**, 505 (1973) [Sov. Phys.—JETP 38, 248 (1974)].
[5]H. Wahlquist and F. Estabrook, J. Math. Phys. **16**, 1 (1975).
[6]V.E. Zakharov and A.V. Shabat, Zh. Eksp. Teor. Fiz. **61**, 118 (1971) [Sov. Phys.—JETP 34, 62 (1972)].
[7]E. Date, Osaka J. Math. **12**, 777 (1975).

ERRATA

Erratum: Wave operators for multichannel scattering by long-range potentials [J. Math. Phys. 17, 1056 (1976)]

W. W. Zachary

Naval Research Laboratory, Washington, D.C. 20375

(Received 2 December 1976)

In (3.5) the quantity $G_{t,12}^{(D)(m)}$ in the exponent should be replaced by $G_t^{(D)(m)}$ as defined in (2.20). A similar replacement is necessary in (3.7). With these modifications the proofs in the paper are not valid in the stated generality, but can be proved if one imposes $m = 1$, $\alpha_1 > \frac{1}{2}$ in addition to the stated conditions. In this case, the methods of Alsholm (Ref. 12 of the paper) can be used to obtain the required estimates.

For this simplified case, a shorter proof will be published in a joint paper with A. W. Sáenz, whom I thank for pointing out the error.

$$\zeta_t = \begin{bmatrix} 2i\lambda^2 + \frac{1}{2i}\sum_\alpha (q_\alpha q_\alpha^*) & \frac{1}{2i} q_{\alpha x} - \lambda q_\alpha \\ \frac{1}{2i} q_{\alpha x}^* \mp \lambda q_\alpha^* & -\frac{1}{2i} q_\alpha^* q_\beta \end{bmatrix} \zeta. \qquad (4.18)$$

Equation (4.16) has the internal symmetry group U(n).

Other multiple component analogs of the standard equations of the AKNS system may also be obtained. For example, one can easily determine from a cubic expansion in λ an inverse scattering problem for the equation

$$q_{\alpha t} + \left(q_{\alpha xx} \pm 2\left(\sum_\gamma q_\gamma^2\right) q_\alpha\right)_x = 0 \quad (\alpha = 1, \ldots, n), \qquad (4.19)$$

which is a multicomponent form of the modified Korteweg–de Vries equation. A more detailed analysis of these equations will be presented elsewhere.

[1]M.J. Ablowitz, D.J. Kaup, A.C. Newell, and H. Segur, Phys. Rev. Lett. **31**, 125 (1973).
[2]R. Hermann, Phys. Rev. Lett. **36**, 835 (1976).
[3]R. Hermann, *Interdisciplinary Mathematics* (Math. Sci. Press, Brookline, Mass., 1975), Vol. X.
[4]S.V. Manakov, Zh. Eksp. Teor. Fiz. **65**, 505 (1973) [Sov. Phys.—JETP 38, 248 (1974)].
[5]H. Wahlquist and F. Estabrook, J. Math. Phys. **16**, 1 (1975).
[6]V.E. Zakharov and A.V. Shabat, Zh. Eksp. Teor. Fiz. **61**, 118 (1971) [Sov. Phys.—JETP 34, 62 (1972)].
[7]E. Date, Osaka J. Math. **12**, 777 (1975).

ERRATA

Erratum: Wave operators for multichannel scattering by long-range potentials [J. Math. Phys. 17, 1056 (1976)]

W. W. Zachary

Naval Research Laboratory, Washington, D.C. 20375

(Received 2 December 1976)

In (3.5) the quantity $G_{t,12}^{(D)(m)}$ in the exponent should be replaced by $G_t^{(D)(m)}$ as defined in (2.20). A similar replacement is necessary in (3.7). With these modifications the proofs in the paper are not valid in the stated generality, but can be proved if one imposes $m = 1$, $\alpha_1 > \frac{1}{2}$ in addition to the stated conditions. In this case, the methods of Alsholm (Ref. 12 of the paper) can be used to obtain the required estimates.

For this simplified case, a shorter proof will be published in a joint paper with A. W. Sáenz, whom I thank for pointing out the error.