Path-integral evaluation of the space–time propagator for quadratic Hamiltonian systems

John T. Marshall and John L. Pell

Department of Physics and Astronomy, Louisiana State University, Baton Rouge, Louisiana 70803
(Received 22 August 1978)

Path-integral methods are used to derive an exact expression for the space–time propagator for systems with quadratic Hamiltonians. For a certain subclass of such systems, the result is reduced to a simplified closed form. The propagators for several illustrative elementary cases are exhibited in detail.

I. INTRODUCTION

It may be fairly widely known that it is possible to solve the equations which describe the quantum behavior of physical systems with quadratic Hamiltonians. We were unable, however, to find in the literature a solution of a general, explicit, and detailed character. Upon setting out to use well-known techniques to obtain such a solution we discovered, moreover, that the task was not nearly so straightforward, nor was the result so simple in its detail structure, as expected. We did succeed in this task by using path-integral methods. Since exact results such as the ones we obtained are potentially broadly useful and since our results apparently are not to be found elsewhere and are not readily derived, we are presenting our results here along with a brief outline of their derivation and some elementary applications to illustrate their use.

The quantum mechanical motion of many physical systems of interest may be derived from Hamiltonian operators which are expressible in the Schrödinger picture, either exactly or as a useful approximation, by the quadratic form

$$H(\hat{q},\hat{p},t) = a_{jk}\hat{p}_j\hat{p}_k + b_{jk}\hat{q}_j\hat{q}_k + \tfrac{1}{2}c_{jk}(\hat{p}_j\hat{q}_k + \hat{q}_k\hat{p}_j) + d_j(t)\hat{p}_j + e_j(t)\hat{q}_j + f(t), \tag{1}$$

where $\hat{q}_j$ is the operator for multiplication by the coordinate q_j associated with the j-th degree of freedom, presumed to be a real variable ranging from $-\infty$ to $+\infty$, and where

$$\hat{p}_j = -i\hbar\frac{\partial}{\partial q_j}. \tag{2}$$

Throughout this paper the summation convention is employed with respect to the indices j, j', k, and k', which are used to index the degrees of freedom and which take on the values 1, 2, 3,...,N. The coefficient matrices a, b, c, d, e, and f occuring in Eq. (1) are taken to be real; a, b, and c are assumed to be independent of time, t; a is assumed to be nonsingular; and, as a notational convenience, both a and b are presumed to have been symmetrized.

In order to explicate the dynamical behavior generally of all systems described by Eq. (1), it is the primary objective of this paper to derive an exact explicit expression for the space–time propagator, $K(q'',t'';q',t')$, which is defined by the requirement that the system's coordinate-space wavefunction, ψ, evolve in time according to the integral transformation

$$\psi(q'',t'') = \int K(q'',t'';q',t')\psi(q',t')\,dV, \tag{3}$$

where dV is the element of volune in coordinate (q) space. The reason for focusing attention upon determination of K is that once K is known, then for any given initial state of the system, all predictions about the subsequent behavior of the system may be expressed directly and explicitly in terms of K and the given initial state.[1] The quadratic systems characterized by Eq. (1) were chosen for investigation because they constitute a fairly broad class of systems for which the method of analysis here used provides a unified, exact explicit description of their quantum behavior.

The resulting expression for K is potentially useful in several respects: (1) It explicates some of the general features common to the dynamical behavior of quadratic systems. Such features may be sufficient in some applications to make some specific predictions of interest, without the need for a complete quantum treatment. An example of such a feature is provided by the well-known fact that for quadratic systems the dependence of $K(q'',t'';q',t')$ upon the coordinates, q'' and q' is contained wholly within the factor $\exp(i\cdot S_c(q'',t'';q',t')/\hbar)$, where $S_c(q'',t'';q',t')$ is the classical action function. (2) In some applications, the Hamiltonian for the system may be adequately approximated by Eq. (1), in which case the solution for K as expressed in detail here may be employed without further approximations to obtain predictions of interest in formats which are explicit and practical for numerical computation. If such an approximation is not in itself entirely adequate, it still may serve as a "zero-th" order approximation in a systematic perturbation calculation in which results are obtained in successively higher orders in the difference between the exact Hamiltonian and its quadratic approximation.[2] To select a suitable quadratic approximate Hamiltonian of the form of Eq. (1), Feynman's path-integral variational method may be helpful.[1] (3) A prospectively powerful use of the result for K arises in connection with the problem of obtaining the propagator for a system with many degrees of freedom such that the Hamiltonian may be expressed in the form of Eq. (1) with the q's and p's representing some (but not all) of the degrees of freedom and with the time-dependent functions, $d(t)$, $e(t)$,

 0022-2488/79/071297-06$01.00

and $f(t)$ representing functions of the motion of the remaining degrees of freedom. In such a case the propagator for the entire system may be expressed as a path integral over all of the degrees of freedom and the integration over some of the degrees of freedom then performed by application of the result obtained here for K. The outcome is an exact path-integral expression requiring path integration of only the remaining degrees of freedom. Exact reduction of the number of degrees of freedom of a system can be a significiant step in the exact or approximate analysis of the system. According to the theory of interactions there are many instances of widely used models of physical systems of great interest for which a set of particles (or a field) interacts directly only with some Bosonic field in such a way that when the Hamiltonian is expressed in terms of the Bosonic field oscillator coordinates and momenta, the result is of the form for which the reduction process just discussed is possible. Feynman's treatment of quantum electrodynamics provides examples of such applications of a general character and his solution of the polaron problem is a specific example carried to approximate numerical conclusion for the polaron self-energy obtained with greater accuracy than had been obtained previously.[1,2]

Section II contains the derivation of the exact expression for the space–time propagator for all systems for which the Hamiltonian is of the form of Eq. (1). For certain special conditions, a simplified closed-form expression for the propagator is also obtained in Sec. II. Some elementary illustrative examples of common interest are displayed in Sec. III. These examples include the general case and some special cases of one-dimensional motion, and the three-dimensional motion of a particle in a constant magnetic field. The present work is summarized in Sec. IV and is compared with other path-integral treatments of quadratic systems in Sec. V.

II. DERIVATION OF THE PROPAGATOR

A. General case

According to Garrod,[3] the phase-space path-integral expression for the propagator for systems described by Eq. (1) is

$$K(q'',t'';q',t') = \int_{(q',t')}^{(q'',t'')} D^N[q(t),p(t)] \times \exp\left\{ \frac{i}{\hbar}\int_{t'}^{t''} [p_j(t)\dot q_j(t) - H(q(t),p(t),t)]dt \right\} \tag{4}$$

$$\equiv \lim_{M\to\infty} \left\{ \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \left[\prod_{j=1}^{N} dq_{j,1}\cdots dq_{j,M-1} \times \frac{dp_{j,1}}{h}\cdots\frac{dp_{j,M}}{h} \right] \exp\left[\frac{i}{\hbar}\left(\sum_{l=1}^{M} p_{j,l}(q_{j,l}-q_{j,l-1}) - \int_{t'}^{t''} H(q(t),p(t),t)dt \right)\right]\right\}, \tag{5}$$

where $q(t)$ is piecewise linear and $p_j(t)$ is piecewise constant on a uniform partition $(t_0 \equiv t', t_1, t_2, \ldots, t_{M-1}, t_M \equiv t'')$ of the time interval $t' \leqslant t \leqslant t''$, where $q_{jl} = q_j(t_l)$, $p_{jl} = p_j(t)$ for $t_{l-1} < t < t_l$, $q' = q(t')$, and $q'' = q(t'')$, and where $H(q,p,t)$ is the classical Hamiltonian function specified by Eq. (1). Integration of the momentum variables p_{jl}, via a linear transformation which uncouples them, yields the following configuration space path integral expression (with the correct normalization constant properly explicated):

$$K(q'',t'';q',t') = \lim_{M\to\infty}\left\{ \int_{-\infty}^{\infty} [\,|a|(2i\hbar T/M)^N]^{-M/2} \times \left[\prod_{j=1}^{N}\prod_{l=1}^{M-1} dq_{j,l} \right] \exp\left[\frac{i}{\hbar} S(q(t))\right]\right\}, \tag{6}$$

or equivalently

$$K(q'',t'';q',t') = \exp[iS_c(q'',t'';q',t')/\hbar]F(T), \tag{7}$$

where $S_c(q'',t'';q',t')$ is the classical limit of the action

$$S(q(t)) = \int_{t'}^{t''} \mathscr{L}(q(t),\dot q(t),t)dt \tag{8}$$

associated with the Lagrangian $\mathscr{L}$ given by

$$\mathscr{L}(q,\dot q,t) \equiv [\tfrac14 a_{jj'}^{-1}(d_j(t)-\dot q_j)(d_{j'}(t)-\dot q_{j'})] + Q_{jj'}\dot q_j q_{j'} - g_{jj'} q_j q_{j'} - \rho_j(t) q_j - f(t), \tag{9}$$

wherein

$$Q = -\tfrac12 a^{-1}c, \tag{10}$$

$$g = b - \tfrac14 \tilde c a^{-1} c, \tag{11}$$

and

$$\rho(t) = e(t) + \tilde Q\, d(t), \tag{12}$$

and where

$$F(T) = \lim_{M\to\infty}\left\{ \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} [(2i\hbar T/M)^N|a|]^{-M/2} \times \prod_{j=1}^{N}\prod_{l=1}^{M-1} dq_{j,l}\right\} \exp\left[\frac{i}{\hbar}\int_{t'}^{t''} \mathscr{L}'(z(t),\dot z(t))dt\right]\Bigg\}, \tag{13}$$

wherein $z(t)$ denotes the deviation of $q(t)$ from its classical limit and

$$\mathscr{L}'(z,\dot z) \equiv [\tfrac14 a_{jj'}^{-1}\dot z_j \dot z_{j'} - Q_{jj'}^a z_j \dot z_{j'} - g_{jj'} z_j z_{j'}], \tag{14}$$

with

$$Q^a = \frac{(Q-\tilde Q)}{2}. \tag{15}$$

The tilde is used to denote transposition. In Eqs. (6), (7), (13) and hereafter $T = t'' - t'$. Also the square root of any complex number, as in Eq. (6), is to be understood to represent the root with an argument, θ, in the range $-(\pi/2) < \theta \leqslant (\pi/2)$.]

In order to evaluate $F(T)$, consider the system for which $L(q,\dot q,t)$ reduces to $L'(q,\dot q)$ and let $q'' = q'$. For this case Eqs. (6) and (7), integrated over q, yield

$$F(T) = A(T)/B(T), \tag{16}$$

where

$$B(T) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \exp[iS_c'(q',q')/\hbar] \prod_{j=1}^{N} dq_j' \tag{17}$$

and

$$A(T) = \lim_{M\to\infty} \left\{ \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} |a|^{-M/2} \exp[iS'(q(t))/\hbar] \times \prod_{j=1}^{N} \left[\left(\prod_{l=1}^{M-1} \frac{dq_{j,l}}{\sqrt{(2ihT/M)}} \right) \frac{dq_j'}{\sqrt{(2ihT/M)}} \right] \right\}. \tag{18}$$

Let $\zeta(n)$ be defined by the expansion

$$\dot q_j(t) = \sum_{n=1}^{\infty} \zeta_j(n) e^{2i\omega_n(t-t')}, \tag{19}$$

where $\omega_n \equiv n\pi/T$. Then,

$$S'(q(t)) \equiv \int_{t'}^{t''} \mathscr{L}'(q(t),\dot q(t))dt$$

$$= \left[\frac{1}{2} \sum_{n=1}^{\infty} \Omega_{jj'}(\omega_n) \zeta_j^*(n) \zeta_{j'}(n) - g_{jj'} Y_j Y_{j'} \right] T, \tag{20}$$

where

$$Y_j = q_j' - \sum_{n=1}^{\infty} (\zeta_j^{\rm I}(n)/\omega_n) \tag{21}$$

and

$$\Omega(\omega_n) = a^{-1} - 2iQ^a/\omega_n - g/\omega_n^2. \tag{22}$$

The notation $\zeta^{\rm R}$ and $\zeta^{\rm I}$ is used to denote the real and imaginary parts of ζ. Extremization of $S'(q(t))$ with respect to ζ holding q' constant yields the classical action

$$S_c'(q',q') = -\Lambda_{jj'} Y_j Y_{j'}, \tag{23}$$

where

$$Y = \Lambda^{-1} g q', \tag{24}$$

$$\Lambda = g - g\chi g, \tag{25}$$

and

$$\chi = \sum_{n=1}^{\infty} \frac{1}{\omega_n^2} (\Omega^{-1}(\omega_n) + \Omega^{*-1}(\omega_n)). \tag{26}$$

Substitution of Eqs. (23)–(26) into Eq. (17) and integration yield

$$B(T) = [\,|g|^{-1}(h/2iT)^N |g^{-1}\Lambda|\,]^{1/2}. \tag{27}$$

Since the infinite expansion in Eq. (19) is equivalent to an expansion

$$\dot q_j(t) = (2h/T)^{1/2} \mu_{jj'} \sum_{n=1}^{\infty} [\beta_{j'}^c(n)\phi_n^c + \beta_{j'}^S(n)\phi_n^S], \tag{28}$$

where μ is a symmetric matrix such that $\mu^2 = a$, and where

$$\phi_0 = 1, \tag{29}$$

$$\phi_n^c = \sqrt{2}\cos 2\omega_n(t-t'), \tag{30}$$

and

$$\phi_n^S = \sqrt{2}\sin 2\omega_n(t-t'), \quad n = 1,2,3,\ldots, \tag{31}$$

then a theorem proved by Davison[4] may be used to evaluate $A(T)$ in Eq. (18). The result is

$$A(T) = |a|^{-1/2} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \exp[iS'(q(t))/\hbar] \times \prod_{j=1}^{N} \left[\left(\prod_{n=1}^{\infty} \frac{d\beta_j^c(n)}{\sqrt{(+i)}} \frac{d\beta_j^s(n)}{\sqrt{(+i)}} \right) \times \frac{dq_j'}{\sqrt{(+2ihT)}} \right] \quad \text{(Ref. 5)}, \tag{32}$$

$$= |a|^{-1/2} \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} \left\{ \prod_{j=1}^{N} \left[\frac{dy_j}{\sqrt{(2ihT)}} \times \prod_n \left(\frac{-iT}{h\alpha_j} d\zeta_j^{\rm R}(n) d\zeta_j^{\rm I}(n) \right) \right] \times \exp(iS'(q(t))/\hbar) \right\} \tag{33}$$

$$= \left[(2iT)^N |a|^{1/2} |g|^{1/2} \prod_{n=1}^{\infty} (|a|\,|\Omega(\omega_n)|) \right]^{-1}, \tag{34}$$

where the α's are the eigenvalues of a.

In summary, the propagator for systems described by Eq. (1) is given by

$$K(q'',t'';q',t') = \exp(iS_c(q'',t'';q',t')/\hbar)\, F(T), \tag{35}$$

where

$$F(T) = \{ [(2iTh)^N |a|\,|\Phi|\,]^{1/2} \Gamma \}^{-1}, \tag{36}$$

$$\Gamma = \prod_{n=1}^{\infty} (|a|\,|\Omega(\omega_n)|), \tag{37}$$

$$\Phi = \mathrm{I} - \sum_{n=1}^{\infty} [\Omega^{-1}(\omega_n) + \Omega^{*-1}(\omega_n)] g/\omega_n^2, \tag{38}$$

$$\Omega(\omega_n) = a^{-1} - 2iQ^a/\omega_n - g/\omega_n^2, \tag{39}$$

$$g = b - \tfrac{1}{4}\tilde c a^{-1} c, \tag{40}$$

$$Q^a = -\tfrac{1}{4}(a^{-1}c - \tilde c a^{-1}), \tag{41}$$

$$\omega_n = n\pi/T, \tag{42}$$

$$T = t'' - t', \tag{43}$$

and $S_c(q'',t'';q',t')$ is the classical action function.

B. Special case

A simplified expression for the propagator $K(q'',t'';q',t')$ may be obtained if a is positive definite, so that a has a unique positive definite square root, μ, and if

$$g'Q' = Q'g', \tag{44}$$

where

$$g' \equiv \mu g \mu, \tag{45}$$

and

$$Q' \equiv \mu Q^a \mu. \tag{46}$$

In this case the infinite product in Eq. (37) and the infinite sum in Eq. (38) can be performed.

Consider the function

$$L = \prod_{n=1}^{\infty} (\Omega'(\omega_n)\Omega'^*(\omega_n)), \tag{47}$$

where

$$\Omega'(\omega_n) = \mu\Omega(\omega_n)\mu \tag{48}$$

$$= \mathrm{I} - 2iQ/\omega_n - g'/\omega_n^2. \tag{49}$$

Note that according to Eqs. (37) and (47), the determinant of L is equal to Γ^2 since $\Omega'(\omega_n)$ is Hermitian and the determinant of a matrix is invariant under transposition. According to Eqs. (47) and (49),

$$\frac{\partial}{\partial g'}\Big|_{Q',\omega_n}(\ln L) = -\sum_{n=1}^{\infty}\frac{1}{\omega_n^2}(\Omega'^{-1}(\omega_n) + \Omega'^{*-1}(\omega_n)). \tag{50}$$

Comparison of Eqs. (38) and (50) indicates that

$$\Phi' \equiv \mu^{-1}\Phi\mu \tag{51}$$

$$= \mathrm{I} + \left[\frac{\partial}{\partial g'}\Big|_{Q',\omega_n}(\ln L)\right]g'. \tag{52}$$

Note that the values of the determinants of Φ and Φ' are equal. Therefore, Eq. (36) may be written as

$$F(T) = [(2iTh)^N|a||\Phi'||L|]^{-1/2}, \tag{53}$$

since, as already noted, the determinant of L is equal to Γ^2.

Substitution of Eq. (49) into Eq. (47), completion of the square in both factors, and factorization of the result yield

$$L = \prod_{n=1}^{\infty}\left(\mathrm{I} + \frac{Q'^2}{\omega_n^2}\right)^2\left[\mathrm{I} - \frac{g' - Q'^2}{(\omega_n\mathrm{I} - iQ')^2}\right]$$

$$\times\left[\mathrm{I} - \frac{g' - Q'^2}{(\omega_n\mathrm{I} + iQ')^2}\right]. \tag{54}$$

Application of the identity

$$\frac{\cos(2x) - \cos(2y)}{2(y^2 - x^2)} \equiv \prod_{k=1}^{\infty}\left(1 - \frac{y^2}{k^2\pi^2}\right)\times\left(1 - \frac{x^2}{(k\pi + y)^2}\right)\left(1 - \frac{x^2}{(k\pi - y)^2}\right) \tag{55}$$

yields

$$L = \frac{\cos(2iTQ') - \cos[2T(g' - Q'^2)^{1/2}]}{2g'T^2}. \tag{56}$$

After substitution of Eqs. (56) and (52) into Eq. (53), one obtains the result that the propagator for the special case is given by Eq. (35) with

$$F(T) = \{(2iTh)^N|a||G[4T^2(g' - Q'^2)]|\}^{-1/2}, \tag{57}$$

where

$$G(x^2) \equiv \frac{\sin(x)}{x}. \tag{58}$$

III. SIMPLE APPLICATIONS

A. One degree of freedom

In one degree of freedom the Hamiltonian function defined by Eq. (1) reduces to the form

$$H(q,p,t) = \frac{1}{2m}p^2 + \frac{1}{2}mbq^2 + \omega_0 pq$$

$$+ d(t)p + e(t)q + f(t), \tag{59}$$

where m, b, ω_0 are given constants and $d(t)$, $e(t)$, $f(t)$ are specified functions of time. This Hamiltonian describes a forced harmonic oscillator of mass, m, for which the classical action is

$$S_c(q'',q')$$

$$= \frac{1}{\sin(\omega T)}\{\tfrac{1}{2}m\omega[(q''^2 + q'^2)\cos(\omega T) - 2q''q']$$

$$+ \left[q''\int_{t'}^{t''}E(t)\sin(\omega(t - t'))dt\right.$$

$$\left. + q'\int_{t'}^{t''}E(t)\sin(\omega(t'' - t))dt\right]$$

$$- \frac{1}{m\omega}\int_{t'}^{t''}dt\int_{t'}^{t}d\tau E(t)E(\tau)$$

$$\times \sin(\omega(t'' - t))\sin(\omega(\tau - t'))\Big\}$$

$$- \tfrac{1}{2}m\omega_0(q''^2 - q'^2) + \int_{t'}^{t''}[\tfrac{1}{2}md^2(t) - f(t)]dt$$

$$- m[q''d(t'') - q'd(t')], \tag{60}$$

where

$$\omega = (b - \omega_0^2)^{1/2} \tag{61}$$

is the angular frequency of oscillation and

$$E(t) = m\dot{d}(t) - e(t) + m\omega_0 d(t) \tag{62}$$

is the driving force on the oscillator. Substitution of Eqs. (40), (41), (45), and (46), into Eq. (57), and use of Eq. (58), gives

$$F(T) = \left(\frac{m\omega}{ih\sin(\omega T)}\right)^{1/2}. \tag{63}$$

According to Eqs. (35) and (63) the propagator is

$$K(q'',t'';q',t') = \left(\frac{m\omega}{ih\sin(\omega T)}\right)^{1/2}\exp(iS_c(q'',q')/\hbar) \tag{64}$$

where $S_c(q'',q')$ is given by Eq. (60).

In the special case of Eq. (59) for which

$$d(t) = e(t) = f(t) = 0, \tag{65}$$

the propagator for a simple harmonic oscillator is recovered as given by

$$K(q'',t'';q',t')$$

$$= \left(\frac{m\omega}{ih\sin(\omega T)}\right)^{1/2}$$

$$\times\exp\left(\frac{im\omega}{2\hbar\sin(\omega T)}[(q''^2 + q'^2)\cos(\omega t) - 2q''q']\right). \tag{66}$$

In the limit $\omega \to 0$, this reduces to the free-particle propagator

$$K(q'',t'';q',t') = \left(\frac{m}{ihT}\right)^{1/2} \exp(im(q''-q')^2/2\hbar T). \tag{67}$$

B. Particle in a magnetic field

The Hamiltonian for a particle in a constant magnetic field may be expressed in the form

$$H(x,p) = \sum_{j=1}^{3} \frac{1}{2m} p_j^2 + \frac{1}{2m}\left(\frac{qB}{c}\right)^2 x_1^2 - \frac{qB}{mc} p_2 x_1, \tag{68}$$

where q is the charge of the particle, m is its mass, and B is the magnitude of the magnetic field, which is taken to be in the x_3 direction. Substitution of the coefficients of Eq. (68) into Eqs. (40), (41), (45), and (46) gives

$$g' = 0 \tag{69}$$

and

$$Q' = \frac{1}{2}\begin{pmatrix} 0 & -\omega & 0 \\ \omega & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \tag{70}$$

where

$$\omega = \tfrac{1}{2} qB/mc. \tag{71}$$

Substitution of Eqs. (69), (70), and (58) into Eq. (57) and use of Eq. (35) yield

$$K(x'',t'';x',t') = \left(\frac{m}{ihT}\right)^{1/2} \frac{m\omega}{ih\sin(\omega T)} \exp(iS_c(x'',x')/\hbar) \tag{72}$$

where $S_c(x'',x')$ is the classical action for a particle in a constant magnetic field, which is given by

$$S_c(x'',x') = \tfrac{1}{2} m\omega\{[(x_1''-x_1')^2 + (x_2''-x_2')^2]\cot(\omega T)$$

$$- 2(x_1''x_2' - x_2''x_1')\} + \frac{m}{2T}(x_3''-x_3')^2. \tag{73}$$

Glasser[6] has performed the path-integral evaluation of this propagator.

IV. SUMMARY

The space–time propagator $K(q'',t'';q',t')$ for a given system is defined by Eq. (3). It represents the probability density amplitude for transition of the system from a given initial configuration, q', at an initial time, t', to a given final configuration, q'', at a final time, t''. Equation (4) is an exact phase-space path-integral expression for $K(q'',t'';q',t')$ for systems whose Hamiltonians have the form specified by Eq. (1). This path integral has been performed exactly to yield the result that $K(q'',t'';q',t') = F(t''-t) \times \exp(iS_c(q'',t'';q',t')/\hbar)$ where $S_c(q'',t'';q',t')$ is the classical action function connecting the initial and final space–time points (q',t') and (q'',t''), and where $F(t''-t')$ is independent of q' and q'' and is given explicitly and in detail by Eqs. (36)–(43) as a function of $(t''-t')$ and the coefficients occurring in only the quadratic part of the Hamiltonian of the system. Under certain special conditions [those stated in the sentence containing Eq. (44)], the infinite sums and products contained in the expression for F may be evaluated and the result for F simplified to the closed form given by Eqs. (57) and (58). As discussed in the introduction, the results obtained here may be useful in the quantum analysis of any system whose Hamiltonian is expressible (exactly or approximately) in such a form that Eq. (1) describes the dependence of the Hamiltonian upon some or all of the system's coordinates and their conjugate momenta.

V. CONCLUSION

In conclusion, the present work will be compared with two other path-integral investigations in each of which the propagators are sought for systems which are quadratic in a more general sense than defined by Eq. (1).

One of these investigations is Feynman's[7] derivation of the form for the propagator for systems with one coordinate, y, and with a quadratic action of the form

$$S(y(t)) = \tfrac{1}{2}\int_{t'}^{t''} y(t)\int_{t'}^{t''} A(t,s)y(s)\,ds\,dt + \int_{t'}^{t''} B(t)\,y(t)\,dt, \tag{74}$$

where $A(t,s)$ and $B(t)$ are independent of path, $y(t)$. His result for the propagator is the same as the form of Eq. (61), where the factor F is independent of q', q'', and B and is to be determined apart from a factor independent of A by the functional differential equation

$$\delta F/\delta A(t,s) = -\tfrac{1}{2}N(t,s)F, \tag{75}$$

where $N(t,s)$ is the reciprocal kernel to $A(t,s)$ subject to appropriate boundary conditions. For some cases this equation may be solved easily.[8] Equations (74) and (75) may be generalized readily to the case of many degrees of freedom, but the functional differential equation remains to be solved for F in the general case. The present work may be viewed as providing the required solution for F in a certain special case (of fairly broad interest) for which the kernel A is a superposition of δ functions and derivatives of δ functions.[9]

The other path-integral investigation of more general quadratic systems to be discussed here is the treatment by DeWitt[10] of a system whose Lagrangian has the form

$$\mathscr{L}(q,\dot q,t) = \tfrac{1}{2}G_{jk}(q,t)\dot q_j\dot q_k + a_j(q,t)\dot q_j - v(q,t). \tag{76}$$

This Lagrangian may be interpreted as describing the motion of a particle moving in a curved multidimensional space. DeWitt uses path-integral methods to derive an expression for the infinitesimal propagator which remains to be iterated to obtain the propagator connecting finitely separated space–time points. Equations (35)–(43) give the result of such an iteration for the special case for which the space is flat (i.e., G_{jk} is constant), and a_j and v are of the forms

$$a_j(q,t) = Q_{jk}g^k + d_j(t) \tag{77}$$

and

$$v(q,t) = g_{jk}q^j q^k + \rho_j(t)q^j + f(t), \tag{78}$$

respectively. The most interesting aspects and applications[11] of DeWitt's paper, however, lie outside of the range of this special case.

[1]R.P. Feynman and A.R. Hibbs, *Quantum Mechanics and Path Integrals* (McGraw–Hill, New York, 1965).
[2]See for example: J.T. Marshall and L.R. Mills, Phys. Rev. B **2**, 3143 (1970).
[3]C. Garrod, Rev. Mod. Phys. **38**, 483 (1966).
[4]B. Davison, Proc. R. Soc. (London) A **225**, 252 (1954).
[5]In Eq. (32) where plus signs occur explicitly, Davison has minus signs. Equation (32) seems to be correct as it stands.
[6]M.L. Glasser, Phys. Rev. **133**, B 831 (1964).
[7]R.P. Feynman, Phys. Rev. **84**, 108 (1951), Appendix C.
[8]See for example: R.W. Hellwarth and P.M. Platzman, Phys. Rev. **128**, 1599 (1962).
[9]B. Friedman, *Principles and Techniques of Applied Mathematics* (Wiley, New York, 1965), Chapter 2, p. 107.
[10]B.S. DeWitt, Rev. Mod. Phys. **29**, 377 (1957).
[11]See for example: Lawrence Schulman, Phys. Rev. **176**, 1558 (1968).

Internal symmetries of the axisymmetric gravitational fields

Cesare Reina

Istituto di Scienze Fisiche, via Celoria 16–Milano, Italy
(Received 19 July 1978; revised manuscript received 12 October 1978)

The group H of the internal symmetries of the axisymmetric field equations in general relativity is known to be isomorphic to SO(2,1), which is the double covering of the conformal group of the hyperbolic complex plane $\mathscr{H}$. The Ernst potential ξ can then be geometrically understood as a map $\xi : R^3/\mathrm{SO}(2) \to \mathscr{H}$. The fact that the hyperbolic plane is split into two connected components is used to introduce an algebraic invariant $n \in Z^+$ for every axisymmetric solution. It is shown that under reasonable hypotheses this invariant is related to the number of S^1 curves where the manifold is intrinsically singular.

INTRODUCTION

The axisymmetric field equations in general relativity contain a large amount of symmetries, which have been extensively discussed by several authors.[1] The main line of research in this field has been directed during the last few years towards the study of the infinite parameter group K, which combines both the coordinate group G and the internal symmetry group H.[2] Nevertheless, there are still some interesting results which can be derived from the study of the group H alone, as shown in the following.

The starting point of the present approach is to note that the most natural geometric interpretation of the Ernst equation is achieved considering the Ernst potential ξ as a map from $R^3/\mathrm{SO}(2)$ to the complex plane with the Poincaré metric. Because of the isomorphism $H \simeq \mathrm{SO}(2,1)$ and of the fact that SO(2,1) is a double covering of the conformal group of the hyperbolic plane, one can interpret the internal symmetries of the Ernst equation as isometries of the hyperbolic plane itself. This amounts to translating into elementary complex geometry the approach by Eris and Nutku.[3]

The map ξ is then studied, and it is shown that one can introduce an algebraic invariant, which classifies the asymptotically flat solutions according to their causal structure.

Finally the particular case in which ξ depends on a single real function is geometrically interpreted as the geodesic problem of the hyperbolic plane.

GEOMETRIC MEANING OF THE ERNST EQUATION

The axisymmetric stationary line element in canonical cylindrical coordinates reads[4]

$$ds = f^{-1}[e^{2\gamma}(dz^2 + d\rho^2) + \rho^2 d\phi^2] - f(dt - \omega d\phi)^2, \quad (1)$$

where f, ω, γ depend on ρ, z only. In this form the field equations for γ decouple, and the relevent problem reduces to two coupled equations for f, ω, which by means of the substitution

$$f = \frac{\xi\bar{\xi} - 1}{|\xi + 1|^2} \quad (2)$$

$$\nabla\omega = \phi \times \nabla \frac{\xi - \bar{\xi}}{|\xi + 1|^2} \quad (3)$$

(where $\hat{\phi}$ is the azimuthal versor pf R^3, and ∇ is the three-dimensional operator) can be transformed into the Ernst equation for the complex potential ξ [5]

$$(\xi\bar{\xi} - 1)\nabla^2\xi = 2\bar{\xi}\,\nabla\xi\cdot\nabla\xi. \quad (4)$$

Equation (4) can be derived from the Lagrangian density

$$L = \frac{\nabla\xi\cdot\nabla\bar{\xi}}{(\xi\bar{\xi} - 1)^2} = g(\nabla\xi, \nabla\bar{\xi}). \quad (5)$$

From Eq. (5) it is apparent that the bilinear operator $g(\ ,\ )$ coincides with the Poincaré metric for the complex hyperbolic plane $\mathscr{H}$ and the Ernst potential can be considered as the map

$$\xi: R^3/\mathrm{SO}(2) \to \mathscr{H}, \quad (6)$$

which in view of the field equation (4) must be extremal.

It is now obvious that the internal symmetries of the problem coincide with the isometries of the hyperbolic plane. These include a continous group (i.e., the conformal group $\mathscr{C}$)[6]

$$\xi \to e^{i\chi}\frac{\xi - p}{1 - \bar{p}\xi}, \quad 0 \leqslant \chi < 2\pi, \ p\bar{p} < 1, \quad (7)$$

and the following discrete transformations:

$$\xi \to -\xi, \quad (8)$$

$$\xi \to \bar{\xi}, \quad (9)$$

$$\xi \to 1/\xi, \quad (10)$$

Equations (8) and (9) are reflections of $\mathscr{H}$, while Eq. (10) arises from the fact that the unit disk $\xi\bar{\xi} < 1$ is an isometric copy of the domain $\xi\bar{\xi} > 1$ under inversion. This explains the origin of the discrete symmetries discovered by Ernst while discussing his equation.[7]

The conformal group (7) coincides with the form of the group H given by Kinnersley.[8] Note that the isotropy subgroup of $\mathscr{C}$ at the origin ($p = 0$) is given by $\xi \to e^{i\chi}\xi$ and generates NUT transformations of the manifold.[9]

From the point of view of elementary group theory, the

 0022-2488/79/071303-03$01.00

present interpretation of the Ernst equation amounts to using the well-known isomorphism $H \simeq SO(2,1)$, and to noting that SO(2,1) is a double covering of the conformal group $\mathscr{C}$.

Incidentally one can emphasize that the Lagrangian (5) presents some formal analogies with the one given by Woo[10] for the σ nonlinear model. In that case, however, the gauge group is SO(3), which is compact, and therefore the conformal factor is the spherical one [i.e., $(\xi\bar{\xi} + 1)^2$] instead of the hyperbolic one appearing in Eq. (5). Moreover, the base space for the σ nonlinear problem is R^2 instead of $R^3/SO(2)$ as in the present case.

TOPOLOGIC AND ALGEBRAIC INVARIANTS

Although one could impose boundary conditions on ξ in order to compactify its domain, the hyperbolic plane is not compact, and therefore it seems irrelevant to investigate the homotopy classes of the map ξ.

There is, however, an interesting invariant, which is related to the algebraic structure of the map ξ. These, in fact, can be classified according to the number of jumps between the two connected components into which the complex plane is split by the Poincaré metric, i.e., according to the number n of rotational bisurfaces in R^3 where $\xi\bar{\xi} = 1$. This number is independent of the coordinates chosen in $R^3/SO(2)$, although it is not invariant with respect to the general group of transformations of the metric (3). Note in fact that the surfaces identified by the equations $\xi\bar{\xi} = 1$ may contain coordinate singularities, the elimination of which will require transformations involving the asymptotically timelike coordinate t.

For instance, in the case of the Schwarzschild and Kerr solutions, for which $\xi_s = x$, $\xi_k = px + iqy, (p^2 + q^2 = 1)$, respectively, in prolate spheroidal coordinates, it turns out that $n = 2$.

Note that as the condition $\xi\bar{\xi} = 1$ is invariant under the action of the conformal group $\mathscr{C}$, also the number n is invariant under its action. This means, for instance, that the NUT generalization of a given field does not change the number n. The interior of the unit disk of the hyperbolic plane is related to the "ergosphere" regions of M, where $f < 0$, the unit circle itself (expect the point $\xi = -1$) being the domain into which ξ maps the "ergosurface." At the point $\xi = -1, f$ diverges, showing that $\xi = -1$ is the image of the intrinsic singularities of M.

A simple interpretation of the meaning of the number n can be obtained under few hypotheses on the map ξ. Choose prolate spheroidal coordinates x, y or $R^3/SO(2)$, and assume that

(A) the gravitational field described by ξ is asymptotically flat. In particular, $\lim_{x\to\infty} |\xi| \neq 1$.

(B) reflecting the space time with respect to the equatorial plane (i.e. $y \to -y$) the angular momentum of the gravitational field changes sign, i.e., $\xi \to \bar{\xi}$. Therefore ξ is real on the equatorial plane ($y = 0$).

(C) ξ is an odd function of x on the equatorial plane.[11]

Then the number m_+ of solutions of the equation $\xi = 1$ is equal to the number m_- of solutions of $\xi = -1$. Obviously on the equatorial plane $m_+ + m_- = n$, and therefore $m_- = n/2$. Since on the equatorial plane ξ is a function of x alone, there will be $n/2$ distinct values $x_1, \ldots, x_{n/2}$, where $\xi = -1$. These points actually represent trajectories of the axisymmetric group, which topologically are S^1 curves (of which one can possibly degenerate to a point) along which $f \to \infty$ and therefore the manifold is singular.

Therefore one can conclude that under the hypotheses (A),(B),(C) the number of ring singularities of the space time described by ξ is exactly $n/2$.

"GEODESIC SOLUTIONS"

In the special case when $\xi = \xi(\tau)$ depends on one real function $\tau: R^3/SO(2) \to R$, the present approach yields a nice geometrical interpretation. Note first that $\xi(\tau)$ is a curve in the hyperbolic plane. The Ernst equation reads

$$(\xi\bar{\xi} - 1)\frac{d^2\xi}{d\tau^2} + 2\bar{\xi}\frac{d\xi}{d\tau}^2 = -(\xi\bar{\xi} - 1)\frac{d\xi}{d\tau}\frac{\nabla^2\tau}{\nabla\tau\cdot\nabla\tau} \tag{11}$$

and coincides with the geodesic equation on the hyperbolic plane if $\nabla^2\tau = 0$, with τ as affine parameter.

If τ is not harmonic, one can introduce a new function $\alpha(\tau)$, in terms of which Eq. (14) becomes

$$(\xi\bar{\xi} - 1)\frac{d\xi}{d\alpha}\alpha'' + \frac{d^2\xi}{d\alpha^2}\alpha'^2 + 2\bar{\xi}\frac{d\xi}{d\alpha}^2\alpha'^2 = -(\xi\bar{\xi} - 1)\frac{d\xi}{d\alpha}\alpha'\frac{\nabla^2\tau}{\nabla\tau\cdot\nabla\tau}.$$

Choosing α such that

$$\frac{\alpha''}{\alpha'} = \frac{\nabla^2\tau}{\nabla\tau\cdot\nabla\tau}, \tag{12}$$

one has

$$(\xi\bar{\xi} - 1)\frac{d^2\xi}{d\alpha^2} + 2\bar{\xi}\frac{d\xi}{d\alpha}^2 = 0,$$

which is again the geodesic equation on the hyperbolic plane. From Eq. (12) one has that $\nabla^2\alpha = 0$, and hence α must be harmonic. Therefore, one can conclude that the geodesics of the hyperbolic plane depending on an affine parameter, which is a harmonic function defined on $R^3/SO(2)$, correspond one to one to the solution of the Ernst equation depending on a single real function. These include the Weyl[12] and Papapetrou[13] solutions.

[1]See W. Kinnersley, in *General Relativity and Gravitation* (Wiley, New York, 1975) and references quoted therein.
[2]For the definition of these groups see Ref. 1 and more recently W. Kinnersley, J. Math. Phys. **18**, 1529 (1977).
[3]A. Eris and Y. Nutku, J. Math. Phys. **16**, 1431 (1975).
[4]A. Papapetrou, Ann. Phys. **12**, 309 (1953).
[5]F.J. Ernst, Phys. Rev. D 7, 2520 (1973).
[6]See, for instance, I.M. Singer and I.A. Thorpe, *Lecture Notes on Elementary Toplogy and Geometry* (Springer-Verlag, New York, 1967).
[7]F.J. Ernst, J. Math. Phys. **15**, 1049 (1974).

[8]W. Kinnersley, J. Math. Phys. **14**, 651 (1973).
[9]C. Reina and A. Treves, J. Math. Phys. **16**, 834 (1975).
[10]A.A. Belavin and A.M. Polyakov, Pis'ma Zh. Eksp. Teor. Fiz, **22**, 503 (1975) [JETPL Lett. **22**, 245 (1976)]. See also G. Woo, J. Math. Phys. **18**, 1264 (1977), where the Lagrangian density for the σ nonlinear model is given in complex stereographic coordinates.
[11]Although this condition could appear a bit "ad hoc," it is satisfied by the entire class of the Tomimatsu and Sato solutions [see A. Tomimatsu and H. Sato, Prog. Theor. Phys. **50**, 95 (1973)].
[12]M. Weyl, Ann. Phys (Leipzig) **54**, 117 (1917).
[13]For more details on these "geodesic" solutions, see V. Benza, S. Morisetti and C. Reina Nuovo Cimento (1979) (in press).

Resolution of Fredholm equations with kernels $K(z-z_0)$ by operational calculus

Do Tan Si[a)]

Faculté des Sciences, Université de l'Etat à Mons, 7000–Mons, Belgium

(Received 27 November 1978)

We show that the solutions of a Fredholm equation with kernels $K(z-z_0)$ is the transform of its second member in a transformation defined by a differential operator. The calculations of these solutions are then a matter of the powerful operational calculus.

I. INTRODUCTION

It is well known[1] that the Fredholm equations,

$$\mu\Psi(z)=\lambda\int_{-\infty}^{\infty}K(z-z_0)\Psi(z_0)\,dz_0+\phi(z), \tag{1}$$

of the first ($\mu=0$) and second kind ($\mu=1$) may be solved by taking the Fourier transforms of both sides of them. This conventional method implies the calculations of three Fourier transforms which may be cumbersome.

The aim of this work is to give another approach to solve Eq. (1) using some operational calculus techniques which may greatly simplify the calculations: One or at most two Fourier transforms are needed. The outline of the method is presented in Sec. II. Some examples are given in Sec. III for comparison with the conventional method's results.

II. THE METHOD

The method we use in this work is based on the following considerations.

Let $f(z)$ be an analytic, square summable function of which a Fourier transform exists,

$$F(z)=\frac{1}{2\pi}\int_{-\infty}^{\infty}e^{-iz_0z}f(z_0)\,dz_0. \tag{2}$$

As $\exp(-iz_0z)$ is an eigenfunction of the derivative operator $D\equiv d/dz$, one can write

$$f(iD)\exp(-iz_0z)=f(z_0)\exp(-iz_0z). \tag{3}$$

Replacing (3) in (2) one obtains the relation

$$F(z)=\frac{1}{2\pi}\int_{-\infty}^{\infty}f(iD)e^{-iz_0z}\,dz_0=f(iD)\delta(z), \tag{4}$$

where $\delta(z)$ is the Dirac delta function. By a change of variable from z to $(z-z_0)$, one finally gets the useful relation

$$F(z-z_0)=f(iD)\delta(z-z_0), \tag{5}$$

which allows us to write Eq. (1) under the form

$$\mu\Psi=\lambda k(iD)\int_{-\infty}^{\infty}\delta(z-z_0)\Psi(z_0)\,dz_0+\phi(z)$$

$$=\lambda k(iD)\Psi(z)+\phi(z). \tag{6}$$

In (6), $k(z)$ is the inverse Fourier transform of $K(z)$, if it exists and is everywhere derivable. The solution of Eq. (1) may then be put into the operational form

$$(\mu-\lambda k(iD))\Psi(z)=\phi(z). \tag{7}$$

From (7) we see that $\Psi(z)$ is the sum of a term $\Psi_h(z)$ such that

$$(\mu-\lambda k(iD))\Psi_h(z)=0, \tag{8}$$

and the particular solution

$$\Psi_p(z)=\frac{1}{\mu-\lambda k(iD)}\phi(z). \tag{9}$$

In order to calculate $\Psi_p(z)$ we note that:

(i) If $\phi(z)$ is an eigenfunction of the operator $k(iD)$ with eigenvalue κ, one gets immediately

$$\Psi_p(z)=(\mu-\lambda\kappa)^{-1}\phi(z),\quad\text{when } \mu\neq\lambda\kappa. \tag{10}$$

(ii) If $\phi(z)$ is not an eigenfunction of $k(iD)$, one has to calculate an expression of the form $f(iD)\phi(z)$. This may be done using the following formula,

$$(F*\phi)(z)=\int_{-\infty}^{\infty}F(z-z_0)\phi(z_0)\,dz_0=f(iD)\phi(z), \tag{11}$$

which is a direct consequence of Eq. (5) and where the left-hand side is the convolution product[2] of $\phi(z)$ and the Fourier transform of $f(z)$.

Another way of calculating an expression of the form $f(iD)\phi(z)$ is based on the remarks that Hermite polynomials[2,3,4] and Laguerre polynomials[4] may readily be put into this form:

$$H_n(z)=\exp(-D^2/4)(2z)^n, \tag{12}$$

$$L_n^a(z)=(-)^n(1-D)^{n+a}z^n/n!, \tag{13}$$

$$z^nL_n^a(z^{-1})=(1+a)_n\,{}_0F_1(-;1+a;-D)z^n/n!. \tag{14}$$

III. EXAMPLES OF APPLICATIONS

Let us solve by the present method some Fredholm equations found in Ref. 1,

$$\phi(z)=\int_{-\infty}^{\infty}\exp[-(z-z_0)^2]H_p(z-z_0)\Psi(z_0)\,dz_0. \tag{15}$$

a)Work supported by the IRIS program sponsored by the Belgian Ministry for Science Policy.

Here we have

$$k(z) = \sqrt{\pi}\, \exp\left(-\frac{z^2}{4}\right)(iz)^p, \tag{16}$$

so that $\Psi_h(z)$ is an arbitrary polynomial of degree $(p-1)$ and the particular solution has the form

$$\Psi_p(z) = \frac{1}{\sqrt{\pi}} \exp\left(-\frac{D^2}{4}\right)(-D)^{-p}\phi(z). \tag{17}$$

Now, by (12) we can write

$$\Psi_p(z) = \frac{1}{\sqrt{\pi}}(-D)^{-p}\sum_{s=0}^{\infty}\frac{\phi^{(s)}(0)}{s!2^s}H_s(z). \tag{18}$$

Besides, from the relation $DH_s(z) = 2sH_{s-1}(z)$, one gets

$$D^{-p}H_s(z) = 2^{-p}\frac{s!}{(p+s)!}H_{s+p}(z) + Q_{p-1}(z), \tag{19}$$

where $Q_{p-1}(z)$ is an polynomial arbitrary of degree $(p-1)$. Finally,

$$\Psi(z) = \frac{1}{\sqrt{\pi}}\sum_{s=0}^{\infty}\frac{\phi^{(s)}(0)}{2^{s+p}(s+p)!}H_{s+p}(z) + \sum_{r=0}^{p-1}Q_r(z), \tag{20}$$

$$\Psi(z) = Ae^{\alpha|z|} + \lambda\int_{-\infty}^{\infty} e^{-|z-z_0|}\Psi(z_0)\,dz_0. \tag{21}$$

Here one has

$$k(iD) = \frac{2}{1-D^2},$$

so that $\Psi_h(z)$ verifies the equation

$$\frac{1-2\lambda-D^2}{1-D^2}\Psi_h(z) = 0. \tag{22}$$

Puting $k_0^2 = 2\lambda - 1$, one gets

$$\Psi_h(z) = C_1e^{+ik_0z} + C_2e^{-ik_0z}. \tag{23}$$

The particular solution is given by the equation

$$\Psi_p(z) = \frac{1-D^2}{1-2\lambda-D^2}Ae^{\alpha|z|} \tag{24}$$

$$= A\left(e^{\alpha|z|} - \frac{2\lambda}{D^2+k_0^2}e^{\alpha|z|}\right). \tag{25}$$

Putting now

$$I(z) = \frac{2\lambda}{D^2+k_0^2}e^{\alpha|z|}, \tag{26}$$

and remarking from (11) that

$$f(iD)\phi(z) = \zeta(iD)F(z), \tag{27}$$

where $\zeta(z)$ is the Fourier transform of $\phi(z)$, one can also write

$$I(z) = \frac{2i\alpha\lambda}{k_0(D^2-\alpha^2)}e^{-ik_0|z|}. \tag{28}$$

Comparing (26) to (28), one gets

$$I(z) = \frac{2\lambda}{\alpha^2+k_0^2}e^{\alpha|z|} - \frac{2i\alpha\lambda}{(\alpha^2+k_0^2)k_0}e^{-ik_0|z|}. \tag{29}$$

Finally, from (23), (25), (29), and the property of symmetry of $\Psi(z)$ with respect to the origin, one can put $\Psi(z)$ into the form

$$\Psi(z) = \frac{\alpha^2-1}{\alpha^2+k_0^2}Ae^{\alpha|z|} + C\cos(k_0|z|-\beta), \tag{30}$$

where the constants C and β are related to C_1, C_2 by the relations $2C_1 = 2C_2 = C\exp(-i\beta)$ and $2(C_1 + 2i\alpha\lambda A/(\alpha^2+k_0^2)k_0) = C\exp(i\beta)$. These relations lead to

$$C\sin\beta = \frac{2\alpha\lambda A}{(\alpha^2+k_0^2)k_0}, \tag{31}$$

i.e.,

$$C = 2\lambda Ak_0^{-1}(\alpha^2+k_0^2)^{-1/2}, \quad \text{and} \quad \tan\beta = \frac{\alpha}{k_0} \tag{32}$$

IV. CONCLUSIONS

We think that the method described in this work is interesting in that it allows the transformation of a type of integral equations into a differential equation of a infinite order which one may solve using the formula (11) or the differential representations of special functions. It seems to us that the expressions of the type $f(D)h(z)$ are worthwhile considering more widely for solving differential and integral equations and defining more special functions.

ACKNOWLEDGMENTS

The author is indebted to Professor R. Dagonnier for his encouragements, and to the referee, and Dr. G. Reidemeister for many useful remarks.

[1] P.M. Morse and H. Feshbach, *Methods of Theoretical Physics* (McGraw-Hill, New York, 1953), pp. 936, 962.
[2] D.V. Widder and I.I. Hirschman, *The Convolution Transform* (Princeton, U.P. Princeton, New Jersey, 1955).
[3] K.B. Wolf, J. Math. Phys. **15**, 1295 (1974).
[4] Do Tan Si, S.I.A.M. J. Math. Anal. **9**, 1068 (1978).

Classical predictive electrodynamics of two charges with radiation: General framework. I

R. Lapiedra and A. Molina

Departamento de Física Teòrica, Facultad de Física, Barcelona, Spain
(Received 7 July 1978; revised manuscript received 6 November 1978)

Outgoing radiation is introduced in the framework of the classical predictive electrodynamics using Lorentz–Dirac's equation as a subsidiary condition. In a perturbative scheme in the charges the first radiative "self-terms" of the accelerations, momentum and angular momentum of a two charge system without external field are calculated.

INTRODUCTION

This is the first of a series of two papers dealing with the classical dynamics of a radiating system consisting of two structureless interacting charges. We assume that each charge is moving in the retarded field of the other according to Lorentz–Dirac's equation.[1]

We take "the absorber" point of view of Wheeler–Feynman[2,3] and use the framework of predictive electrodynamics.[4,5] This theory is seen to be consistent with the phenomena of classical radiation and more precisely with the Lorentz–Dirac equation.

In Sec. 2 we show within a perturbative scheme in the charges how to construct the dynamical predictive system (the accelerations) of two classical interacting charges when radiation is present and there is no external field. Then in Sec. 3 we give explicitly the first radiative "self-terms" of the accelerations.

To fourth order in the charges ($n + m \leqslant 4$, e_1^n, e_2^m) the other terms in the accelerations, i.e., terms in $e_1 e_2$ and $e_1^2 e_2^2$, are shown to be those of Refs. 4 and 5.

Section 4 contains a review of the definitions of Hamilton–Jacobi coordinates, momentum and angular momentum in predictive relativistic mechanics together with some techniques to calculate them in our perturbative scheme. Proofs and explanations are omitted and the reader is referred to the work of Bel and Martin.[6]

Next, in Sec. 5, we calculate the first radiative "self-term" of Hamilton–Jacobi's coordinates, momentum and angular momentum in the perturbative scheme. For all those magnitudes the terms in $e_1 e_2$ can be found in Ref. 6, while terms in $e_1^2 e_2^2$ for the Hamilton–Jacobi's momenta will be given in paper II (this issue). Our calculations show that our radiative system is not conservative in the sense of Ref. 6: Angular momentum does not recover its free particle expression after the two particles have undergone mutual interaction. This fact allows us to compute the lower "self-term" of the total intrinsic angular momentum radiated by the system. This is done by calculating the limit for the "future infinity of the first radiative "self-term" in the intrinsic angular momentum.

Finally we calculate the 3-accelerations of the two charges to third order in $1/c$. This gives us the first correction to the accelerations which are derived from Darwin's Lagrangian,[7] when outgoing radiation is accounted for.

More detailed calculations, including scattering cross sections and the 4-momentum balance of a scattering process, will be given in paper II.

1. LORENTZ–DIRAC EQUATION FROM THE POINT OF VIEW OF PREDICTIVE RELATIVISTIC MECHANICS (PRM)

Let us consider a system of n point structureless classical particles. In PRM the dynamics of such a system is governed by a differential system of the form

$$\frac{dx_a^\alpha}{ds_a} = u_a^\alpha,$$

$$\frac{du_a^\alpha}{ds_a} = \xi_a^\alpha(x_b^\beta, u_c^\gamma)$$

$$(\alpha,\beta,\gamma,\dots = 0,1,2,3, a,b,c,\dots = 1,2,\dots,n), \tag{1.1}$$

where x_a^α, u_a^α, and s_a^α stand for the 4-position, 4-velocity, and proper time of the particle a. The functions ξ_a^α (the accelerations) are Poincaré invariant 4-vectors which are the solution of the system

$$u_{a'}^\rho \frac{\partial \xi_a^\alpha}{\partial x^{a'\rho}} + \xi_{a'}^\rho \frac{\partial \xi_a^\alpha}{\partial u^{a'\rho}} = 0, \tag{1.2}$$

where we have raised index a' to invalidate the summation convention, which will only work in the case where equal indices stand in covariant and contravariant positions, respectively. According to this, the index ρ is summed in (1.2). Let us note that we will raise and lower latin indices without change of sign. Finally a' means "different from" a.

The functions ξ_a^α also satisfy the constraints

$$\xi_a^\alpha u_{\alpha a} = 0. \tag{1.3}$$

This guarantees that solutions of (1.1) initially satisfying $u_a^2 = -1$ (We choose signature $+2$) will maintain this relation forever.

We summarize here the main results of PRM. Further details can be found in the original papers.[6,8]

 0022-2488/79/071308-08$01.00

Using perturbation methods, and imposing the compatibility of PRM with classical electrodynamics Bel and Martin and co-workers have singled out the unique accelerations ξ^{α}_{a} describing the classical particle-particle electromagnetic interaction.[4,5,9] Classical electrodynamics, through the Lorentz force law and the formula of retarded (alternatively advanced or time-reversal) potentials, specifies the values of functions ξ^{α}_{a} for arguments x^{β}_{b} standing on null configurations,

$$(x^{\alpha}_{\alpha} - x^{\alpha}_{a'})(x_{\alpha a} - x_{a'\alpha}) = 0. \tag{1.4}$$

Because of (1.1) PRM is a dynamical theory of "Newtonian type" in the sense that a finite number of initial conditions (more precisely initial positions and velocities) are enough to determine the trajectories. Hence the word predictive in the name (predictive relativistic mechanics) of the theory. Following this terminology we will speak about *predictive electrodynamics* as it refers to the electrodynamics built into the framework of PRM.

PRM, as it has been described here, concerns itself with isolated systems of particles. In this sense it seems at first sight that predictive electrodynamics is unable to account for the fundamental phenomena of electromagnetic radiation. However, that this is not the case can be clearly seen if we take the point of view of Wheeler–Feynman[2,3] and others. According to these authors the theory of classical electromagnetic radiation is only a way to account for the interaction of a given system of charges with all other charges of the entire universe (theory of the absorber). In particular, the Lorentz–Dirac equation for an accelerated radiating charge is given by

$$\xi^{\alpha} = \frac{e}{m} F^{\alpha}{}_{\beta}u^{\beta} + \frac{2e^2}{3m}(\dot{\xi}^{\alpha} - \xi^{2}u^{\alpha}), \tag{1.5}$$

where u^{α},ξ^{α}, e, m are the 4-velocity, 4-acceleration, charge, and mass of the electric charge, respectively. $F^{\alpha\beta}$ is the given external *retarded* electromagnetic field acting on the charge. Finally $\dot{\xi}^{\alpha}$ stands for $(d/ds)\xi^{\alpha}$ and s is the proper time of the charge.

If we take Eq. (1.5) as the differential equation describing the charges' motion we would have to abandon predictive relativistic mechanics, since Eq. (1.5) is a third-order differential equation. Thus if we want to keep predictive electrodynamics, Eq. (1.5) has to be taken as a differential equation for the acceleration. In fact, we always have differential equations for the accelerations in PRM; e.g., Eqs. (1.2), which are by no means the differential equations of the particle motion. In each physical situation we must supply (1.2) with the good asymptotic conditions in order to get a unique acceleration ξ^{α}_{a}. Then if we plug into (1.1) the acceleration obtained in this way, we will be able to write down the equation of motion. Analogously, Eq. (1.5) has many solutions. (For example those accelerations which allow for the run away trajectories.) The problem consists in finding the right asymptotic conditions in order to select the physical ones. We assume now that the physical solutions, ξ^{α}, of (1.5) can be expanded in powers of e,

$$\xi^{\alpha} = e\ {}^{1}\xi^{\alpha} + e^{2}\ {}^{2}\xi^{\alpha} + \cdots. \tag{1.6}$$

It is obvious that (1.5) has a unique solution of the form (1.6). Furthermore, accelerations such as (1.6) exclude automatically the pathological solutions of (1.5) called run away solutions,[10] since these solutions are not analytical functions in e (see for instance Ref. 11). For instance we get for the first terms of (1.6)

$${}^{1}\xi^{\alpha} = \frac{1}{m}F^{\alpha}{}_{\beta}u^{\beta}, \quad {}^{2}\xi^{\alpha} = 0, \quad {}^{3}\xi^{a} = \frac{2}{3m^2}\dot{F}^{\alpha}{}_{\beta}u^{\beta}. \tag{1.7}$$

2. THE CASE OF TWO PARTICLES

We will consider, within the framework of predictive electrodynamics, the case of two charged particles mutually interacting, without external fields, but taking account of their electromagnetic radiation. We write an equation similar to (1.5) for each particle. (Remember that we raise and lower latin indices $a,b,\cdots$ without any change: $\xi^{a\alpha} \equiv \xi^{\alpha}_{a}$. Here again we have raised the index a to invalidate the summation convention.)

$$\xi^{\alpha}_{a} = \frac{e_a}{m_a} F^{\alpha}_{a'\beta}u^{\beta}_{a} + \frac{2e^2_a}{3m_a}(\dot{\xi}^{\alpha}_{a} - \xi^{2}_{a}u^{\alpha}_{a})$$
$$\left(\dot{\xi}^{\alpha}_{a} \equiv \frac{d\xi^{\alpha}_{a}}{ds_a}\right), \tag{2.1}$$

where a relates to the particle we are dealing with and $F^{\alpha}_{a'\beta}$ is the *retarded* electromagnetic field created by particle $a'(a' \neq a)$ on the particle a. The problem with equations (2.1) is that they are not differential equations—since the term $(e_a/m_a)F^{\alpha}_{a'\beta}u^{\beta}$ is not defined for any x^{α}_{1} and x^{α}_{2}, but only for null configurations

$$(x^{\alpha}_{2} - x^{\alpha}_{1})(x_{2\alpha} - x_{1\alpha}) = 0. \tag{2.2}$$

A similar problem must be faced when one considers the Lorentz equations of two interacting charges

$$\hat{\xi}^{\alpha}_{a} = (e_a/m_a)F^{\alpha}_{a'\beta}u^{\beta}, \tag{2.3}$$

where we have written $\hat{\xi}^{\alpha}_{a}$ instead of ξ^{α}_{a} to denote the accelerations. Using (2.3) as asymptotic conditions and making the assumption that accelerations can be expanded in powers of e_1e_2, a unique acceleration to be used in (1.1) can be obtained,[4,5] as explained before. This acceleration $\hat{\xi}^{\alpha}_{a}$, could, in principle, be determined by using a perturbative scheme in the coupling constant $g = e_1e_2$. The first two terms in the series, which we write in evident notation ${}^{(1,1)}\hat{\xi}^{\alpha}_{a}$, ${}^{(2,2)}\hat{\xi}^{\alpha}_{a}$, are given explicitly in Refs. 4 and 5. [Attention must be paid to the fact that the acceleration notation is slightly ambiguous. The acceleration $\hat{\xi}^{\alpha}_{a}$ about which we are speaking here is not the same as in (2.3). In (2.3), it is only defined for null configurations (2.2) and here it is defined for any pair of four positions x^{α}_{1} and x^{α}_{2} and coincides with $\hat{\xi}^{\alpha}_{a}$ in (2.3) for null configurations. Something similar can be said about functions ξ^{α}_{a} in (2.4) and (2.1), respectively.] Now we substitute Eq. (2.1) for the new equations,

$$\xi^{\alpha}_{a} = \hat{\xi}^{\alpha}_{a} + \frac{2e^2_a}{3m_a}(\dot{\xi}^{\alpha}_{a} - \xi^{2}_{a}u^{\alpha}_{a}). \tag{2.4}$$

That is to say

$$\xi^{\alpha}_{a} = \frac{2e_a^2}{3m_a}(\dot{\xi}^{\alpha}_{a} - \xi^{2}_{a}u^{\alpha}_{a}) + {}^{(1,1)}\widehat{\xi}^{\alpha}_{a} + {}^{(2,2)}\widehat{\xi}^{\alpha}_{a} + \cdots. \quad (2.5)$$

Assuming that ξ^{α}_{a} can be expanded in powers of e_1, e_2 we get for the first terms

$$^{(1,1)}\xi^{\alpha}_{a} = {}^{(1,1)}\widehat{\xi}^{\alpha}_{a}, \quad {}^{(2,2)}\xi^{\alpha}_{a} = {}^{(2,2)}\widehat{\xi}^{\alpha}_{a}, \quad (2.6)$$

$$^{(3,1)}\xi^{\alpha}_{a} = \frac{2}{3m_a}\,{}^{(1,1)}\dot{\xi}^{\alpha}_{a'}, \quad (2.7)$$

where $^{(3,1)}\xi^{\alpha}_{a}$ is the coefficient of $e_a^3 e_{a'}$ in the expansion of ξ^{α}_{a}.

Then we have a crucial point to clarify: Does ξ^{α}_{a}, as determined by (2.4), satisfy Eqs. (1.2)? It can be easily seen that this is the case for all orders. Furthermore, ξ^{α}_{a} in (2.4) satisfies (1.3), so it is also easy to see that (1.3) is satisfied by ξ^{α}_{a} given by (2.4).

Summing up, we have two accelerations, ξ^{α}_{a}, which describe a predictive relativistic dynamical interaction because Eqs. (1.2) and (1.3) are satisfied. Also ξ^{α}_{a} satisfies Eqs. (2.4) with Eq. (2.1) as asymptotic conditions; hence this predictive interaction describes two mutually interacting charged particles with outgoing radiation and without external fields. [For a more rigorous approach to Eq. (2.4) see Sanz,[12] who first and independently of us has adopted most of the points of view that we have developed in Secs. 1 and 2.]

3. THE FIRST RADIATIVE "SELF-TERM" IN THE ACCELERATIONS

As it has been pointed out in the last section, the accelerations ξ^{α}_{a} corresponding to the predictive electrodynamics of two isolated particles (no Dirac term, no external field) are in principle calculable within a perturbative scheme on the coupling constant. $g \equiv e_1 e_2$ and $^{(1,1)}\widehat{\xi}^{\alpha}_{a}, {}^{(2,2)}\widehat{\xi}^{\alpha}_{a}$ are explictly given in Refs. 4 and 5. Here we calculate the first radiative "self-term" in the accelerations; this is $^{(3,1)}\xi^{\alpha}_{a}$. According to (2.6), [(2.7)] we only need $^{(1,1)}\xi^{\alpha}_{a}$ ($= {}^{(1,1)}\widehat{\xi}^{\alpha}_{a}$) to calculate $^{(3,1)}\xi^{\alpha}_{a}$. For $^{(1,1)}\xi^{\alpha}_{a}$ we have the expression[4]

$$^{(1,1)}\xi^{\alpha}_{a} = \frac{\eta_a}{m_a r_a^3}[(\chi u_a)u^{\alpha}_{a'} - (u_1u_2)x^{\alpha}] \quad (\eta_1 \equiv 1, \ \eta_2 \equiv -1), \quad (3.1)$$

where $x^{\alpha}_1 - x^{\alpha}_2 \equiv x^{\alpha}$ and $(xu_a) \equiv x^{\alpha}u_{a\alpha}$, $(u_1u_2) \equiv u^{\alpha}_1 u_{2\alpha}$, and

$$r_a \equiv [x^2 + (xu_{a'})^2]^{1/2}. \quad (3.2)$$

Now we have to compute $\dot{\xi}^{\alpha}_{a}$. In order to make the calculations easier, let us introduce a system of new variables which will replace $x^{\alpha}, u^{\alpha}_1, u^{\alpha}_2$. We define the three linearly independent 4-vectors

$$h^{\alpha} \equiv x^{\alpha} - z_1 u^{\alpha}_1 + z_2 u^{\alpha}_2, \quad t^{\alpha}_{a} \equiv u^{\alpha}_a + (u_1u_2)u^{\alpha}_{a'} \quad (3.3)$$

where

$$z_a \equiv \frac{\eta_a[(xu_a) + (u_1u_2)(xu_{a'})]}{[(u_1u_2)^2 - 1]}. \quad (3.4)$$

These two 4-scalars, z_a, with $h^{\alpha}h_{\alpha}$ and $(u_1u_2)^2 - 1$ constitute a set of four independent variables:

$$h^2 \equiv h^{\alpha}h_{\alpha}, \quad \Lambda^2 \equiv (u_1u_2)^2 - 1, \quad z_a, \quad (3.5)$$

and they replace the 4-scalars x^2, (xua), (u_1u_2). With these new variables, using the definition of (3.2), r_a can be written

$$r_a = [h^2 + \Lambda^2 z_a^2]^{1/2}. \quad (3.6)$$

From (3.1) we get

$$^{(1,1)}\xi^{\alpha}_{a} = \frac{\eta_a k}{m_a r_a^3}h^{\alpha} - \frac{z_a}{m_a r_a^3}t^{\alpha}_{a'}, \quad (3.7)$$

where

$$k \equiv -(u_1u_2). \quad (3.8)$$

Now since we want the first order in $\dot{\xi}^{\alpha}_{a}$, $^{(1,1)}\dot{\xi}^{\alpha}_{a}$ in our notation, we can write

$$^{(1,1)}\dot{\xi}^{\alpha}_{a} = u^{\rho}_{\alpha}\frac{\partial^{(1,1)}\xi^{\alpha}_{a}}{\partial x^{a\rho}}. \quad (3.9)$$

On the other hand, it can be seen that

$$u^{\rho}_{a}\frac{\partial h^{\alpha}}{\partial x^{a\rho}} = 0, \quad u^{\rho}_{a}\frac{\partial t^{\rho}_{b}}{\partial x^{a\rho}} = 0, \quad (3.10)$$

$$u^{\rho}_{a}\frac{\partial \Lambda^2}{\partial x^{a\rho}} = 0, \quad u^{\rho}_{a}\frac{\partial z_b}{\partial x^{a\rho}} = \delta_{ab}, \quad (3.11)$$

thus

$$u^{\rho}_{a}\frac{\partial}{\partial x^{a\rho}} = \frac{\partial}{\partial z^{a}} \quad (3.12)$$

in the system of variables (3.5). Taking into account (3.10), (3.11), and (3.7), the calculation of (3.9) is straightforward. We get

$$^{(1,1)}\dot{\xi}^{\alpha}_{a} = -\frac{3\eta_a k\Lambda^2 z_a}{m_a r_a^5}h^{\alpha} + \left(\frac{3\Lambda^2 z_a^2}{m_a^2 r_a^5} - \frac{1}{m_a^2 r_a^3}\right)t^{\alpha}_{a'} \quad (3.13)$$

and using (2.7)

$$^{(3,1)}\xi^{\alpha}_{a} = -\frac{2\eta_a k\Lambda^2 z_a}{m_a^2 r_a^5}h^{\alpha} + \left(\frac{2\Lambda^2 z_a^2}{m_a^2 r_a^5} - \frac{2}{3m_a^2 r_a^3}\right)t^{\alpha}_{a'}. \quad (3.14)$$

4. MOMENTUM AND ANGULAR MOMENTUM FOR TWO PARTICLES

In this section we review the definitions of momentum and angular momentum in PRM, the asymptotic conditions and calculational techniques. Proofs and detailed explanations are omitted. We refer the reader to Ref. 6 for them.

The momentum P^{α} is a 4-vector *invariant by M_4 translations* such that

$$\frac{dP^{\alpha}}{ds_a} = 0 \quad \left(a = 1,2; \frac{d}{ds_a} = u^{\alpha}_{a}\frac{\partial}{\partial x^{a\alpha}} + \xi^{\alpha}_{a}\frac{\partial}{\partial u^{a\alpha}}\right). \quad (4.1)$$

The angular momentum is an antisymmetric 4-tensor, $J^{\alpha\beta}$, such that

$$\frac{dJ^{\alpha\beta}}{ds_a} = 0 \quad (4.2)$$

and its behavior under M_4 translations is given by

$$\frac{\partial J^{\alpha\beta}}{\partial x^{\gamma}_1} + \frac{\partial J^{\alpha\beta}}{\partial x^{\gamma}_2} = \delta^{\beta}_{\gamma}P^{\alpha} - \delta^{\alpha}_{\gamma}P^{\beta}. \quad (4.3)$$

Let us consider a canonical coordinate system q_a^α, p_a^α in PRM. If $q_a^\alpha - x_a^\alpha$, p_a^α are invariant vectors under the Poincaré group they are said to form an *adapted* canonical coordinate system. For such a coordinate system we have

$$P^\alpha = p_1^\alpha + p_2^\alpha, \tag{4.4}$$

$$J^{\alpha\beta} = q_a^\alpha p^{a\beta} - q_a^\beta p^{a\alpha}, \tag{4.5}$$

where P^α, $J^{\alpha\beta}$ given by these expressions are a solution of (4.1) and (4.2), (4.3), respectively. A particular kind of adapted canonical coordinates are the so-called Hamilton–Jacobi coordinates, which are characterized by the supplementary relations

$$\frac{dp_a^\alpha}{ds_b} = 0, \quad \frac{dq_a^\alpha}{ds_b} = \delta_{ab} p_a^\alpha. \tag{4.6}$$

We will first calculate a system of Hamilton–Jacobi coordinates and then (4.4) and (4.5) will give us P^α and $J^{\alpha\beta}$. In this way when determining momentum and angular momentum we get a coordinate canonical system and this is interesting if one wants to quantize the system. Calculations will be made in the perturbative scheme that we have mentioned in paragraph two. So, we make the assumption that p_a^α, q_a^α can be expanded in a power series of e_1, e_2. Accelerations ξ_a^α are known explicitly in this perturbative scheme up to fourth order: ${}^{(1,1)}\xi_a^\alpha$ is given in (3.1), ${}^{(3,1)}\xi_a^\alpha$ (term relative to $e_a^3 e_{a'}$) in (3.6) and ${}^{(2,2)}\xi_a^\alpha$ is not given here because we will not use it (it can be found in Ref. 4).

To the same fourth order we have to calculate ${}^{(1,1)}p_a^\alpha$, ${}^{(3,1)}p_a^\alpha$, and ${}^{(2,2)}p_a^\alpha$ and the same for q_a^α. The first terms, ${}^{(1,1)}p_a^\alpha$, and ${}^{(1,1)}q_a^\alpha$, can be found in Ref. 6: Obviously they are the same as in the more conventional case where the Dirac term is absent. We will limit ourselves to the calculation of the terms (3.1) which represent the first radiative "self-terms." We are able to do this because in the differential equations (4.6) they are not coupled to terms ${}^{(2,2)}p_a^\alpha$, ${}^{(2,2)}q_a^\alpha$. Neither are they coupled to ${}^{(2,2)}\xi_a^\alpha$.

In order to get a family of Hamilton–Jacobi coordinates giving us to this order one unique momentum and one unique intrinsic angular momentum, we need to define appropriate asymptotic conditions for Eqs. (4.6). To do so we work with the system of new variables (3.3).

The asymptotic conditions that we are going to define are of two different kinds, corresponding, roughly speaking, to the assumption that we have a free particle system when (a) $x^2 \to +\infty$ or when (b) $e_a e_{a'} \to 0$. As far as case (a) is concerned it can be seen that in the system of variables h^2, z_a, Λ^2, the limit $x^2 \to +\infty$ corresponds to one or both of these two different situations,

$$x^2 \to +\infty \Rightarrow \begin{cases} \text{(I)} & h^2 \to \infty, \quad \forall\, z_a, \\ \text{(II)} & z_a \to \gamma\infty, \quad \forall\, h^2, \end{cases} \tag{4.7}$$

where γ takes the values $+1$ or -1. Situation (I) means that we consider successive pairs of trajectories more and more further away. Situation (II) means that we go to future infinity ($\gamma = 1$) or past infinity ($\gamma = -1$) along the straight lines defined by the two initial 4-velocities. To distinguish both cases—future or past infinity—we will put $x^2 \to \infty_f$ or $x^2 \to \infty_p$, respectively. With these notations the first group of asymptotic conditions to be attached to (4.6) is

$$\lim_{x^2 \to \infty_p} p_a^\alpha = m_a u_a^\alpha, \quad \lim_{x^2 \to \infty_p} \frac{1}{X}(q_a^\alpha - x_a^\alpha) = 0, \tag{4.8}$$

or

$$\lim_{x^2 \to \infty_f} p_a^\alpha = m_a u_a^\alpha, \quad \lim_{x^2 \to \infty_f} \frac{1}{X}(q_a^\alpha - x_a^\alpha), \tag{4.9}$$

while the second group leads to

$${}^{(0,0)}p_a^\alpha = m_a u_a^\alpha, \quad {}^{(0,0)}q_a^\alpha = x_a^\alpha. \tag{4.10}$$

Because of (4.8) and (4.9), p_a^α, q_a^α are called "regular" in past infinity or future infinity, respectively.

It can be proved that if for a PRM system there exist invariant Poincaré vectors $p_a^\alpha, q_a^\alpha - x_a^\alpha$, such that (4.6), (4.8) [or alternatively (4.9)] are satisfied, then p_a^α, q_a^α are a set of canonical coordinates and so a set of Hamilton–Jacobi coordinates. Finally it can also be shown in a perturbative framework that Hamilton–Jacobi coordinates regular at infinity really do exist.[13]

Under supplementary assumptions which roughly speaking reduce again to the general assumption that we get a free particle system when $x^2 \to +\infty$, it can be proved that

$$p_a^\alpha p_{a\alpha} = -m_a^2. \tag{4.11}$$

These identities show that, in the language of Dirac,[14] we are in the presence of a dynamical system with primary constraints. These constraints have their origin in the identities $u_a^2 = -1$. Primary constraints introduce difficulties when one tries to quantize classical dynamical systems. One way of getting round this problem is to substitute the dynamical system (1.1) by a new auxiliary one, the so-called "auxiliary dynamical system":

$$\frac{dx_a^\alpha}{d\lambda} = \pi_a^\alpha, \quad \frac{d\pi_a^\alpha}{d\lambda} = \theta_a^\alpha(x_b^\beta, \pi_c^\gamma), \tag{4.12}$$

where λ is a 4-scalar parameter and ($\pi_a^2 \equiv -\pi_a^\alpha \pi_{a\alpha}$)

$$\theta_a^\alpha(x_b^\beta, \pi_c^\gamma) = \pi_a^2 \xi_a^\alpha(x_b^\beta, \pi_c^{-1}\pi_c^\gamma, m_d \to \pi_d). \tag{4.13}$$

In this way we go round the constraints $u_a^2 = -1$ since the π_a^2 are now to be considered as two new independent variables to add to the other four: x^2, $(x\pi a)$, $(\pi_1\pi_2)$. Definitions (3.3) and (3.4) now become

$$\tilde{h}^\alpha \equiv x^\alpha - \tilde{z}_1 \pi_1^\alpha + \tilde{z}_2 \pi_2^\alpha, \quad \tilde{t}_a^\alpha \equiv \pi_{a'}^2 \pi_a^\alpha + (\pi_1\pi_2)\pi_{a'}^\alpha, \tag{4.14}$$

and

$$\tilde{z}_a \equiv \frac{\eta_a[\pi_{a'}^2(x\pi_a) + (\pi_1\pi_2)(x\pi_{a'})]}{\tilde{\Lambda}_2}, \tag{4.15}$$

with

$$\tilde{\Lambda}^2 \equiv (\pi_1\pi_2)^2 - \pi_a^2 \pi_{a'}^2. \tag{4.16}$$

Now we can determine Hamilton–Jacobi coordinates, $\tilde{p}_a^\alpha, \tilde{q}_a^\alpha$, regular at past (or future) infinity for the dynamical system (4.12), taking 4-vectors $\tilde{p}_a^\alpha, \tilde{q}_a^\alpha$ such that $\tilde{p}_a^\alpha$, $q_a^\alpha - x_a^\alpha$ are Poincaré invariant and such that they are solutions of (4.6) satisfying the asymptotic conditions

$$\lim_{x^{i}\to\infty_{\mu}} \tilde p_a^\alpha = \pi_a^\alpha, \quad \lim_{x^{i}\to\infty_{\mu}} \frac{1}{x}(\tilde q_a^\alpha - x_a^\alpha), \tag{4.17}$$

or the equivalent ones for the infinite future, all of which are the subsidiary conditions corresponding to (4.8) or (4.9), respectively. On the other hand, conditions (4.10) become

$$^{(0,0)}\tilde p_a^\alpha = \pi_a^\alpha, \quad ^{(0,0)}\tilde q_a^\alpha = x_a^\alpha. \tag{4.18}$$

The differential operator d/ds_a is now

$$\frac{d}{ds_a} = \pi_a^\alpha \frac{\partial}{\partial x^{a\alpha}} + \theta_a^\alpha \frac{\partial}{\partial \pi^{a\alpha}}$$

instead of the similar expression given in (4.1). Finally the identity (4.11) becomes

$$\tilde p_a^\alpha \tilde p_{a\alpha} = -\pi_a^\alpha \pi_{a\alpha} \equiv \pi_a^2. \tag{4.19}$$

From $\tilde p_a^\alpha$, $\tilde q_a^\alpha$ we get the momentum, $\widetilde P^\alpha$, and angular momentum, $\widetilde J^{\alpha\beta}$ of the auxiliary dynamical system (4.11) through expressions like (4.4), (4.5). Then it can be proved that we obtain P^α and $J^{\alpha\beta}$, the momentum and angular momentum, respectively, of the original dynamical system (1.1) by "mass geometrization" of $\widetilde P^\alpha$ and $\widetilde J^{\alpha\beta}$, that is, making the substitution $\pi_a^\alpha \to m_a u_a^\alpha$.

5. CALCULATION OF THE LOWEST RADIATIVE "SELF-TERM" IN MOMENTUM AND ANGULAR MOMENTUM

Taking (3.14) into account we get from (4.13)

$$^{(3,1)}\theta_a^\alpha = -\frac{2\eta_a \pi_{a'}^2 \widetilde k \widetilde A^2 \tilde z_a}{\pi_a^2 \tilde r_a^5}\widetilde h^\alpha + \frac{2\pi_{a'}^2}{\pi_a^2 \tilde r_a^3}\left(\frac{\widetilde A^2 \tilde z_a^2}{\tilde r_a^2} - \frac{1}{3}\right)\widetilde t_{a'}^\alpha, \tag{5.1}$$

where

$$\tilde r_a \equiv (\pi_{a'}^2 \widetilde h^2 + \widetilde A^2 \tilde z_a^2)^{1/2}, \quad \widetilde k \equiv -(\pi_1 \pi_2). \tag{5.2}$$

Let us write $\tilde p_a^\alpha$ in the general form

$$\tilde p_a^\alpha = \eta_a \tilde\alpha_a \widetilde h^\alpha + \tilde\mu_{aa} \tilde t_a^\alpha + \tilde\mu_{aa'} \widetilde t_{a'}^\alpha. \tag{5.3}$$

In Ref. 6 the reader will find the first order expressions for $\tilde p_a^\alpha$ and $\tilde q_a^\alpha$, which in our notations are $^{(1,1)}\tilde p_a^\alpha, {}^{(1,1)}\tilde q_a^\alpha$. Here we are interested in $^{(3,1)}\tilde p_a^\alpha$ (and later on in $^{(3,1)}\tilde q_a^\alpha$). From the first group of Eqs. (4.6) we get to this order

$$\pi_b^\rho \frac{\partial^{(3,1)}\tilde p_a^\alpha}{\partial x^{b\rho}} + {}^{(3,1)}\theta_b^\rho \frac{\partial^{(0,0)}\tilde p_a^\alpha}{\partial \pi^{b\rho}} = 0, \tag{5.4}$$

or taking (4.19) into account

$$D_b\, {}^{(3,1)}\tilde p_a^\alpha = -\delta_{ab}\, {}^{(3,1)}\theta_b^\alpha, \tag{5.5}$$

where δ_{ab} is the Kronecker delta and D_b is the differential operator

$$D_b \equiv \pi_b^\rho \frac{\partial}{\partial x^{b\rho}}. \tag{5.6}$$

Now it can be seen that

$$D_b \widetilde h^\alpha = D_b \tilde t_a^\alpha = 0. \tag{5.7}$$

From (4.19) and the first equation of (4.18) we get

$$^{(3,1)}\tilde\mu_{aa} = 0. \tag{5.8}$$

Taking this and (5.7) into account, Eqs. (5.4) become

$$\frac{\partial^{(3,1)}\tilde\alpha_a}{\partial \tilde z_b} = \frac{2\pi_{a'}^2 \widetilde k \widetilde A^2 \tilde z_a}{\pi_a^2 \tilde r_a^5}\delta_{ab}, \tag{5.9}$$

$$\frac{\partial^{(3,1)}\tilde\mu_{aa'}}{\partial \tilde z_b} = \left(\frac{2\pi_{a'}^2}{3\pi_a^2 \tilde r_a^3} - \frac{2A^2 \tilde z_a^2 \pi_{a'}^2}{\pi_a^2 \tilde r_a^5}\right)\delta_{ab}, \tag{5.10}$$

where use has been made of the following results:

$$D_b \pi_a^2 = D_b \widetilde h^2 = D_b \widetilde A^2 = 0, \quad D_b \tilde z_a = \delta_{ab}. \tag{5.11}$$

[We can get (3.10) and (3.11) from (5.7) and (5.11), respectively, by making the substitutions: $\pi_a^\alpha \to m_a u_a^\alpha$.] On the other hand, we have the asymptotic conditions (4.20) (or the equivalent ones for the future). The only solution to (5.8), (5.9) "regular" at infinity is

$$^{(3,1)}\tilde\alpha_a = \frac{2\pi_{a'}^2 \widetilde k \widetilde A^2}{\pi_a^2}\int_{\gamma\infty}^{\tilde z_a} \frac{\tilde z_a d\tilde z_a}{\tilde r_a^5} = -\frac{2\widetilde k \pi_{a'}^2}{3\pi_a^2 \tilde r_a^3}, \tag{5.12}$$

$$^{(3,1)}\tilde\mu_{aa'} = \frac{2\pi_{a'}^2}{3\pi_a^2}\int_{\gamma\infty}^{\tilde z_a} \frac{d\tilde z_a}{\tilde r_a^3} - \frac{2\widetilde A^2 \pi_{a'}^2}{\pi_a^2} \times \int_{\gamma\infty}^{\tilde z_a} \frac{\tilde z_a^2 d\tilde z_a}{\tilde r_a^5} = \frac{2\pi_{a'}^2 \tilde z_a}{3\pi_a^2 \tilde r_a^3}. \tag{5.13}$$

So we have

$$^{(3,1)}\tilde p_a^\alpha = -\frac{2\eta_a \pi_{a'}^2 \widetilde k}{3\pi_a^2 \tilde r_a^3}\widetilde h^\alpha + \frac{2\pi_{a'}^2 \tilde z_a}{3\pi_a^2 \tilde r_a^3}\widetilde t_{a'}^\alpha. \tag{5.14}$$

Let us calculate $^{(3,1)}\tilde q_a^\alpha$. From the second equation (4.6) we get

$$D_b\, {}^{(3,1)}\tilde q_a^\alpha = {}^{(3,1)}\tilde p_a^\alpha \delta_{ab} \tag{5.15}$$

and if we write

$$^{(3,1)}\tilde q_a^\alpha = \eta_a\, {}^{(3,1)}\tilde\gamma_a \widetilde h^\alpha + {}^{(3,1)}\tilde v_{aa} \tilde t_a^\alpha + {}^{(3,1)}v_{aa'} \widetilde t_{a'}^\alpha, \tag{5.16}$$

Eqs. (5.15) become

$$\frac{\partial^{(3,1)}\tilde\gamma_a}{\partial \tilde z_b} = \frac{-2\pi_{a'}^2 \widetilde k}{3\pi_a^2 \tilde r_a^3}\delta_{ab}, \tag{5.17}$$

$$\frac{\partial^{(3,1)}\tilde v_{aa'}}{\partial \tilde z_b} = \frac{2\pi_{a'}^2 \tilde z_a}{3\pi_a^2 \tilde r_a^3}\delta_{ab}, \tag{5.18}$$

$$\frac{\partial^{(3,1)}\tilde v_{aa'}}{\partial \tilde z_b} = 0. \tag{5.19}$$

The general solution of these equations satisfying the asymptotic condition given by the second of the equations (4.17) (or the equivalent ones for the future) gives the following expression for $^{(3,1)}\tilde q_a^\alpha$,

$$^{(3,1)}\tilde q_a^\alpha = \frac{2\eta_a \widetilde k}{3\pi_a^2 \widetilde h^2}\left(\frac{\gamma}{\widetilde A} - \frac{\tilde z_a}{\tilde r_a}\right)h^\alpha - \frac{2\pi_{a'}^2}{3\pi_a^2 \widetilde A^2 \tilde r_a} t_{a'}^\alpha + {}^{(3,1)}_{\ *}v_{aa} \tilde t_a^\alpha + {}^{(3,1)}_{\ *}v_{aa'} \widetilde t_{a'}^\alpha, \tag{5.20}$$

where ${}^{(3,1)}_{\bullet}\nu_{ab}$ are arbitrary functions of $\widetilde{h}^2, \widetilde{A}^2, \pi_b$ except for the condition that $\lim \widetilde{h}^{-1} {}_\bullet\nu_{ab} = 0$. We can see that at variance with ${}^{(3,1)}\widetilde{p}^\alpha_a$, we do not have uniqueness for the coordinates ${}^{(3,1)}\widetilde{q}^\alpha_a$.

Using expressions (5.14), the formula (4.4), and making the substitution $\pi^\alpha_a \to m_a u^\alpha_a$ we get one unique radiative "self-term" in linear momentum,

$$
{}^r P^\alpha \equiv e_1^3 e_2 \, {}^{(3,1)}p_1^\alpha + e_2^3 e_1 \, {}^{(3,1)}p_2^\alpha = \tfrac{2}{3} e_1 e_2 k \left(\frac{e_2^2}{m_2 r_2^3} - \frac{e_1^2}{m_1 r_1^3} \right) h^\alpha + \tfrac{2}{3} e_1 e_2 \left(\frac{e_1^3 z_1}{m_1 r_1^3} t_2^\alpha + \frac{e_2^3 z_2}{m_2 r_2^3} t_1^\alpha \right). \tag{5.21}
$$

Since the ${}^{(3,1)}\widetilde{q}^\alpha_a$ given by (5.20) are not unique, the radiative "self-term" in angular momentum ${}^r J^{\alpha\beta}$, given by (4.5), (5.14), (5.20) is not unique. Nevertheless it can be seen that the radiative "self-term," ${}^r W^\alpha$, of the intrinsic angular momentum W^α, defined by

$$
W^\alpha \equiv \frac{1}{2P} \delta^{\alpha\beta\lambda\mu} P_\beta J_{\lambda\mu} [P \equiv (-P^\alpha P_\alpha)^{1/2}, \ \delta^{0123} = 1] \tag{5.22}
$$

is unique. We first get

$$
e_1^3 e_2 \, {}^{(3,1)}\widetilde{W}^\alpha + e_2^3 e_1 \, {}^{(1,3)}\widetilde{W}^\alpha = \frac{1}{[-({}^{(0,0)}\widetilde{P}^\alpha \, {}^{(0,0)}\widetilde{P}_\alpha)]^{1/2}} \Bigg[e_1^3 e_2 ({}^{(3,1)}\gamma_1 - \tilde{z}_1 \, {}^{(3,1)}\widetilde{\alpha}_1) + e_2^3 e_1 ({}^{(3,1)}\gamma_2 - \tilde{z}_2 \, {}^{(3,1)}\widetilde{\alpha}_2)
$$

$$
- \left(\frac{\widetilde{A}^2}{{}^{(0,0)}\widetilde{P}^\alpha \, {}^{(0,0)}\widetilde{P}_\alpha} + k \right) (e_1^3 e_2 \, {}^{(3,1)}\widetilde{\mu}_{12} + e_2^3 e_1 \, {}^{(3,1)}\widetilde{\mu}_{21}) \Bigg] \delta^{\alpha\beta\gamma\delta} x_\beta \pi_{1\gamma} \pi_{2\delta} \tag{5.23}
$$

and then

$$
{}^r W^\alpha \equiv e_1^3 e_2 \, {}^{(3,1)}W^\alpha + e_2^3 e_1 \, {}^{(1,3)}W^\alpha = \frac{1}{(m_1^2 + m_2^2 + 2m_1 m_2 k)^{1/2}} \Bigg\{ e_1^3 e_2 \Bigg[\frac{2m_2 k}{3m_1 h^2} \left(\frac{\gamma}{A} - \frac{z_1}{r_1} \right)
$$

$$
+ \frac{2m_2^2 A^2 z_1}{3(m_1^2 + m_2^2 + 2m_1 m_2 k) r_1^3} \Bigg] + e_2^3 e_1 \Bigg[\frac{2m_1 k}{3m_2 h^2} \left(\frac{\gamma}{A} - \frac{z_2}{r_2} \right)
$$

$$
+ \frac{2m_1^2 A^2 z_2}{3(m_1^2 + m_2^2 + 2m_1 m_2 k) r_2^3} \Bigg] \Bigg\} n^\alpha, \tag{5.24}
$$

where we have put

$$
\delta^{\alpha\beta\lambda\mu} x_\beta u_{1\lambda} u_{2\mu} \equiv n^\alpha. \tag{5.25}
$$

One could ask why we have not considered the terms ${}^{(1,3)}\widetilde{p}^\alpha_a, {}^{(1,3)}\widetilde{q}^\alpha_a$. The answer is that if one considers them, then a similar calculation such as the one that has been used to calculate ${}^{(3,1)}\widetilde{p}^\alpha_a, {}^{(3,1)}\widetilde{q}^\alpha_a$, gives ${}^{(1,3)}\widetilde{p}^\alpha_a = 0$, ${}^{(1,3)}\widetilde{q}^\alpha_a = {}^{(1,3)}_{\bullet}\nu_{aa'}\widetilde{t}^\alpha_a + {}^{(1,3)}_{\bullet}\nu_{aa'}\widetilde{t}^\alpha_{a'}$ and these expressions do not change either (5.21) or (5.24). ${}^r W^\alpha$ depends on γ and so it has different values at past or future infinity (it goes to zero at one of these two infinities depending on γ). In the language of Ref. 6, when radiation is present, W^α is not "conservative." Of course, W^α maintains its numerical values along a given pair of trajectories, but it does not keep the form of the expressions corresponding to free particles, which W^α takes at the past (or future) infinity. On the other hand, P^α does not depend on γ, so it is "conservative."

When the Dirac term is not considered, P^α, W^α are conservative to first order[6]: That is, ${}^{(1,1)}P^\alpha$, ${}^{(1,1)}W^\alpha$ do not depend on γ. [It can be easily recognized that terms like ${}^{(n,n)}P^\alpha$, ${}^{(n,n)}W^\alpha$ are the same if we consider the Dirac term (Lorentz–Dirac equation) as if we do not (Lorentz equation with retarded potentials).].

So, to fourth order in $e_1^n e_2^m$ $(n + m \leqslant 4)$, we have that the momentum P^α

$$
P^\alpha = m_1 u_1^\alpha + m_2 u_2^\alpha + e_1 e_2 \, {}^{(1,1)}P^\alpha + e_1^2 e_2^2 \, {}^{(2,2)}P^\alpha + {}^r P^\alpha, \tag{5.26}
$$

with ${}^r P^\alpha$ given by (5.21), is numerically conserved along any given pair of trajectories and the same can be said about W^α. In the spirit of field theory and to this order the first part of (5.26)— ${}^p P^\alpha \equiv m_1 u_1^\alpha + m_2 u_2^\alpha + e_1 e_2 \, {}^{(1,1)}P^\alpha + e_1^2 e_2^2 \, {}^{(2,2)}P^\alpha$— could be interpreted as the total momentum of the system, consisting of the two charges plus their electromagnetic field, excluding self-interactions, while ${}^r P^\alpha$ would represent the momentum of the radiated electromagnetic field that takes the self-interaction into account. Neither ${}^p P^\alpha$, nor ${}^r P^\alpha$ are conserved numerically by themselves; but their sum to this order is conserved. In a similar way we could define ${}^p W^\alpha$ and then decompose W^α as ${}^p W^\alpha + {}^r W^\alpha$.

Now we make the choice $\gamma = -1$ in (5.24), so we take the intrinsic angular momentum W^α to be "regular" at past infinity.

Let us introduce ${}^r W$ such that ${}^r W^\alpha = {}^r W n^\alpha$. Then obviously we have

$$
\lim_{x^2 \to \infty_p} {}^r W = 0
$$

and for future infinity

$$ {}^rW_{+\infty} \equiv \lim_{x^0\to\infty} {}^rW = -\frac{4m_1m_2e_1e_2(e_1^2/m_1^2+e_2^2/m_2^2)k}{3(m_1^2+m_2^2+2m_1m_2k)^{1/2}\Lambda h^2}. \quad (5.27) $$

Since

$$ \lim_{x^0\to\infty} {}^{(1,1)}W^\alpha = 0, $$

we have the evident notation

$$ \lim_{x^0\to\infty} W^\alpha = {}^{(0,0)}W^\alpha + e_1^2e_2^2\,{}^{(2,2)}W_{+\infty}n^\alpha + {}^rW_{+\infty}n^\alpha, \quad (5.28) $$

where it is clear that having $W^\alpha = {}^{(0,0)}W^\alpha$ at past infinity we do not recover the form of this expression in the infinite future. [From the definition (5.25) we get

$$ {}^{(0,0)}W^\alpha = m_1m_2n^\alpha/(m_1^2+m_2^2+2m_1m_2k)^{1/2}.] $$

On the other hand, if we calculate

$$ {}^rP^\alpha_{+\infty} \equiv \lim_{x^0\to\infty} {}^rP^\alpha $$

we will get zero since ${}^rP^\alpha$ does not depend on γ.

From the point of view of the "theory of the absorber" (see Sec. 1) $e_1^2e_2^2\,{}^{(2,2)}W_{+\infty}n^\alpha + {}^rW_{+\infty}n^\alpha$ could be interpreted as all the intrinsic angular momentum that our two particles, have delivered during all their history to the charges of the entire universe. The term ${}^rW_{+\infty}n^\alpha$ accounts for the intrinsic angular momentum which corresponds, in the language of field theory to the "self-terms" of the radiated electromagnetic field, while the other term $e_1^2e_2^2\,{}^{(2,2)}W_{+\infty}n^\alpha$ belongs to the system consisting of the two charges plus their electromagnetic field (self-interaction excluded). This term would be absent if the two charges were to interact through time-reversal invariant potentials instead of through retarded ones, as it is actually the case in the Lorentz–Dirac equation (2.1).

Similar considerations can be made about the momentum P^α but now the term ${}^rP^\alpha_{+\infty}$ cancels, that is to say, in our approximation, the all radiated "self-term" in the momentum by the two interacting particles, is zero.

Probably we would have to go on with our expansion in $e_1^ne_2^m$ up to order $n+m=6$ to get ${}^rP^\alpha_{+\infty}\neq 0$. In fact dipolar radiation in conventional electrodynamics begins at sixth order in the charges involved.

We end this paper by making some physical considerations. First of all, we could recover the formal expansions in the charges of this article by making the reasonable assumption that the magnitudes we have calculated can be expanded in powers of the dimensionless quantities e_ae_b/c^2m_cx ($x\equiv 3$ – distance between the two charges). These are not Lorentz scalar quantities. So, we cannot attach an invariant meaning to the fact that these quantities were small in a given inertial system. In spite of this, expansions in e_ae_b/c^2m_cx are meaningful since, because of the Poincaré invariance, accelerations, momentum, etc., are always the same functions of positions and velocities, no matter which inertial system we are talking about. Then each inertial observer must only make sure that quantities e_ae_b/c^2m_cx are small enough for him in order to get a fast convergence of the expansions.

For electrons we have $e_ae_b/c^2m_cx\sim 1$ when $x\sim 3$ Fermi. So we can see that e_ae_b/c^2m_cx will be very small in all the physical situations where the classical theory developed here can be used.

In the case of the accelerations, where the terms ${}^{(2,2)}\xi^\alpha_a$ are known,[4] we can easily get the 3-accelerations $\mu^i_a = d^2x^i_a/dt^2$ up to third order in $1/c$. (see, for instance, Ref. 8 about the three-dimensional formalism of the PRM.) From ${}^{(1,1)}\xi^\alpha_a$, ${}^{(3,1)}\xi^\alpha_a$, given by (3.1), (3.14), respectively, from ${}^{(2,2)}\xi^\alpha_a$,[4] and making use of the relation[7]

$$ \mu^i_a = \left(1-\frac{v_a^2}{c^2}\right)\left[(\xi^i_a)_{t_1=t_2} - \frac{1}{c}(\xi^0_a)_{t_1=t_2}V^i_a\right] \quad (5.29) $$

(where V^i_a is the 3-velocity of the particle a) we get to third order in $1/c$

$$ \begin{aligned} \mu^i_a = {} & \frac{\eta_ae_1e_2}{m_ar^3}r^i + \frac{\eta_ae_1e_2}{c^2m_ar^3}\left[\left(\frac{1}{2}V^2_{a'} - \mathbf{V}_1\cdot\mathbf{V}_2 - \frac{3}{2}\frac{(\mathbf{r}\cdot\mathbf{V}_{a'})^2}{r^2}\right)r^i \right.\\ & \left. - \eta_a(\mathbf{r}\cdot\mathbf{V}_a)V^i\right] + \frac{\eta_ae_1^2e_2^2}{c^2m_1m_2r^4}r^i + \frac{2\eta_ae_1^2e_2^2\mathbf{r}\cdot\mathbf{V}}{c^3m_1m_2r^5}r^i \\ & - \frac{2\eta_ae_a^3e_{a'}\mathbf{r}\cdot\mathbf{V}}{c^3m_a^2r^5}r^i - \frac{2\eta_ae_1^2e_2^2}{3c^3m_1m_2r^3}V^i + \frac{2\eta_ae_a^3e_{a'}}{3c^3m_a^2r^3}V^i, \end{aligned} \quad (5.30) $$

where we have put

$$ V^i \equiv V^i_1 - V^i_2,\quad r^i\equiv x^i,\quad V\equiv(\mathbf{V}\cdot\mathbf{V})^{1/2},\quad r\equiv(\mathbf{r}\cdot\mathbf{r})^{1/2} $$

and $\mathbf{V}\cdot\mathbf{V}$, $\mathbf{r}\cdot\mathbf{r}$ are R^3 scalar products.

In (5.30), to zeroth order, we recognize, Coulomb's law. The second order is the correction of Darwin's Lagrangian to Coulomb's law.[7] In the third we have two kind of terms: Only those in $e_a^3e_{a'}$, come from the Dirac term in the Lorentz–Dirac equation. We can think about it as a radiative term. The other two terms in $1/c^3$ have nothing to do with the Dirac term. They would be absent if we had taken time-reversal potentials instead of the retarded ones that must be used in the case of the Lorentz–Dirac equation.

Notice the fact that the term of μ^i_α in $1/c^3$ vanishes when $e_1/m_1 = e_2/m_2$.

ACKNOWLEDGMENTS

Helpful suggestions from J. Martin, J.L. Sanz, and X. Fustero are gratefully acknowledged.

[1]There is extensive literature on the Lorentz–Dirac equation, see for instance, the review article of G.N. Plass. [Rev. Mod. Phys. **33**, 37 (1961)]. References to later works can be found in E.G.P. Rowe, Phys. Rev. D **12**, 1576 (1975); R.G. Beil, *ibid.* **12**, 2266 (1975); and J.C. Herrera, *ibid.* **15**, 453 (1977), and in the papers to which these authors refer.

[2]J.A. Wheeler and R.P. Feynman, Rev. Mod. Phys. **17**, 157 (1945).

[3]J.A. Wheeler and R.P. Feynman, Rev. Mod. Phys. **21**, 425 (1949).
[4]L. Bel, A. Salas, and J.M. Sanchez, Phys. Rev. D **7**, 1099 (1973).
[5]A. Salas and J.M. Sanchez, Nuovo Cimento B **20**, 209 (1974).
[6]L. Bel and J. Martin, Ann. Inst. H. Poincaré, **22**, 173 (1975).
[7]R. Lapiedra and A. Mussons, Ann. Fis. **72**, 109 (1976).
[8]L. Bel, Ann. Inst. H. Poincaré, **14**, 189 (1971).
[9]L. Bel and J. Martin, Phys. Rev. D **8**, 4347 (1973).
[10]H.J. Bhabha, Phys. Rev. **70**, 759 (1946).
[11]F. Rohrlich, *Classical Charged Particles* (Addison-Wesley, Reading, Massachussetts, 1965).
[12]J.L. Sanz, thesis, Universidad Autónoma de Madrid (Cantoblanco) (1976).
[13]L. Bel and X. Fustero, Ann. Inst. H. Poincaré A **25**, 411 (1976).
[14]P.A.M. Dirac, *Lectures on Quantum Mechanics* (Belfer Graduate School of Science, New York, 1964).

Classical predictive electrodynamics of two charges with radiation: Energy and 3-momentum balance and scattering cross sections. II

R. Lapiedra
Departamento de Física Teórica, Facultad de Ciencias, Santander, Spain
F. Marqués and A. Molina
Departamento de Física Teórica, Facultad de Física, Barcelona, Spain

(Received 7 July 1978; revised manuscript received 6 November 1978)

We deal with a classical predictive mechanical system of two spinless charges where radiation is considered and there are no external fields. The terms ${}^{(2,2)}P_a^\alpha$ of the expansion in the charges of the Hamilton–Jacobi momenta are calculated. Using these, together with known previous results, we can obtain the p_a^α up to the fourth order. Then we have calculated the "radiated" energy and the 3-momentum in a scattering process as functions of the impact parameter and the incident energy for the former and 3-momentum for the latter. Scattering cross-sections are also calculated. Good agreement with well known results, including those of quantum electrodynamics, has been found.

INTRODUCTION

In this paper we pursue the calculations of the preceding paper (I) (this issue) in order to obtain physical results such as the cross sections and the "radiated" energy and 3-momentum. [We explain what is to be considered as radiated energy and 3-momentum in paper I (Sec. 5) and again at the end of Sec. 3 in this paper.] We use the notation and general scheme developed in Ref. 1, i.e., a classical predictive mechanical system consisting of two spinless charges, each one moving in the field of the other according to the appropriate Lorentz–Dirac equation.

Let us summarize the relevant results of (I): In a perturbative scheme in the charges, e_a, the 4-accelerations of the "auxiliary dynamical system" (I.4.12) are determined up to fourth order (included). [(I.4.12) means the formula (4.12) of Ref. 1.] The terms ${}^{(1,1)}\theta_a^\alpha$ and ${}^{(2,2)}\theta_a^\alpha$ of the expansion are the same as if radiation (that is, the Dirac term in the Lorentz–Dirac equation) was not present and they were already known. The term ${}^{(1,1)}\theta_a^\alpha$ can be found in Ref. 1. On the other hand, the terms ${}^{(2,2)}\xi_a^\alpha$, of the original dynamical system's accelerations, can be found in Ref. 2. From them, the terms ${}^{(2,2)}\theta_a^\alpha$ are easily calculated. They are given in Appendix B. The terms ${}^{(3,1)}\theta_a^\alpha$, calculated in (I), are new. The terms ${}^{(1,3)}\theta_a^\alpha$ can be seen to be absent.

The Hamilton–Jacobi coordinates $\tilde{p}_a^\alpha,\tilde{q}_a^\alpha$, can be computed taking into account the 4-accelerations, θ_a^α, of "the auxiliary dynamical system." The first terms ${}^{(1,1)}\widetilde{p}_a^\alpha$, ${}^{(1,1)}\tilde{q}_a^\alpha$, in the expansions of $\tilde{p}_a^\alpha,\tilde{q}_a^\alpha$ are obtained in Ref. 1. In (I) we have obtained the terms ${}^{(3,1)}\widetilde{p}_a^\alpha$, ${}^{(3,1)}\tilde{q}_a^\alpha$ (the terms ${}^{(1,3)}\widetilde{p}_a^\alpha$, ${}^{(1,3)}\tilde{q}_a^\alpha$ can be taken equal to zero). In order to have all terms to fourth order in $\tilde{p}_a^\alpha,\tilde{q}_a^\alpha$, the terms ${}^{(2,2)}\widetilde{p}_a^\alpha$, ${}^{(2,2)}\tilde{q}_a^\alpha$ should also be calculated. Their calculation follows the same lines as that of ${}^{(3,1)}\widetilde{p}_a^\alpha$, ${}^{(3,1)}\tilde{q}_a^\alpha$, but it is very cumbersome. Nevertheless if any more concrete results are to be obtained from our mechanical system, knowledge of ${}^{(2,2)}p_a^\alpha$ is needed.

In Sec. 1 we will compute ${}^{(2,2)}p_a^\alpha$ and so we will get the total 4-momentum, $P^\alpha = p_1^\alpha + p_2^\alpha$, of the system to fourth order. In Sec. 2 the future (respectively past) infinite limit of the calculated Hamilton–Jacobi momenta, p_a^α, are calculated. In this way we can establish the 4-momentum balance of the interaction and so, in our approximation. We see that there is no energy and 3-momentum "radiated". In Sec. 3 we use the asymptotic behavior of p_a^α in order to calculate the scattering cross section of the classical process, to sixth order in e_1, e_2. In the limit where sixth order terms are negligible, our results agree with the previous one of Bel[3] and they are also shown to agree with those of quantum electrodynamics.

1. THE CALCULATION OF THE TERMS ${}^{(2,2)}p_a^\alpha$ IN THE HAMILTON–JACOBI MOMENTA EXPANSION

We start with the first equation in (I.4.6), written for the "auxiliary dynamical system" (I.4.12),

$$\frac{d\tilde{p}_a^\alpha}{ds_b} = 0, \tag{1.1}$$

where

$$\frac{d}{ds_b} = \pi_b^\rho \frac{\partial}{\partial x^{b\rho}} + \theta_b^\rho \frac{\partial}{\partial \pi^{b\rho}}. \tag{1.2}$$

Fourth order terms of the form (2.2) (terms in $e_a^2 e_{a'}^2$) in (1.1) give

$$\pi_b^\rho \frac{\partial\, {}^{(2,2)}\tilde{p}_a^\alpha}{\partial x^{b\rho}} = -{}^{(1,1)}\theta_b^\rho \frac{\partial\, {}^{(1,1)}\tilde{p}_a^\alpha}{\partial \pi^{b\rho}} - {}^{(2,2)}\theta_b^\rho \frac{\partial\, {}^{(0,0)}\tilde{p}_a^\alpha}{\partial \pi^{b\rho}}, \tag{1.3}$$

or taking into account (I.4.18) and (I.5.6),

$$D_b\, {}^{(2,2)}\tilde{p}_a^\alpha = -{}^{(1,1)}\theta_b^\rho \frac{\partial\, {}^{(1,1)}\widetilde{p}_a^\alpha}{\partial \pi^{b\rho}} - {}^{(2,2)}\theta_a^\alpha \delta_{ab}, \tag{1.4}$$

that is,

$$D_a\, {}^{(2,2)}\tilde{p}_a^\alpha = -{}^{(1,1)}\theta_a^\rho \frac{\partial\, {}^{(1,1)}\widetilde{p}_a^\alpha}{\partial \pi^{a\rho}} - {}^{(2,2)}\theta_a^\alpha \tag{1.5}$$

and

$$D_{a'}\,{}^{(2,2)}\tilde{p}^{\alpha}_{a} = -\,{}^{(1,1)}\theta^{\rho}_{a'}\frac{\partial\,{}^{(1,1)}\tilde{p}^{\alpha}_{a}}{\partial\pi^{a'\rho}}. \qquad (1.6)$$

Now, for ${}^{(1,1)}\theta^{\rho}_{a}$ we get [see (1.3.7) and (1.4.13)]

$${}^{(1,1)}\theta^{\rho}_{a} = \frac{\eta_a\pi^2_{a'}\tilde{k}}{\tilde{r}^3_a}\tilde{h}^{\rho} - \frac{\pi^2_{a'}\tilde{z}_a}{\tilde{r}^3_a}\tilde{t}^{\rho}_{a'}, \qquad (1.7)$$

and from Ref. 1 one has for ${}^{(1,1)}\tilde{p}^{\alpha}_{a}$

$${}^{(1,1)}\tilde{p}^{\alpha}_{a} = \frac{\eta_a\tilde{k}}{\tilde{h}^2}\left(\frac{\gamma}{\tilde{\Lambda}} - \frac{\tilde{z}_a}{\tilde{r}_a}\right)\tilde{h}^{\alpha} - \frac{\pi^2_{a'}}{\tilde{\Lambda}^2\tilde{r}_a}\tilde{t}^{\alpha}_{a'}, \qquad (1.8)$$

where γ stands for $+1$ or -1, according to whether we want ${}^{(1,1)}\tilde{p}^{\alpha}_{a}$ to be zero at the future infinite or at the past infinite. (When comparing with Ref. 1 attention must be paid to the fact that there is a slight change of notation for the symbol $\tilde{r}_a$.) It can readily be verified from (1.8) that

$$\lim_{z_{a'}\to\gamma\infty}\left[\frac{\eta_a\tilde{k}}{\tilde{h}^2}\left(\frac{\gamma}{\tilde{\Lambda}} - \frac{\tilde{z}_a}{\tilde{r}_a}\right)\tilde{h}^{\alpha} - \frac{\pi^2_{a'}}{\tilde{\Lambda}^2\tilde{r}_a}\tilde{t}^{\alpha}_{a'}\right] = 0. \qquad (1.9)$$

Now, as we know, the symbol $\sim$ in the notation $\tilde{k}$, $\tilde{h}$, $\tilde{\Lambda}$,...,etc., means that we refer to "the auxiliary dynamical system" (I.4.12) and then we have the definitions (I.4.14), (I.4.15),$\cdots$, while k, h, Λ,$\cdots$ refer to the definitions (I.3.3), (I.3.4),$\cdots$. All through the rest of this article some very lengthy expressions will appear, we will omit from now on the symbol $\sim$ when refering to "the auxiliary dynamical system." So we will write $h^{\alpha}, t^{\alpha}_{a}, z_a, \cdots$, for $\tilde{h}^{\alpha}, \tilde{t}^{\alpha}_{a}, \tilde{z}_a, \cdots$. Later on we will have to work with the original dynamical system. Then, symbols without a tilde will again have their original meaning. This will be pointed out when necessary.

Now let us come to Eq. (1.5) and (1.6), which we must integrate in order to find ${}^{(2,2)}\tilde{p}^{\alpha}_{a}$. For this we first need to compute ${}^{(1,1)}\theta^{\rho}_{b}(\partial^{(1,1)}\tilde{p}^{\alpha}_{a}/\partial\pi^{b\rho})$. Taking into account (1.7) and (1.8) we find after some lengthy calculations (see Appendix A for details)

$${}^{(1,1)}\theta^{\rho}_{a}\frac{\partial\,{}^{(1,1)}\tilde{p}^{\alpha}_{a}}{\partial\pi^{a\rho}} = \frac{\eta_a\pi^2_{a'}\Lambda^2 z_a}{h^2 r^3_a}\left(\frac{\gamma}{\Lambda} - \frac{z_a}{r_a}\right) + \frac{\pi^2_{a'}}{r^3_a}\left(\frac{z_a}{r_a} - \frac{\gamma k^2}{\Lambda^3}\right)t^{\alpha}_{a'}, \qquad (1.10)$$

$$\begin{aligned}
{}^{(1,1)}\theta^{\rho}_{a'}&\frac{\partial\,{}^{(1,1)}\tilde{p}^{\alpha}_{a}}{\partial\pi^{a'\rho}}\\
&= \frac{\eta_a\pi^2_a}{r^3_{a'}}\left(\frac{kz_a}{h^2 r^3_a}(kz_a - \pi^2_{a'}z_{a'})(kh^2 + \Lambda^2 z_a z_{a'})\right.\\
&\quad - \frac{1}{h^2 r_a}[kz_{a'}(kz_a - \pi^2_{a'}z_{a'}) + (kh^2 + \Lambda^2 z_a z_{a'})]\\
&\quad \left. + \frac{\gamma\Lambda z_{a'}}{h^2}\right)h^{\alpha} + \frac{\pi^2_{a'}}{r^3_a}\left(\frac{\pi^2_{a'}}{\Lambda^2 r^3_a}(kz_a - \pi^2_{a'}z_{a'})(kh^2 + \Lambda^2 z_a z_{a'})\right.\\
&\quad \left. - \frac{k}{\Lambda^2 r_a}(kz_a - \pi^2_{a'}z_{a'}) + \frac{\gamma k^2}{\Lambda^3}\right)t^{\alpha}_{a'}, \qquad (1.11)
\end{aligned}$$

that is, ${}^{(1,1)}\theta^{\rho}_{b}(\partial\,{}^{(1,1)}\tilde{p}^{\alpha}_{a}/\partial\pi^{b\rho})$ can be written in the form

$${}^{(1,1)}\theta^{\rho}_{b}\frac{\partial\,{}^{(1,1)}\tilde{p}^{\alpha}_{a}}{\partial\pi^{b\rho}} = \eta_a{}^{(2,2)}b_b h^{\alpha} + {}^{(2,2)}C_b t^{\alpha}_{b}, \qquad (1.12)$$

where the functions ${}^{(2,2)}b_b$ and ${}^{(2,2)}C_b$ are those appearing in (1.10) and (1.11).

On the other hand, the terms ${}^{(2,2)}\xi^{\alpha}_{a}$ in the accelerations ξ^{α}_{a} of our predictive dynamical system can be found in Ref. 4 or, in a more compact writing, in Ref. 2. from ${}^{(2,2)}\xi^{\alpha}_{a}$ the term ${}^{(2,2)}\theta^{\alpha}_{a}$ can be calculated according to (I.4.13). It can be found in Appendix B. Its expression has the form

$${}^{(2,2)}\theta^{\alpha}_{a} = \eta_a{}^{(2,2)}a_a h^{\alpha} + {}^{(2,2)}l_{aa'}t^{\alpha}_{a'}, \qquad (1.13)$$

with ${}^{(2,2)}a_a, {}^{(2,2)}l_{aa'}$ being the appropriate functions.

Finally, according to the notation of (I.5.3) we write

$${}^{(2,2)}\tilde{p}^{\alpha}_{a} = \eta_a{}^{(2,2)}\tilde{\alpha}_a h^{\alpha} + {}^{(2,2)}\tilde{\mu}_{aa}t^{\alpha}_{a} + {}^{(2,2)}\tilde{\mu}_{aa'}t^{\alpha}_{a'}. \qquad (1.14)$$

Now, putting (1.12), (1.13), (1.14), into (1.5), (1.6) and keeping in mind (I.5.7), we have

$$D_a{}^{(2,2)}\tilde{\alpha}_a = -{}^{(2,2)}b_a - {}^{(2,2)}a_a,\quad D_a{}^{(2,2)}\tilde{\mu}_{aa} = -{}^{(2,2)}C_a,\quad D_a{}^{(2,2)}\tilde{\mu}_{aa'} = -{}^{(2,2)}l_{aa'}, \qquad (1.15)$$

$$\begin{aligned}
&D_{a'}{}^{(2,2)}\tilde{\alpha}_a = -{}^{(2,2)}b_{a'},\quad D_{a'}{}^{(2,2)}\tilde{\mu}_{aa} = 0,\\
&D_{a'}{}^{(2,2)}\tilde{\mu}_{aa'} = -{}^{(2,2)}C_{a'}. \qquad (1.16)
\end{aligned}$$

Here, Eqs. (1.15) are equivalent to (1.5), and equations (1.16) to (1.6)

Because of (I.5.11), taking $\pi^2_a, h^2, \Lambda^2, z_a$ as independent variables, E qs. (1.15) and (1.16) can also be written

$$\frac{\partial\,{}^{(2,2)}\alpha_a}{\partial z_a} = -{}^{(2,2)}b_a - {}^{(2,2)}a_a,\quad \frac{\partial\,{}^{(2,2)}\mu_{aa}}{\partial z_a} = -{}^{(2,2)}C_a,\quad \frac{\partial\,{}^{(2,2)}\mu_{aa'}}{\partial z_a} = -{}^{(2,2)}l_{aa'}, \qquad (1.17)$$

$$\begin{aligned}
&\frac{\partial\,{}^{(2,2)}\alpha_a}{\partial z_{a'}} = -{}^{(2,2)}b_{a'},\quad \frac{\partial\,{}^{(2,2)}\mu_{aa}}{\partial z_{a'}} = 0,\\
&\frac{\partial\,{}^{(2,2)}\mu_{aa'}}{\partial z_{a'}} = -{}^{(2,2)}C_{a'}. \qquad (1.18)
\end{aligned}$$

We will use these equations to determine ${}^{(2,2)}\alpha_a$ and ${}^{(2,2)}\mu_{aa'}$, while ${}^{(2,2)}\mu_{aa}$ will be obtained from (I.4.19). Let us begin with ${}^{(2,2)}\alpha_a$. We will have

$${}^{(2,2)}\tilde{\alpha}_a = \int_{z_a}^{\gamma\infty}({}^{(2,2)}b_a + {}^{(2,2)}a_a)dz_a + \int_{z_{a'}}^{\gamma\infty}\left(\lim_{z_a\to\gamma\infty}{}^{(2,2)}b_{a'}\right)dz_{a'} \qquad (1.19)$$

and depending on whether we take $\gamma = +1$ or $\gamma = -1$ we will have $\lim {}^{(2,2)}\alpha_a = 0$ at the future or past infinite respectively. In a similar way we have

$$ {}^{(2,2)}\widetilde{\mu}_{aa'} = \int_{z_a}^{\gamma\infty} {}^{(2,2)}l_{aa'}dz_a + \int_{z_a}^{\gamma\infty} \Big(\lim_{z_a\to\gamma\infty} {}^{(2,2)}C_{a'}\Big) dz_{a'}. \tag{1.20} $$

Taking into account (1.12) and (1.11), it is easy to see that

$$ \lim_{z_a\to\gamma\infty} {}^{(2,2)}b_{a'} = 0, \quad \lim_{z_a\to\gamma\infty} {}^{(2,2)}C_{a'} = 0, \tag{1.21} $$

so that (1.19) and (1.20) become:

$$ {}^{(2,2)}\widetilde{\alpha}_a = \int_{z_a}^{\gamma\infty} ({}^{(2,2)}b_a + {}^{(2,2)}a_a)dz_a, \tag{1.22} $$

$$ {}^{(2,2)}\widetilde{\mu}_{aa'} = \int_{z_a}^{\gamma\infty} {}^{(2,2)}l_{aa'}dz_{a'}. \tag{1.23} $$

Now, as it can be seen in Appendix B, ${}^{(2,2)}a_a, {}^{(2,2)}l_{aa'}$, are rather involved expressions. Consequently these integrals, though expressible as elementary functions, are involved too. The final result is:

$$ \begin{aligned} {}^{(2,2)}\widetilde{\alpha}_a &= \frac{\pi_a^2\pi_{a'}^4 k}{(kr_a - \Lambda^2 z_a)(k^2r_a^2 - \Lambda^4 z_a^2)} + \frac{\pi_a^2\pi_{a'}^2}{2\Lambda^2h^2} \\ &\times \frac{(2k^2-\Lambda^2)r_a + k\Lambda^2 z_a}{k^2r_a^2 - \Lambda^4 z_a^2} + \frac{\pi_{a'}^2(2\pi_a^2k - \Lambda^2)z_a}{2\Lambda^2h^2r_a^2} \\ &- \frac{\pi_a^2}{\Lambda^2h^2r_{a'}r_a}(\pi_{a'}^2 r_{a'} + \pi_a^2 kz_{a'} + \Lambda^2 z_a) + \frac{k\pi_{a'}^2}{\Lambda^2 r_{a'}r_a^3} \\ &\times [kr_{a'} + (2k^2 - \Lambda^2)z_{a'} - 2\pi_a^2 kz_a] + \frac{\gamma}{\Lambda h^2} \\ &\times \left(\frac{\pi_a^2}{r_{a'}} + \frac{\pi_{a'}^2}{r_a}\right) - \frac{\pi_a^2\pi_{a'}^4 k}{\Lambda^3 r_a^3}\log\frac{(r_a - \Lambda z_a)(r_{a'} + \Lambda z_{a'})}{(k-\Lambda)h^2} \\ &+ \frac{\pi_a^2\pi_a' k}{\Lambda^3h^3}\arctan\frac{r_a - \Lambda z_a}{\pi_{a'}h} + \frac{\pi_a'(\pi_a^2k + \Lambda^2)}{2\Lambda^3h^3} \\ &\times \arctan\frac{\Lambda z_a}{\pi_{a'}h} + \frac{\pi_a}{2\Lambda h^3}\left(\arctan\frac{\pi_a r_a}{\Lambda h} + \arctan\frac{\pi_a\Lambda z_a}{kh}\right) \\ &- \frac{\pi\pi_a^2\pi_{a'}k}{4\Lambda^3h^3} - \frac{\pi\pi_a}{4\Lambda h^3} - \frac{\gamma\pi(\pi_a + \pi_{a'})}{4\Lambda h^3}, \end{aligned} \tag{1.24} $$

$$ \begin{aligned} {}^{(2,2)}\widetilde{\mu}_{aa'} &= \frac{\pi_a^2\pi_{a'}^2}{2\Lambda^2h^2}\,\frac{\pi_a^2 h^2 - 2kr_a z_a}{k^2r_a^2 - \Lambda^4 z_a^2} - \pi_{a'}^4\,\frac{\pi_a^2 kr_a z_a + k^2h^2}{(k^2r_a^2 - \Lambda^4 z_a^2)^2} \\ &+ \frac{\pi_a^2\pi_{a'}^4}{\Lambda^4 r_a^2} + \frac{2\pi_a^2\pi_{a'}^2 k}{\Lambda^4 r_a r_{a'}} - \frac{2\pi_a^2\pi_{a'}^4 kh^2}{\Lambda^4 r_a^3 r_{a'}} \\ &+ \frac{kz_a kz_{a'} + r_{a'})}{\Lambda^2h^2r_ar_{a'}} - \frac{z_a\pi_{a'}^2[kr_{a'} + (2k^2 - \Lambda^2)z_{a'}]}{\Lambda^2r_a^3r_{a'}} \\ &- \gamma\frac{k^2z_{a'}}{\Lambda^3h^2r_{a'}} - \frac{\pi_a^2\pi_{a'}^4 z_a}{\Lambda^3 r_a^3}\log\frac{(r_a + \Lambda z_a)(k-\Lambda)}{\pi_{a'}^2(r_{a'} + \Lambda z_{a'})}. \end{aligned} \tag{1.25} $$

Let us now calculate ${}^{(2,2)}\mu_{aa'}$. Keeping fourth order terms in (I.4.19) we obtain

$$ 2^{(2,2)}\tilde{p}_a^\alpha \pi_{a\alpha} + {}^{(1,1)}\tilde{p}_a^\alpha\,{}^{(1,1)}\tilde{p}_{a\alpha} = 0 \tag{1.26} $$

and from this equation and (I.5.3)

$$ {}^{(2,2)}\widetilde{\mu}_{aa} = -\frac{h^2\,{}^{(1,1)}\widetilde{\alpha}_a^2 + \pi_a^2\Lambda^2\,{}^{(1,1)}\widetilde{\mu}_{aa'}^2}{2\Lambda^2}. \tag{1.27} $$

Then, because of (1.8), we find

$$ {}^{(2,2)}\widetilde{\mu}_{aa} = \frac{\pi_{a'}^2}{2\Lambda^2r_a^2} + \frac{k^2}{\Lambda^3h^2}\left(\frac{\gamma z_a}{r_a} - \frac{1}{\Lambda}\right). \tag{1.28} $$

Finally, substitution of (1.28), (1.25), and (1.24) in (1.14) gives us ${}^{(2,2)}\tilde{p}_a^\alpha$. Then from (I.4.18) and (I.5.14) we can write

$$ \begin{aligned} \tilde{p}_a^\alpha &= \pi_a^\alpha + e_1e_2\left[\frac{\eta_a k}{h^2}\left(\frac{\gamma}{k} - \frac{z_a}{r_a}\right)h^\alpha - \frac{\pi_{a'}^2}{\Lambda^2 r_a}t_{a'}^\alpha\right] \\ &+ e_a^3e_{a'}\left[-\frac{2\eta_a\pi_{a'}^2k}{3\pi_a^2r_a^3}h^\alpha + \frac{2\pi_{a'}^2z_A}{3\pi_a^2r_a^3}t_{a'}^\alpha\right] \\ &+ e_1^2e_2^2[\eta_a\,{}^{(2,2)}\widetilde{\alpha}_a h^\alpha + {}^{(2,2)}\widetilde{\mu}_{aa}t_a^\alpha + {}^{(2,2)}\widetilde{\mu}_{aa'}t_{a'}^\alpha] \end{aligned} \tag{1.29} $$

with ${}^{(2,2)}\widetilde{\alpha}_a, {}^{(2,2)}\widetilde{\mu}_{aa}$, and ${}^{(2,2)}\widetilde{\mu}_{aa'}$, given by (1.24), (1.28), and (1.25), respectively. So we have calculated the Hamilton–Jacobi momenta, $\tilde{p}_a^\alpha$, up to terms in $e_a^n e_{a'}^m$ ($n + m \leqslant 4$). Should we need the total 4-momentum, $\widetilde{P}^\alpha$, to this order, we only would have to put $\widetilde{P}^\alpha = \tilde{p}_1^\alpha + \tilde{p}_2^\alpha$.

In the next sections some approximated physical conclusions will be worked out from the expression (1.29). Actually, we will obtain the "radiated" energy, the 3-momentum, and the scattering cross section.

2. THE "RADIATED" ENERGY AND 3-MOMENTUM

The approximated Hamilton–Jacobi momenta, $\tilde{p}_a^\alpha$, given in (1.29) have been defined in such a way that the condition $\lim_{z_a,z_{a'}\to\gamma\infty}\tilde{p}_a^\alpha = \pi_a^\alpha$, is satisfied, as can be verified directly in (1.29). So, according to whether we take $\gamma = -1$ or $\gamma = +1$, these approximated momenta reduce to the free momenta π_a^α, in the infinite past or infinite future, respectively. Nevertheless, in the language of Ref. 1, these momenta, $\tilde{p}_a^\alpha$, are not conserved, that is, we do not recover the free momenta, π_a^α, when taking the limit $z_a, z_{a'} \to -\gamma\infty$ as it also can be seen from (1.29). Similar considerations can be made for the total 4-momentum, $\widetilde{P}^\alpha = \tilde{p}_1^\alpha + \tilde{p}_2^\alpha$. Of course, $\widetilde{P}^\alpha$ and also $\tilde{p}_a^\alpha$ are conserved numerically along a given pair of trajectories.

So, let us take the limit $z_a, z_{a'} \rightarrow -\gamma\infty$ in (1.29). Taking into account (1.28), (1.25), and (1.24) we obtain

$$\widetilde{p_a^\alpha}|_{-\gamma\infty} \equiv \lim_{\tilde{z}_a, \tilde{z}_{a'} \rightarrow -\gamma\infty} \tilde{p}_a^\alpha = \pi_a^\alpha + e_1 e_2 \frac{2\eta_a \gamma \widetilde{k}}{\widetilde{h}^2 \widetilde{A}} \widetilde{h}^\alpha + e_1^2 e_2^2$$

$$\times \left(- \frac{\eta_a \gamma \pi (\pi_a + \pi_{a'})}{2\widetilde{A}\widetilde{h}^3} \widetilde{h}^\alpha - \frac{2\widetilde{k}^2}{\widetilde{A}^4 \widetilde{h}^2} \tilde{t}_a^\alpha + \frac{2\widetilde{k}^2}{\widetilde{A}^4 \widetilde{h}^2} \tilde{t}_{a'}^\alpha \right), \tag{2.1}$$

where we have restored the symbol $\sim$ because this expression, as compared with (1.29), is a relatively short one.

By making the substitutions[1] $\pi_a^\alpha \rightarrow m_a u_a^\alpha$, we obtain $p_a^\alpha|_{-\gamma\infty}$. This is analogous to (2.1), now refering to the original dynamical system. That is,

$$\begin{aligned} p_a^\alpha&|_{-\gamma\infty} \\ &\equiv \lim_{z_a, z_{a'} \rightarrow -\gamma\infty} p_a^\alpha = m_a u_a^\alpha + e_1 e_2 \frac{2\eta_a \gamma k}{h^2 A} h^\alpha + e_1^2 e_2^2 \\ &\times \left(- \frac{\eta_a \gamma \pi (m_1 + m_2)}{2 m_1 m_2 A h^3} h^\alpha - \frac{2k^2}{m_a A^4 h^2} t_a^\alpha \right. \\ &\left. + \frac{2k^2}{m_{a'} A^4 h^2} t_{a'}^\alpha \right). \end{aligned} \tag{2.2}$$

Here $k, A, \cdots$ are the quantities of the original system which were given in (I.3.3), (I.3.5),and (I.3.8).

From (2.2) and the total 4-momentum, $P^\alpha = p_1^\alpha + p_2^\alpha$, we get in our approximation

$$P^\alpha|_{-\gamma\infty} \equiv \lim_{z_a, z_{a'} \rightarrow -\gamma\infty} P^\alpha = m_1 u_1^\alpha + m_2 u_2^\alpha.$$

Then the "radiated" 4-momentum up to order $e_1^a e_2^b$ with $a + b < 6$ is zero. But the intrinsic angular momenta radiated by the two charges is different from zero to order $e_a e_{a'}^3$ (I.5.27) due to the presence of the Dirac term.

We do not know presently if they are "radiation" due to the retarded potential term (I.2.1), because we have not computed the intrinsic angular momentum to order $e_1^2 e_2^2$. But, if this contribution exists, it would not cancel the $e_a e_{a'}^3$ term (except perhaps for equal charges).

Let us note that the first contribution to radiation comes from the intrinsic angular momentum. Up to order [SI:e:a:1][SI:e:b:2], $a + b = 4$, energy and momentum are conserved, but not the intrinsic angular momentum which is tranferred by the system to the universe (from the point of view of the Wheeler–Feynman absorber's theory).

3. SCATTERING CROSS SECTIONS

Let us consider (2.2) for $\gamma = 1$. Then since the Hamilton–Jacobi momenta, p_a^α are numerically conserved and since for $\gamma = 1$ we have $\lim_{z_a, z_{a'} \rightarrow \infty} p_a^\alpha = m_a u_a^\alpha$, we can establish the following equation,

$$\begin{aligned} m_a u_a^\alpha &+ e_1 e_2 \frac{2\eta_a k}{A h^2} h^\alpha + e_1^2 e_2^2 \left(- \frac{\eta_a \pi (m_1 + m_2)}{2 m_1 m_2 A h^3} h^\alpha \right. \\ &\left. - \frac{2k^2}{m_a A^4 h^2} t_a^\alpha + \frac{2k^2}{m_{a'} A^4 h^2} t_{a'}^\alpha \right) \\ &= m_a u_{aF}^\alpha, \end{aligned} \tag{3.1}$$

where symbols on the left side refer to the initial state and those on the right side to the final state.

From (3.1) and in the laboratory frame, $\mathbf{v}_{a'} = 0$, we obtain the 3-vector equation

$$\begin{aligned} m_a \gamma_a \mathbf{v}_a &+ e_1 e_2 \frac{2\eta_a}{v_a \mathbf{h}^2} \mathbf{h} + e_1^2 e_2^2 \\ &\times \left(- \frac{\eta_a \pi (m_1 + m_2)}{2 m_1 m_2 \gamma_a v_a |\mathbf{h}|^3} \mathbf{h} - 2 \frac{m_{a'} + m_a \gamma_a}{m_a m_{a'} \gamma_a v_a^4 \mathbf{h}^2} \mathbf{v}_a \right) \\ &= m_a \gamma_{aF} \mathbf{v}_{aF}. \end{aligned} \tag{3.2}$$

From the identity $h^\alpha u_{a\alpha} = 0$, we have $\mathbf{h} \cdot \mathbf{v}_a = 0$ for $x_1^0 = x_2^0$. Then Eq. (3.2) is equivalent to the two equations

$$m_a \gamma_a v_a - 2 e_1^2 e_2^2 \frac{m_{a'} + m_a \gamma_a}{m_a m_{a'} \gamma_a v_a^3 \mathbf{h}^2} = m_a \gamma_{aF} v_{aF} \cos\theta, \tag{3.3}$$

$$\frac{2 e_1 e_2}{v_a |\mathbf{h}|} - e_1^2 e_2^2 \frac{\pi (m_1 + m_2)}{2 m_1 m_2 \gamma_a v_a \mathbf{h}^2} = m_a \gamma_{aF} v_{aF} \epsilon \sin\theta, \tag{3.4}$$

where ϵ is the sign of the product $e_1 e_2$ and θ is the scattering angle between $\mathbf{v}_{aF}$ and $\mathbf{v}_a$. From (3.3), (3.4) we have to fourth order

$$\tan\theta = \frac{2 e_1 e_2 \epsilon}{m_a \gamma_a v_a^2 |\mathbf{h}|} \left(1 - \frac{\pi e_1 e_2}{4 m \gamma_a |\mathbf{h}|} \right), \tag{3.5}$$

where m is the reduced mass: $m \equiv m_1 m_2 / (m_1 + m_2)$. Then

$$\frac{1}{|\mathbf{h}|} = \frac{2 m \gamma_a}{\pi e_1 e_2} \left[1 \pm \left(1 - \frac{\epsilon \pi m_a v_a^2 \tan\theta}{2m} \right)^{1/2} \right]. \tag{3.6}$$

Here we want that $\lim_{\theta \rightarrow 0} |\mathbf{h}|^{-1} = 0$ and so only the solution

$$\frac{1}{|\mathbf{h}|} = \frac{2 m \gamma_a}{\pi e_1 e_2} \left[1 - \left(1 - \frac{\epsilon \pi m_a v_a^2 \tan\theta}{2m} \right)^{1/2} \right] \tag{3.7}$$

must be retained.

Now the expansion in e_1, e_2, θ, as a function of $|\mathbf{h}|$ and γ_a, begins with $e_1 e_2$ as it can be seen from (3.5). Taking this into account we can approximate (3.7) to get

$$|\mathbf{h}| = \frac{2 e_1 e_2 \epsilon}{m_a \gamma_a v_a^2 \sin\theta} - \frac{\pi e_1 e_2}{4 m \gamma_a}, \tag{3.8}$$

where the term $2 e_1 e_2 \epsilon / m_a \gamma_a v_a^2 \sin\theta$ in fact beings with zero order in $e_1 e_2$ because of the $\sin\theta$ appearing in it. Only the other term, $\pi e_1 e_2 / 4 m \gamma_a$, really behaves as an $e_1 e_2$ term.

From (3.8) we can evaluate the scattering cross section in the LAB frame, $\mathbf{v}_{a'} = 0$,

$$d\sigma = \frac{e_1^2 e_2^2}{4 m_a^2 \gamma_a^2 v_a^2 \sin^4(\theta/2)} \left(1 - \frac{\epsilon \pi m_a v_a^2 \sin(\theta/2)}{4m} \right) d\Omega, \tag{3.9}$$

with $d\Omega$ the differential solid angle. To lowest order, when

$[\epsilon\pi m_a v_a^2 \sin(\theta/2)]/4m \ll 1$, we obtain

$$d\sigma = \frac{e_1^2 e_2^2 d\Omega}{4m_a^2\gamma_a^2 v_a^2 \sin^4(\theta/2)} \quad (3.10)$$

which is in fact the cross section that would have been obtained if we had neglected terms in $e_1^2e_2^2$ in (3.1).

Let us calculate the cross section (3.10) in the center of mass system. This can be easily done by neglecting terms in $e_1^2e_2^2$ in (3.1). We find

$$d\sigma = \frac{e_1^2 e_2^2 k^2 d\Omega}{4m_a^2\gamma_a^2 v_a^2 \Lambda^2 \sin^4(\theta/2)}. \quad (3.11)$$

It can be seen that this expression agrees with Ref. 3. As it is noted there, (3.11) reduces to Rutherford's formula in the low energy limit and when one of the masses becomes infinite it reduces to Mott's formula for a spinless particle.

Now let us suppose that $m_a = m_{a'}$ and that after the scattering we cannot distinguish between the two particles. From (3.11) we then find for the scattering cross section

$$d\sigma = \frac{e_1^2 e_2^2 (2\gamma_a^2 - 1)^2}{m_a^2 v_a^4 \gamma_a^6 \sin^4\theta}. \quad (3.12)$$

Akhiezer and Berestetskii, in "Quantum Electrodynamics," p. 838, give (in our notation) to lowest order

$$d\sigma = \frac{e_1^2 e_2^2 (2\gamma_a^2 - 1)^2}{4m_a^2 v_a^4 \gamma_a^6}\left(\frac{2}{\sin^2\theta} - \frac{\gamma_a^2 - 1}{2\gamma_a^2 - 1}\right)^2 d\Omega \quad (3.13)$$

for the scattering of two indistinguishable spinless particles in the center of mass system. This formula does not depend on Planck's constant. In (3.13) the expression $|(\gamma_a^2 - 1)/(2\gamma_a^2 - 1)|$ is always less than $\frac{1}{2}$. Then we have $2/\sin^2\theta \gg (\gamma_a^2 - 1/2)(2\gamma_a^2 - 1)$ and (3.13) reduces to (3.12). So, in the appropriate limit, our results agree with those of quantum electrodynamics, at least as far as (3.13) is concerned.

ACKNOWLEDGMENT

We thank Dr. J. Martin for supplying us with the useful expressions given in Appendices A and B.

APPENDIX A

The calculation of ${}^{(1,1)}\theta^{\rho}_{a'}(\partial^{(1,1)}\bar{p}^{\alpha}_a/\partial\pi^{a'\rho})$ becomes easier by systematically using the following relations (J. Martin, private communication):

$$N_a h^\alpha = -z_a h^\alpha - \eta_a \Lambda^{-2} h^2 t^\alpha_a,$$

$$N_a t^\alpha_a = \eta_a \pi^2_{a'} h^\alpha,$$

$$N_a t^\alpha_{a'} = \eta_{a'} k h^\alpha,$$

$$N_a z_a = \pi^2_{a'} h^2 \Lambda^{-2}, \quad N_a z_{a'} = k h^2 \Lambda^{-2},$$

$$N_a h^2 = -2h^2 z_a, \quad N_a k = N_a \pi_b^2 = 0,$$

$$Q_a h^\alpha = 0, \quad Q_a t^\alpha_a = -k t^\alpha_a, \quad Q_a t^\alpha_a = -\pi^2_a t^\alpha_a - 2k t^\alpha_{a'},$$

$$Q_a z_a = k z_a, \quad Q_a z_{a'} = \pi^2_a z_a,$$

$$Q_a k = -\Lambda^2, \quad Q_a h^2 = Q_a \pi_b^2 = 0,$$

where N_a and P_a mean the differential operators

$$N_a \equiv \eta_a h^\rho \frac{\partial}{\partial\pi^{a\rho}}, \quad Q_a \equiv t^\rho_a \frac{\partial}{\partial\pi^{a\rho}}.$$

APPENDIX B

Writing

$${}^{(2,2)}\theta^\alpha_a = \eta_a {}^{(2,2)}a_a h^\alpha + {}^{(2,2)}l_{aa'} t^\alpha_{a'},$$

we have for the functions ${}^{(2,2)}a_a$ and ${}^{(2,2)}l_{aa'}$ the following expressions (J. Martin, private communication):

$$\begin{aligned}
{}^{(2,2)}a_a = & -\pi_a^2\pi_{a'}^4 \Lambda^2 r_a^{-2}(r_a - kz_a)(kr_a - \Lambda^2 z_a)^{-3} \\
& + \pi_a^2\pi_{a'}^2(3k^2h^2 - \pi_a^2 r_a^2) r_a^{-5} \\
& \times\{(kh^{-2}\pi_a^{-2} z_a z_{a'} + \pi_{a'}^2\Lambda^{-2}) r_{a'}^{-1} \\
& - [kh^{-2}(\pi_a^2\pi_{a'}^2)^{-1} z_a(kz_a - r_a) + \pi_{a'}^2\Lambda^{-2}] \\
& \times\pi_{a'}^2(kr_a - \Lambda^2 z_a)^{-1}\} \\
& + 3\pi_a^2\pi_{a'}^2 k\Lambda^2 z_a r_a^{-5}\Big(-\Lambda^{-2}(kz_a - \pi_{a'}^2 z_{a'}) r_{a'}^{-1} \\
& + \pi_{a'}^2\Lambda^{-2} r_a (kr_a - \Lambda^2 z_a)^{-1} + \pi_{a'}^2\Lambda^{-3}\log\frac{\Lambda z_a + r_a}{\Lambda z_{a'} + r_{a'}} \\
& + \pi_{a'}^2\Lambda^{-3}\log\frac{k - \Lambda}{\pi_{a'}^2}\Big),
\end{aligned}$$

$$\begin{aligned}
{}^{(2,2)}l_{aa'} = & \pi_a^2\pi_{a'}^6 h^2 r_a^{-2}(kr_a - \Lambda^2 z_a)^{-3} - 3\pi_a^2\pi_{a'}^2 kh^2 z_a r_a^{-5} \\
& \times\{[kh^{-2}\pi_{a'}^{-2} z_a z_{a'} + \pi_{a'}^2\Lambda^{-2}] r_{a'}^{-1} \\
& - [kh^{-2}(\pi_a^2\pi_{a'}^2)^{-1} z_a(kz_a - r_a) \\
& + \pi_{a'}^2\Lambda^{-2}]\pi_{a'}^2(kr_a - \Lambda^2 z_a)^{-1}\} \\
& - \pi_a^2\pi_{a'}^2(3\Lambda^2 z_a^2 - r_a^2) r_a^{-5} \\
& \times\{-\Lambda^{-2}(kz_a - \pi_{a'}^2 z_{a'}) r_{a'}^{-1} \\
& + \Lambda^{-2}\pi_{a'}^2 r_a (kr_a - \Lambda^2 z_a)^{-1} \\
& + \pi_{a'}^2\Lambda^{-3}\log\frac{\Lambda z_a + r_a}{\Lambda z_{a'} + r_{a'}} + \pi_{a'}^2\Lambda^{-3}\log\frac{k - \Lambda}{\pi_{a'}^2}\Big\}.
\end{aligned}$$

[1]L. Bel and J. Martin, Ann. Inst. H. Poincaré **22**, 173 (1975).
[2]L. Bel, "Journées relativistes de Toulouse," Université de Toulouse (1974).
[3]L. Bel, *Contribution to Differential Geometry and Relativity*, edited by Cohen and Flato (Reidel, Dordrecht, Holland, 1976).
[4]A. Salas and J.M. Sánchez, Nuovo Cimento B **20**, 209 (1974).

Current responses of all orders in a collisionless plasma. I. General theory

J. Larsson

Department of Plasma Physics, Umeå University, S–901 87 Umeå, Sweden

(Received 1 May 1978)

An arbitrary electromagnetic perturbation of a general solution of the relativistic Vlasov–Maxwell equations is considered. The nonlinear current responses are expressed in a form which in particular is an all order manifestation of the Manley–Rowe relations. A coordinate free formalism is used, starting with a representation of Minkowski space in terms of abstract linear algebra, and all formulas are intrinsically covariant. In the method used to derive the current responses the perturbation of particle orbits rather than of distribution functions is calculated.

1. INTRODUCTION

Consider an electromagnetic perturbation of an arbitrary solution to the (relativistic) Vlasov–Maxwell equations. We denote the change in the electromagnetic 4-potential by ϕ and in the plasma 4-current by $\delta J[\phi]$. Here δJ is a nonlinear function of ϕ, and if ϕ is sufficiently small we may obtain $\delta J[\phi]$ as an expansion

$$\delta J[\phi] = \delta J^{(1)}[\phi] + \delta J^{(2)}[\phi] + \cdots, \tag{1.1}$$

where $\delta J^{(1)}$ is linear and $\delta J^{(n)}$, for $n \geqslant 2$, nonlinear of order n in ϕ. In this paper formulas for $\delta J^{(n)}[\phi]$ are derived for all $n \geqslant 2$. The results are formally valid for a relativistic multicomponent space–time dependent plasma in an arbitrary external electromagnetic field.

One would like the general formulas for the nonlinear currents to:

(a) Provide useful formulas for different particular situations of interest,

(b) Exhibit structure and symmetry properties.

The only way to discover how well the formulas in this paper satisfy property (a) is to actually work them out for different particular cases. For a homogeneous magnetized plasma this has been done[1] and relativistic wave coupling coefficients of arbitrary order may be obtained in forms which generalize the result in Ref. 2 (where the 3-wave coupling coefficients for a nonrelativistic plasma are given as a series containing Bessel functions, just as in the standard expression for the linear conductivity tensor). Concerning (b) our formulas show an expected approximate symmetry which is an all order manifestation of the Manley–Rowe relations.[3] Structure and symmetry properties are discussed in Sec. 5.

The method of derivation used in this paper is "the dual" of the standard method; the perturbed motions of the particles are calculated instead of the perturbation in the distribution function. Thus instead of the Vlasov equations we use the equivalent system of equations given in Lemma 2(a), Sec. 4. It is demonstrated how to obtain general formulas with satisfactory symmetry properties from this starting point, and the final form obtained for $\delta J^{(n)}[\phi]$ indicates why these results probably would be much more difficult to derive by the usual method where one iterates the Vlasov equation (see Sec. 5).

The generality of the formulas in this paper causes notational problems; the standard notation of index calculus normally used in special relativity has turned out to be embarrassingly tedious to use in the derivations and also clearly unsuitable if we want to write transparent formulas. Therefore, a coordinate-free system of notation has been used and this is advantageous from at least three points of view, firstly we get rid of indices, secondly the coordinate-free approach is much more readily capable of geometric interpretations, and thirdly the formulas are intrinsically covariant. The mathematics of coordinate-free analysis may be found in books on differential geometry but these are in general unnecessarily advanced for our needs. In Ref. 4 the coordinate-free approach to special relativity and Maxwell's equation is discussed, and this is the best reference for us. However, since we are doing kinetic and not gravitation theory, it will be practical for us to use a slightly different notation which is closer to standard advanced calculus. Unfortunately, there is no reference quite suitable for us and this is why Sec. 2 is somewhat long. The reader not familiar with the method of coordinate-free analysis is advised to study Chapters 2, 3, and 4 in Ref. 4 before reading Sec. 2 of this paper.

The coordinate-free language is introduced in Sec. 2. In Sec. 3 the main results of this paper are formulated, in Sec. 4 the derivations are given, and in Sec. 5 we discuss the results.

2. COORDINATE-FREE NOTATION AND SPECIAL RELATIVITY

A. General remarks and our conventions concerning units and bold face letters

In subsection B below we give in (i)–(v) the same abstract description of the Minkowski space as in Ref. 5, with the only difference that we state the orientation in space and time more explicitly with (iv) and (v). For such a space there is a well developed coordinate-free calculus[6] and this is a suitable starting point for us. In Sec. 2 B we also give the relationship between this algebraic description and the stan-

 0022-2488/79/071321-10$01.00

dard one on R^4. The scalar product given in (iii) induces natural measures on 1, 2, 3, and 4-surfaces and also certain operations on the tensor spaces (for the definition of these spaces, see Ref. 4) defined in (2.3)–(2.8). We define the gradient operators (2.12)–(2.14) and here we may note some differences as compared with Ref. 4. We make use of the scalar product in the definition of the gradients and these operators appear much like vectors in our description $[(u\cdot\nabla_E)T = u\cdot(\nabla_E T)$ for us, compare with (3.39), p. 82 in Ref. 4]. The form of coordinate-free calculus used in this paper is closer to standard advanced calculus: We identify the dual space V^* with our vector space V and make no further use of it, the gradient operators behave as vectors and we have the familiar form of the Taylor expansion (2.19).

Unit convention: It will be convenient not to associate any units with our abstract spaces. In order to achieve this formally the formulas and equations in this paper only concern the numerical parts of the physical quantities expressed in MKSA units. Thus, the velocity of light is c meter/second, the particle rest mass and charge (with particle specie index omitted) m_0 kg and q coulomb, the electromagnetic field tensor is F volt/meter, the 4-potential Φ volt, and the 4-current J ampere/second·meter^2.

The use of bold face letters: Bold face letters will be used only when a particular Lorentz system has been chosen in order to simplify the transition from coordinate-free formalism to conventional "space-plus–time" statements. Our convention is understood from (2.9) and (2.26)–(2.30).

B. The Minkowski space of special relativity

We are given $(E,V,\cdot,S,V^{\wedge 4}_{+})$ such that

(i) E is a four-dimensional real affine space,[7]

(ii) V is the four-dimensional real vector space of translations belonging to the affine space E,

(iii) $(\cdot)$ denotes a symmetric bilinear form, which we will call the "scalar product" on V, with positive index of inertia 3 and negative index 1. We write $x\cdot y\in R$ for the scalar product of $x,y\in V$.

(iv) S is (a choice of) one of the two components (i.e., maximal connected subsets) in the set $\{u\in V\,|\,u\cdot u = -1\}$

(v) $V^{\wedge 4}_{+}$ is one of the two components in $V^{\wedge 4}\setminus\{0\}$, where $V^{\wedge 4}$ is the set of alternating tensors of order 4 (see subsection C below).

We interpret E as the set of events and S the set of 4-velocities, i.e., unit vectors in the future direction. Thus (iv) gives us an orientation in time and the orientation in space is defined from (iv) and (v): Three spacelike linearly independent vectors v_1, v_2, v_3 (spacelike means $v_i\cdot v_i > 0$) are positively oriented iff $u\wedge v_1\wedge v_2\wedge v_3\in V^{\wedge 4}_{+}$, where u is the (unique) element in S determined by $u\cdot v_i = 0$ for $i = 1$, 2, and 3. If we only make use of the structure given in (i)–(v) and avoid using any particular coordinate systems we will automatically obtain coordinate free, and thus intrinsically covariant, formulas. Contact with standard index notation is obtained by the (proper) Lorentz (coordinate) systems. Such a system is defined by a choice of origin $O\in E$ together with four vectors (e_0,e_1,e_2,e_3) such that

$$|e_i\cdot e_j| = \delta_{ij}$$

and

$$e_0\in S \quad\text{and}\quad e_0\wedge e_1\wedge e_2\wedge e_3\in V^{\wedge 4}_{+}.$$

A Lorentz system L defines bijections

$$L_V: V\to R^4 \quad\text{and}\quad L_E: E\to R^4 \text{ by}$$

$$L_V(y) = (y^i) \quad\text{iff}\quad y = \sum y^i e_i,$$

$$L_E(P) = (x^i) \quad\text{iff}\quad P = O + \sum x^i e_i,$$

where (y^i) and (x^i) are the index notations for y and P. If L and K are two Lorentz systems, then the composed mapping $K_E\circ L_E^{-1}$ is a Poincaré transformation with the Lorentz transformation $K_V\circ L_V^{-1}$ as its homogeneous part.

C. Calculus on Minkowski space

Tensor spaces may be constructed from the vector space V and they will inherit some structure from the scalar product. We may identify V with its dual space V^* by associating with each $v\in V$ the mapping v: $V\to R$ defined by $v(w) = v\cdot w$, thus nothing is gained by also using V^* in building tensor spaces. Accordingly we consider only the tensor spaces $V^{\otimes n}$ and their subspaces of n-vectors (i.e., alternating tensors) $V^{\wedge n}$. A base $\{e_0,e_1,e_2,e_3\}$ of V induces the base

$$\{e_{i_1}\otimes\cdots\otimes e_{i_n}\,|\,0\leqslant i_k\leqslant 3,\ 1\leqslant k\leqslant n\} \tag{2.1}$$

of $V^{\otimes n}$ and the base

$$\{e_{i_1}\wedge\cdots\wedge e_{i_n}\,|\,0\leqslant i_1<\cdots<i_n\leqslant 3,\} \tag{2.2}$$

of $V^{\wedge n}$. The dimensions of $V^{\otimes n}$ and $V^{\wedge n}$ are 4^n and $\binom{4}{n}$, respectively. The space $V^{\wedge n}$ is nontrivial only for $n\in\{0,1,2,3,4\}$ and by definition $V^{\otimes 0} = V^{\wedge 0} = R$ and $V^{\otimes 1} = V^{\wedge 1} = V$. The scalar product on V induces the mapping $(\overset{k}{\cdot})$ and $\langle\ ,\ \rangle$ such that

$$(\overset{k}{\cdot}): V^{\otimes l}\times V^{\otimes m}\to V^{\otimes(l+m-2k)} \quad\text{for } l,m\geqslant k\geqslant 1, \tag{2.3}$$

$$\langle\ ,\ \rangle: V^{\wedge m}\times V^{\wedge m}\to R \quad\text{for } m\geqslant 1, \tag{2.4}$$

defined by

$$v_1\otimes\cdots\otimes v_l(\overset{k}{\cdot})w_1\otimes\cdots\otimes w_m = \left(\prod_{i=1}^{k} v_{l-k+i}\cdot w_i\right)v_1\otimes\cdots\otimes v_{l-k}\otimes w_{k+1}\otimes\cdots\otimes w_m \tag{2.5}$$

and

$$\langle v_1\wedge\cdots\wedge v_m, w_1\wedge\cdots\wedge w_m\rangle = \frac{1}{m!}\,v_1\otimes\cdots\otimes v_m(\overset{m}{\cdot})w_1\otimes\cdots\otimes w_m. \tag{2.6}$$

We use the notation $\cdot$ for $(\overset{1}{\cdot})$, and : for $(\overset{2}{\cdot})$. The operator $\langle\ ,\ \rangle$ is a scalar product on $V^{\wedge m}$ such that the base (2.2) for $V^{\wedge m}$ is orthonormal if $\{e_0,e_1,e_2,e_3\}$ is orthonormal, i.e., if $|e_i\cdot e_j| = \delta_{ij}$. It may be shown that

$$\langle v_1\wedge\cdots\wedge v_m, w_1\wedge\cdots\wedge w_m\rangle = \det(v_i\cdot w_j) \tag{2.7}$$

and this is the scalar product used in Ref. 8.

Given a 2-surface Σ in E we define its area $A(\Sigma)$ as

$$A(\Sigma) = \int_{\Omega} |\langle \gamma_s'(s,t) \wedge \gamma_t'(s,t), \gamma_s'(s,t) \wedge \gamma_t'(s,t)\rangle|^{1/2} ds\, dt, \tag{2.8}$$

where the mapping $\gamma: R^2 \supset \Omega \to E$ is a parametrization of Σ. It is a standard result of advanced calculus that this area is independent of the particular parametrization if we only consider sufficiently well behaved parametrizations. It is straightforward to generalize (2.8) and obtain natural length, area, and volume measures on 1, 2, 3, or 4-surfaces in E or V.

Example: S has the parametrization $\gamma: R^3 \to V$, where

$$\gamma(\mathbf{u}) = u^0 e_0 + u^1 e_1 + u^2 e_2 + u^3 e_3, \tag{2.9}$$

where $u^0 = (1 + \mathbf{u}\cdot\mathbf{u})^{1/2}$ and where e_0, e_1, e_2, e_3 belongs to a Lorentz frame and $\mathbf{u} = (u^1, u^2, u^3)$. We have

$$\frac{\partial \gamma}{\partial u^i} = \frac{u^i}{u^0} e_0 + e_i \quad \text{for } i = 1,2,3$$

and thus

$$\begin{aligned} &\left|\left\langle \frac{\partial \gamma}{\partial u^1} \wedge \frac{\partial \gamma}{\partial u^2} \wedge \frac{\partial \gamma}{\partial u^3}, \frac{\partial \gamma}{\partial u^1} \wedge \frac{\partial \gamma}{\partial u^2} \wedge \frac{\partial \gamma}{\partial u^3}\right\rangle\right|^{1/2} du^1 du^2 du^3 \\ &= \left|\det(-u^i u^j/(u^0)^2 + \delta_{ij})_{\substack{1 \leqslant i \leqslant 3 \\ 1 \leqslant j \leqslant 3}}\right|^{1/2} du^1 du^2 du^3 \\ &= \frac{1}{u^0} du^1\, du^2\, du^3. \end{aligned} \tag{2.10}$$

Thus

$$\int_S g(u)\, du = \int_{R^3} g \circ \gamma(\mathbf{u}) \frac{1}{u^0} du^1\, du^2\, du^3$$

for a function $g: S \to R$. In the same way we obtain for $h: E \to R$,

$$\int_E h(P)\, dP = \int_E h \circ L_E^{-1}(x^i)\, dx^0\, dx^1\, dx^2\, dx^3, \tag{2.11}$$

where $L_E: E \to R^4$ is a map belonging to a Lorentz system as in Section B above. It is practical to write $g(\mathbf{u})$ for $g \circ \gamma(\mathbf{u})$ and $h(x^i)$ for $h \circ L_E^{-1}(x^i)$.

We now define the operators ∇_E, ∇_V, and ∇_S by

(a) Let $T: E \to V^{\otimes k}$, then $\nabla_E T: E \to V^{\otimes(k+1)}$ such that

$$\|T(P) - T(P_0) - (P - P_0)\cdot\nabla_E T(P_0)\| = o(\|P - P_0\|). \tag{2.12}$$

(b) Let $T: V \to V^{\otimes k}$, then $\nabla_V T: V \to V^{\otimes(k+1)}$ such that

$$\|T(y) - T(y_0) - (y - y_0)\cdot\nabla_V T(y_0)\| = o(\|y - y_0\|). \tag{2.13}$$

(c) Let $T: S \to V^{\otimes k}$, then $\nabla_S T: S \to V^{\otimes(k+1)}$ such that

$$\|T(u) - T(u_0) - (u - u_0)\cdot\nabla_S T(u_0)\| = o(\|u - u_0\|). \tag{2.14}$$

and $u_0\cdot\nabla_S T(u_0) = 0$.

In (a), (b), and (c) $\|\ \|$ denotes some norm on the relevant vector space. The definitions do not depend on the particular norms chosen since all norms on a finite-dimensional vector space are equivalent.[9] In the right-hand side of (2.12)–(2.14) o has the property $\lim_{t\to 0}[o(t)/t] = 0$.

In a Lorentz system L we obtain from (a), (b), and (c):

$$\nabla_E = -e_0 \frac{\partial}{\partial x^0} + \sum_{i=1}^{3} e_i \frac{\partial}{\partial x^i}, \quad \text{where } L_E(P) = (x^i), \tag{2.15}$$

$$\nabla_V = -e_0 \frac{\partial}{\partial y^0} + \sum_{i=1}^{3} e_i \frac{\partial}{\partial y^i}, \quad \text{where } L_V(y) = (y^i), \tag{2.16}$$

$$\nabla_S = \sum_{i=1}^{3} e_i \frac{\partial}{\partial u^i} + u \sum_{i=1}^{3} u^i \frac{\partial}{\partial u^i}, \quad \text{where } u = \sum_{i=1}^{3} u^i e_i, \tag{2.17}$$

Example: Let T be as in (a) above. Express T in the system L as $T = \Sigma_{\alpha\in I} T_\alpha \lambda_\alpha$, where $T_\alpha: R^4 \to R$ and $\{\lambda_\alpha | \alpha \in T\}$ is the base of $V^{\otimes k}$ obtained as in (2.1). Then

$$\begin{aligned} \nabla_E T = &- \sum_{\alpha\in I}\left(\frac{\partial}{\partial x^0} T_\alpha\right) e_0 \otimes \lambda_\alpha \\ &+ \sum_{\alpha\in I} \sum_{i=1}^{3} \left(\frac{\partial}{\partial x^i} T_\alpha\right) e_i \otimes \lambda_\alpha. \end{aligned} \tag{2.18}$$

For $\phi: E \to V$ we define the four-dimensional curl operator $\nabla_E \wedge \phi = \nabla_E \phi - (\nabla_E \phi)^t$, where t means $(v \otimes w)^t = w \otimes v$. We will also need the Taylor expansion of a function $h: E \times V \to V^{\otimes k}$,

$$\begin{aligned} h(P_0 + \Delta x, y_0 + \Delta y) = &\sum_{n=0}^{\infty} \frac{1}{n!} (\Delta x\cdot\nabla_E + \Delta y\cdot\nabla_V)^n \\ &\times h(P_0, y_0) \end{aligned} \tag{2.19}$$

which is a practical but formal notation for

$$\begin{aligned} h&(P_0 + \Delta x, y_0 + \Delta y) \\ &= \sum_{n=0}^{\infty} \frac{1}{n!}\binom{n}{m} \Delta x^{\otimes m} \otimes \Delta y^{\otimes(n-m)} \\ &\quad \times \binom{n}{\cdot} \nabla_E^m \nabla_V^{(n-m)} h(P,y)|_{(P_0,y_0)}. \end{aligned} \tag{2.20}$$

D. The Vlasov and Maxwell equations

The electromagnetic field tensor F and 4-current J are mappings $F: E \to V^{\wedge 2}$ and $J: E \to V$. The Maxwell's equations are given by

$$\nabla_E \cdot F = \left(\frac{\mu_0}{\epsilon_0}\right)^{1/2} J, \tag{2.21}$$

$$\nabla_E \cdot {}^* F = 0, \tag{2.22}$$

where $*$ is the Hodge star operator defined in Ref. 8. We will in this paper assume the existence of a four-potential $\Phi: E \to V$ such that $-\nabla_E \wedge \Phi = F$. This is a natural assumption in a Vlasov plasma where discrete particle effects are neglected.

Let $f: E \times S \to R$ be the distribution function of some particle species in the plasma with charge q and rest mass m_0. The Vlasov equation for f may be written

$$Df = 0 \tag{2.23}$$

where the Vlasov operator is

$$D = u\cdot\nabla_E + \frac{q}{m_0c^2}(\nabla_E\wedge\Phi\cdot u)\cdot\nabla_S. \quad (2.24)$$

The electromagnetic wave equation

$$\Box\Phi = \left(\frac{\mu_0}{\epsilon_0}\right)^{1/2} J, \quad (2.25)$$

where $\Box\Phi = \nabla_E\cdot(\nabla_E\wedge\Phi)$, together with the Vlasov equation for all particle species constitutes, if the eventual external 4-current is known, a closed system of equations.

In a Lorentz system L we keep the standard notation for electric and magnetic field $\mathbf{E}$ and $\mathbf{B}$, the ordinary current density $\mathbf{J}$, the scalar and vector potentials U and $\mathbf{A}$, and the charge density ρ. We have:

$$F = \mathbf{E}\wedge e_0 + c^*(\mathbf{B}\wedge e_0), \quad (2.26)$$

$$^*F = {}^*(\mathbf{E}\wedge e_0) - c\mathbf{B}\wedge e_0, \quad (2.27)$$

$$\Phi = Ue_0 + c\mathbf{A}, \quad (2.28)$$

$$J = c\rho e_0 + \mathbf{J}, \quad (2.29)$$

and the operators

$$D = \sum_{i=0}^{3} u^i \frac{\partial}{\partial x^i} + \frac{q}{m_0c^2}(u_0\mathbf{E} + c\mathbf{u}\times\mathbf{B})\cdot\frac{\partial}{\partial\mathbf{u}} \quad (2.30)$$

and

$$\Box = -\frac{\partial^2}{\partial(x^0)^2} + \frac{\partial^2}{\partial(x^1)^2} + \frac{\partial^2}{\partial(x^2)^2} + \frac{\partial^2}{\partial(x^3)^2}. \quad (2.31)$$

Substitution of (2.26)–(2.27) in (2.21)–(2.22) yields the "space–plus–time" form of Maxwells equations. If we use

$$\nabla_E\cdot{}^*(\mathbf{B}\wedge e_0) = \mathrm{curl}\,\mathbf{B}, \quad (2.32)$$

$$\nabla_E\cdot(\mathbf{B}\wedge e_0) = -c^{-1}\frac{\partial}{\partial t}\mathbf{B} + (\mathrm{div}\,\mathbf{B})e_0, \quad (2.33)$$

and the corresponding relation for $\mathbf{E}$. The Lorentz condition $\nabla_E\cdot\Phi = 0$ is only used in the derivation of (2.31).

3. THE MAIN RESULTS

The plasma is described by the relativistic Vlasov equation (2.23) and the electromagnetic wave equation (2.25). The unperturbed plasma state may be inhomogeneous nonstationary and space–time dependent external electromagnetic fields may be present. The state is determined by the distribution functions $f_0^\sigma: E\times S\to V$, for all particle species σ, and the 4-potential $\Phi_0: E\to V$. Here Φ_0 satisfies the electromagnetic wave equation with plasma and external 4-currents included.

Omission of index σ: For notational reason we will frequently omit the particle species index σ and also summations over σ. It is an easy task to fill in the missing σ and Σ_σ.

"Space-plus-time" statements: A statement involving space or time separately without reference to any Lorentz frame means that there exists a Lorentz frame such that the statement is true.

Formal calculations: The results formulated in this section are obtained by formal calculations and their domain of validity are undetermined. Thus we must not take the rather formal way of presentation too seriously, the use of a mathematical terminology clarifies the structure of the content but we prove no mathematically rigorous results (see however remark 2 below).

Definition 1: $L_0(E,V)$ and $L^0(E,V)$ are spaces of functions from E to V such that

(i) $\phi\in L_0(E,V)\Rightarrow\phi\to 0$ uniformly sufficiently fast towards the past,

(ii) $\phi\in L^0(E,V)\Rightarrow\phi\to 0$ uniformly sufficiently fast towards the future.

Remark 1: Towards the limit of the infinite past we want the perturbed system to approach the unperturbed one (see Definition 4 and Lemma 1). This is why the function in $L_0(E,V)$ have to vanish sufficiently fast towards the past. For mathematical reasons it is convenient to allow perturbations which do not vanish exactly prior to any finite time. It is for example sometimes suitable to calculate the response of a perturbation which exponentially approaches zero towards the past.[10]

Definition 2: For $\phi,\phi_1,\ldots,\phi_m\in L_0(E,V)$ and $m = 1,2,\cdots$ we define

(i) $\delta J[\phi]: E\to V$ is the change in plasma 4-current due to the perturbation ϕ of the 4-potential,

(ii) $\delta J^{(m)}[\phi]$ is defined from the expansion (1.1) of $\delta J[\phi]$,

(iii) $\delta J^{(m)}[\phi_1,\ldots,\phi_m]$ is determined by

(a) it is linear in each variable ϕ_i,

(b) it is symmetric with respect to permutations of its m arguments ϕ_i,

(c) $\delta J^{(m)}[\phi_1,\ldots,\phi_m] = \delta J^{(m)}[\phi]$ if $\phi_1 = \cdots = \phi_m = \phi$.

Definition 3: Some set-theoretic notations will be used.

(i) $N = \{0,1,\ldots\}$, $N_m = \{0,1,\ldots,m\}$, $N^+ = \{1,2,\ldots\}$, $N_m^+ = \{1,2,\ldots,m\}$;

(ii) Let B, C, and D denote finite sets:

$P(B) = \{$all partitions of $B\}$
$= \{\Gamma \mid \Gamma$ is a set of disjoint nonempty subsets in B, such that $\bigcup_{C\in\Gamma} C = B\}$,

$n(B)$ = number of elements in B,

$$|\Gamma| = \max_{B\in\Gamma} n(B) \quad\text{and}\quad \Gamma! = \frac{[\Sigma_{B\in\Gamma} n(B)]!}{\Pi_{B\in\Gamma}[n(B)!]}$$

for $\Gamma\in P(B)$.

We will use upper and lower indices and a second set in the argument, in any combination, to denote certain subsets of $P(B)$

$$P_k^l(B,C) = \{\Gamma\in P(B) \mid n(\Gamma)\geqslant k, |\Gamma|\leqslant l, C\in\Gamma\}.$$

Thus in (3.1)–(3.2)

$$P_3^l(N_m) = \{\Gamma\in P(N_m) \mid n(\Gamma)\geqslant 3, |\Gamma|\leqslant l\},$$

(iii) If $\Gamma\in P(B)$ and $k\in B$ and $\{k\}\in\Gamma$, then $\Gamma\setminus\{k\}$ is denoted $(\Gamma|\setminus|k)$. In general we use the notation $(\alpha|\beta|\cdots|\delta|\setminus|\gamma|\cdots|\theta)$, where in the place of a letter we have one or several elements of N or one finite subset of N or some partition of a finite subset of N. The definition is understood from an example:

$$(B\mid \Gamma_1\mid \Gamma_2\mid 1,2\mid\setminus\mid \Gamma_3\mid 3)\overset{\text{def}}{=}\{B\}\cup\Gamma_1\cup\Gamma_2\cup\{\{1,2\}\}\setminus\Gamma_3\cup\{\{3\}\},$$

where Γ_i are partitions of some subsets in N and $B\subset N$.

Result 1: Take $\phi_1,\dots,\phi_m\in L_0(E,V)$, and $\phi_0\in L^0(E,V)$. Then

$$\int_E \phi_0(P)\cdot\delta J^{(m)}[\phi_1,\dots,\phi_m](P)\,dP$$

$$=\begin{cases}\sum_{\Gamma\in P_3^l(N_m)} A(\Gamma), & \text{for } m=2l, \quad (3.1)\\ \frac{1}{2}\sum_{\Gamma\in I_m} A(\Gamma)+\sum_{\Gamma\in P_3^l(N_m)} A(\Gamma), & \text{for } m=2l+1, \quad (3.2)\end{cases}$$

for $l=1,2,\cdots$ and $I_m=P_3^{l+1}(N_m)\setminus P_3^l(N_m)$. Here $m\geqslant 2$ and the $m=1$ use is treated in (3.14) part II. We define

$$A(\Gamma)=\sum_\sigma qc(m+1)(\Gamma!)^{-1}\int_{E\times S} f_0(P,u)\Big[\Delta(\Gamma)(\Phi_0(P)\cdot u)+\sum_{\{k\}\in\Gamma}\Delta(\Gamma\mid\setminus\mid k)(\phi_k(P)\cdot u)\Big]\,dP\,du \quad (3.3)$$

and [if we use formal notations analogous with (2.19)]

$$\Delta(\Gamma)=\prod_{B\in\Gamma}[\delta x(B)\cdot\nabla_E+\delta u(B)\cdot\nabla_V], \quad (3.4)$$

where $\delta x(B)$ and $\delta u(B)$ are functions from $E\times S$ to V and determined from the hierachy of equations

$$D_0\delta x(B)=\delta u(B), \quad (3.5)$$

$$D_0\delta u(B)=qm_0^{-1}c^{-2}\sum_{\Gamma\in P(B)}(\Gamma!)^{-1}\Big[\Delta(\Gamma)(\nabla_E\wedge\Phi_0\cdot u)+\sum_{\{k\}\in\Gamma}\Delta(\Gamma\mid\setminus\mid k)(\nabla_E\wedge\phi_k\cdot u)\Big], \quad (3.6)$$

together with the boundary conditions:

$\delta x(B),\ \delta u(B)\to 0$ towards the past if $0\notin B$, (3.7)

$\delta x(B),\ \delta u(B)\to 0$ towards the future if $0\in B$. (3.8)

Remark 2: In Result 1 the conditions we may impose on ϕ_i for $i\geqslant 1$ in a meaningful way is restricted by the fact that we are interested in $\delta J^{(m)}[\phi]$ for solving the electromagnetic wave equation selfconsistently for ϕ. Thus ϕ is somewhat out of our control and it is for $\phi_1=\cdots=\phi_m=\phi$ that $\delta J^{(m)}[\phi_1,\dots,\phi_m]$ has a physical interpretation. On the other hand, the choice of ϕ_0 is ours and we may rather freely impose conditions on ϕ_0 to compensate for bad behavior in ϕ_i, $i\geqslant 1$. In fact we may use ϕ_0 much as a test function in distribution theory and, just as in that theory, formal calculations on even badly behaving functions may turn out to be justified when we consider integrated quantities like (3.1).

Remark 3: The restriction $\phi_0\in L^0(E,V)$ is made to make it probable that $\phi_0(P)\cdot\delta J^{(m)}[\phi_1,\dots,\phi_m](P)$ approaches zero in all directions (we consider the plasma to be spatially finite at finite times) so that we may integrate this quantity over E and so that surface contributions vanish when we perfrom partial integrations (cf. Lemma 6 in Sec. 4).

Definition 4: For ϕ, $\phi_1,\dots,\phi_m\in L^0(E,V)$ we define

(i) $\delta x[\phi]$ and $\delta u[\phi]$ such that they describe the change in the particles positions in $E\times S$ due to the perturbation ϕ such that given the orbit $(P(s),u(s))$, where s/c denotes proper time, of a plasma particle in the unperturbed plasma the orbit of this particle in the presence of the perturbation ϕ would be $(P(s)+\delta x[\phi](P(s),u(s)),\ u(s)+\delta u[\phi](P(s),u(s)))$, where s/c still denotes proper time.

(ii) $\delta x^{(m)}[\phi]$, $\delta u^{(m)}[\phi]$ are defined from the expansion of $\delta x[\phi]$ and $\delta u[\phi]$ analogous to (ii) in Definition 2.

(iii) $\delta x^{(m)}[\phi_1,\dots,\phi_m]$ and $\delta u^{(m)}[\phi_1,\dots,\phi_m]$ are determined by

(a) They are linear in each variable ϕ_i

(b) They are symmetric with respect to permuations of their m arguments ϕ_i

(c) $\delta x^{(m)}[\phi_1,\dots,\phi_m]=\delta x^{(m)}[\phi]$ and

$\delta u^{(m)}[\phi_1,\dots,\phi_m]=\delta u^{(m)}[\phi]$ if $\phi_1=\cdots=\phi_m=\phi$.

Result 2: Let ϕ_i, $\delta x(B)$, and $\delta u(B)$ be as in Result 1 and $0\notin B=\{i(1),\dots,i(m)\}$ and $m=n(B)$, then

$$\delta x(B)=\delta x^{(m)}[\phi_{i(1)},\dots,\phi_{i(m)}], \quad (3.9)$$

$$\delta u(B)=\delta u^{(m)}[\phi_{i(1)},\dots,\phi_{i(m)}]. \quad (3.10)$$

Corollary 1: If ϕ_0, $\phi_1,\dots,\phi_m\in L_0(E,V)\cap L^0(E,V)$ and if $\delta x(B),\delta u(B)\to 0$ both towards the past and the future, the quantity

$$\int_{E\times S}\phi_0(P)\cdot\delta J^{(m)}[\phi_1,\dots,\phi_m](P)\,dP \quad (3.11)$$

is symmetric with respect to permutations of ϕ_0, $\phi_1,\dots,\phi_m$.

Proof: It is only (3.7), (3.8), and the requirement that $\phi_0\in L^0(E,V)$ while other $\phi_i\in L_0(E,V)$ that makes our expression for (3.11) in Result 1 not symmetric in general. These asymmetries are eliminated by the assumptions in Corollary 1.

Remark 4: In Result 1 we consider $\phi_0,\dots,\phi_m$ as given in order to simplify notation. However, we are in reality interested in varying these functions and we will now introduce some more operator minded notations.

Definition 5: We define on operators A_m and A_j: $L^0(E,V)\times(L_0(E,V))^{\times m}\to R$ for $j\in T_m$ [see Definition 6(a) below] and a subset $T_3^l(m)$ of T_m:

(i) $T_3^l(m)=\{j\in T_m\mid\Sigma_{i=}^\infty j(i)\geqslant 3$ and $j(i)=0$ for $i>l\}$,

(ii) $A_j(\phi_0,\dots,\phi_m)=\Sigma_{\Gamma:j_\Gamma=j}A(\Gamma)$, where j_Γ is defined in (4.16),

(iii) $A_m(\phi_0,\dots,\phi_m)=\int_E\phi_0(P)\cdot\delta J^{(m)}[\phi_1,\dots,\phi_m](P)\,dP$.

Remark 5: In (ii) above we add all $A(\Gamma)$ for all Γ having a particular structure. (With the structure of Γ we mean the information of how many sets in Γ have one element, two elements, and so on. The structure of Γ is determined by the function j_Γ where $j_\Gamma(i)$ is the number of sets in Γ with i elements.) It is easy to see that A_j ($j\in T_m$) is symmetric with respect to permutation of its last m variables, and linear in each variable.

Corollary 2:

$$A_m = \begin{cases} \sum_{j\in T_3^l(m)} A_j, & \text{for } m = 2l, \\ \frac{1}{2}\sum_{j\in j(I_m)} A_j + \sum_{j\in T_3^l(m)} A_j, & \text{for } m = 2l+1, \end{cases}$$

where $j(I_m) = T_3^{l+1}(m)\setminus T_3^l(m)$ and $l = 1,2,\cdots$.

Corollary 3: If $\phi_0, \phi_1,\ldots,\phi_m \in L_0(E,V)\cap L^0(E,V)$ and if $\delta x(B), \delta u(B)\to 0$ both towards the past and the future, then $A_j(\phi_0,\ldots,\phi_m)$ is symmetric with respect to permutation of $\phi_0,\ldots,\phi_m$.

Proof of Corollary 3: The same as for Corollary 1.

4. DERIVATION OF THE RESULTS

The unperturbed plasma (f_0,Φ_0) and $\phi_0,\ldots,\phi_m$ are taken as in Result 1. We consider $\phi\in L_0(E,V)$ as given. The Vlasov equation for the unperturbed plasma is $D_0 f_0 = 0$.

Lemma 1: The transformation of $E\times S$ defined by $(P,u)\to(P+\delta x[\phi](P,u), u+\delta u[\phi](P,u))$ is measure preserving.

Proof: A 4-potential Φ: $E\to V$ determines a one parameter group of transformations $\Upsilon_s[\Phi]$: $E\times S\to E\times S$ by the vector field $v[\Phi]$: $E\times S\to V\times V$ defined as

$$v[\Phi](P,u) = \left(u, \frac{q}{m_0c^2}\nabla_E\wedge\Phi(P)\cdot u\right).$$

Take $(P(s),u(s)) = \Upsilon_s[\Phi](P,u)$, then $(P(s),u(s))$ is the path of a plasma particle in $E\times S$, where s/c is the lapse of proper time and $(P(0),u(0)) = (P,u)$. Since $(P+\delta x[\phi](P,u), u + \delta u[\phi](P,u)) = \lim_{s\to\infty}\Upsilon_s[\Phi_0+\phi]\circ\Upsilon_{-s}[\Phi_0](P,u)$, it is sufficient to prove that $\Upsilon_s[\Phi]$ is measure preserving for all s. Take a Lorentz frame L. Then with notation introduced in Sec. 2, the map $L_E\times L_V$: $E\times S\to R^7$ is seen to transform our problem on $E\times S$ into the following on R^7: Prove that the flow w: $R^7\to R^7$ defined by

$$\mathbf{w}((x^i),\mathbf{u}) = \left(\mathbf{u}, \frac{q}{m_0c^2}(u^0\mathbf{E}+\mathbf{u}\times\mathbf{B})\right)$$

preserves the measure $(u^0)^{-1}dx^0\,dx^1\,dx^2\,dx^3\,du^1\,du^2\,du^3$. A sufficient condition for this is $\operatorname{div}[(u^0)^{-1}\mathbf{w}] = 0$, where div is the seven-dimensional divergence. We obtain by straightforward computation

$$\operatorname{div}[(u^0)^{-1}\mathbf{w}] = \frac{q}{m_0c^2}\frac{\partial}{\partial\mathbf{u}}\cdot((u^0)^{-1}\mathbf{u}\times\mathbf{B}) = 0.$$

Lemma 2: (a) The function $\delta x[\phi]$ and $\delta u[\phi]$ are determined from

$$D_0\delta x[\phi](P,u) = \delta u[\phi](P,u), \tag{4.1}$$

$$D_0\delta u[\phi](P,u) = \frac{q}{m_0c^2}[\nabla_E\wedge(\Phi_0+\phi)|_{P+\delta x(P,u)}\cdot(u + \delta u[\phi](P,u)) - \nabla_E\wedge\Phi_0(P)\cdot u], \tag{4.2}$$

with the boundary condition $\delta x[\phi]$ and $\delta u[\phi]\to 0$ towards the past.

$$\text{(b)} \int_E \phi_0(P)\cdot\delta J[\phi](P)\,dP = qc\int_{E\times S}[\phi_0(P+\delta x(P,u))\cdot(u+\delta u(P,u)) - \phi_0(P)\cdot u] \times f_0(P,u)\,dP\,du. \tag{4.3}$$

Proof: Part (a) follows straightforwardly from the definition of $\delta x[\phi]$ and $\delta u[\phi]$. Part (b) follows from the equality

$$\int_{E\times S}\phi_0(P)\cdot uf(P,u)\,dP\,du = \int_{E\times S}\phi_0(P+\delta x(P,u))\cdot(u+\delta u(P,u))f_0(P,u)\,dP\,du. \tag{4.4}$$

To prove (4.4) we make the variable substitution $(P,u)\to(P+\delta x(P,u), u+\delta u(P,u))$ in the left-hand side and make use of Lemma 1 and

$$f_0(P,u) = f(P+\delta x(P,u), u+\delta u(P,u)). \tag{4.5}$$

The relation (4.5) is a consequence of the Vlasov equation.

Definition 6:

(a) $T_m = \left\{j \mid j: N^+\to N \text{ and } \sum_{i=1}^\infty ij(i) = m\right\}$ for $m\in N$;

(b) $\Delta(j) = \prod_{i=1}^\infty (\delta x^{(i)}[\phi]\cdot\nabla_E + \delta u^{(i)}[\phi]\cdot\nabla_S)^{j(i)}$

for $j\in T_m$ and $m\in N$. Note that $\Delta(j)$ is the identity operator for $j\in T_0$;

(c) $j! = \prod_{i=1}^\infty (j(i)!)$ for $j\in T_m$.

Lemma 3: (a) The functions $\delta x^{(m)}[\phi]$ and $\delta u^{(m)}[\phi]$ are determined from

$$D_0\delta x^{(m)}[\phi] = \delta u^{(m)}[\phi], \tag{4.6}$$

$$D_0\delta u^{(m)}[\phi] = \frac{q}{m_0c^2}\left[\sum_{j\in T_m}(j!)^{-1}\Delta(j)(\nabla_E\wedge\Phi_0\cdot u) + \sum_{j\in T_{m-1}}(j!)^{-1}\Delta(j)(\nabla_E\wedge\phi\cdot u)\right], \tag{4.7}$$

with the boundary conditions $\delta x[\phi]$ and $\delta u[\phi]\to 0$ towards the past.

$$\text{(b)} \int_E \phi_0(P)\cdot\delta J^{(m)}[\phi](P)\,dP = qc\sum_{j\in T_m}(j!)^{-1}\int_{E\times S} f_0(P,u)\Delta(j)(\phi_0(P)\cdot u)\,dP\,du. \tag{4.8}$$

Proof: Taylor expansion of the right-hand sides in (a) and (b) Lemma 2 yields

$$D_0\delta u[\phi] = \frac{q}{m_0c^2}$$

$$\left[\sum_{m=1}^{\infty}\frac{1}{m!}(\delta x[\phi]\cdot\nabla_E + \delta u[\phi]\cdot\nabla_V)^m(\nabla_E\wedge\Phi_0\cdot u)\right.$$

$$\left. + \sum_{m=0}^{\infty}\frac{1}{m!}(\delta x[\phi]\cdot\nabla_E + \delta u[\phi]\cdot\nabla_V)^m(\nabla_E\wedge\phi\cdot u)\right] \quad (4.9)$$

and

$$\int_E \phi_0(P)\cdot\delta J[\phi](P)\,dP$$

$$= qc\int_{E\times S} f_0(P,u)\sum_{m=1}^{\infty}(m!)^{-1}$$

$$\times(\delta x[\phi]\cdot\nabla_E + \delta u[\phi]\cdot\nabla_V)^m(u\cdot\phi_0)\,dP\,du. \quad (4.10)$$

We here regard $u\cdot\phi_0(P)$ as a function on $E\times V$ (not only on $E\times S$) and accordingly $\nabla_V(u\cdot\phi_0) = \phi_0$. We substitute $(\delta J[\phi],\delta x[\phi],\delta u[\phi]) = \Sigma_{m=1}^{\infty}(\delta J^{(m)}[\phi],\delta x^{(m)}[\phi],\delta u^{(m)}[\phi])$ and get equations for each order of nonlinearity.

We obtain for example: Terms nonlinear of order m in $\Sigma_{k=1}^{\infty}(k!)^{-1}[\Sigma_{n=1}^{\infty}(\delta x^{(n)}[\phi]\cdot\nabla_E + \delta u^{(n)}\cdot\nabla_V)]^k$ sum up to

$$\sum_{j\in T_m}(j!)^{-1}\prod_{i=1}^{\infty}(\delta x^{(i)}[\phi]\cdot\nabla_E + \delta u^{(i)}[\phi]\cdot\nabla_V)^{j(i)}, \quad (4.11)$$

where thus $j(i)$ is the number of $(\delta x^{(i)}[\phi]\cdot\nabla_E + \delta u^{(i)}[\phi]\cdot\nabla_V)$ factors in a term and $\Sigma_{i=1}^{\infty}ij(i) = m$ is the order. It is an easy combinatorial problem to find the factor $(j!)^{-1}$ in (4.11).

Proof of Result 2: We define and $\delta\tilde{x}^{(k)}[\phi_{i_1},...,\phi_{i_k}] = \delta x(B)$ and $\delta\tilde{u}^{(k)}[\phi_{i_1},...,\phi_{i_k}] = \delta u(B)$, where $B = \{i_1,...,i_k\}$ and $0<i_1<\cdots<i_k$. It is easy to see that $\delta\tilde{x}^{(k)}$ and $\delta\tilde{u}^{(k)}$ is linear in each variable and symmetric, and if we also prove that

$$\delta x^{(k)}[\phi] = \delta\tilde{x}^{(k)}[\phi] \overset{\text{def}}{=} \delta\tilde{x}^{(k)}[\phi,...,\phi] \quad (k \text{ times } \phi), \quad (4.12)$$

$$\delta u^{(k)}[\phi] = \delta\tilde{u}^{(k)}[\phi] \overset{\text{def}}{=} \delta\tilde{u}^{(k)}[\phi,...,\phi] \quad (k \text{ times } \phi), \quad (4.13)$$

it follows from (iii) in Definition 4 that the $\sim$ (overtilde) may be removed without changing the functions and thus Result 2 follows.

We first prove, with the obvious definition of $\tilde{\Delta}(j)$, that $\phi_1 = \cdots = \phi_m = \phi$ implies:

$$\sum_{\Gamma\in P(N_m^+)}(\Gamma!)^{-1}\Delta(\Gamma) = \sum_{j\in T_m}(j!)^{-1}\tilde{\Delta}(j) \quad (4.14)$$

and

$$\sum_{\Gamma,k}(\Gamma!)^{-1}\Delta(\Gamma\mid\setminus\mid k) = \sum_{j\in T_{m-1}}(j!)^{-1}\tilde{\Delta}(j), \quad (4.15)$$

where $\Sigma_{k,\Gamma}$ the sum is taken over all k, Γ such that $k\in N_m^+$ and $\{k\}\in\Gamma\in P(N_m^+)$.

We use the variable substitution $P(N_m^+)\ni\Gamma\to j_\Gamma\in T_m$ in the left-hand side of (4.14) where

$$j_\Gamma(i) = n\{B\in\Gamma \mid n(B) = i\}. \quad (4.16)$$

This substitution make sense since $\Delta(\Gamma)$ and $\Gamma!$ may be expressed in terms of j_Γ: $\Delta(\Gamma) = \tilde{\Delta}(j_\Gamma)$ and $\Gamma! = m![\Pi_{i=1}^{\infty}(i!)j_\Gamma(i)]^{-1}$. In this way we obtain

$$\sum_{\Gamma\in P(N_m^+)}(\Gamma!)^{-1}\Delta(\Gamma) = \sum_{j\in T_m}(m!)^{-1}\prod_{i=1}^{\infty}(i!)^{j(i)}$$

$$\times n\{\Gamma\in P(N_m^+)\mid j_\Gamma = j\}\tilde{\Delta}(j_\Gamma). \quad (4.17)$$

The number $n\{\Gamma\in P(N_m^+)\mid j_\Gamma = j\}$ is calculated by counting the number of ways we may construct Γ for a given j by the following procedure:

(1) Divide N_m^+ in disjoint sets B_i such that $N_m^+ = \cup_{i=1}^{\infty}B_i$ and $n(B_i) = ij(i)$.

(2) Divide each B_i in $j(i)$ subsets with i elements in each.

If the subsets in (2) are used as the elements of Γ, then $j_\Gamma = j$. The number of ways of performing (1) and (2), respectively is

$$n_1 = \binom{m}{j(1)}\binom{m-j(1)}{2j(2)}\binom{m-j(1)-2j(2)}{3j(3)}\cdots, \quad (4.18)$$

$$n_2 = \prod_{i=1}^{\infty}\left[\frac{1}{j(i)!}\binom{ij(i)}{i}\binom{i(j(i)-1)}{i}\binom{i(j(i)-2)}{i}\cdots\right], \quad (4.19)$$

and we have

$$n\{\Gamma\in P(N_m^+)\mid j_\Gamma = j\} = n_1\cdot n_2 = m!\left[j!\prod_{i=1}^{\infty}(i!)^{j(i)}\right]^{-1}. \quad (4.20)$$

Substitutions of (4.20) in (4.17) yields (4.14). We obtain (4.15) from (4.14) and $(\Gamma\mid\setminus\mid k)! = m^{-1}\Gamma!$ From (4.14) and (4.15) we now see that (3.5) and (3.6) becomes exactly the same equations for $\delta\tilde{x}^{(k)}[\phi]$ and $\delta\tilde{u}^{(k)}[\phi]$ as (4.6) and (4.7) are for $\delta x^{(k)}[\phi]$ and $\delta u^{(k)}[\phi]$ and thus (4.12) and (4.13) are true and so Result 2 follows.

Lemma 4:

$$\int_E \phi_0(P)\cdot\delta J^{(m)}[\phi_1,...,\phi_m](P)\,dP$$

$$= qc\sum_{\Gamma\in P(N_m^+)}(\Gamma!)^{-1}\int_{E\times S} f_0(P,u)\Delta(\Gamma)u\cdot\phi_0(P)\,dP\,du. \quad (4.21)$$

Proof: It is sufficient to prove that (4.21) is consistent with (iii) in Definition 2. Here (a) and (b) are easily checked and (c) follows from (4.14) substituted in 4.21 when $\phi_1 = \cdots = \phi_m = \phi$ and comparison with (4.8), remembering that $\tilde{\Delta}(j) = \Delta(j)$ as proved above.

Lemma 5: Let B and C be two nonempty finite and disjoint subsets of N and $\Gamma\in P(C)$, then

$$\text{(a) } \Delta(B)(\phi\cdot u) = \delta x(B)\cdot\nabla_E\wedge\phi\cdot u + D_0[\delta x(B)\cdot\phi], \quad (4.22)$$

$$\text{(b) } \Delta(\Gamma\mid B)(\phi\cdot u) = \delta x(B)\cdot\Delta(\Gamma)(\nabla_E\wedge\phi\cdot u) + D_0[\delta x(B)\cdot\Delta(\Gamma)\phi]. \quad (4.23)$$

Proof: The proof is a straightforward computation where we made use of identities like

$$\Delta(\Gamma)(\phi\cdot u) = (\Delta(\Gamma)\phi)\cdot u + \sum_{D\in\Gamma} \delta u(D)\cdot\Delta(\Gamma \mid \setminus \mid D)\phi, \tag{4.24}$$

$D_0\delta x(B) = \delta u(B)$, and $D_0\phi = u\cdot\nabla_E\phi$.

Lemma 6: Let $h: E\times S \to R$ be vanishing sufficiently fast towards the past and the future, then

$$\int_{E\times S} f_0(P,u) D_0 h(P,u)\, dP\, du = 0. \tag{4.25}$$

Proof: We have $f_0 D_0 h = D_0(f_0 h)$ since $D_0 f_0 = 0$ To prove $\int_{E\times S} D_0(f_0 h) dPdu = 0$ we transform the integral to a Lorentz frame and the result follows by partial integration since the surface contribution vanishes.

Lemma 7: Let B and C be two nonempty disjoint and finite subsets of N and $0\in B\cup C$, then

$$\sum_{\Gamma\in P(B)} (\Gamma!)^{-1} \int_{E\times S} \Big[\Delta(\Gamma \mid C)\Phi_0\cdot u + \sum_{\{k\}\in\Gamma} \Delta(\Gamma \mid C \mid\setminus\mid k)(\phi_k\cdot u)\Big] f_0 dPdu$$
$$= \sum_{\Gamma\in P(C)} (\Gamma!)^{-1} \int_{E\times S} \Big[\Delta(\Gamma \mid B)\Phi_0\cdot u + \sum_{\{k\}\in\Gamma} \Delta(\Gamma \mid B \mid\setminus\mid k)(\phi_k\cdot u)\Big] f_0 dPdu \tag{4.26}$$

Proof: From (3.6) we obtain

$$\frac{m_0c^2}{q}\int_{E\times S} \delta x(C)\cdot[D_0\delta u(B)] f_0 dPdu = \sum_{\Gamma\in P(B)} (\Gamma!)^{-1} \int_{E\times S} \delta x(C)\cdot\Big[\Delta(\Gamma)(\nabla_E\wedge\Phi_0\cdot u)$$
$$+ \sum_{\{k\}\in\Gamma} \Delta(\Gamma \mid\setminus\mid k)(\nabla_E\wedge\phi_k\cdot u)\Big] f_0 dPdu. \tag{4.27}$$

Substitution of (4.22) and (4.23) in the right-hand side of (4.27) yields

$$\frac{m_0c^2}{q}\int_{E\times S} \delta x(C)\cdot D_0\delta u(B) f_0 dPdu = \sum_{\Gamma\in P(B)} (\Gamma!)^{-1} \int_{E\times S} \Big[\Delta(\Gamma|C)(\Phi_0\cdot u) + \sum_{\{k\}\in\Gamma} \Delta(\Gamma|C \mid\setminus\mid k)(\phi_k\cdot u)\Big] f_0 dPdu. \tag{4.28}$$

The D_0 parts from (4.22) and (4.23) vanishes due to Lemma 6, which may be used since $0\in B\cup C$. We obtain form Lemma 6 that

$$\int_{E\times S} \delta x(C)\cdot D_0\delta u(B) f_0 dPdu = \int_{E\times S} \delta x(B)\cdot D_0\delta u(C) f_0 dPdu. \tag{4.29}$$

Lemma 7 now follows from (4.28) and (4.29).

Lemma 8: Let $m\in N^+$ and $m\geqslant 2$, then

$$\int_{E\times S} \Delta(N_m^+)(\phi_0\cdot u) f_0 dPdu = \sum_{\Gamma\in P_2(N_m^+)} (\Gamma!)^{-1} \int_{E\times S} \Big[\Delta(\Gamma \mid 0)(\Phi_0\cdot u) + \sum_{\{k\}\in\Gamma} \Delta(\Gamma \mid 0 \mid\setminus|k)(\phi_k\cdot u)\Big] f_0 dPdu. \tag{4.30}$$

Proof: Choose $B=\{0\}$ and $C = N_m^+$ in Lemma 7 and Lemma 8 easily follows.

Lemma 9: Let B and C be disjoint subsets of N with $n(B)\geqslant 2$, $n(C)\geqslant 2$, and $0\in B\cup C$, then

$$\sum_{\Gamma\in P_3(B\cup C,C)} A(\Gamma) = \sum_{\Gamma\in P_3(B\cup C,B)} A(\Gamma). \tag{4.31}$$

Proof: For $\Gamma\in P(B)$ we have $(\Gamma \mid C)! = (n(B)+n(c))! \times (n(b)!n(c)!)^{-1}\Gamma!$ and since $n(C)\geqslant 2$ we have $\Sigma_{\{k\}\in\Gamma} = \Sigma_{\{k\}\in(\Gamma\mid c)}$ and thus it follows from (4.26) and (3.3) that

$$\sum_{\Gamma\in P_2(B\cup C,C)} A(\Gamma) = \sum_{\Gamma\in P_2(B\cup C,B)} A(\Gamma) \tag{4.32}$$

and Lemma 9 follows.

Lemma 10: Let $m\in N$ and $m\geqslant 2$, then

$$\int_E \phi_0\cdot\delta J^{(m)}[\phi_1,\ldots,\phi_m]dP = \sum_{\Gamma\in P_3(N_m,\{0\})} A(\Gamma). \tag{4.33}$$

Proof: Rewrite the term corresponding to $\Gamma = \{N_m^+\}$ in Lemma 4 according to Lemma 8, then it is straightforward to derive Lemma 10.

Definition 7: $M(m,k) = \{\Gamma\in P_3^{m-k}(N_m) \mid 0\in B\in\Gamma \Rightarrow n(B)\leqslant k\}$.

Lemma 11: Let $k,m\in N^+$ and $2+2k\leqslant m$, then

$$\sum_{\Gamma\in M(m,k)} A(\Gamma) = \sum_{\Gamma\in M(m,k+1)} A(\Gamma). \tag{4.34}$$

Proof: Trivially (4.34) is equivalent to

$$\sum_{\Gamma\in I_1} A(\Gamma) = \sum_{\Gamma\in I_2} A(\Gamma), \tag{4.35}$$

where

$$I_1 = M(m,k+1)\setminus M(m,k),$$
$$I_2 = M(m,k)\setminus M(m,k+1). \tag{4.36}$$

We will prove that

$$I_1 = \bigcup_{C\in I_3} P_3(N_m,C),\quad I_2 = \bigcup_{D\in I_4} P_3(N_m,D), \tag{4.37}$$

where

$$I_3 = \{C\subset N_m \,|\, 0\in C \text{ and } n(C) = k+1\},$$
$$I_4 = \{D\subset N_m \,|\, 0\notin D \text{ and } n(D) = m-k\}. \tag{4.38}$$

In order to do this we define

$$M_1 = \{\Gamma \in P_3(N_m) \mid (1){:}\ |\Gamma| \geqslant k+1, (2){:}\ |\Gamma| \leqslant m-k-1,$$

$$(3){:} 0 \in C \in \Gamma \Rightarrow n(C) = k+1\}, \tag{4.39}$$

$$M_2 = \{\Gamma \in P_3(N_m) \mid (a){:}\ I_4 \cap \Gamma \neq \emptyset, (b){:}\ |\Gamma| = m-k, \tag{4.40}$$

$$(c){:}\ 0 \in C \in \Gamma \Rightarrow n(C) \neq k+1\}.$$

Now (4.37) easily follows from (4.41)–(4.44) below,

$$M_1 \subset I_1 \cap \Big[\bigcup_{C \in I_3} P_3(N_m, C) \Big], \tag{4.41}$$

$$M_2 \subset I_2 \cap \Big[\bigcup_{D \in I_4} P_3(N_m, D) \Big], \tag{4.42}$$

$$I_1 \cup \Big[\bigcup_{C \in I_3} P_3(N_m, C) \Big] \subset M_1, \tag{4.43}$$

$$I_2 \cup \Big[\bigcup_{D \in I_4} P_3(N_m, D) \Big] \subset M_2. \tag{4.44}$$

It is mainly a matter of checking to prove (4.41)–(4.44). We will just demonstrate (4.43) since it is rather similar to show (4.44) and (4.41)–(4.42) are trivial. It is easy to see that $I_1 \subset M_1$. For a $\Gamma \in P_3(N_m, C)$ where $C \in I_3$ property (1) and (3) in (4.39) follows directly. We prove that Γ also have property (2) by contradiction: $|\Gamma| > m-k-1 \Rightarrow \exists D \in \Gamma{:}\ n(D) > m-k-1$ and $0 \notin D$ (since $0 \in D \Rightarrow n(D) = k+1 > m-k-1 \Rightarrow m < 2k+2$ but we have assumed $m \geqslant 2k+2$) $\Rightarrow$we have $\Gamma = \{D\} \cup \{C\} \cup \Gamma_1$ and since $n(\Gamma) \geqslant 3$ we have $\Gamma_1 \neq \emptyset$ and $n(N_m) \geqslant n(D) + n(C) + 1 \Rightarrow m+1 > (m-k-1) + (k+1) + 1 \Rightarrow m+1 > m+1$.

It is easily proved that I_1 and I_2 are expressed as unions of pairwise disjoint sets by (4.37), thus

$$\sum_{\Gamma \in I_1} A(\Gamma) = \sum_{C \in I_3} \Big[\sum_{\Gamma \in P_3(N_m, C)} A(\Gamma) \Big], \tag{4.45}$$

$$\sum_{\Gamma \in I_2} A(\Gamma) = \sum_{D \in I_4} \Big[\sum_{\Gamma \in P_3(N_m, D)} A(\Gamma) \Big]. \tag{4.46}$$

The mapping $C \to D = N_m \setminus C$ defines a bijection $I_3 \to I_4$ and may be used as a variable substitution in (4.45). From Lemma 9 we obtain

$$C \in I_3 \text{ and } D = N_m \setminus C \Rightarrow \sum_{\Gamma \in P_3(N_m, C)} A(\Gamma) = \sum_{\Gamma \in P_3(N_m, D)} A(\Gamma). \tag{4.47}$$

By the variable substitution and (4.47) we transform the right-hand side in (4.45) to the expression on the right-hand side of (4.46), and now Lemma 11 is proved.

Proof of Result 1: We have $P_3(N_m, \{0\}) = M(m,1)$ and thus from Lemmas 10 and 11 we obtain

$$\int_E \phi_0 \cdot \delta J^{(m)}[\phi_1, \ldots, \phi_m]\, dP = \sum_{\Gamma \in M(m,l)} A(\Gamma),$$

$$m = 2l \text{ or } 2l+1, \quad l = 1, 2, \cdots. \tag{4.48}$$

The set $M(m,l)$ may be expressed as

$$M(m,l) = P_3^l(N_m), \quad m = 2l, \tag{4.49}$$

$$M(m,l) = P_3^l(N_m) \cup \Big[\bigcup_{D \in I_6} P_3(N_m, D) \Big], \quad m = 2l+1. \tag{4.50}$$

We define I_5 and I_6 as

$$I_5 = \{C \subset N_m \mid 0 \in C \text{ and } n(C) = l+1\},$$

$$I_6 = \{D \subset N_m^+ \mid n(D) = l+1\}. \tag{4.51}$$

We use the notation

$$A(I) = \sum_{\Gamma \in I} A(\Gamma) \tag{4.52}$$

and in a similar way as in the proof of Lemma 11 we obtain

$$A\Big(\bigcup_{C \in I_5} P_3(N_m, C) \Big) = A\Big(\bigcup_{D \in I_6} P_3(N_m, D) \Big). \tag{4.53}$$

From (4.53) we now get

$$A\Big(\bigcup_{D \in I_6} P_3(N_m, D) \Big) = \tfrac{1}{2} A(I_m) \tag{4.54}$$

since I_m, defined in (3.1), may be expressed as

$$I_m = \Big[\bigcup_{C \in I_5} P_3(N_m, C) \Big] \cup \Big[\bigcup_{D \in I_6} P_3(N_m, D) \Big]. \tag{4.55}$$

From (4.48), (4.49), (4.50), and (4.54) we obtain Result 1.

5. DISCUSSION

There is a formal relationship between our expressions (3.1)–(3.2) and the Mayer cluster expansion of the $m+1$ particle distribution function. In both cases we obtain a sum of terms each depending on a correlation index[11] [in this paper we have chosen to represent the set of correlation indices by $P(N_m)$], however in the Mayer cluster expansion all correlation indices in $P(N_m)$ appear, while in expression (3.1)–(3.2) for $A_m(\phi_0, \ldots, \phi_m)$ only some of them appear. That, in the Mayer expansion, all indices $0, 1, \ldots, m$ should be treated in a perfectly symmetric way is evident from the outset but it is a remarkable fact that this happens to be the case also in (3.1)–(3.2). Clearly the index 0 would be expected to play a role of its own and indeed it does but not as a summation index in (3.1)–(3.2). The different treatment of 0 as compared with 1, $2, \ldots, m$ enters in the boundary conditions (3.7)–(3.8). This peculiar quasisymmetry in 0, $1, \ldots, m$ becomes a "true" symmetry, if the perturbation asymptotically leaves particle orbits unperturbed (Corollary 1). In other words, if we neglect particles which are resonant with the perturbation this symmetry is obtained.

Corollary 1 is a general form of an important class of symmetry relations among plasma response functions. The simplest example is the anti-Hermitian property of the *principal part* (thus neglecting resonant wave-particle interaction) of the linear conductivity tensor. Another important example is the well-known symmetry of the principal part of the three wave coupling coefficients implying conservation of wave energy and momentum in the coherent interaction of three waves and also in the weak turbulence equations. Recently corresponding three-wave results were derived for inhomogenous relativistic plasmas.[12] For a homogeneous magnetized plasma a more detailed discussion of these symmetries will be given in part II of this paper and Ref. 1 where the expressions for the response tensors of arbitrary order are obtained from Result 1. From Corollary 3 we observe that not only the operator A_m possesses this symmetry but also each part A_j has it. The symmetries of the coupling coefficients are often used as an indication that no mistake has been made in their normally very lengthy algebraic derivations. For 4-wave and higher order interactions a possible mistake, which will not be discovered by symmetry violations, is the omission of some A_j term in A_m.

In the Introduction it was said that the results of this paper would probably be much more difficult to obtain by straightforward iterations in the Vlasov equation. This is simply due to the fact that while the quantities $\delta x(B)$ and $\delta u(B)$ naturally appear in the method used in this paper, they do not in the other approach. One may object that there may very well be other equally useful formulas for the operator $\delta J^{(n)}$ and that we obtained the particular one in Result 1 is due to the method of calculation used. An indication that this is not the case is the result for the three-wave coupling coefficient in Ref. 13. The structure of the expression obtained is in good agreement with the results in this paper and the iterative method was used.

[1]J. Larsson, "Conductivity tensors of all orders in a collisionless plasma," to be published in J. Plasma Phys. (1979).
[2]J. Larsson and L. Stenflo, Beitr. Plasma Phys. **16**, 79 (1976).
[3]A.N. Kaufman and L. Stenflo, Plasma Phys. **17**, 403 (1975).
[4]C.W. Misner, K.S. Thorn, and J.A. Wheeler, *Gravitation* (Freeman, San Fransisco, 1973).
[5]H. Weyl, *Space Time Matter*, translated by H.L. Brose (Dover, New York, 1952).
[6]F.R. Nevanlinna, *Absolute Analysis*, translated by P. Emig (Springer-Verlag, Berlin, 1967).
[7]S. Mac Lane and G. Birkhoff, *Algebra* (Macmillan, New York, 1967).
[8]H. Flanders, *Differential Forms with Applications to the Physical Sciences* (Academic, New York, 1963).
[9]G.F. Simmons, *Introduction to Topology and Modern Analysis* (McGraw-Hill, New York, 1963).
[10]S. Ichimaru, *Basic Principles of Plasma Physics* (Benjamin, London, 1973).
[11]R. Balescu, *Equilibrium and Nonequilibrium Statistical Mechanics* (Wiley, New York, 1975).
[12]S. Johnston and A.N. Kaufman, Nobel foundation symposium 36, *Plasma Physics, Nonlinear Theory and Experiments*, edited by H. Wilhelmsoon (Plenum, New York, 1977).
[13]J. Larsson, J. Plasma Phys. **14**, 467 (1975).

Current responses of all orders in a collisionless plasma. II. Homogeneous plasma

J. Larsson

Department of Plasma Physics, Umeå University, S-901 87 Umeå, Sweden

(Received 7 July 1978)

Expressions for the admittance tensors of all orders are obtained for a relativistic magnetized Vlasov–Maxwell plasma. Symmetries, from which the Manley–Rowe relations follow, are explicit in the expressions obtained. The treatment is covariant.

1. INTRODUCTION

In the first part of this paper,[1] which we denoted by "I," only the most general background plasma was considered. In this Part II we apply the general formulas in I to the particular case of a homogeneous and stationary background plasma. We then obtain covariant expressions for the admittance tensors of all orders with the important symmetries related to the Manley–Rowe relations explicitly exhibited. In the particular case of the nonrelativistic second order response we compare and get agreement with a result in Ref. 2.

We will use the notation in I. A slight extension of the formalism is needed since we want to make the calculations in Fourier space and coordinate-free. We thus introduce the complexification $V^+ = V + iV$ of the real vector space V and the admittance tensor of order m is an element in the tensor space $(V^+)^{\otimes(m+1)}$.

2. THE ADMITTANCE TENSORS OF ALL ORDERS

We consider a homogeneous stationary background plasma and thus $f_0(P,u) = f_0(u)$ (i.e., f_0 is independent of $P \in E$) and $\nabla_E \wedge \Phi_0$ is independent of $P \in E$. We choose an event O as origin and this gives rise to a vector space structure on E, since each $P \in E$ now determines the vector $P - O$ in V. We will use the somewhat sloppy notation $x \in E$ where $x \in V$ and the event we have in mind is $x + O \in E$.

The Fourier transform of a function $G(x)$ where $x \in E$ is

$$\tilde{G}(\kappa) = \int_E G(x) \exp(-i\kappa\cdot x)\, dx, \tag{2.1}$$

where $\kappa \in V$ or sometimes $\kappa \in V^+ = V + iV$. (In a Lorentz frame we recognize $\kappa = \omega c^{-1} e_0 + k_1 e_1 + k_2 e_2 + k_3 e_3$.)

The inverse transform is

$$G(x) = (2\pi)^{-4} \int_V \tilde{G}(\kappa) \exp(i\kappa\cdot x)\, d\kappa. \tag{2.2}$$

Remark 1: In Sec. 2 of I it was shown how to define a natural measure on E and V so the integrals (2.1) and (2.2) are well defined.

Remark 2: Depending on the properties of $G(x)$ it may be advantageous to allow κ to take values in V^+. When $G(x)$ is a perturbation we have the boundary condition $G(x) \to 0$ when $x \to$ the infinite past. If we take $\kappa \in V^+$ and $\mathrm{Im}\kappa$ (imaginary part of κ) in the future direction, then $|\exp(-i\kappa\cdot x)| = \exp(\mathrm{Im}\kappa\cdot x) \to 0$ exponentially towards the future and this might improve the convergence of (2.1). The validity of (2.2) may then be increased if we integrate over a plane $V + i\,\mathrm{Im}\kappa$ in V^+.

Remark 3: In the particular case when $G(x)$ vanishes in a half-space in the past we take $\mathrm{Im}\kappa$ perpendicular to the spacelike plane defining the half-space and the origin O on this plane. Then (2.1) and (2.2) is the Fourier–Laplace transform and its inversion. This is easily seen if we take a Lorentz frame $(0, e_0, e_1, e_2, e_3)$ with e_0 parallel to $\mathrm{Im}\kappa$:

$$\kappa = \omega c^{-1} e_0 + \mathbf{k}, \quad x = cte_0 + \mathbf{x}, \quad \mathrm{Im}\kappa = c^{-1}\gamma e_0,$$

$$\tilde{G}(\omega,\mathbf{k}) = \int_0^\infty c\, dt \int_{R^3} G(t,\mathbf{x}) e^{-i\mathbf{k}\cdot\mathbf{x}} e^{i\omega t}\, d\mathbf{x}, \tag{2.3}$$

$$G(t,\mathbf{x}) = (2\pi)^{-4} \int_{R+i\gamma} c^{-1}\, d\omega \int_{R^3} G(\omega,\mathbf{k}) e^{i\mathbf{k}\cdot\mathbf{x}} e^{-i\omega t}\, d\mathbf{k}. \tag{2.4}$$

Definition 1: The admittance tensors $\Lambda^{(m)}_{\kappa_1,\ldots,\kappa_m} \in (V^+)^{\otimes(m+1)}$, where $m = 1,2,\cdots$ and $V^+ = V + iV$, are defined by [we take $\phi \in L_0(E,V)$ and $\tilde{\phi}_\kappa = \int_E \phi(x) \times \exp(-i\kappa\cdot x)\, dx$]:

$$\text{(i)}\quad (2\pi)^{-4m} \int \Lambda^{(m)}_{\kappa_1,\ldots,\kappa_m} \binom{m}{\cdot} \tilde{\phi}_{\kappa_1} \otimes \cdots \otimes \tilde{\phi}_{\kappa_m}$$

$$\times \exp[i(\kappa_1 + \cdots + \kappa_m)\cdot x]\, d\kappa_1 \cdots d\kappa_m$$

$$= \delta J^{(m)}[\phi](x), \tag{2.5}$$

$$\text{(ii)}\quad \Lambda^{(m)}_{\kappa_{\pi(1)},\ldots,\kappa_{\pi(m)}} \binom{m}{\cdot} a_{\pi(1)} \otimes \cdots \otimes a_{\pi(m)}$$

$$= \Lambda^{(m)}_{\kappa_1,\ldots,\kappa_m} \binom{m}{\cdot} a_1 \otimes \cdots \otimes a_m \tag{2.6}$$

for arbitrary vectors $a_1,\ldots,a_m \in V^+$ and permutations π of $\{1,\ldots,m\}$.

Remark 4: From (2.5) we obtain

$$\delta J_\kappa = \sum_{m=1}^{\infty} (2\pi)^{-4(m-1)} \int \Lambda^{(m)}_{\kappa_1,\ldots,\kappa_m} \binom{m}{\cdot} \tilde{\phi}_{\kappa_1} \otimes \cdots \otimes \tilde{\phi}_{\kappa_m}$$

 0022-2488/79/071331-05$01.00

$$\times\delta(\kappa-\kappa_1-\cdots-\kappa_m)d\kappa_1\cdots d\kappa_m, \tag{2.7}$$

which combined with the transformed Maxwells equations is a suitable starting point for the study of weak interaction of modes or quasimodes.

Remark 5: To get numbers from our formulas we have to leave the coordinate-free formalism and express our quantities as coordinates. In a given Lorentz system $(0,e_0,e_1,e_2,e_3)$ we express the mth order conductivity tensor as complex numbers $\Lambda(\kappa_1,...,\kappa_m)^{i_0,...,i_m}$, where $i_j=0,1,2$, or 3, and the transformed 4-potential and κ_ν as numbers $\phi(\kappa)^i$ and $\kappa(\nu)^i$, $i=0,1,2,3$. These numbers are coefficients for our abstract tensor space quantities so that

$$\Lambda^{(m)}_{\kappa_1,...,\kappa_m}=\Lambda(\kappa_1,...,\kappa_m)^{i_0,...,i_m}e_{i_0}\otimes\cdots\otimes e_{i_m}, \tag{2.8}$$

$$\phi(\kappa)=\phi(\kappa)^ie_i,\quad \kappa_\nu=\kappa(\nu)^ie_i, \tag{2.9}$$

where we use the Einstein summation convention. In index calculus we have rules for raising and lowering indices. If we lower the index i_ν on Λ in (2.8) and raise it on e $(e_{i_\nu}\rightarrow e^{i_\nu})$ the equality in (2.8) is still valid. The vectors e^i are defined from e_j by $e^i\cdot e_j=\delta_{i,j}$. Since we are working in a Lorentz system, $e_0=-e^0$ and $e_i=e^i$ for $i=1,2$, and 3. Raising or lowering index 0 on a number thus results in multiplication with -1, raising or lowering 1, 2, or 3 leave the number unchanged.

Remark 6: In (2.3) a factor c appears and in (2.4) a factor c^{-1}. In the usual definition of the Laplace–Fourier transform these factors do not appear and in order to avoid them we could have replaced dx and $d\kappa$ in (2.1) and (2.2) with $c^{-1}dx$ and $cd\kappa$. It is easy to see that $\Lambda^{(m)}_{\kappa_1,...,\kappa_m}$ is not changed if we make this replacement everywhere above where dx or $d\kappa$ appears.

Result 1: Let m be a positive integer and $\kappa_0,...,\kappa_m\in V$, where $\kappa_0+\cdots+\kappa_m=0$. Then for arbitrary vectors $\tilde\phi_1,...,\tilde\phi_m\in V^+$ we have

$$\Lambda^{(m)}_{\kappa_1,...,\kappa_m}=\begin{cases}\frac12\Lambda^{(0|1)}_{\kappa_1}, & \text{if } m=1, \quad (2.10)\\ \sum_{\Gamma\in P_3^l(N_m)}\Lambda^{\Gamma}_{\kappa_1,...,\kappa_m}, & \text{if } m=2l,\quad l=1,2,..., \quad (2.11)\\ \frac12\sum_{\Gamma\in I_m}\Lambda^{\Gamma}_{\kappa_1,...,\kappa_m}+\sum_{\Gamma\in P_3^l(N_m)}\Lambda^{\Gamma}_{\kappa_1,...,\kappa_m}, & \text{if } m=2l+1,\quad l=1,2,\cdots, \quad (2.12)\end{cases}$$

where $I_m=P_3^{l+1}(N_m)\setminus P_3^l(N_m)$ and $\Lambda^{\Gamma}_{\kappa_1,...,\kappa_m}$ is defined from

$$\tilde\phi_0\cdot\Lambda^{\Gamma}_{\kappa_1,...,\kappa_m}\binom{m}{\cdot}\tilde\phi_1\otimes\cdots\otimes\tilde\phi_m=c(m+1)(\Gamma!)^{-1}\sum_\sigma m_0\sum_{\{k\}\in\Gamma}\sum_{B\in(\Gamma|\setminus|k)}\int_S f_0(u)\Big[\prod_{C\in(\Gamma|\setminus|k|B)}i\kappa_k\cdot\delta\tilde x(C)\Big]$$

$$\times[qm_0^{-1}\delta\tilde u(B)\cdot\nabla_E\wedge\Phi_0\cdot\delta\tilde x(\{k\})-c^2\delta\tilde u(B)\cdot\delta\tilde u(\{k\})]du. \tag{2.13}$$

The quantities $\delta\tilde x(B)$ and $\delta\tilde u(B)$ depend linearly on each $\tilde\phi_i$ for $i\in B$ and are determined by the hierachy of equations

$$D[\kappa_B]\delta\tilde x(B)=\delta\tilde u(B), \tag{2.14}$$

$$(D[\kappa_B]-qm_0^{-1}c^{-2}\nabla_E\wedge\Phi_0\cdot)\delta\tilde u(B)=qm_0^{-1}c^{-2}\sum_{\Gamma\in P(B)}(\Gamma!)^{-1}\sum_{\{k\}\in\Gamma}\Big[\prod_{C\in(\Gamma|\setminus|k)}i\kappa_k\cdot\delta\tilde x(C)\Big]$$

$$\times\Big\{i\kappa_k\wedge\tilde\phi_k\cdot u+\sum_{C\in(\Gamma|\setminus|k)}[\kappa_k\cdot\delta\tilde x(C)]^{-1}\kappa_k\wedge\tilde\phi_k\cdot\delta\tilde u(C)\Big\}, \tag{2.15}$$

where $\kappa_B=\Sigma_{i\in B}\kappa_i$ and

$$D[\kappa_B]=iu\cdot\kappa_B+qm_0^{-1}c^{-2}(\nabla_E\wedge\Phi_0\cdot u)\cdot\nabla_S. \tag{2.16}$$

Poles appearing in (2.13) shall be treated in accordance with the Landau prescription (see Remark 9).

Remark 7: In a Lorentz frame $(O,e_0,e_1,e_2,e_3,)$, chosen such that the external field is purely magnetic and $\mathbf{B}_0=B_0e_3$, we obtain

$$D[\kappa]=u^0c^{-1}\Big(-i\omega+i\mathbf{k}\cdot\mathbf{v}-\omega_c\frac{\partial}{\partial\phi}\Big), \tag{2.17}$$

$$D[\kappa]-qm_0^{-1}c^{-2}\nabla_E\wedge\Phi_0\cdot=u^0c^{-1}\Big(-i\omega+i\mathbf{k}\cdot\mathbf{v}-\omega_c\frac{\partial}{\partial\phi}-\omega_ce_1\wedge e_2\cdot\Big). \tag{2.18}$$

Here $\mathbf{v}$ is the ordinary velocity $(u=u^0e_0+c^{-1}\mathbf{v})$ and ϕ is defined from $\mathbf{v}=v_\perp\cos\phi e_1+v_\perp\sin\phi e_2+v_ze_3$. In (2.17) and (2.18) we recognize the relativistic versions of g_κ^{-1} and h_κ^{-1} appearing in Ref. 2. We observe that $-B_0e_1\wedge e_2$ is a four-dimensional analog of $\mathbf{B}_0\times$.

Remark 8: The hierachy of equations (2.14) and (2.15) defines $\delta\tilde x(B)$ and $\delta\tilde u(B)$ recursively. First calculate $\delta\tilde x(B)$ and $\delta\tilde u(B)$ for $n(B)=1$; then for $n(B)=2$ and so on. It is easy to see that $\delta\tilde x(B)$ and $\delta\tilde u(B)$ will depend linearly on each $\tilde\phi_i$ for $i\in B$.

Remark 9: It follows directly from (2.13) substituted in (2.10)–(2.12) that

$$\tilde{\phi}_0 \cdot PA^{(m)}_{\kappa_1,\ldots,\kappa_m}\binom{m}{\cdot}\tilde{\phi}_1 \otimes \cdots \otimes \tilde{\phi}_m = \tilde{\phi}_{\pi(0)} \cdot PA^{(m)}_{\kappa_{\pi(1)},\ldots,\kappa_{\pi(m)}}\binom{m}{\cdot}\tilde{\phi}_{\pi(1)} \otimes \cdots \otimes \tilde{\phi}_{\pi(m)}, \tag{2.19}$$

where π is an arbitrary permutation of $\{0,\ldots,m\}$ and P indicates that only the principal part of the integral in (2.13) is included. The symmetry (2.19) is obvious, due to the perfectly symmetric treatment of the indices $0,1,\ldots,m$ in Result 1, apart from the Landau prescription which of course has no effect on the principal parts appearing in (2.19). The pole contributions are correctly obtained if we, while performing the velocity space integration in (2.13), take $\kappa_i \in V^+$ with $\operatorname{Im}\kappa_i$ directed towards the future for $i = 1,2,\ldots,m$ and κ_0 still such that $\kappa_0 + \cdots + \kappa_m = 0$. Clearly $\operatorname{Im}\kappa_0$ is directed towards the past, which introduces an asymmetry. If pole contributions may be neglected, the symmetry relations (2.19) imply conservation of wave energy and momentum and the Manley–Rowe relations.[3]

Result 2: If we in Result 1(I) take $\phi_j(x) = \tilde{\phi}_j \exp(i\kappa_j \cdot x)$, where $j \in B$ and $\tilde{\phi}_j \in V^+$, then

$$\delta\tilde{x}(B) = \delta x(B)\exp(-i\kappa_B \cdot x), \tag{2.20}$$

$$\delta\tilde{u}(B) = \delta u(B)\exp(-i\kappa_B \cdot x), \tag{2.21}$$

where $\kappa_B = \Sigma_{j\in B}\kappa_j$ and $\delta\tilde{x}(B)$ and $\delta\tilde{u}(B)$ are defined by (2.14) and (2.15).

Result 3: If $\phi_j(x) = \tilde{\phi}_j \exp(i\kappa_j \cdot x)$, $j \in N_m$ and $\tilde{\phi}_j \in V^+$, then

$$A(\Gamma) = (2\pi)^4\delta(\kappa_0 + \cdots + \kappa_m)\tilde{\phi}_0 \cdot A^{\Gamma}_{\kappa_1,\ldots,\kappa_m}\binom{m}{\cdot}\tilde{\phi}_1 \otimes \cdots \otimes \tilde{\phi}_m. \tag{2.22}$$

Remark 10: In Result 3 κ_0 is a free variable and not defined as $\kappa_0 = -\kappa_1 - \cdots - \kappa_m$. However, when $\kappa_0 \neq -\kappa_1 - \cdots - \kappa_m$ both sides of (2.22) vanish.

Corollary 1:

$$\tilde{\phi}_0 \cdot A^{(2)}_{\kappa_1,\kappa_2} : \tilde{\phi}_1 \otimes \tilde{\phi}_2 = \frac{im_0 c}{2}\int_S f_0(u)\Big(c^2\kappa_0 \cdot \delta\tilde{x}(0)\,\delta\tilde{u}(1)\cdot\delta\tilde{u}(2) + c^2\kappa_1 \cdot \delta\tilde{x}(1)\delta\tilde{u}(0)\cdot\delta\tilde{u}(2) + c^2\kappa_2 \cdot \delta\tilde{x}(2)\,\delta\tilde{u}(0)\cdot\delta\tilde{u}(1)$$
$$+ \frac{q}{m_0}\langle\kappa_2 \wedge \nabla_E \wedge \Phi_0, \delta\tilde{x}(0) \wedge \delta\tilde{u}(1) \wedge \delta\tilde{x}(2)\rangle + \frac{q}{m_0}\langle\kappa_1 \wedge \nabla_E \wedge \Phi_0, \delta\tilde{x}(0) \wedge \delta\tilde{u}(2) \wedge \delta\tilde{x}(1)\rangle\Big)\,du, \tag{2.23}$$

where $\delta\tilde{x}(j) = \delta\tilde{x}(\{j\})$, etc.

Remark 11: Corollary 1 is a relativistic generalization of (6d) in Ref. 2. Note that $\delta\tilde{x}(i)$ corresponds to $m_0^{-1}g_{\kappa_i}[h_{\kappa_i}(\mathbf{F}_{\kappa_i})]$ and $c\delta\tilde{u}(i)$ to $m_0^{-1}h_{\kappa_i}(\mathbf{F}_{\kappa_i})$ in that paper.

3. DERIVATION OF THE RESULTS IN SEC. 2

We take f_0 and $\nabla_E \wedge \Phi_0$ independent of $P \in E$ and $\phi_0,\ldots,\phi_m$ as in Result 1 of I. In Lemma 2 the accordingly simplified version of Result 1 (I) is given. Result 1 (of this paper) is just the κ-space version of Lemma 2.

Definition 2: $\Delta_E(\Gamma) = \prod_{B\in\Gamma}\delta x(B)\cdot\nabla_E$ in formal notation analogous with (2.19) in I. (3.1)

Lemma 1:

$$\Delta(\Gamma|\setminus|k)(u\cdot\phi_k) = D_0[\Delta_E(\Gamma|\setminus|k)D_0^{-1}(u\cdot\phi_k)] + \sum_{B\in(\Gamma|\setminus|k)}\delta u(B)\cdot\Delta_E(\Gamma|\setminus|k|B)\{\phi_k - D_0^{-1}[\nabla_E(u\cdot\phi_k)]\} \tag{3.2}$$

and

$$\phi_k - D_0^{-1}\nabla_E(u\cdot\phi_k) = -m_0c^2q^{-1}\delta u(k) + \nabla_E \wedge \Phi_0 \cdot \delta x(k). \tag{3.3}$$

Proof:

$$\Delta(\Gamma|\setminus|k)(u\cdot\phi_k) = \Delta_E(\Gamma|\setminus|k)(u\cdot\phi_k) + \sum_{B\in(\Gamma|\setminus|k)}\delta u(B)\cdot\Delta_E(\Gamma|\setminus|k|B)\phi_k, \tag{3.4}$$

$$D_0[\Delta_E(\Gamma|\setminus|k)D_0^{-1}(u\cdot\phi_k)] = \sum_{B\in(\Gamma|\setminus|k)}\delta u(B)\cdot\Delta_E(\Gamma|\setminus|k|B)\nabla_E D_0^{-1}(u\cdot\phi_k) + \Delta_E(\Gamma|\setminus|k)(u\cdot\phi_k). \tag{3.5}$$

Substitution of (3.4) and (3.5) in (3.2) and making use of $\nabla_E D_0^{-1} = D_0^{-1}\nabla_E$, which is valid due to homogeneity, proves (3.2). Apply the operator D_0 to both sides of (3.3) and we obtain (3.6) of I for $B = \{k\}$ if we use the fact that $\nabla_E \wedge \Phi_0$ is constant.

Lemma 2: Take $\phi_1,\ldots\phi_m\in L_0(E,V)$ for sufficiently nice $\phi_0\in L^0(E,V)$; we have

$$\int_E \phi_0(P)\cdot\delta J^{(m)}[\phi_1,\ldots,\phi_m](P)\,dP=\begin{cases}\tfrac{1}{2}A(0|1), & \text{for } m=1, \quad (3.6)\\ \sum_{\Gamma\in P_3^l(N_m)} A(\Gamma), & \text{for } m=2l,\quad l=1,2,\cdots, \quad (3.7)\\ \tfrac{1}{2}\sum_{\Gamma\in I_m} A(\Gamma)+\sum_{\Gamma\in P_3^l(N_m)} A(\Gamma), & \text{for } m=2l+1,\quad l=1,2,\cdots, \quad (3.8)\end{cases}$$

where $I_m=P_3^{l+1}(N_m)\setminus P_3^l(N_m)$. Here

$$A(\Gamma)=\sum_\sigma m_0c(m+1)(\Gamma!)^{-1}\sum_{\{k\}\in\Gamma}\sum_{B\in(\Gamma|\setminus|k)}\int_{E\times S} f_0(u)\,\delta u(B)\cdot\Delta_E(\Gamma|\setminus|k|B)[m_0^{-1}q\nabla_E\wedge\Phi_0\cdot\delta x(k)-c^2\delta u(k)]\,dP\,du, \tag{3.9}$$

where

$$D_0\delta x(B)=\delta u(B), \tag{3.10}$$

$$D_0\delta u(B)-qm_0^{-1}c^{-2}\nabla_E\wedge\Phi_0\cdot\delta u(B)=qm_0^{-1}c^{-2}\sum_{\Gamma\in P(B)}(\Gamma!)^{-1}\sum_{\{k\}\in\Gamma}\Delta(\Gamma|\setminus|k)(\nabla_E\wedge\phi_k\cdot u), \tag{3.11}$$

with boundary conditions

$$\delta x(B),\ \delta u(B)\to 0 \text{ towards the past if } 0\notin B, \tag{3.12}$$

$$\delta x(B),\ \delta u(B)\to 0 \text{ towards the future if } 0\in B. \tag{3.13}$$

Proof: We first give an extension of Result 1(I) in order to also cover the linear ($m=1$) case. By combining Lemma 4(I) for $m=1$ with Lemma 7(I) for $B=\{0\}$ and $C=\{1\}$ we obtain

$$\int_E \phi_0(P)\cdot\delta J^{(1)}[\phi_1](P)\,dP=(qc/2)\int_{E\times S} f_0(P,u)\{\Delta(1)[u\cdot\phi_0(P)]+\Delta(0)u\cdot\phi_1(P)\}\,dP\,du=\tfrac{1}{2}A(0|1), \tag{3.14}$$

where the last equality is a definition. Expression (3.14) is valid in the general inhomogeneous and nonstationary situation. Note that formula (I.3.3) [(3.3) of I] gives $A(\Gamma)$ only for $n(\Gamma)\geqslant 3$. However, in the homogeneous stationary case we obtain a single formula (3.9) valid for $n(\Gamma)\geqslant 2$.

From Result 1(I) and (3.14) we now derive Lemma 2. It is easy to see that (3.10) and (3.11) follows from (I.3.5) and (I.3.6) and the homogeneity in space–time. From (I.3.3) and (3.14) and due to homogeneity,

$$\Delta(\Gamma)[u\cdot\Phi_0(P)]=0 \quad \text{for} \quad n(\Gamma)\geqslant 3,$$

we obtain

$$A(\Gamma)=qc(m+1)(\Gamma!)^{-1}\sum_{\{k\}\in\Gamma}\int_{E\times S} f_0(u)\Delta(\Gamma|\setminus|k)[\phi_k(P)\cdot u]\,dP\,du. \tag{3.15}$$

Now (3.9) follows from substitution of Lemma 1 in (3.15) and an application of Lemma 6 of I. This finishes the proof of Lemma 2.

Proof of Result 2: $\delta x(B)$ and $\delta u(B)$ are determined from (3.10) and (3.11). Substitute $\phi_j(x)=\tilde{\phi}_j\exp(i\kappa_j\cdot x)$ in these equations. It easily follows [by induction on $n(B)$, first take $n(B)=1$ in (3.10) and (3.11), then $n(B)=2$, and so on] that $\delta x(B)$ and $\delta u(B)$ vary in space–time as $\exp(i\kappa_B\cdot x)$. Thus $D_0\delta x(B)=D[\kappa_B]\delta x(B)$ and $D_0\delta u(B)=D[\kappa_B]\delta u(B)$ and it is easy to show that $\delta x(B)\exp(-i\kappa_B\cdot x)$ and $\delta u(B)\exp(-i\kappa_B\cdot x)$ satisfy (2.14) and (2.15) and Result 2 follows.

Proof of Result 3: Substitute $\phi_j(x)=\tilde{\phi}_j\exp(i\mathbf{k}_j\cdot x)$ in (3.9) and make use of Result 2, then

$$A(\Gamma)=m_0c(m+1)(\Gamma!)^{-1}\sum_{\{k\}\in\Gamma}\sum_{B\in(\Gamma|\setminus|k)}\int_{E\times S} f_0(u)\exp[i(\kappa_0+\cdots+\kappa_m)\cdot x]\Big[\prod_{C\in(\Gamma|\setminus|k|B)} i\kappa_k\cdot\delta\tilde{x}(C)\Big]$$

$$\times(qm_0^{-1}\delta\tilde{u}\cdot(B)\nabla_E\wedge\Phi_0\cdot\delta\tilde{x}(k)-c^2\delta\tilde{u}(B)\cdot\delta\tilde{x}(k)\,dx\,du. \tag{3.16}$$

We now use

$$\int_E \exp[i(\kappa_0+\cdots+\kappa_m)\cdot x]\,dx=(2\pi)^4\delta(\kappa_0+\cdots+\kappa_m) \tag{3.17}$$

in (3.16) and perform the x-integration. Comparison with the definition of A^Γ in (2.13) now gives Result 3.

Proof of Result 1: We prove that $A^{(m)}_{\kappa_1,\ldots,\kappa_m}$ as given in Result 1 have the properties in Definition 1. Here only (2.5) is nontrivial. We will prove that for $\phi_0\in L^0(E,V)$ and $\phi_j\in L_0(E,V)$, $j=1,\ldots,m$ that

$$\int_E \phi_0(x)\cdot\delta J^{(m)}[\phi_1,\ldots,\phi_m](x)\,dx = (2\pi)^{-4m}\int \phi_0(x)\cdot A^{(m)}_{\kappa_1,\ldots,\kappa_m}\binom{m}{\cdot}\tilde\phi_1(\kappa_1)\otimes\cdots\otimes\tilde\phi_m(\kappa_m)\exp[i(\kappa_1+\cdots+\kappa_m)\cdot x]\,d\kappa_1\cdots d\kappa_m\,dx, \tag{3.18}$$

which implies (2.5), since (3.18) is true for all $\phi_0\in L^0(E,V)$. From (2.10)–(2.12) and (3.6)–(3.8) we see that it is sufficient to show

$$A(\Gamma) = (2\pi)^{-4m}\int \phi_0(x)\cdot A^{\Gamma}_{\kappa_1,\ldots,\kappa_m}\binom{m}{\cdot}\tilde\phi_1(\kappa_1)\otimes\cdots\otimes\tilde\phi_m(\kappa_m)\exp[i(\kappa_1+\cdots+\kappa_m)\cdot x]\,d\kappa_1\cdots d\kappa_m\,dx, \tag{3.19}$$

for arbitrary $\Gamma\in P_2(N_m)$. Substitution first of $\phi_0(x) = (2\pi)^{-4}\int\tilde\phi_0(\kappa_0)\exp(i\kappa_0\cdot x)\,d\kappa_0$ and then of (3.17) in (3.19) yields

$$A(\Gamma) = (2\pi)^{-4(m+1)}\int (2\pi)^4\delta(\kappa_0+\cdots+\kappa_m)\tilde\phi_0(\kappa_0)A^{\Gamma}_{\kappa_1,\ldots,\kappa_m}\binom{m}{\cdot}\tilde\phi_1(\kappa_1)\otimes\cdots\otimes\tilde\phi_m(\kappa_m)\,d\kappa_0\cdots d\kappa_m \tag{3.20}$$

and since (3.20) easily follows from Result 3, which we have already proved, this finishes the proof of Result 1.

Proof of Corollary 1: From Result 1 we obtain

$$\tilde\phi_0\cdot A^{(2)}_{\kappa_1,\kappa_2}:\tilde\phi_1\otimes\tilde\phi_2 = \frac{m_0 c}{2}\sum_{\substack{\alpha,\beta,\gamma=0\\ \alpha\neq\beta\neq\gamma\neq\alpha}}^{2}\int f_0(u)i\kappa_\alpha\cdot\delta\tilde x(\gamma)\Big(\frac{q}{m_0}\delta\tilde u(\beta)\cdot\nabla_E\wedge\Phi_0\cdot\delta\tilde x(\alpha) - c^2\delta\tilde u(\beta)\cdot\delta\tilde u(\alpha)\Big)\,du. \tag{3.21}$$

It is very easy to compare the terms not containing Φ_0 in (2.23) and (3.21), we only need to use

$$\kappa_\alpha+\kappa_\beta+\kappa_\gamma = 0. \tag{3.22}$$

It is rather tedious to compare the Φ_0-terms in these expressions. Note that the Φ_0-terms in (2.23) contain no $\delta u(0)$, in (3.22) we substitute $D[\kappa_0]\delta x(0)$ for $\delta u(0)$ and integrate partially {i.e., we use Lemma 6(I) in κ-space; $\int_{E\times S} f_0(u)D[\kappa]h(u)\,du = 0$} and so we may avoid $\delta u(0)$ also in (3.22). It is now straightforward but tedious to compare the Φ_0-terms in (2.23) and (3.22) so we omit these details.

4. CONCLUSIONS

Expressions for response tensors in magnetized kinetic plasmas have been given in this paper, which give new insight into their structure and symmetry properties. We refer to the discussion in Sec. 5 of I in which some statements, in view of this Part II, now are seen more explicitly.

Very little was previously known concerning third and higher order response tensors, so essentially everything is new for $m\geqslant 3$ in this paper. The case $m = 2$ has received much attention since it concerns the lowest order nonlinear effects (i.e., typically three-wave interaction). Expressions exhibiting the important symmetries (7.19) have been derived in this case by different methods.[4,5]

A most important aspect of the response tensor formulas in Result 1 is of course whether they can be useful in actual numerical calculations. The need for such formulas is evident from the present literature on nonlinear plasma theory. Typically nonlinear investigations start from the basic equations, in our case the Vlasov–Maxwell equations, and lengthy derivations are needed to obtain for example a three-wave coupling coefficient valid for three particular normal modes. The situation is quite different for linear investigations; in this case we do not start from the Vlasov–Maxwell equations but instead from the existing standard expression for the linear response,[6] and this saves us from much tedious work and many possibilities of making mistakes. Actually there exists a corresponding expression for the second-order response tensor,[7] which is thus much better to begin with than the basic equations, when we consider quadratic processes in a plasma. This expression deserves particular attention since at this time there is no alternative formula of comparable simplicity. It is thus most promising that this formula was derived from a particular case of Result 1 in this paper, since then it might be possible to derive corresponding results to all orders and relativistically from Result 1. Indeed this turns out to be possible. The algebra is, however, simplified considerably if we do not insist on writing covariant formulas but instead choose a Lorentz frame in which the constant external field is purely magnetic. These results will be presented in a separate paper without recourse to the coordinate-free formalism used in this paper.[1]

The homogeneous plasma is just the simplest application of Result 1(I). It remains to be seen how useful this result will be in cases where the unperturbed plasma has some particular geometry and/or is time dependent. It may for example be toroidal and turbulent. At least the potential area of applications of Result 1(I) is extremely large; mathematically it is a general alternative starting point to perturbation problems for a Vlasov–Maxwell plasma.

In some cases the Vlasov equation is unnecessarily advanced and we may use for example the collisionless two-fluid model with scalar pressure; it is almost evident that formulas for the response tensors, which are closely related to those in this paper, in such a plasma can be derived.

[1]J. Larsson, "Current responses of all orders in a collisionless plasma. I. General Theory," J. Math. Phys. **20**, 1321 (1979).
[2]J. Larsson, J. Plasma Phys. **14**, 467 (1975). The result in this paper was rederived from a more general formula by S. Johnston and A.N. Kaufman, in Ref. 12 in part I.
[3]A.N. Kaufman and L. Stenflo, Plasma Phys. **17**, 403 (1975).
[4]V.V. Pustovalov and V.P. Silin, *Theory of Plasmas*, edited by D.V. Skobel'tsyn (Consultants Bureau, New York, 1975). A symmetric expression was given on p. 191. Another symmetric result was obtained in Ref. 2 above and Ref. 12 in part I.
[5]H. Kim "Lagrangian Description of Warm Plasmas," Stanford University, SUIPR Report No. 470 (1972). A symmetric expression was given on p. 57.
[6]S. Ichimaru, *Basic Principles of Plasma Physics* (Benjamin, London, 1973).
[7]J. Larsson and L. Stenflo, Beitr. Plasmaphys. **16**, 79 (1976).

Weak quantization in a nonperturbative model

Gérard G. Emch[a] and Kalyan B. Sinha[b]

Departamento de Física Matemática, Instituto de Física, Universidade de Sao Paulo, Brazil
(Received 1 September 1978)

The concepts of extended operator convergence and of spectral concentration are used to study rigorously a class of simple models for the tunnel effect and the laser. We compute exactly the asymptotic decay times of the eigenmodes, and we prove their link with the line width of the corresponding resonances.

I. DESCRIPTION OF THE MODEL

A massive quantum particle is restricted to move on the one-dimensional half-space $[0,\infty)$ with a rigid wall at $x = 0$. Its motion is free, except for a "square" potential barrier, starting at $x = \pi$, with height a and width b; for simplicity, we shall first assume that b is independent of a, and we will only later (Sec. IV) indicate the modifications to be brought to the theory when b is allowed to go to zero as a approaches infinity. The Hamiltonian of the system for finite $a \geqslant 0$ is thus $H_a = H_0 + aV$, where $(Vf)(x) = \chi_{[\pi,\pi+b]}(x)f(x)$ for all f in $\mathscr{H} = \mathscr{L}^2([0,\infty),dx)$; H_0 is the self-adjoint operator $-\Delta$, where Δ is the Laplacian, with domain [see X.3. in Ref. 1] $\mathscr{D}_0 = \{\phi\in\mathscr{H} \mid \phi$ and ϕ' absolutely continuous; $\phi''\in\mathscr{H}$; and $\phi(0) = 0\}$. Since V is bounded, $\mathscr{D}_0$ is also (see V.4.1 in Ref. 2) the domain of self-adjointness of H. Note that the spectrum of H_a is $[0,\infty)$ and is absolutely continuous with respect to Lebesgue measure; in particular, $H_a \geqslant 0$ for every finite $a \geqslant 0$.

This system is thus the simplest possible, and is a well-known (e.g., Ex. III.3 in Ref. 3) model for the tunnel effect. The purpose of this paper is to present a precise mathematical analysis of the asymptotic behavior of this system as a tends to infinity.

In the limit of infinitely large a, the physicist's intuition is that the wall decouples the inside region $\mathrm{I} = [0,\pi]$ from the outside region $\mathrm{III} = [\pi + b,\infty)$, and that the evolution is free in both of these regions, which are then limited by rigid walls at $x = 0$, π and $\pi + b$. The Hilbert space of the system thus becomes $\mathscr{H}_\infty = \mathscr{H}^{\mathrm{I}} \oplus \mathscr{H}^{\mathrm{III}}$ with $\mathscr{H}^{\mathrm{I}} = \mathscr{L}^2\{[0,\pi],dx\}$, $\mathscr{H}^{\mathrm{III}} = \mathscr{L}^2\{[\pi + b,\infty),dx\}$ and the evolution is governed by the self-adjoint operator $H_\infty = H^{\mathrm{I}} \oplus H^{\mathrm{III}}$ given by $-\Delta$ in both regions, with respective domains[1]: $\mathscr{D}^{\mathrm{I}} = \{\phi\in\mathscr{H}^{\mathrm{I}} \mid \phi$ and ϕ' absolutely continuous; $\phi''\in\mathscr{H}^{\mathrm{I}}$; and $\phi(0) = 0 = \phi(\pi)\}$ and $\mathscr{D}^{\mathrm{III}} = \{\phi\in\mathscr{H}^{\mathrm{III}} \mid \phi$ and ϕ' absolutely continuous; $\phi''\in\mathscr{H}^{\mathrm{III}}$; and $\phi(\pi + b) = 0\}$. Note that H_∞ is the Friedrichs extension [see for instance VI.2.3 in *Ref.* 2] in $\mathscr{H}_\infty$ of the restriction of H_0 to the dense domain $\mathscr{D}_1 = \mathscr{D}_0 \cap P\mathscr{H}$, where P is the projector from $\mathscr{H}$ onto $\mathscr{H}_\infty$.

H^{III} is clearly unitarily equivalent to H_0, and therefore has absolutely continuous spectrum. H^{I} on the other hand has a purely discrete spectrum $\{m^2 \mid m = 1,2,\ldots\}$.

Whereas H_a is obviously a perturbation of H_0, it is not a small perturbation away from H_∞. Our aim is to describe how H_∞ is nevertheless the limit of H_a as a tends to infinity, and to control this limit well enough to allow an understanding of the exponential decay which one expects on physical grounds in the tunnel effect.

II. OPERATOR CONVERGENCE

Before addressing the problem of exponential decay we want in this section to elucidate the sense in which H_∞ is the limit of H_a as a tends to infinity.

*Theorem II.*1: For every ϕ in $\mathscr{D}$, a domain of essential self-adjointness of H_∞, there exists $\{\phi_a \mid a\in(0,\infty)\} \subset D_0$ such that, as $a\to\infty$:

(i) $\|\phi_a - \phi\| = 0\ (a^{-1/2})$,

(ii) $\|H_a\phi_a - H_\infty\phi\| = 0\ (a^{-1/2})$.

Proof: We can deal with regions I and III separately. Let first $\phi\in\mathscr{D}^{\mathrm{III}}$, which we embed in $\mathscr{H}$ by setting $\phi(x) = 0$ for all $x \leqslant X = \pi + b$. If $\phi'(X) = 0$, ϕ belongs to the domain of H_a as well, so that (i) and (ii) are trivially satisfied by $\phi_a = \phi$ for all $a\in(0,\infty)$. We can therefore suppose, without loss of generality, that $\phi'(X) = A \neq 0$, and that there exists $\epsilon > 0$ for which ϕ does not vanish in $(X, X + \epsilon]$. Let ξ and ζ be two nonincreasing functions in $\mathscr{C}^\infty(-\infty,\infty)$ with $\xi(x) = 1$ for all $x \leqslant 0$, $\xi(x) = 0$ for all $x \geqslant \pi$; $\zeta(x) = 1$ for all $x \leqslant X$, $\zeta(x) = 0$ for all $x \geqslant X + \epsilon$. We further define, for every $a > 0$ and every x in $[0,X]$:

$$\psi_a(x) = Aa^{-1/2}\exp[(x - X)a^{1/2}].$$

One then verifies that an approximating net $\{\phi_a \mid a\in(0,\infty)\}$, in the sense of the theorem, is obtained by setting $\phi_a(x)$ equal to:

$$\begin{array}{ll} -Aa^{-1/2}\exp(-Xa^{1/2})\,\xi(x) + \psi_a(x), & \text{for } x\in\mathrm{I} = [0,\pi] \\ \psi_a(x), & \text{for } x\in\mathrm{II} = [\pi,\pi + b] \\ Aa^{-1/2}\zeta(x) + \phi(x), & \text{for } x\in\mathrm{III} = [\pi + b,\infty) \end{array}$$

Hence $\mathscr{D}^{\mathrm{III}} \subseteq \mathscr{D}$. A similar argument could be made for region I. We, however, find it more instructive to construct explicitly one approximating net $\{\phi_a^{(m)} \mid a\in(0,\infty)\}$ for each eigenvector $\phi^{(m)}$ $(m = 1,2,\cdots)$ of H^{I}. We chose $\phi^{(m)}(x) = \sin mx$, and embed $\phi^{(m)}$ in $\mathscr{H}$ be setting $\phi^{(m)}(x) = 0$ for

---

[a]Permanent address: Departments of Mathematics and of Physics, University of Rochester, N.Y. 14627. Research supported in part by the US-NSF grant MCS 76-07286.

[b]Permanent address: Dept. of Mathematics and Statistics, Indian Statistical Institute, New Delhi 110029 (India).

 0022-2488/79/071336-05$01.00

all $x \geqslant \pi$. For each m fixed, we define m_a, with $m_a \to m$ as $a \to \infty$, by $m_a^{-1} \tan(m_a\pi) = -a^{-1/2}$. We further introduce $M_a^{(m)} = a^{1/2}\sin(m_a\pi)$. Upon noticing that $(m - m_a)$ and $M - M_a^{(m)}$ are both $O(a^{-1/2})$ as $a \to \infty$, one verifies that an approximating net $\{\phi_a^{(m)} | a \in (0,\infty)\}$, in the sense of the theorem, is obtained for $\phi^{(m)}$ by setting $\phi_a^{(m)}(x)$ equal to:

$$\sin(m_a x), \qquad \text{for } x \in \mathrm{I}$$

$$M_a^{(m)} a^{-1/2} \exp[(\pi - x)a^{1/2}], \qquad \text{for } x \in \mathrm{II} \cup \mathrm{III}.$$

Note that $\{\phi^{(m)} | m = 1,2,\ldots\}$ is an orthogonal basis in $\mathscr{H}^{\mathrm{I}}$, consisting of eigenvectors of H^{I}; H^{I} is thus essentially self-adjoint on the linear span of these vectors. The above argument shows that this manifold is contained in $\mathscr{D}$. We can therefore prove the assertion of the theorem with

$$\mathscr{D} = \mathrm{span}\{\phi^{(m)} | m = 1,2,\cdots\} \oplus \mathscr{D}^{\mathrm{III}}. \qquad \text{Q.E.D.}$$

Let now $\{U_a(t) | t \in (-\infty, +\infty)\}$ (resp. $\{U_\infty(t) | t \in (-\infty, +\infty)\}$) be the unitary group on $\mathscr{H}$ (resp. $\mathscr{H}_\infty$) generated by H_a (resp. H_∞).

Corollary II.2: For every $\phi \in \mathscr{H}_\infty$ and every $T \in [0,\infty)$

$$\lim_{a\to\infty} \sup_{0 \leqslant t \leqslant T} \| U_a(t)\phi - U_\infty(t)\phi \| = 0$$

Proof: With $\mathscr{D}$ as in Thm. II.1, we have [see V.3.4 in Ref. 2] $\{(H_\infty - i\mathrm{I})\phi | \phi \in \mathscr{D}\}$ dense in $\mathscr{H}_\infty$. The corollary then follows directly from Kurtz' criterion[4] in his theory of extended operator convergence.

Hence on the Hilbert space $\mathscr{H}_\infty$ corresponding to the limit of an infinitely high wall, the time-evolution $U_a(t)$ converges strongly to the limiting time-evolution $U_\infty(t)$, uniformly in t on compacts. The latter result (for b independent of a) is not new.[5,6] As a particular case of these papers, one has indeed, as $a \to \infty$, that the resolvant $R_a(z)$ of H_a converges strongly on $\mathscr{H}_\infty$ to the resolvant $R_\infty(z)$ of H_∞ for every $z \in \mathbb{C} - [0,\infty)$ [in conformity with Corollary II.2, by a slight modification of the classical argument (see IX.2.5 in Ref. 2)]. Moreover,[6] it follows from the strong resolvant convergence that, as $a \to \infty$, the semigroup $\{S_a(t) = \exp(-H_a t) | t \in [0,\infty)\}$ converges strongly on $\mathscr{H}_\infty$ to the semigroup $\{S_\infty(t) = \exp(-H_\infty t) | t \in [0,\infty)\}$. The estimate of Theorem II.1 however is new, and is of some independent interest [see in particular Sec. III and IV below].

III. DECAY

We saw in Sec. II that the limiting dynamics corresponds to the hard wall condition in the Hamiltonian H_∞. The limiting process has drastically changed the spectrum from continuous (H_a) to discrete (H_∞). We now turn around and think of the initial system as the one with infinitely high walls, and then bring down the wall to a finite, albeit very large, height. In such a scenario the spectrum of the relevant Hamiltonian makes a transition from discrete to continuous, a transition we want to investigate in detail. For this purpose we use the spectral transformation (or generalized Fourier transform) of H_a (for these notions the reader may consult Ref. 7).

In this simple model, we can solve the Schrödinger equation exactly and obtain the eigenfunctions $\{\psi_\lambda | \lambda \in [0,\infty)\}$:

$$\psi_\lambda(x) = \begin{cases} \alpha(k)\sin kx, \\ \beta_-(k)\exp[-\xi(x-\pi)] + \beta_+(k)\exp[\xi(x-\pi)], \\ \gamma(k)\sin[k(x-\pi-b)+\delta], \end{cases}$$

where we have written $k^2 = \lambda$ and $\xi = (a - k^2)^{1/2}$. The coefficients α, β are determined in terms of γ by the requirement that ψ_λ be locally in the domain $\mathscr{D}_0$ of H_a, i.e., ψ_λ and ψ_λ' be locally absolutely continuous. For most of our calculations we shall need the details of only α, the latter turning out to be

$$\alpha(k)^2 = \gamma(k)^2\{\sin^2(k\pi) - k^2\xi^{-2}\cos^2(k\pi) + a[k^{-1}\eta_-(k)\sin(k\pi) + \xi^{-1}\eta_+(k)\cos(k\pi)]^2\}^{-1}, \quad (1)$$

where $\eta_\pm(k) = [\exp(\xi b) \pm \exp(-\xi b)]/2$. We choose the normalization $\gamma(k) = (\pi k)^{-1/2}$.

We take the initial situation to be the one with infinitely high walls and begin with an eigenmode ϕ_n $[\phi_n(x) = (2/\pi)^{1/2}\sin(nx)]$ of H^{I} trapped in region I. The wall is then "lowered" from $a = \infty$ to some finite, but large a. We want the asymptotic behavior, as $a \to \infty$, of the probability $|(\phi_n, \exp[-iH_a t]\phi_n)|^2$ that the eigenmode ϕ_n (now evolving under the group $\exp[-iH_a t]$) will remain in the same mode after a time t has elapsed. From Sec. II, $U_a(t)$ converges strongly to $U_\infty(t)$ on $\mathscr{H}^{\mathrm{I}}$; from this it follows that the above probability converges (uniformly in t on compacts) to $\|\phi_n\|^2 = 1$, as $a \to \infty$. Further information on the rate of this convergence is of physical interest for the description of the tunnel effect. We observe that

$$(\phi_n, \exp(-iH_a t)\phi_n) = \int_0^\infty d\lambda \exp(-i\lambda t)|\phi_n(\lambda)|^2, \quad (2)$$

where

$$\phi_n(\lambda) = \int_0^\pi dx\, \psi_\lambda(x)^* \phi_n(x) \quad (3)$$

is the spectral representative (or generalized Fourier transform) of ϕ_n. We have

$$|\phi_n(\lambda)|^2 = (2\pi)^{-1}\alpha(k)^2[(k-n)^{-1}\sin(k-n)\pi - (k+n)^{-1}\sin(k+n)\pi]^2. \quad (4)$$

The term in square brackets is bounded in k and converges to π^2 as $k \to n$; therefore, the major contribution will come from $\alpha(k)^2$. From (1) one concludes that this contribution originates from the neighborhoods of the zeros of the *a priori* larger term (as $a \to \infty$). This leads to the resonance equation (or approximate eigenvalue equation):

$$F(k,a) = 0, \quad 0 < k < a^{1/2}, \quad (5)$$

where

$$F(k,a) = [\eta_-(k)/\eta_+(k)]\tan k\pi + [k/\xi] \quad (6)$$

We observe that F is $\mathscr{C}^\infty$ in a neighborhood of (n,∞), with $F(n,\infty) = 0$ and $F_k(n,\infty) \equiv (\partial F/\partial k)(n,\infty) \neq 0$. From the implicit function theorem[8] there exists a positive a_0, large enough such that for all $a > a_0$, the resonance equation (5) has a unique solution $k = k_n(a)$ [i.e., $\lambda = \lambda(n,a) = k_n(a)^2$]

with $\lambda(n,\infty) = n^2$. An asymptotic expansion of $\lambda(n,a)$ in terms of $a^{-1/2}$ can now be derived:

$$\lambda(n,a) = n^2 - 2n^2\pi^{-1}a^{-1/2}[\eta_+/\eta_-]_{a=\infty} + O(a^{-1}). \quad (7)$$

We remark that, as $a \to \infty$, $\lambda(n,a)$ approaches the nth eigenvalue of H^1. Thus the resonance equation (5), by itself, asymptotically selects H^1 from the one-dimensional[1] manifold spanned by the self-adjoint extensions of the symmetric operator obtained as the restriction of H_0 (or H_a) to $\mathscr{D}\cap P^1\mathscr{H}$

The phase-shift δ in the eigenfunction in region III is

$$\tan\delta = k\xi^{-1}[\tan(k\pi) + k\xi^{-1}(\eta_-/\eta_+)]F(k,a)^{-1}. \quad (8)$$

From this follows that the phase-shift at resonance is $\delta_n = \pi/2$. Moreover

$$\frac{d\delta}{dk}[k = k_n(a)] = \xi_n^2 k_n^{-2}\eta_+(k_n)\eta_-(k_n)F_k(k_n,a), \quad (9)$$

which is a large positive number. This is in conformity with the conventional definition of a resonance. The amplitude at resonance is

$$\alpha(k_n)^2 = \gamma(k_n)^2\eta_+(k_n)^2(\sin k_n\pi)^{-2}. \quad (10)$$

From Sec. II, recall that for every $z\in\mathbb{C} - [0,\infty)$, $(z - H_a)^{-1}\phi \to (z - H_\infty)^{-1}\phi$ for all $\phi\in\mathscr{H}^1$. We thus expect (see VIII.5.2 in Ref. 2) to have a "spectral concentration," expressing that the spectral measure of H_a concentrates, as a becomes large, in some neighborhoods of the eigenvalues n^2 of $H_\infty^{\ 1}$. We now want to compute the details of this concentration, i.e., in physical terms, the asymptotic line shape as $a \to \infty$.

Let us denote by $\{E_a(\lambda)|\lambda\in[0,\infty)\}$ the spectral family of H_a. Since H_a is spectrally absolutely continuous, there exist positive, integrable functions $f_a(n,\lambda)$ such that for every real c and d:

$$(\phi_n, E_a([c,d])\phi_n) = \int_c^d d\lambda\, f_a(n,\lambda). \quad (11)$$

The next theorem states the asymptotic properties of $f_a(n,\cdot)$ as $a \to \infty$.

Theorem III.1: Let $\lambda(n,a)$ and $\{f_a(n,\cdot)|a\in[0,\infty)\}$ be defined as above, and let

$$\Gamma(n,a) = 2\pi^{-1}\lambda(n,a)^{3/2}[a - \lambda(n,a)]^{-1}\eta_-[k_n(a)]^2. \quad (12)$$

Then $f_a(n,\lambda) = |\phi_n(\lambda)|^2$ [see (4)], and the function $g_a(n,\cdot)$, defined on $(-\infty, +\infty)$ by

$$g_a(n,h) = \Gamma(n,a)f_a(n,\lambda(n,a) + h\Gamma(n,a)) \quad (13)$$

converge as $a \to \infty$, pointwise and in $\mathscr{L}^1$-norm to $g(\cdot)$ with $g(h) = [\pi(1 + h^2)]^{-1}$.

Proof: The first assertion follows directly from (11). Upon using (6), we rewrite (1) as

$$\alpha(k)^2/\gamma(k)^2 = (\sec^2 k\pi)[k^2\xi^{-2}\eta_-^{-2} - 2k\xi^{-1}\eta_+^2\eta_-^{-2}F + (ak^{-2}\eta_+^2 + \eta_+^2\eta_-^{-2})F^2]^{-1}. \quad (14)$$

Since $F[\lambda(n,a),a] = 0$ and F is a C^∞ function in a neighborhood of (n,∞), we have the Taylor expansion

$$\begin{aligned} F(k,a) &= [\lambda - \lambda(n,a)]F_\lambda[\lambda(n,a),a] + 2^{-1}[\lambda - \lambda(n,a)]^2F_{\lambda\lambda}(\lambda,a) \\ &= (2k)^{-1}F_k(k_n(a),a)\Gamma(n,a)h + 2^{-1}F_{\lambda\lambda}(\lambda,a)\Gamma(n,a)^2h^2, \end{aligned} \quad (15)$$

where we defined h by

$$\lambda = \lambda(n,a) + h\Gamma(n,a). \quad (16)$$

From (12), (14), and (15), we see that

$$\Gamma(n,a)\alpha(k)^2 = 2\pi^{-2}\{1 - 2\Delta(n,a)h + h^2 + O[\Gamma(n,a)]\}^{-1}, \quad (17)$$

where

$$\Delta(n,a) = (\eta_+/\eta_-)a^{-1/2}k_n(a). \quad (18)$$

Since for any fixed real h, we can find $a > 0$ large enough so that $\lambda = \lambda(n,a) + \Gamma(n,a)h > 0$ we have, for such h, that $\Gamma(n,a)f_a(n,\lambda(n,a) + \Gamma(n,a)h)$ approaches $g(h)$ as a tends to infinity, pointwise in h. On the other hand, upon setting $g_a(n,h) = 0$ for $h < -\lambda(n,a)/\Gamma(n,a)$, we have

$$\int_{-\infty}^{+\infty} dh\, g_a(n,h) = \int_0^\infty d\lambda\, f_a(n,\lambda) = 1 \quad \text{for all } a > 0. (19)$$

Since $\mathscr{L}^1$-norm of g is also 1, we have[9] that $g_a(n,\cdot)$ converges to $g(\cdot)$ in $\mathscr{L}^1$-norm. Q.E.D.

The theorem has two corollaries, both of which can be obtained as in Ref. 9.

Corollary III.2: For any $h_1 < h_2$ real:

$$\lim_{a\to\infty}(\phi_n, E_a[\lambda(n,a) + \Gamma(n,a)h_1, \lambda(n,a) + \Gamma(n,a)h_2]\phi_n) = \int_{h_1}^{h_2} dh\,[\pi(1 + h^2)]^{-1}.$$

This result gives the explicit form of the spectral concentration: For large a, the resonance approaches a Lorentzian, centered around $\lambda(n,a)$, and of width $\Gamma(n,a)$ given by (12).

Corollary III.3: For any $\tau \geqslant 0$:

$$\lim_{a\to\infty}(\phi_n, \exp\{-i[H_a - \lambda(n,a)]\Gamma(n,a)^{-1}\tau\}\phi_n) = \exp(-\tau),$$

and the convergence is uniform in $0 \leqslant \tau < \infty$. Consequently, the probability $|(\phi_n, \exp[-iH_a\Gamma(n,a)^{-1}\tau]\phi_n)|^2$ behaves asymptotically as $\exp(-2\tau)$ when $a \to \infty$. Upon reintroducing the unscaled time $t = \Gamma(n,a)^{-1}\tau$, we thus find that, for large a, $|(\phi_n, U_a(t)\phi_n)|^2$ behaves as $\exp[-2\Gamma(n,a)t]$. In other words, as we "lower" the barrier from an infinite to a finite but large height, we can interpret $[2\Gamma(n,a)]^{-1}$ as the half-life of the eigenmode ϕ_n. This confirms the usual relation between the half-line width of a resonance and the half-life time of its decay. The above calculation indeed shows in a precise manner how the scaling in energy is inversely related to the scaling in time. The rescaled time τ is of the order of $\Gamma(n,a)^{-1}$, i.e., $a\exp(a^{1/2}b)$, which is very large; hence the decay of the eigenmodes indeed takes place very slowly; equivalently the resonances are very sharp, with a very small linewidth.

At this point it is worth mentioning that since $a^{1/2}\Gamma(n,a) \to 0$ as $a \to \infty$, there is no contradiction between

Theorem II.1 and Theorem III.1 (or Corollaries III.2 and III.3), thus bypassing the objection raised by Davies [compare indeed these results with conditions (2) and (4)–(7) in Ref. 9].

A computation, similar to that carried above, can be made for the off-diagonal elements of $\exp[-iH_a\tau/\Gamma(n,a)]$ in $\mathscr{H}^1$, indicating that the eigenmode n not only decays, but actually leaks out of region I.

IV. GENERALIZATION OF THE MODEL

The generalization consists in allowing $b\to 0$ as $a\to\infty$, i.e., more precisely: $b = 0(a^\nu)$ with $\nu<0$. If $0>\nu>-\frac{1}{2}$ (or $a^{1/2}b\to\infty$ as $a\to\infty$), the construction and the proof of Theorem II.1 remain essentially unchanged. In this case however, the concept of extended operator convergence[4] takes full force and goes beyond the case studied in Ref. 6. Also $(\Delta\lambda)(n,a)\equiv\lambda(n,a)-n^2 = -2n^2\pi^{-1}a^{-1/2}+O(a^{-1})$ while $\Gamma(n,a)\simeq 8n^3\pi^{-1}a^{-1}\exp(-2a^{1/2}b)$, showing that the half-width still is exponentially small compared to the shift, If $\nu=-\frac{1}{2}$ (i.e., $a^{1/2}b\to\beta>0$), $(\Delta\lambda)(n,a) = -2n^2\pi^{-1}a^{-1/2}\coth\beta + O(a^{-1})$ and $\Gamma(n,a)\simeq 2n^3\pi^{-1}a^{-1}\csc^2\beta$. If however $-\frac{1}{2}>\nu>-1$ (i.e., $ab=\Lambda\to\infty$ as $a\to\infty$), the resonance equation (5.6) has no solution, and it should be modified to read

$$G(k,a)\equiv a^{1/2}F(k,a)=0. \tag{20}$$

This modified resonance equation has a unique solution $\lambda(n,a)$ in the neighborhood of n^2, and we have $(\Delta\lambda)(n,a)\simeq -2n^2\pi^{-1}\Lambda^{-1}$, while $\Gamma(n,a)\simeq 2n^3\pi^{-1}\Lambda^{-2}$. In all these cases, one has spectral concentration and decay in the sense of Sec. III. Moreover, upon using the estimates of Sec. III, one proves again that $\exp(-iH_a t)f$ converges strongly to $\exp(-iH^1 t)f$ as $a\to\infty$, for all f in $\mathscr{H}^1$. Finally, if $\nu\leqslant -1$ (i.e., $ab\to$finite, possibly zero, limit), none of the considerations of Sec. III applies, and even the modified resonance equation (20) fails to have a solution near n^2; in fact, in the extreme case where $b=O(a^{-2})$, H_a[resp. $U_a(t)$] clearly converges strongly to H_0 [resp. $U_0(t)$]: The wall has become completely transparent.

V. CONCLUSIONS

The model is nonperturbative by nature. Yet it is simple enough to be exactly solvable, and to allow a precise control of its asymptotic behavior as a approaches infinity. It is moreover sophisticated enough to exhibit a host of interesting features, both physical and mathematical, which we briefly review on the basis of our analysis.

First of all, the model exhibits exponential decay, although all the Hamiltonians occuring in the problem are uniformly bounded below, namely by zero. This should be contrasted with the situation encountered in nonequilibrium statistical mechanics, where the presence of an infinite bath at finite temperature allows the generator of the time-evolution to have Lebesgue spectrum, covering the whole real line (for general arguments to this effect, as well as for models, see for instance Refs. 10 and 11). The exponential decay found in the present model emphasizes the role of the rescaling in time, which allows to bypass the usual no-go theorems (e.g., 7.3.3 in Ref. 10) by the mechanism described in Sec. III. This mechanism appears to be quite different from that occuring in the van Hove limit of statistical mechanics.[10,12]

We might remark here that the exact asymptotic lifetime and width found in this model coincide with the value found in the WKB approximation (see for instance Ref. 3); a similar feature has been noticed also in Ref. 13. This coincidence with the exact result, found by an unperturbative approach, seems to have a status similar to that of the Born approximation in the master equation theory (see e.g., Ref. 12).

The decay found in the present model can be related to the phenomenon known in physics as "weak quantization" (see for instance p. 251 in Ref. 14, or pp. 403–408 in Ref. 15). The physical picture is given a firm mathematical basis in this model; we indeed saw that the point spectrum, encountered when the inside region I is decoupled from the outside by an infinitely high hard wall, only persists, as the wall is lowered, in the form of Lorentzian resonances: the higher the wall, the sharper the resonances; still for any finite height of the wall, the spectrum of the Hamiltonian remains absolutely continuous with respect to Lebesgue measure. This phenomena is also known in the mathematical literature (e.g., Ref. 2) as "spectral concentration." It should be, however, noticed that, for $b=O(a^\nu)$ with $0\geqslant\nu\geqslant-\frac{1}{2}$, the spectral concentration found in the present model is much stronger than the usual concentration of polynomial type.[9,16]

Whereas the present model describes very well the qualitative features of the tunnel effect, its one-dimensional character should be removed for a realistic theory of α-decay. On the other hand, the model as it stands presents some instructive analogy with the laser, its finite high wall playing the role of a semitransparent mirror. Some of the qualitative asymptotic features of the model are also found,[17] upon using the techniques of S-matrix theory, when the semitransparent mirror is mimicked by a "δ-function potential of strength Λ"; in the latter case, the limit of large Λ plays the role of our limit of large a. Incidentally, the form-sum $H_0+\Lambda\delta_\pi$ can be obtained as the form-limit, when $a\to\infty$, of H_a with $b=\Lambda a^{-1}$ ($\Lambda\neq 0$). When $ab\to 0$, one finds H_0 back. A true theory of the laser would, however, require two modifications of the present model. Firstly, the Maxwell equation, rather than the Schrödinger equation for a massive particle, should be taken as the starting point; secondly, a second-quantization, rather than first-quantization, formalism should be used. Nevertheless, it seems likely that the phenomenon of "weak quantization," or "spectral concentration," would persist in such a complete theory, and that it could provide a useful basis for its discussion.

ACKNOWLEDGMENTS

Professor H.M. Nussenzveig and Mr. B. Baseia should be thanked here for useful discussions on the laser problems. It is a pleasure for us to acknowledge Professor Nussenzveig's warm hospitality at the Departamento de Física Matemática of the University of São Paulo. Our collaboration was made possible by a grant from the Fundacão de Amparo à Pesquisa do Estado de São Paulo (FAPESP).

[1]M.H. Stone, *Linear Transformations in Hilbert Space and their Applications to Analysis* (AMS Colloquium Publications, New York, 1932), Vol. XV.
[2]T. Kato, *Perturbation Theory for Linear Operators* (Springer, Berlin, 1966).
[3]A. Messiah, *Mécanique Quantique* (Dunod, Paris, 1959).
[4]T.G. Kurtz, Funct. Anal. **3**, 354 (1969); or **12**, 55 (1973).
[5]E.B. Davies, Helv. Phys. Acta **48**, 365 (1975).
[6]H. Baumgärtel and M. Demuth, preprint, Berlin, 1977.
[7]W.O. Amrein, J.M. Jauch, and K.B. Sinha, *Scattering Theory in Quantum Mechanics* (Benjamin, Reading, Massachusetts, 1977).
[8]E. Hille, *Methods in Classical and Functional Analysis*, (Addison-Wesley, Reading, Massachusetts, 1972).
[9]E.B. Davies, Lett. Math. Phys. **1**, 31 (1975).
[10]E.B. Davies, *Quantum Theory of Open Systems* (Academic, London, 1976).
[11]G.G. Emch, Commun. Math. Phys. **49**, 191 (1976).
[12]Ph. Martin and G.G. Emch, Helv. Phys. Acta **48**, 59 (1975).
[13]K. Sinha, Lett. Math. Phys. **1**, 251 (1976).
[14]L.D. Landau and E.M. Lifshitz, *Quantum Mechanics* (Pergamon, London, 1958).
[15]E.C. Kemble, *Fundamental Principles of Quantum Mechanics* (McGraw-Hill, New York, 1937).
[16]C.C. Conley and P.A. Rejto, in *Perturbation Theory and its Applications in Quantum Mechanics*, edited by C.H. Wilcox (Wiley, New York, 1966).
[17]H.M. Nussenzveig, *Causality and Dispersion Relations*, (Academic, New York, 1972).

Representations of the Poincaré group for relativistic extended hadrons

Y. S. Kim

Center for Theoretical Physics, Department of Physics and Astronomy, University of Maryland, College Park, Maryland 20742

Marilyn E. Noz

Department of Radiology, New York University, New York, New York 10016

S. H. Oh

Laboratory of Nuclear Science and Department of Physics, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139
(Received 7 June 1978; revised manuscript received 23 October 1978)

Representations of the Poincaré group are constructed from the relativistic harmonic oscillator wave functions which have been effective in describing the physics of internal quark motions in the relativistic quark model. These wave functions are solutions of the Lorentz-invariant harmonic oscillator differential equation in the "cylindrical" coordinate system moving with the hadronic velocity in which the time-separation variable is treated separately. This result enables us to assert that the hadronic mass spectrum is generated by the internal quark level excitation, and that the hadronic spin is due to the internal orbital angular momentum. An addendum relegated to PAPS contains discussions of detailed calculational aspects of the Lorentz transformation, and of solutions of the oscillator equation which are diagonal in the Casimir operators of the homogeneous Lorentz group. It is shown there that the representation of the homogeneous Lorentz group consists of solutions of the oscillator partial differential equation in a "spherical" coordinate system in which the Lorentz-invariant Minkowskian distance between the constituent quarks is the radial variable.

I. INTRODUCTION

In building models of relativistic extended hadrons, we have to keep in mind the fundamental fact that the overall space–time symmetry structure is that of the Poincaré group.[1] In our previous papers on physical applications of the relativistic harmonic oscillator,[2] our primary purpose was to devise a calculational scheme for explaining experimental observations. As was pointed out by Biedenharn *et al.*,[3] the question of the Poincaré symmetry has not been systematically discussed.

The purpose of the present paper is to address this symmetry problem. We are considering a model hadron consisting of two spinless quarks bound together by a harmonic oscillator potential. In this case, we are led to consider the center-of-mass coordinate which specifies the space–time location of the hadron, and the relative coordinate which specifies the internal space–time separation between the quarks.

Both the hadronic and internal coordinates are subject to Poincaré transformations consisting of translations and Lorentz transformations. The hadronic coordinate undergoes Poincaré transformation in the usual manner. However, the internal coordinate is invariant under translations. This coordinate should, nonetheless, satisfy the Poincaré symmetry as a whole. We discuss in this paper the role of this internal coordinate, and show that internal excitations generate the hadronic mass spectrum, and that the internal angular momentum corresponds to the spin of the hadron.

In Sec. II, we formulate the problem using a model hadron consisting of two spinless quarks bound together by a harmonic oscillator potential of unit strength, and then discuss the generators of the Poincaré group applicable to the entire system. In Sec. III, we present the oscillator wave functions which are diagonal in the invariant Casimir operators of the Poincaré group.

II. FORMULATION OF THE PROBLEM

In our previous papers on physical applications of the relativistic harmonic oscillators, we started with the following Lorentz-invariant differential equation:

$$\{2[\Box_1 + \Box_2] - \tfrac{1}{16}(x_1 - x_2)^2 + m_0^2\}\phi(x_1,x_2) = 0, \quad (1)$$

where x_1 and x_2 are the space–time coordinates for the two spinless quarks bound together by a harmonic oscillator potential with unit spring constant. In order to simplify the above equation, let us define new coordinate variables

$$X = \tfrac{1}{2}(x_1 + x_2), \quad x = (1/2\sqrt{2})(x_1 - x_2). \quad (2)$$

The X coordinate represents the space–time specification of the hadron as a whole, while the x variable measures the relative space–time separation between the quarks. In terms of these variables, Eq. (1) can be written as

$$\left[\frac{\partial^2}{\partial X_\mu^{\,2}} + m_0^2 + \frac{1}{2}\left(\frac{\partial^2}{\partial x_\mu^{\,2}} - x_\mu^{\,2}\right)\right]\phi(X,x) = 0. \quad (3)$$

The above equation is separable in the X and x variables. Thus we write

$$\phi(X,x) = f(X)\psi(x), \quad (4)$$

where $f(X)$ and $\psi(x)$ satisfy the following differential equations respectively:

 0022-2488/79/071341-04$01.00

$$\left[\frac{\partial^2}{\partial X_\mu^2} + m_0^2 + (\lambda + 1)\right] f(X) = 0, \tag{5}$$

$$\frac{1}{2}\left(\frac{\partial^2}{\partial x_\mu^2} - x_\mu^2\right)\psi(x) = (\lambda + 1)\psi(x). \tag{6}$$

The differential equation of Eq. (5) is a Klein–Gordon equation, and its solutions are well known. $f(X)$ takes the form

$$f(X) = \exp(\pm ip{\cdot}X), \tag{7}$$

with

$$p^2 = m_0^2 + (\lambda + 1), \tag{8}$$

where p is the 4-momentum of the hadron. p^2 is, of course, the mass of the hadron and is numerically constrained to take the values allowed by Eq. (8). The separation constant λ is determined from the solutions of the harmonic oscillator differential equation of Eq. (6). The physical solutions of the oscillator equation satisfy the subsidiary condition

$$p^\mu a_\mu^\dagger \psi_\beta(x) = 0, \tag{9}$$

where

$$a_\mu^\dagger = x_\mu + \frac{\partial}{\partial x^\mu}.$$

The physics of this subsidiary condition has been extensively discussed in the literature.[2,4]

The space–time transformation of the total wave function of Eq. (4) is generated by the following ten generators of the Poincaré group. The operators

$$P_\mu = i\frac{\partial}{\partial X^\mu} \tag{10}$$

generate space–time translations. Lorentz transformations, which include boosts and rotations, are generated by

$$M_{\mu\nu} = L^*_{\mu\nu} + L_{\mu\nu}, \tag{11}$$

where

$$L^*_{\mu\nu} = i\left(X_\mu \frac{\partial}{\partial X^\nu} - X_\nu \frac{\partial}{\partial X^\mu}\right),$$

$$L_{\mu\nu} = i\left(x_\mu \frac{\partial}{\partial x^\nu} - x_\nu \frac{\partial}{\partial x^\mu}\right).$$

The translation operators P_μ act only on the hadronic coordinate, and do not affect the internal coordinate. The operators $L^*_{\mu\nu}$ and $L_{\mu\nu}$ Lorentz-transform the hadronic and internal coordinates respectively. The above ten generators satisfy the commutation relations for the Poincaré group.

In order to consider irreducible representations of the Poincaré group, we have to construct wave functions which are diagonal in the invariant Casimir operators of the group, which commute with all the generators of Eqs. (10) and (11). The Casimir operators in this case are

$$P^\mu P_\mu \quad \text{and} \quad W^\mu W_\mu, \tag{12}$$

where

$$W_\mu = \epsilon_{\mu\nu\alpha\beta} P^\nu M^{\alpha\beta}.$$

The eigenvalues of the above P^2 and W^2 represent respectively the mass and spin of the hadron.

III. PHYSICAL WAVE FUNCTIONS AND REPRESENTATIONS OF THE POINCARÉ GROUP

In constructing wave functions diagonal in the Casimir operators of the Poincaré group, we note first that the operator which acts on the wave function in the subsidiary condition of Eq. (9) commutes with these invariant operators:

$$[P^2, p^\mu a_\mu^\dagger] = 0, \tag{13}$$

$$[W^2, p^\mu a_\mu^\dagger] = 0. \tag{14}$$

Therefore, the wave functions satisfying the condition of Eq. (9) can be diagonal in the Casimir operators.

In order to obtain the solutions explicitly, let us assume without loss of generality that the hadron moves along the z direction with the velocity parameter β. Then we are led to consider the Lorentz frame where the hadron is at rest, and the coordinate variables are given by

$$\begin{aligned} x' &= x, \quad y' = y, \\ z' &= (z - \beta t)/(1 - \beta^2)^{1/2}, \\ t' &= (t - \beta z)/(1 - \beta^2)^{1/2}. \end{aligned} \tag{15}$$

The Lorentz-invariant oscillator equation of Eq. (6) is separable in the above variables. In terms of these primed variables, we can construct a complete set of wave functions

$$\psi_\beta(x) = f_b(x')f_s(y')f_n(z')f_k(t'), \tag{16}$$

where

$$f_n(z') = (\sqrt{\pi}\, 2^n n!)^{-1/2} H_n(z') \exp(-z'^2/2),$$

$$f_k(t') = (\sqrt{\pi}\, 2^k k!)^{-1/2} H_k(t') \exp(-t'^2/2).$$

If the excitation numbers, $b,\ldots,k$ are allowed to take all possible nonnegative integer values, the solutions in Eq. (16) form a complete set. However, the eigenvalues λ takes the form

$$\lambda = b + s + n - k. \tag{17}$$

Because the coefficient of k is negative in the above expression, λ has no lower bound, and there is an infinite degeneracy for a given value of λ.

In terms of the primed coordinates, the subsidiary condition of Eq. (9) takes the simple form

$$\left(\frac{\partial}{\partial t'} + t'\right)\psi_\beta(x) = 0. \tag{18}$$

This limits $f_k(t')$ to $f_0(t')$, and the eigenvalue λ becomes

$$\lambda = b + s + n, \tag{19}$$

The physical wave functions satisfying the subsidiary condition of Eq. (9) or (18) have nonnegative values of λ.

As far as the x', y', z' coordinates are concerned, they form an orthogonal Euclidean space, and $f_b(x'), f_s(y'), f_n(z')$ form a complete set in this three-dimensional space. The Hermite polynomials in these Cartesian wave functions can then be combined to form the eigenfunctions of W^2 which, in terms of the primed coordinate variables, takes the form

$$W^2 = M^2(\mathbf{L}')^2, \tag{20}$$

where

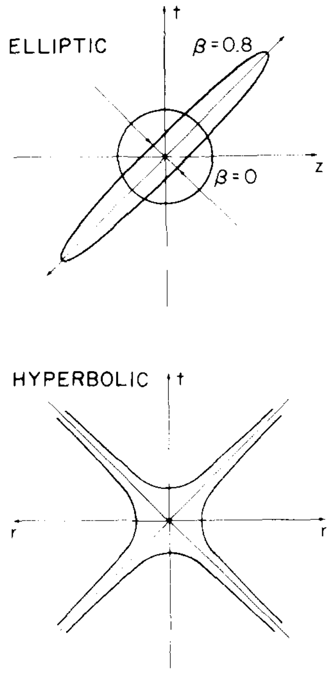

FIG. 1. Elliptic and hyperbolic localizations in space–time. The wave functions in the present paper are elliptically localized, and undergo Lorentz deformation as the hadron moves. The Lorentz invariant form $x^{\mu}x_{\mu}$, to which we are accustomed, is hyperbolically localized, and is basically different from the form used in the present paper.

$$L_i' = -i\epsilon_{ijk}x_j' \frac{\partial}{\partial x_k'},$$

and M is the hadronic mass.

The physical wave functions now take the form

$$\psi_{\beta}^{\lambda lm}(x) = (1/\pi)^{1/4}[\exp(-t'^2)]R_{\lambda l}(r')Y_{lm}(\theta',\phi'), \quad (21)$$

where r', θ', ϕ' are the radial and spherical variables in the three-dimensional space spanned by x', y', z'. $R_{\lambda l}(r')$ is the normalized radial wave function for the three-dimensional isotropic harmonic oscillator, and its form is well known. The above wave function is diagonal in W^2 for which the eigenvalue is $l(l+1)M^2$, and l represents the total spin of the hadron in the present case. The quantum number m corresponds to the helicity.

Since the eigenvalue p^2 of the Casimir operator P^2 is constrained to take the numerical values allowed by Eq. (8), the hadronic mass is given by

$$M^2 = m_0^2 + (\lambda + 1). \quad (22)$$

If we relax the subsidiary condition of Eq. (18), we indeed obtain a complete set. In this case, λ of Eq. (17) can become negative for sufficiently large values of k. For $\lambda > 0$, the solutions become

$$\psi_{\beta}^{\lambda lmk}(x) = [\sqrt{\pi}\,2^k k!]^{-1/2}H_k(t')[\exp(-t'^2/2)] \times R_{\lambda+k,l}(r')Y_{lm}(\theta',\phi'). \quad (23)$$

For $\lambda < 0$, the solutions take the form

$$\psi^{\lambda lmk}(x) = [\sqrt{\pi}\,2^{(k-\lambda)}(k-\lambda)!]^{-1/2}H_{k-\lambda}(t') \times [\exp(-t'^2)]R_{k,l}(r')Y_{lm}(\theta',\phi'). \quad (24)$$

The eigenvalues of P^2 and W^2 are again $m_0^2 + (\lambda + 1)$ and $l(l+1)M^2$ respectively. In both of the above cases, k is allowed to take all possible integer values.

The functional forms of Eqs. (23) and (24) are relatively simple, and they suggest that this representation of the Poincaré group corresponds to the solution of the Lorentz-invariant oscillator differential equation in a "cylindrical" coordinate system moving with the hadronic velocity where the t' variable is treated separately. We are then led to the question of why this fact was not known.

Even though the above representations take simple forms, the wave functions contain the following nonconventional features. The first point to note is that they are written as functions of the x', y', z', t' variables. The transverse variables x', y' are simply x and y respectively. However, z' and t' are linear combinations of z and t. Because the physical meaning of the time-separation variable was not clearly understood, the t dependence discouraged us in the past from using it explicitly in representation theory. The explicit use of this variable in the present paper is based on the progress that has been made in our physical understanding of this time-separation variable in terms of measurable quantities, and in terms of the relativistic wave functions carrying a covariant probability interpretation.[2]

Another factor which used to discourage the use of the t variable was that we are accustomed to its appearance through the form

$$x^{\mu}x_{\mu} = t^2 - r^2, \quad (25)$$

where

$$r^2 = x^2 + y^2 + z^2.$$

In terms of this form, it is very inconvenient, if not impossible, to describe functions which are localized in a finite space–time region.

In contrast to the above hyperbolic case, the wave functions which we constructed in this paper are well localized within the region

$$(z'^2 + t'^2) < 2, \quad (26)$$

due to the Gaussian factor appearing in the wave functions. This elliptic form was obtained from the covariant expression

$$-x^{\mu}x_{\mu} + 2(x\cdot p/M)^2 = x'^2 + y'^2 + z'^2 + t'^2. \quad (27)$$

The x' and y' variables have been omitted in Eq. (26) because they are trivial. In terms of z and t, the above inequality takes the form

$$\left[\frac{1-\beta}{1+\beta}(z+t)^2 + \frac{1+\beta}{1-\beta}(z-t)^2\right] < 2. \quad (28)$$

We are therefore dealing with the function localized within an elliptic region defined by this inequality, and can control the t variable in the same manner as we do in the case of the spatial variables appearing in nonrelativistic quantum mechanics. This localization property together with the hyperbolic case is illustrated in Fig. 1.

IV. CONCLUDING REMARKS

We have shown in this paper that the wave functions used in our previous papers are diagonal in the Casimir operators of the Poincaré group, which specify covariantly the mass and total spin of the hadron. These wave functions are well localized in a space–time region, and undergoes elliptic Lorentz deformation.

An addendum to this paper containing a discussion of Lorentz transformation of the physical wave function and a construction of the representation of the homogeneous Lorentz group is relegated to PAPS.[5] It is shown there that solutions of the oscillator equation diagonal in the Casimir operators of the homogeneous Lorentz group are localized within the Lorentz-invariant hyperbolic region illustrated in Fig. 1.

[1]E.P. Wigner, Ann. Math. **40**, 149 (1939).
[2]Y.S. Kim and M.E. Noz, Phys. Rev. D **8**, 3521 (1973), **12**, 129 (1975); **15**, 335 (1977); Y.S. Kim, J. Korean Phys. Soc. **9**, 54 (1976); **11**, 1 (1978); Y.S. Kim and M.E. Noz, Prog. Theor. Phys. **57**, 1373 (1977); **60**, 801 (1978); Y.S. Kim and M.E. Noz, Found. Phys. **9**, 375 (1979); Y.S. Kim, M.E. Noz, and S.H. Oh, "Lorentz Deformation and the Jet Phenomenon," Found. Phys. (to be published). For review articles written for teaching purposes, see Y.S. Kim and M.E. Noz, Am. J. Phys. **46**, 480, 486 (1978). For a review written for the purpose of formulating a field theory of extended hadrons, see T.J. Karr, Ph.D. thesis (University of Maryland, 1976).
[3]L.C. Biedenharn and H. van Dam, Phys. Rev. D **9**, 471 (1974).
[4]T. Takabayashi, Phys. Rev. **139**, B1381 (1965); S. Ishida and J. Otokozawa, Prog. Theor. Phys. **47**, 2117 (1972).
[5]See AIP document no. PAPS JMAPA-20-1336-12 for twelve pages of discussions of the Lorentz transformation of the physical wave functions, and of the representations of the homogeneous Lorentz group. Order by PAPS number and journal reference from American Institute of Physics, Physics Auxiliary Publication Service, 335 East 45th Street, New York, N.Y. 10017. The price is $1.50 for each microfiche (98 pages), or $5 for photocopies of up to 30 pages with $0.15 for each additional page over 30 pages. Airmail additional. Make checks payable to the American Institute of Physics. This material also appears in *Current Physics Microfilm*, the monthly microfilm edition of the complete set of journals published by AIP, on the frames immediately following this journal article.

Bifurcate Killing horizons

J. G. Miller

Department of Mathematics, Texas A&M University, College Station, Texas 77843
(Received 5 April 1978)

It is shown that an analytic spacetime with a bifurcate Killing horizon is locally symmetric with respect to the axis of rotation. It is also shown that if the surface gravity of a Killing horizon is a nonzero constant, then there exists a local prolongation (extension) of the spacetime that contains a bifurcate Killing horizon.

The main result of this paper is a proof that if the surface gravity of a Killing horizon is a nonzero constant, then there exists a local prolongation (extension) of the spacetime that contains a bifurcate Killing horizon. First, Killing vector fields are reviewed in Sec. 1. In Sec. 2, the existence of canonical coordinates on a neighborhood of a zero of a Killing vector field is established. In Sec. 3, Killing horizons are reviewed. The main result is proved in Sec. 4.

1. KILLING VECTOR FIELDS

Let (M,g) be a spacetime of class C^∞ with signature $(+,+,+,-)$ for the Lorentz metric g. In some cases, analyticity assumptions will be made. A vector field ξ defined on the manifold M or any open subset of M is a Killing vector field if and only if the Lie derivative $L_\xi g$ is zero. The Lie derivative of g with respect to a vector field ξ is given by

$$(L_\xi g)(\eta,\zeta)=\xi\,(g(\eta,\zeta))-g([\xi,\eta],\zeta)-g(\eta,[\xi,\zeta\,]) \tag{1}$$

for arbitrary vector fields η and ζ. The Lie derivative of a vector field η with respect to a vector field ξ is given by $L_\xi\eta=[\xi,\eta]$, the Lie bracket of ξ and η. For a given vector field ξ, define a tensor field A_ξ of type (1,1) by

$$A_\xi\eta=-\nabla_\eta\xi \tag{2}$$

for arbitrary vector field η. The pseudo-Riemannian connection is denoted by ∇. If ξ vanishes at a point p in M, then A_ξ evaluated at p is an endomorphism of the tangent space T_pM which indicates how the local flow of ξ is rotating around p. The torsion tensor field T of the pseudo-Riemannian connection ∇ vanishes identically. Thus,

$$T(\eta,\zeta)=\nabla_\eta\zeta-\nabla_\zeta\eta-[\eta,\zeta\,]=0 \tag{3}$$

for arbitrary vector fields η and ζ; and the tensor field A satisfies

$$A_\xi=L_\xi-\nabla_\xi. \tag{4}$$

The metric tensor field g is covariantly constant with respect to ∇. Thus,

$$(\nabla_\xi g)(\eta,\zeta)=\xi\,(g(\eta,\zeta))-g(\nabla_\xi\eta,\zeta)-g(\eta,\nabla_\xi\zeta)=0; \tag{5}$$

and the Lie derivative (1) may be expressed in terms of A_ξ:

$$(L_\xi g)(\eta,\zeta)=-g(A_\xi\eta,\zeta)-g(\eta,A_\xi\zeta). \tag{6}$$

Hence, ξ is a Killing vector field if and only if A_ξ is skew symmetric with respect to g.

Assume that ξ is a Killing vector field. Let $\{U_\epsilon\}$ be the local one-parameter group of isometries generated by ξ. The orbit of ξ through a point q in M is $\{U_\epsilon(q)\}$. Some results of Boyer[1] are rederived. Since A_ξ is skew symmetric, $\xi\,(g(\xi,\xi))=2g(\nabla_\xi\xi,\xi)=-2g(A_\xi\xi,\xi)=0$. Thus,

Lemma 1: $\xi^2\equiv g(\xi,\xi)$ is constant on any orbit of ξ; the one-dimensional orbits of ξ can be classified as spacelike, timelike or null according to the sign or vanishing of ξ^2. Also,

$$\begin{aligned}d\xi^2(\eta)&=\eta(g(\xi,\xi))=2g(\xi,\nabla_\eta\xi)=-2g(\xi,A_\xi\eta)\\&=2g(A_\xi\xi,\eta)=-2g(\nabla_\xi\xi,\eta).\end{aligned}$$

Hence,

Lemma 2: $-\tfrac{1}{2}d\xi^2=\xi^\flat\circ A_\xi=(\nabla_\xi\xi)^\flat$. If ζ is a vector field, $\zeta^\flat$ is the 1-form defined by $\zeta^\flat(\eta)=g(\zeta,\eta)$ for arbitrary vector field η. The notation $\zeta^\flat$ is due to Abraham.[2] Using Lemma 1,

$$\begin{aligned}(\nabla_\xi d\xi^2)(\eta)&=\xi\,(d\xi^2(\eta))-d\xi^2(\nabla_\xi\eta)\\&=\xi\,(\eta(\xi^2))-(\nabla_\xi\eta\xi)\xi^2\\&=[\xi,\eta]\xi^2-(\nabla_\xi\eta)\xi^2=-(\nabla_\eta\xi)\xi^2\\&=(A_\xi\eta)\xi^2=d\xi^2(A_\xi\eta).\end{aligned}$$

Thus,

Lemma 3: $\nabla_\xi d\xi^2=d\xi^2\circ A_\xi$.

2. CANONICAL COORDINATES

Assume that ξ is a Killing vector field which vanishes at a point p in M but does not vanish identically on some open neighborhood of p. Let $\{U_\epsilon\}$ be the local one-parameter group of isometries generated by ξ. Then $\{L_\epsilon\}\equiv\{(dU_\epsilon)_p\}$ is a local one-parameter group of Lorentz transformations of the tangent space T_pM and $-A_\xi(p)$ is the infinitesimal generator of $\{L_\epsilon\}$; i.e., $(dL_\epsilon/d\epsilon)|_{\epsilon=0}=-A_\xi(p)$. Since U_ϵ maps geodesics into geodesics, there exist open neighborhoods V and W of p with $U_\epsilon(V)\subset W$ for sufficiently small ϵ such that

$$\begin{tikzcd} V \arrow[r, "U_\epsilon"] \arrow[d, "\exp^{-1}"'] & W \arrow[d, "\exp^{-1}"] \\ T_pM \arrow[r, "L_\epsilon"'] & T_pM \end{tikzcd}$$

is a commutative diagram. The exponential mapping[3] at p is denoted by exp.

Let (e_1,e_2,e_3,e_4) be an ordered basis of T_pM and let

 0022-2488/79/071345-04$01.00

(x^1,x^2,x^3,x^4) be the normal coordinate system[3] on W with respect to (e_1,e_2,e_3,e_4). Then $\exp^{-1}(q) = x^\mu(q)e_\mu$ for q in W (summation implied over repeated indices) or $x^\mu(q) = [\exp^{-1}(q)]^\mu$, the μth component of the vector $\exp^{-1}(q)\in T_pM$ with respect to the ordered basis (e_1,e_2,e_3,e_4). In this normal coordinate system, $\xi = \xi^\mu\partial/\partial x^\mu$. For q in V,

$$\begin{aligned}\xi^\mu(q) &= (d\,[x^\mu(U_\epsilon(q))]/d\epsilon)|_{\epsilon=0}\\ &= (d\,[\exp^{-1}(U_\epsilon(q))]^\mu/d\epsilon)|_{\epsilon=0}\\ &= (d\,[L_\epsilon(\exp^{-1}(q))]^\mu/d\epsilon)|_{\epsilon=0}\\ &= (d\,[L_\epsilon(x^\nu(q)e_\nu)]^\mu/d\epsilon)|_{\epsilon=0}\\ &= x^\nu(q)(d\Lambda^\mu{}_\nu/d\epsilon)|_{\epsilon=0},\end{aligned}$$

where $(\Lambda^\mu{}_\nu)$ is the matrix representation of L_ϵ with respect to the ordered basis (e_1,e_2,e_3,e_4). Since $-A_\xi(p)$ is the infinitesimal generator of L_ϵ, $(d\Lambda^\mu{}_\nu/d\epsilon)|_{\epsilon=0}\,e_\mu = -A_\xi(p)\cdot e_\nu$. This result is summarized in

Lemma 4: In a normal coordinate system (x^1,x^2,x^3,x^4) with origin at a zero of ξ, the components of ξ are given by $\xi^\mu = x^\nu(d\Lambda^\mu{}_\nu/d\epsilon)|_{\epsilon=0}$.

The type of the local one-parameter group of Lorentz transformations $\{L_\epsilon\}$ is characterized by two invariants

$$\begin{aligned}I_1&\equiv -\tfrac{1}{2}\mathrm{trace}({*A_\xi}(p)\cdot A_\xi(p))\equiv 2\mu\kappa,\\ I_2&\equiv -\tfrac{1}{2}\mathrm{trace}(A_\xi(p)\cdot A_\xi(p))\equiv\mu^2-\kappa^2.\end{aligned}\tag{7}$$

The star is the duality operator. The classification of local one-parameter groups of Lorentz transformations is reproduced in Table I from Ref. 1. Table I also classifies the fixed points of $\{U_\epsilon\}$ or the zeros of ξ. The Killing vector field ξ is called an infinitesimal 4-screw, spacelike rotation, timelike rotation or null rotation depending on the type of zero of ξ. The invariants μ and κ are defined by (7) except for an overall sign.

For a local one-parameter group $\{L_\epsilon\}$ of Lorentz transformations, it is well known[4] that there exists an orthonormal basis (e_1,e_2,e_3,e_4) with signature $(+,+,+,-)$ such that

$$(\Lambda^\mu{}_\nu) = \begin{pmatrix}\cos(\mu\epsilon) & \sin(\mu\epsilon) & 0 & 0\\ -\sin(\mu\epsilon) & \cos(\mu\epsilon) & 0 & 0\\ 0 & 0 & \cosh(\kappa\epsilon) & \sinh(\kappa\epsilon)\\ 0 & 0 & \sinh(\kappa\epsilon) & \cosh(\kappa\epsilon)\end{pmatrix}\tag{8}$$

if L_ϵ is a 4-screw ($\mu\kappa\neq0$), spacelike rotation ($\kappa=0,\mu\neq0$) or timelike rotation ($\mu=0$, $\kappa\neq0$); and

TABLE I. Classification of local one-parameter groups of Lorentz transformations.

I_1	I_2	Type of Lorentz transformation L_ϵ
$\neq0$		4-screw
0	>0	spacelike rotation
0	<0	timelike rotation
0	0	null rotation

$$(\Lambda^\mu{}_\nu) = \begin{pmatrix}1 & 0 & 0 & 0\\ 0 & 1 & -\epsilon & \epsilon\\ 0 & \epsilon & 1-\epsilon^2/2 & \epsilon^2/2\\ 0 & \epsilon & -\epsilon^2/2 & 1+\epsilon^2/2\end{pmatrix}\tag{9}$$

if L_ϵ is a null rotation.

In Riemann normal coordinates $(x^1,x^2,x^3,x^4) = (x,y,z,t)$ with respect to the orthonormal basis (e_1,e_2,e_3,e_4),

$$\xi = \mu\Big(y\frac{\partial}{\partial x} - x\frac{\partial}{\partial y}\Big) + \kappa\Big(t\frac{\partial}{\partial z} + z\frac{\partial}{\partial t}\Big)\tag{10}$$

if L_ϵ is a 4-screw ($\mu\kappa\neq0$), spacelike rotation ($\kappa=0,\mu\neq0$) or timelike rotation ($\mu=0$, $\kappa\neq0$); and

$$\xi = \Big(y\frac{\partial}{\partial z} - z\frac{\partial}{\partial y}\Big) + \Big(t\frac{\partial}{\partial y} + y\frac{\partial}{\partial t}\Big)\tag{11}$$

if L_ϵ is a null rotation. An infinitesimal null rotation is the sum of an infinitesimal spacelike rotation and an infinitesimal timelike rotation that do not commute.

Note that the zeros of ξ are isolated only if L_ϵ is a 4-screw. In the other cases, the axis of rotation (zeros of ξ) is a totally geodesic two-dimensional submanifold of M and the invariants μ and κ are constant on the axis of rotation. Any coordinate system in which a Killing vector field ξ is given by (10) or (11) will be called a canonical coordinate system. A canonical coordinate system is not necessarily a Riemann normal coordinate system. The Riemann normal coordinate system was used only to establish the existence of a canonical form for ξ on a neighborhood of a zero of ξ.

3. KILLING HORIZONS

A Killing horizon with respect to a Killing vector field ξ is a null hypersurface K on which ξ is null. Assume K is a Killing horizon with respect to a Killing vector field ξ. Since $\xi^2\equiv g(\xi,\xi)$ is constant on K (zero, in fact), there exists a real-valued differentiable function κ on K such that $-\tfrac{1}{2}d\xi^2 = \kappa\xi^\flat$. By Lemma 2, $\nabla_\xi\xi = \kappa\xi$ on K. This implies that the null orbits of ξ on K are null geodesics. On K, $-\tfrac{1}{2}\nabla_\xi d\xi^2 = \nabla_\xi(\kappa\xi^\flat) = (\xi\kappa)\xi^\flat + \kappa\nabla_\xi\xi^\flat = (\xi\kappa)\xi^\flat + \kappa(\nabla_\xi\xi)^\flat = (\xi\kappa)\xi^\flat + \kappa^2\xi^\flat$. Also by Lemmas 2 and 3, $-\tfrac{1}{2}\nabla_\xi d\xi^2 = -\tfrac{1}{2}d\xi^2\circ A_\xi = \kappa\xi^\flat\circ A_\xi = -\tfrac{1}{2}\kappa\,d\xi^2 = \kappa^2\xi^\flat$ on K. Hence, $\xi\kappa=0$ and κ is constant on any null orbit of ξ on K.

Define $\xi(\epsilon)$ to be ξ evaluated along an orbit $\{U_\epsilon(q)\}$. For $q\in K$ and $\kappa(q)=0$, $\xi(\epsilon)$ is parallelly propagated on the null orbit $\{U_\epsilon(q)\}$ and the group parameter ϵ is an affine parameter. For $q\in K$ and $\kappa(q)\neq0$, $e^{-\kappa\epsilon}\xi(\epsilon)$ is parallelly propagated on the null orbit $\{U_\epsilon(q)\}$ since $D(e^{-\kappa\epsilon}\xi(\epsilon))/d\epsilon \equiv\nabla_{\xi(\epsilon)}(e^{-\kappa\epsilon}\xi(\epsilon)) = -\kappa e^{-\kappa\epsilon}\xi(\epsilon) + e^{-\kappa\epsilon}\nabla_{\xi(\epsilon)}\xi(\epsilon) = 0$; and $u = e^{\kappa\epsilon}$ is an affine parameter. Thus, a null orbit of ξ on K is geodesically incomplete if κ is nonzero on the orbit. These results, due to Boyer,[1] are summarized in

Lemma 5: There exists a real-valued differentiable function κ on K such that $-\tfrac{1}{2}d\xi^2 = \kappa\xi^\flat = (\nabla_\xi\xi)^\flat$ and $\xi\kappa = 0$. The null orbits of ξ on K are null geodesics. The function κ is constant on any null orbit of ξ on K and the orbit is geodesically incomplete if κ is nonzero.

A Killing horizon K is said to be nondegenerate if and only if κ is nonzero on K. Hence, every null orbit on a nondegenerate Killing horizon is geodesically incomplete.

4. BIFURCATE KILLING HORIZONS

If ξ is an infinitesimal timelike rotation, then there are four Killing horizons with respect to ξ that have the axis of rotation as a common boundary.[1] The union of the four horizons and the axis of rotation is called a bifurcate Killing horizon. An analytic spacetime with a bifurcate Killing horizon is locally symmetric with respect to the axis of rotation in the following sense:

Theorem 1: If $\xi = t\partial/\partial z + z\partial/\partial t$ is an analytic Killing vector field of an analytic spacetime, then $(x,y,z,t) \rightsquigarrow (x,y,-z,-t)$ is an isometry.

Proof: The calculation is easier in coordinates $u = (t+z)/(2)^{1/2}$ and $v = (t-z)/(2)^{1/2}$. In these coordinates, $\xi = u\partial/\partial u - v\partial/\partial v$ and $U_\epsilon(x,y,u,v) = (x,y,ue^\epsilon,ve^{-\epsilon})$. Let $(x^1,x^2,x^3,x^4) = (x,y,u,v)$ and $1 \leqslant i,j \leqslant 2$. The fact that the metric components $g_{\mu\nu}$ are analytic functions and that U_ϵ is an isometry imply $g_{ij}(x,y,u,v) = \tilde{g}_{ij}(x,y,uv)$, $g_{34}(x,y,u,v) = \tilde{g}_{34}(x,y,uv)$, $g_{i3}(x,y,u,v) = v\tilde{g}_{i3}(x,y,uv)$, $g_{i4}(x,y,u,v) = u\tilde{g}_{i4}(x,y,uv)$, $g_{33}(x,y,u,v) = v^2\tilde{g}_{33}(x,y,uv)$, and $g_{44}(x,y,u,v) = u^2\tilde{g}_{44}(x,y,uv)$, where the $\tilde{g}_{\mu\nu}$ are analytic functions of three variables. It follows immediately that $(x,y,u,v)\ \ (x,y,-u,-v)$ is an isometry.

The theorem is not true if the analyticity assumption is dropped. The metric given by

$$ds^2 = dx^2 + dy^2 + f^2(z,t)[dz^2 - dt^2], \tag{12}$$

$$f^2(z,t) = \begin{cases} 1, & \text{when } t \geqslant z \text{ or } t \geqslant -z. \\ 1 + \exp[-(z^2 - t^2)^{-2}], & \\ & \text{when } t < z \text{ and } t < -z \end{cases}$$

is of class C^∞ and $\xi = t\,\partial/\partial z + z\,\partial/\partial t$ is a Killing vector field. The transformation $(x,y,z,t) \rightsquigarrow (x,y,-z,-t)$ is not an isometry.

Boyer[1] has shown that if an incomplete null geodesic orbit of a Killing vector field ξ on a Killing horizon K is extendible, then K is a branch of a bifurcate Killing horizon and ξ is an infinitesimal timelike rotation. The axis of rotation consists of limit points of the incomplete null orbits of ξ. This implies that the κ of Lemma 5 is constant on K since it can be shown that its limit on the axis of rotation is equal to the invariant κ (or possibly $-\kappa$) in (7). The main result of this paper is to prove that if κ is constant on K, then there exists a local prolongation (extension) of the space–time that contains a bifurcate Killing horizon.

Theorem 2: Let (M,g) be an analytic spacetime with nondegenerate Killing horizon K with respect to an analytic Killing vector field ξ. If κ is a constant function on K, then there exists an open submanifold U of M that intersects the horizon K and an analytic prolongation of (U,g) that contains a bifurcate Killing horizon.

Proof: By an application of the theorem that states that a nonzero vector field can be represented locally as a coordinate vector field,[2] there exists a coordinate system (x^1,x^2,x^3,x^4) such that the Killing horizon K is given locally by $x^3 = 0$ and $\xi = \partial/\partial x^4$. Let U be the domain of the coordinate system (x^1,x^2,x^3,x^4). Since ξ is a Killing vector field, $\partial g_{\mu\nu}/\partial x^4 = 0$. The range of the coordinate x^4 in U may not be the whole real line. First, consider the prolongation of (U,g) that extends the range of the coordinate x^4 to the whole real line, if it is not so already. This local prolongation may not be an open submanifold of M. Taub-NUT space, in which the orbits of ξ are closed, is such an example.[5]

Since ξ is null on K, $\xi^2 = g_{44}(x^1,x^2,0) = 0$; and since K is a null hypersurface $g^{33}(x^1,x^2,0) = 0$. On K, $-\frac{1}{2}d\xi^2 = \kappa\xi^\flat$, $d\xi^2 = (\partial g_{44}/\partial x^3)(x^1,x^2,0)\,dx^3$ and $\xi^\flat = g_{14}(x^1,x^2,0)\,dx^1 + g_{24}(x^1,x^2,0)\,dx^2 + g_{34}(x^1,x^2,0)\,dx^3$. Thus, $g_{14}(x^1,x^2,0) = g_{24}(x^1,x^2,0) = 0$. Since the metric g has Lorentz signature, $g_{34}(x^1,x^2,0) \neq 0$. Thus,

$$-\tfrac{1}{2}(\partial g_{44}/\partial x^3)(x^1,x^2,0) = \kappa g_{34}(x^1,x^2,0). \tag{13}$$

Now consider the coordinate transformation

$$x = x^1, \quad y = x^2, \quad u = e^{\kappa x^4}, \quad v = x^3 e^{-\kappa x^4} \tag{14}$$

with inverse

$$x^1 = x, \quad x^2 = y, \quad x^3 = uv, \quad x^4 = \kappa^{-1}\ln u. \tag{15}$$

Then

$$\begin{aligned} ds^2 = g_{\mu\nu}\,dx^\mu\,dx^\nu &= g_{44}(x,y,uv)(\kappa u)^{-2}\,du^2 \\ &\quad + 2g_{34}(x,y,uv)(\kappa u)^{-1}(u\,dv + v\,du)\,du \\ &\quad + 2g_{24}(x,y,uv)(\kappa u)^{-1}\,dy\,du \\ &\quad + 2g_{14}(x,y,uv)(\kappa u)^{-1}\,dx\,du + \cdots \\ &= [g_{44}(x,y,uv)(\kappa u)^{-1} + 2g_{34}(x,y,uv)\,v] \\ &\quad \times(\kappa u)^{-1}\,du^2 + \cdots. \end{aligned} \tag{16}$$

The transformation (14) maps the extended domain of (x^1,x^2,x^3,x^4) onto the half-space $u > 0$; and the terms omitted in (16) are clearly analytic on the symmetric extension of the half-space, $(x,y,u,v) \rightsquigarrow (x,y,-u,-v)$, including $u = 0$. The term in brackets in (16) is also analytic on the symmetric extension.

From (13),

$$\lim_{u\to 0}[g_{44}(x,y,uv)(\kappa u)^{-1} + 2g_{34}(x,y,uv)\,v]$$
$$= (\partial g_{44}/\partial x^3)(x,y,0)\kappa^{-1}\,v + 2g_{34}(x,y,0)\,v = 0.$$

Thus, the metric components in the (x,y,u,v) coordinate system are analytic functions on the symmetric extension of the half-space $u > 0$. It is easy to check that this prolongation of (U,g) has Lorentz signature, even when $u = 0$. In this prolongation, $\xi = \kappa(u\partial/\partial u - v\partial/\partial v)$ is an infinitesimal timelike rotation. Canonical coordinates are given by $z = (u-v)/(2)^{1/2}$ and $t = (u+v)/(2)^{1/2}$.

Remark 1: Theorem 2 is a generalization of previous work on analytic extensions[6,7] that does not require the existence of special two-dimensional timelike submanifolds. In Theorem 2, no local symmetry other than that generated by ξ is assumed to exist and no field equations have been imposed on the metric. The Einstein field equations and the dominant energy condition imply that κ is a constant and κ is called the surface gravity of the horizon.[8]

Remark 2: Theorem 2 is a local theorem. Taub-NUT space is an example of a spacetime that satisfies the hypothesis of Theorem 2 and yet there is no prolongation of Taub-NUT space that contains a bifurcate Killing horizon.[5] The reason for this is that the orbits of ξ are closed. Thus, a global version of Theorem 2 is not true. Theorem 2 implies that there exists a local prolongation of Taub-NUT space that contains a bifurcate Killing horizon. Such a local prolongation of Taub-NUT space has already been exhibited.[5]

Remark 3: The analyticity assumption in Theorem 2 was made in order to guarantee uniqueness of the local prolongation. If the analyticity assumption is dropped, the theorem still goes through to yield a local symmetric prolongation that contains a bifurcate Killing horizon. However, in the nonanalytic case, there also exist nonsymmetric prolongations that contain a bifurcate Killing horizon. For example, the Minkowski metric $ds^2 = dx^2 + dy^2 + dz^2 - dt^2$ defined on the half-space $t > -z$ admits the analytic (symmetric) prolongation to Minkowski space and the C^∞ (nonsymmetric) prolongation to metric (12).

[1]R.H. Boyer, Proc. R. Soc. London A **311**, 245 (1969).
[2]R. Abraham, *Foundations of Mechanics* (Benjamin, New York, 1967).
[3]S. Kobayashi and K. Nomizu, *Foundations of Differential Geometry* (Interscience, New York, 1963), Vol. I.
[4]J. Ehlers, W. Rindler, and I. Robinson, *Quaternions, Bivectors, and the Lorentz Group, in Perspectives in Geometry and Relativity* (Indiana U.P., Bloomington, Indiana, 1966).
[5]J.G. Miller, M.D. Kruskal, and B. Godfrey, Phys. Rev. D **4**, 2945 (1971).
[6]M. Walker, J. Math. Phys. **11**, 2280 (1970).
[7]B. Godfrey, J. Math. Phys. **12**, 606 (1971).
[8]J.M. Bardeen, B. Carter, and S.W. Hawking, Commun. Math. Phys. **31**, 161 (1973).

Decay of local correlations and absence of phase transitions

Joachim Messer

Institut für Theoretische Physik, Universität Göttingen, D-3400 Göttingen, Federal Republic of Germany
(Received 25 April 1978)

If the local truncated correlation functions of a system of statistical mechanics decay in a prescribed manner the limiting pressure becomes differentiable with respect to the activity. Criterions of clustering of the local truncated m-point correlation functions are shown to lead to a pressure being element of C^n for arbitrary n and $m = 1,\ldots,n+1$.

1. INTRODUCTION

It is a well-known law of thermodynamics that the local fluctuations $(\Delta n)_\Lambda$ of the local particle number n_Λ behave for large n_Λ like

$$\frac{(\Delta n)_\Lambda}{n_\Lambda} \sim \frac{1}{\sqrt{n_\Lambda}}.$$

If this behavior is expressed in a rigorous inequality,

$$\frac{(\Delta n)_\Lambda}{n_\Lambda} \geqslant \frac{c}{\sqrt{n_\Lambda}}, \tag{1.1}$$

for some small constant $c > 0$, then for the infinite system the pressure turns out to be a continuous function of the specific volume ρ^{-1}. It was proved by Ruelle for classical continuous systems that (1.1) generally holds for a large class of potentials.[1,2]

Assuming the reversed inequality

$$\frac{(\Delta n)_\Lambda}{n_\Lambda} \leqslant \frac{C}{\sqrt{n_\Lambda}}. \tag{1.2}$$

for some large constant $C > 0$, we shall show that this decay property for the local truncated two-point correlation function (in the grand canonical ensemble) forces the pressure of the infinite system to be differentiable with respect to the activity.

The task of the following two sections is to give a rigorous proof and formulation of the connection of the limited growth of the local particle number fluctuations like (1.2), and the absence of phase transitions. In Sec. 2 we concentrate on simple differentiability, mainly using convexity arguments, whereas in Sec. 3 we generalize to n-fold differentiability for arbitrary n. For the infinite system, n-fold differentiability of the pressure is interpreted according to Ehrenfest as absence of a phase transition of the nth kind.

We are not concerned in this article with a special type of system of statistical mechanics. The results are obligatory for all types of systems, classical and quantum as well as lattice and continuous systems, provided the local pressures exist, and the interaction is either stable[1] in case of a continuous system, or an element of the usual Banach space[1] for lattice systems. For lattice systems the term "stable" used in the text has to be replaced by "element of the Banach space," and the constant B in an appropriate way by the Banach space norm $\|\phi\|$. All notations and definitions used are based on Ref. 1.

2. PHASE TRANSITION OF FIRST ORDER

In this section we are concerned with the cluster properties of the local truncated two-point correlation function. For convenience and to elucidate the connection of the derivatives of the pressure with respect to the chemical potential and the truncated correlation functions, we treat classical continuous systems. The propositions are also valid for other systems of statistical mechanics, in particular Proposition (2.1) without change and Proposition (2.2) as a special case of Proposition (3.1).

Let p_Λ be the pressure of a statistical system confined to a bounded Lebesgue measurable set $\Lambda \subset \mathbb{R}^\nu$ with measure $|\Lambda|$, in the grand canonical ensemble. If ζ denotes the chemical potential, β the inverse temperature, and $\rho_{\Lambda,k}(x_1,\ldots,x_k)$ the local k-point correlation function for classical continuous systems,[1] then $p_\Lambda \in C^\infty$ and

$$\frac{\partial^2 \beta p_\Lambda}{\partial \zeta^2} = \frac{1}{|\Lambda|}\int_{\Lambda^2} (\rho_{\Lambda,2}(x,y) - \rho_{\Lambda,1}(x)\rho_{\Lambda,1}(y))dx\,dy + \frac{1}{|\Lambda|}\int_\Lambda \rho_{\Lambda,1}(x)dx.$$

Proposition (2.1): If there exists $k > 0$ such that

$$\frac{\partial^2 \beta p_{\bar\Lambda}}{\partial \zeta^2} \geqslant k > 0$$

for a subnet $\bar\Lambda \subset \mathbb{R}^\nu$ with $\bar\Lambda \to \infty$ and each $\zeta \in \mathbb{R}$, and if the thermodynamic limit $\lim_{\Lambda\to\infty} p_\Lambda = p$ exists, then p is a strictly convex function of ζ for each $\zeta \in \mathbb{R}$.

Proof: Suppose there exist $\zeta_1,\zeta_2 \in \mathbb{R}$, $\zeta_1 \neq \zeta_2$ and p is affine for all $\zeta \in [\zeta_1,\zeta_2]$, then $\rho = (\partial\beta p/\partial\zeta) = \rho_0 = \text{const} \geqslant 0$ exists. Since $\rho_{\bar\Lambda} = (\partial\beta p_{\bar\Lambda}/\partial\zeta)$ is monotone increasing and converges to ρ_0 for $\zeta \in [\zeta_1,\zeta_2]$, one can choose $\epsilon > 0$ so small, that $\epsilon/(|\zeta_2 - \zeta_1|) < k$. $\rho_{\bar\Lambda}$ is a C^∞ function, and there exists $\tilde\Lambda \in \{\bar\Lambda \mid \bar\Lambda \subset \mathbb{R}^\nu, \bar\Lambda \to \infty\}$ and $\zeta_0 \in [\zeta_1,\zeta_2]$, such that

$$\frac{\partial p_{\tilde\Lambda}(\zeta_0)}{\partial \zeta_0} = \frac{\epsilon}{|\zeta_2 - \zeta_1|} < k$$

 0022-2488/79/071349-03$01.00

which contradicts our presupposition. Q.E.D.

A direct consequence of the strict convexity of the pressure is the continuity of the pressure as a function of the density in its domain of existence.

An example for classical continuous systems is given by superstable, lower regular, and regular pair interactions. It is an easy consequence of Ref. 2 and Proposition (2.1) that the pressure is a strictly convex function of the chemical potential.

Proposition (2.2): Let the thermodynamic limit

$$p(\zeta,\beta) = \lim_{\Lambda\to\infty} p_\Lambda(\zeta,\beta)$$

exist for $(\zeta,\beta)\in D\subset\mathbb{R}^2$ with $\beta>0$ and a stable but arbitrary interaction. Let $\rho^{(\lambda)}_{\Lambda,k}(x_1,\ldots,x_k)$ be the local k-point correlation function for the inverse temperature β and chemical potential $\zeta_\lambda = (1-\lambda)\zeta_1 + \lambda\zeta_2$, $\lambda\in[0,1]$, such that for all $\lambda\in[0,1]$ and all $(\beta,\zeta_\lambda)\in D$ and for a subnet $\{\bar\Lambda\,|\,\bar\Lambda\subset\mathbb{R}^\nu, \bar\Lambda\to\infty\}$ we have

$$\sup_{\bar\Lambda}\frac{1}{|\bar\Lambda|}\int_0^1 d\lambda\int_{\bar\Lambda^2} dx\,dy[\rho^{(\lambda)}_{\bar\Lambda,2}(x,y)-\rho^{(\lambda)}_{\bar\Lambda,1}(x)\rho^{(\lambda)}_{\bar\Lambda,1}(y)] \leqslant g(\beta,\zeta_1,\zeta_2), \tag{2.1}$$

with $g(\beta,\cdot,\cdot)$ being a continuous function of ζ_1 and ζ_2. Then $p(\cdot,\beta)$ is differentiable for each $\zeta\in\mathbb{R}$ with $(\beta,\zeta)\in D$ and there is no phase transition of first order.

Proof: The local pressure is differentiable with respect to ζ, and for each $\bar\zeta_1,\bar\zeta_2\in\mathbb{R}$, each $\zeta_1\in[\bar\zeta_1,\bar\zeta_2]$, and $\zeta_2\in[\bar\zeta_1,\bar\zeta_2]$ such that $(\beta,\bar\zeta_1)\in D$, $(\beta,\bar\zeta_2)\in D$ we have

$$\left|\frac{\partial}{\partial\zeta}p_{\bar\Lambda}(\zeta,\beta)|_{\zeta_1}-\frac{\partial}{\partial\zeta}p_{\bar\Lambda}(\zeta,\beta)|_{\zeta_2}\right| \leqslant\left|\int_0^1 d\lambda\left(\frac{\partial}{\partial\lambda}\left(\frac{\partial}{\partial\zeta}p_{\bar\Lambda}(\zeta,\beta)\Big|_{(1-\lambda)\zeta_2+\lambda\zeta_1}\right)\right)\right| \leqslant K|\zeta_1-\zeta_2|\beta^{-1},$$

with

$$K=\sup_{\bar\Lambda}\left[\frac{1}{|\bar\Lambda|}\int_0^1 d\lambda\int_{\bar\Lambda^2}[\rho^{(\lambda)}_{\bar\Lambda,2}(x,y) -\rho^{(\lambda)}_{\bar\Lambda,1}(x)\rho^{(\lambda)}_{\bar\Lambda,1}(y)]\,dx\,dy+\frac{1}{|\bar\Lambda|}\int_0^1 d\lambda\int_{\bar\Lambda}\rho^{(\lambda)}_{\bar\Lambda,1}(x)\,dx\right].$$

The interchanging of limits is justified by the stability of the interaction and the application of the theorems of Lebesgue, Fubini–Tonelli, and the Weierstrass majorization criterion. $p_{\bar\Lambda}(\zeta,\beta)$ is a convex function of ζ with positive derivatives $(\partial/\partial\zeta)p_{\bar\Lambda}(\zeta,\beta)\geqslant 0$. Let $\bar z>0$ ($\bar\zeta=\ln\bar z$), then convexity implies

$$\frac{1}{2}\frac{1}{|\Lambda|}\int_\Lambda\rho^{(\bar z)}_{\Lambda,1}(x)\,dx\leqslant\int_{\bar z}^{2\bar z}\frac{\beta}{z}\frac{\partial}{\partial\zeta}p_\Lambda(\zeta,\beta)\Big|_{\ln z}dz = \beta p_\Lambda(\ln 2\bar z,\beta)-\beta p_\Lambda(\ln\bar z,\beta) \leqslant 2\bar z e^{\beta B}.$$

B denotes the stability lower bound, and $\rho^{(\bar z)}_{\Lambda,1}(x)$ is the local one-point correlation function for activity $\bar z$ and inverse temperature β. Since g is a continuous function there exists a constant $G>0$ such that

$$K\leqslant\max_{\zeta_1,\zeta_2\in[\bar\zeta_1,\bar\zeta_2]}K(\zeta_1,\zeta_2)\leqslant G,$$

and the local densities are uniformly Lipschitz; in particular $\rho_{\bar\Lambda}(\zeta)$ is equicontinuous,

$$|\rho_{\bar\Lambda}(\zeta_1,\beta)-\rho_{\bar\Lambda}(\zeta_2,\beta)|\leqslant G|\zeta_1-\zeta_2|. \tag{2.2}$$

Convexity of the pressure as a function of ζ implies that $(\partial/\partial\zeta)p_{\bar\Lambda}(\zeta,\beta)$ converges to $(\partial/\partial\zeta)p(\zeta,\beta)$ for each $\zeta\in\mathbb{R}\setminus Q$, where Q is at most a countably infinite subset of $\mathbb{R}$ [see Ref. 3, Ref. 4 (Lemma III), Ref. 5. (Appendix A) for sequences and Ref. 6 for generalized sequences]. Since $\mathbb{R}\setminus Q$ is dense in $\mathbb{R}$ and the family $((\partial/\partial\zeta)p_{\bar\Lambda}(\zeta,\beta))_{\bar\Lambda\subset\mathbb{R}^\nu}$ is equicontinuous because of (2.2), $(\partial/\partial\zeta)p_{\bar\Lambda}(\zeta,\beta)$ converges for each $\zeta\in\mathbb{R}$. The convex function $p(\cdot,\beta)$ possesses right and left derivatives at each point $\zeta\in\mathbb{R}$ (see Ref. 7). Let A be a directed set. $(\zeta_\alpha)_{\alpha\in A}$ with $\lim_{\alpha\to\infty}\zeta_\alpha=\zeta_0$ denote a generalized sequence of chemical potentials. For each generalized subsequence $(\bar\Lambda_\alpha)_{\alpha\in A}$ with $\lim_{\alpha\to\infty}\bar\Lambda_\alpha=\infty$ it follows from (2.2) that

$$\lim_{\alpha\to\infty}\frac{\partial}{\partial\zeta}p_{\bar\Lambda_\alpha}(\zeta,\beta)\Big|_{\zeta_\alpha}=\lim_{\alpha\to\infty}\frac{\partial}{\partial\zeta}p_{\bar\Lambda_\alpha}(\zeta,\beta)\Big|_{\zeta_0} =\lim_{\bar\Lambda_\alpha\to\infty}\frac{\partial}{\partial\zeta}p_{\bar\Lambda}(\zeta,\beta)\Big|_{\zeta_0}.$$

To each $\zeta_0\in Q$ there exist generalized sequences $(\zeta_\alpha)_{\alpha\in A}$, $\zeta_\alpha\in\mathbb{R}$, $\lim_{\alpha\to\infty}\zeta_\alpha=\zeta_0$, and $(\zeta_{\beta'})_{\beta'\in A}$, and $\zeta_{\beta'}\in\mathbb{R}$, $\lim_{\beta'\to\infty}\zeta_{\beta'}=\zeta_0$ which approximate the right and left derivatives of the pressure[6]:

$$\lim_{\alpha\to\infty}\frac{\partial}{\partial\zeta}p_{\bar\Lambda_\alpha}(\zeta,\beta)\Big|_{\zeta_\alpha}=\frac{\partial^+}{\partial\zeta}p(\zeta,\beta)\Big|_{\zeta_0},$$

$$\lim_{\beta'\to\infty}\frac{\partial}{\partial\zeta}p_{\bar\Lambda_{\beta'}}(\zeta,\beta)\Big|_{\zeta_{\beta'}}=\frac{\partial^-}{\partial\zeta}p(\zeta,\beta)\Big|_{\zeta_0}.$$

The equation

$$\frac{\partial^+}{\partial\zeta}p(\zeta,\beta)\Big|_{\zeta_0}=\lim_{\alpha\to\infty}\frac{\partial}{\partial\zeta}p_{\bar\Lambda_\alpha}(\zeta,\beta)\Big|_{\zeta_\alpha} =\lim_{\bar\Lambda\to\infty}\frac{\partial}{\partial\zeta}p_{\bar\Lambda}(\zeta,\beta)\Big|_{\zeta_0} =\lim_{\beta'\to\infty}\frac{\partial}{\partial\zeta}p_{\bar\Lambda_{\beta'}}(\zeta,\beta)\Big|_{\zeta_{\beta'}} =\frac{\partial^-}{\partial\zeta}p(\zeta,\beta)\Big|_{\zeta_0}$$

concludes the proof. Q.E.D.

An immediate consequence of Proposition (2.2) is the continuity of the density as a function of the pressure $\rho(p)$. Since ρ is bounded by $4\,e^{\zeta}e^{\beta B}$, and assuming that $p(\zeta)$ is everywhere differentiable, $\rho(\zeta)$ is continuous for all $\zeta\in\mathbb{R}$.

The condition $(\beta,\zeta)\in D$ places no serious restriction and shall be omitted in this argument. If $p\neq 0$, the pressure is strictly monotonic-increasing and continuous as a function of ζ. Let ζ_0 be chosen such that $p(\zeta)>0$ for all $\zeta>\zeta_0$ and $p(\zeta)=0$ for all $\zeta\leqslant\zeta_0$, then $p(\zeta)$ is invertible for $\zeta>\zeta_0$ and $\zeta(p)$ is even differentiable. Thus $\rho(p)$ is continuous.

3. PHASE TRANSITION OF ARBITRARY ORDER

In the theorem of this section a generalization of Proposition (2.2) to an arbitrary degree of differentiability of the pressure is obtained. This demonstrates the exclusion of phase transitions of nth order, provided the first $n+1$ truncated local correlation functions decay sufficiently strongly. This is expressed in Presupposition (3.1).

Proposition (3.1): Let $f_\beta:K\rightarrow\mathbb{R}$ with $f_\beta\in C^\infty(K,\mathbb{R})$ be a generalized subsequence of $f_\alpha\in C^\infty(K,\mathbb{R})$ and K a compact space. For $n\geqslant 1$ and $n\in\mathbb{N}$ the following assumption shall be satisified: To each $k\in\mathbb{N}$ with $1\leqslant k\leqslant n+1$ we assume that

$$\left|\frac{\partial^k}{\partial\zeta^k}f_\beta(\zeta)\right|\leqslant g_k(\zeta), \tag{3.1}$$

where g_k is a continuous function on K for each $1\leqslant k\leqslant n+1$, and furthermore $f_\alpha(\zeta)$ converges pointwise to $f(\zeta)$ for each $\zeta\in K$.

Then $(\partial^k/\partial\zeta^k)f_\alpha(\zeta)$ converges uniformly to $f_k^*(\zeta)$ for all $0\leqslant k\leqslant n$ [with $(\partial^0 f_\alpha/\partial\zeta^0)=f_\alpha$ and $f_0^*=f$] and $f_k^*(\zeta) = (\partial^k/\partial\zeta^k)f(\zeta)$ for each $\zeta\in K$. In particular $f\in C^n(K,\mathbb{R})$.

Proof: Let $G_{k,K}=\max\limits_{\zeta\in K} g_k(\zeta)$, then equicontinuity (or uniform Lipschitz continuity) holds for $0\leqslant k\leqslant n$:

$$\left|\frac{\partial^k}{\partial\zeta^k}f_\beta(\zeta_1)-\frac{\partial^k}{\partial\zeta^k}f_\beta(\zeta_2)\right|\leqslant G_{k+1,K}\,|\zeta_1-\zeta_2| \tag{3.2}$$

for all $\zeta_1,\zeta_2\in K$. Since we have

$$\left|\frac{\partial^k}{\partial\zeta^k}f_\beta(\zeta)\right|\leqslant G_{k,K}$$

the theorem of Arzela–Ascoli implies that there exists a subsequence $(\partial^k/\partial\zeta^k)f_\gamma(\zeta)$ of $(\partial^k/\partial\zeta^k)f_\beta(\zeta)$ which converges pointwise to a function $f_k^{**}(\zeta)$ for all $\zeta\in K$. As a consequence of (3.2), $(\partial^k/\partial\zeta^k)f_\gamma(\zeta)$ converges uniformly on K. Since f_γ converges to f, it follows immediately by uniform convergence of the $(\partial^k/\partial\zeta^k)f_\gamma(\zeta)$ for $0\leqslant k\leqslant n$ that $f_k^{**}(\zeta)=f_k^*(\zeta)=(\partial^k/\partial\zeta^k)f(\zeta)$.

Q.E.D.

In particular Proposition (3.1) holds if f_α is the local pressure, $\zeta\in\mathbb{R}$ the chemical potential, and K some interval $[\zeta_1,\zeta_2]$ depending on the temperature such that $(\beta,\zeta_1)\in D$ and $(\beta,\zeta_2)\in D$.

Corollary (3.2): In each system of statistical mechanics with stable interactions for which the pressure converges pointwise with respect to the chemical potential, it also converges uniformly on each compact set. This is a consequence of convexity.

In classical statistical mechanics of continuous systems we meet the following illustrating cases. Let the potential be a positive pair interaction, which decreases at infinity weaker than $1/|x|^{\nu-\delta}$, where $\nu\in\mathbb{N}$ is the dimension and $\delta>0$; the density and pressure is then zero. If it decreases exactly like $a/|x|^\nu$, then Presupposition (2.1) is fulfilled and the absence of a phase transition of first order follows in the region.

$$e^\zeta<\epsilon e^{\beta a\epsilon},$$

because the pressure converges according to Ref. 2. ϵ is the (nonzero) density at low activities. [If the density is zero everywhere, the relation is valid for arbitrary ϵ.] If the potential decreases stronger than $1/|x|^{\nu+\delta}$, then it is regular, and the Kirkwood–Salsburg region and its extension for positive potentials is known as the domain of analyticity of the pressure. In this case the assumptions of Proposition (3.1) are also fulfilled for arbitrary n, because there is at least a subnet $(\partial^k/\partial\zeta^k)p_{\tilde\Lambda}(\zeta)$ that converges uniformly to the continuous function $(\partial^k/\partial\zeta^k)p(\zeta)$ for a compact interval $[\zeta_1,\zeta_2]$ with $\zeta\in[\zeta_1,\zeta_2]$.

ACKNOWLEDGMENTS

The author would like to thank Professor Dr. H. J. Borchers and Dr. H. Roos for valuable discussions.

[1] D. Ruelle, *Statistical Mechanics. Rigorous Results* (Benjamin, New York, 1974).
[2] D. Ruelle, Commun. Math. Phys. **18**, 127–59 (1970).
[3] G.H. Hardy, J.E. Littlewood, and G. Pólya, *Inequalities* (Cambridge U.P., Cambridge, 1952), Chapter 3.1.8, p.94.
[4] M.E. Fisher, J. Math. Phys. **6**, 1643–53 (1965).
[5] R.B. Griffith, J. Math. Phys. **5**, 1215–25 (1964).
[6] H. Roos, "Consequences of convexity: A variational principle for the pressure and related results," to be published in Rep. Math. Phys. (1978).
[7] R.T. Rockafellar, *Convex Analysis* (Princeton U.P., Princeton, New Jersey, 1970).

On stationary, axially symmetric solutions of Jordan's five-dimensional, unified theory

J. Burzlaff and D. Maison

Max-Planck-Institut für Physik und Astrophysik, München, Federal Republic of Germany
(Received 1 December 1978)

It is shown that the association of a linear eigenvalue problem for solutions of Einstein's equations admitting a two-parameter Abelian group of isometries can be extended to Jordan's five-dimensional, unified theory admitting three commuting Killing vectors. The reduction to a two-dimensional problem, the derivation of infinitely many conservation laws and the generation of one-parameter families of solutions can thereby be transcribed almost literally.

Jordan's[1] five-dimensional, unified theory can serve as a first step towards a geometrical unification of gravitation with other interactions. Thus one should try to learn as much as possible about the structure and especially about the solutions of this theory. To reach this aim, the easiest approach could be a generalization of techniques which are known within Einstein's theory of gravitation.

For Einstein's theory of gravitation admitting two commuting Killing vectors one of the authors[2] has recently constructed a linear eigenvalue problem similar to the ones known for completely integrable systems. In this paper we transcribe this construction to Jordan's theory admitting three commuting Killing vectors. We find that in spite of increasing technical difficulties the generalization is straightforward.

The derivation of the linear eigenvalue problem proceeds as in the four-dimensional theory. In Sec. 1 we reduce the problem of finding solutions to Jordan's equations with three commuting Killing vectors to a two-dimensional one. In Sec. 2 we make use of the existence of infinitely many conservation laws to derive a symmetry transformation. With the help of this transformation we construct a linear eigenvalue problem in Sec. 3.

1. REDUCTION TO A TWO-DIMENSIONAL PROBLEM

As in the case of four-dimensional Einstein space admitting two commuting Killing vectors, looking for solutions of the five-dimensional generalization admitting a three-parameter Abelian group G_3 can be reduced to a problem on a two-dimensional manifold S. To see this, we define with the three commuting Killing vectors ξ_i $(i = 1,2,3)$ the projection

$$*\Pi^{\mu}{}_{\nu} \equiv \delta^{\mu}{}_{\nu} - \lambda^{ik} \xi_i^{\mu} \xi_{k\nu} \tag{1.1}$$

$(\mu,\nu,\cdots = 1,\ldots,5)$, where λ^{ik} is the inverse of the 3×3 matrix

$$\lambda_{ik} \equiv \xi_i^{\mu} \xi_k^{\nu} g_{\mu\nu}. \tag{1.2}$$

$g_{\mu\nu}$ is the metric tensor in five dimensions with signature $(+,+,+,-,+)$.

The projection of $g_{\mu\nu}$ defines a metric tensor h_{ab} $(a,b,\cdots = 1,2)$ on S. Two cases have to be distinguished:

(A) The Killing vector fields are all spacelike; hence $\operatorname{sgn}(\lambda_{ik}) = (+,+,+)$ and $\operatorname{sgn}(h_{ab}) = (-,+)$.

(B) One of the Killing vector fields is time-like (stationarity) and hence $\operatorname{sgn}(\lambda_{ik}) = (+,+,-)$ and $\operatorname{sgn}(h_{ab}) = (+,+)$.

Whenever necessary we shall discuss the two cases separately referring to them as (A) and (B).

Because of $R_{\mu\nu} = 0$ the vector fields

$$\Omega_{ijk\mu} \equiv \epsilon_{\mu\nu\kappa\rho\sigma} \xi_i^{\nu} \xi_j^{\kappa} \nabla^{\rho} \xi_k^{\sigma} \tag{1.3}$$

are curl-free,

$$\Omega_{ijk\,[\mu,\nu]} = 0, \tag{1.4}$$

and can thus be derived from potentials Ω_{ijk},

$$\Omega_{ijk\mu} = \partial_{\mu} \Omega_{ijk}. \tag{1.5}$$

In order to be well defined on S, the Ω_{ijk} have to be constant along the orbits of G_3, i.e.,

$$_{\xi_i}\mathscr{L} \Omega_{jkl} = \xi_i^{\mu} \Omega_{jkl\mu} = 0, \tag{1.6}$$

which will be assumed from now on. For

$$\Omega_i = \epsilon_{\mu\nu\kappa\rho\sigma} \xi_1^{\mu} \xi_2^{\nu} \xi_3^{\kappa} \nabla^{\rho} \xi_i^{\sigma} \tag{1.7}$$

we now get

$$\xi_i^{\mu} \Omega_{jkl\mu} = \epsilon_{ijk} \Omega_l \tag{1.8}$$

and with Eq. (1.6)

$$\Omega_l = 0. \tag{1.9}$$

Because

$$\partial_{\mu} \Omega_l = 0 \tag{1.10}$$

holds anyway, this is not an essential restriction.

For the covariant derivative of the Killing vector fields we have

$$\nabla_{\mu} \xi_{i\nu} = \tfrac{2}{3} \lambda^{jk} \xi_{j[\nu} \nabla_{\mu]} \lambda_{ki} - \tfrac{1}{12} \lambda^{jk} \lambda^{lm} \epsilon_{\mu\nu\kappa\rho\sigma} \xi_j^{\kappa} \xi_l^{\rho} \Omega^{\sigma}_{kmi}, \tag{1.11}$$

from which

$$\nabla^{\mu} \nabla_{\mu} \lambda_{ik} = \tfrac{2}{3} \nabla^{\mu} \lambda_{ij} \lambda^{jl} \nabla_{\mu} \lambda_{lk} - \tfrac{1}{6} \lambda^{jm} \lambda^{ln} \Omega^{\mu}_{mni} \Omega_{jlk\mu} - 2 R_{\mu\nu} \xi_i^{\mu} \xi_k^{\nu} \tag{1.12}$$

 0022-2488/79/071352-04$01.00

follows. With the formula

$$\lambda^{jm}\lambda^{ln}\Omega^{\mu}_{mni}\Omega_{jlk\mu}$$

$$= -2D^{\mu}\lambda_{ij}\lambda^{jl}D_{\mu}\lambda_{lk} + 6(\det\lambda)^{-1}\Omega_i\Omega_k \qquad (1.13)$$

Eq. (1.12) now reads

$$D^aD_a\lambda_{ik} = D^a\lambda_{ij}\lambda^{jl}D_a\lambda_{lk} - \tfrac{1}{2}(\det\lambda)^{-1}(D_a\det\lambda)D_a\lambda_{ik}$$

$$-(\det\lambda)^{-1}\Omega_i\Omega_k - 2R_{\mu\nu}\xi^{\mu}_i\xi^{\nu}_k, \qquad (1.14)$$

where D_a is the covariant derivative corresponding to the metric h_{ab}. Because of $R_{\mu\nu}=0$ and the subsidiary condition (1.9) we get

$$D^a(\tau\lambda^{-1}D_a\lambda)=0 \qquad (1.15)$$

with

$$\tau^2 \equiv |\det\lambda|. \qquad (1.16)$$

Moreover, the Ricci tensor on S obeys the equation

$$R^{(h)}_{ab} = \tfrac{1}{2}\operatorname{Tr}(\lambda^{-1}D_aD_b\lambda) - \tfrac{1}{4}\operatorname{Tr}(\lambda^{-1}D_a\lambda\lambda^{-1}D_b\lambda). \qquad (1.17)$$

We thus obtain equations which are equal in form to the ones of Einstein's theory derived by Geroch.[3]

We now choose coordinates in which

$$h_{ab} = h\eta_{ab}$$

with

$$(\eta_{ab}) = \begin{pmatrix} -1 & 0 \\ 0 & 1\end{pmatrix}, \quad \text{case (A)},$$

$$(\eta_{ab}) = \begin{pmatrix} 1 & 0 \\ 0 & 1\end{pmatrix}, \quad \text{case (B)}, \qquad (1.18)$$

holds. In these special coordinates Eq. (1.15) decouples from Eq. (1.17), and Eq. (1.17) can be easily integrated if we have solved Eq. (1.15). That is why we study only Eq. (1.15) in the following.

2. THE GENERATION OF ONE-PARAMETER FAMILIES OF SOLUTIONS

Up to the fact that λ is a symmetric 3×3 matrix instead of a 2×2 matrix the equation of motion for λ is the same in Jordan's theory as in Einstein's theory. So one can conjecture that the results can be taken over from the four-dimensional theory if they are formulated for λ. This is indeed the case for the infinitely many conservation laws and the generation of one parameter families of solutions which we now turn to.

Equation (1.15) is the first conservation law or can be read as the integrability condition for ($\tilde\partial_a \equiv \epsilon_{ab}\partial^b$)

$$\partial_a\omega \equiv \mp\tau\lambda^{-1}\tilde\partial_a\lambda. \qquad (2.1)$$

The different signs refer to cases (A) and (B) respectively. Taking the trace, we get ($\sigma \equiv -\tfrac{1}{2}\operatorname{Tr}\omega$)

$$\partial_a\sigma = \pm\tilde\partial_a\tau. \qquad (2.2)$$

Further potentials λ_n and ω_n can be defined recursively

$$\partial\lambda_{n+1} = -\partial(\lambda\omega_n)\mp\tau\tilde\partial\lambda\lambda^{-1}\lambda_n - \lambda\partial\omega_n + 2\partial(\sigma\lambda_n) \qquad (2.3a)$$

$$\partial\omega_{n+1} = \mp\partial(\tau^2\lambda^{-1}\lambda_n) - \partial\omega\omega_n \mp\tau^2\lambda^{-1}\partial\lambda_n + 2\sigma\partial\omega_n \qquad (2.3b)$$

with the initial data

$$\lambda_0 = 0, \quad \lambda_1 = \lambda, \quad \omega_0 = -1, \quad \omega_1 = \omega. \qquad (2.3c)$$

The integrability conditions follow by induction.

With the generating functions

$$V(s) = \sum_{n=0}^{\infty} s^n\lambda_n, \quad U(s) = \sum_{n=0}^{\infty} s^n\omega_n \qquad (2.4)$$

Eqs. (2.3a) and (2.3b) read

$$s^{-1}\partial V(s) = -\partial(\lambda U(s))\mp\tau\tilde\partial\lambda\lambda^{-1}V'(s) - \lambda\partial U(s) + 2\partial(\sigma V(s)), \qquad (2.5a)$$

$$s^{-1}\partial U(s) = \mp\partial(\tau^2\lambda^{-1}V(s)) - \partial\omega U(s) \mp\tau^2\lambda^{-1}\partial V(s) + 2\sigma\partial U(s). \qquad (2.5b)$$

The ansatz

$$V(s) = f(s,\tau,\sigma)\lambda U(s) \qquad (2.6)$$

with

$$f(s,\tau,\sigma) = \frac{1}{2s\tau^2}\{\mp 1 \pm 2s\sigma \pm [(1-2s\sigma)^2 \mp 4s^2\tau^2]^{1/2}\} \qquad (2.7)$$

solves Eq. (2.5a) identically. Then putting

$$U'(s) \equiv (1\mp\tau^2f^2)U \qquad (2.8)$$

Eq. (2.5b) yields

$$\partial U' = \pm\frac{\tau f}{1\mp\tau^2f^2}(-\lambda^{-1}\tilde\partial\lambda + \tau f\lambda^{-1}\partial\lambda)U'. \qquad (2.9)$$

In order to decouple these equations, we choose new coordinates

$$\xi = \frac{x^1+x^2}{2}, \quad \eta = \frac{x^1-x^2}{2}, \quad \text{case (A)}, \qquad (2.10a)$$

$$\xi = x^1 + ix^2, \quad \bar\xi = x^1 - ix^2, \quad \text{case (B)}, \qquad (2.10b)$$

and get

$$\partial_\xi U' = -\tfrac{1}{2}(1-1/\gamma)\lambda^{-1}\partial_\xi\lambda U', \qquad (2.11a)$$

$$\partial_\eta U' = -\tfrac{1}{2}(1-\gamma)\lambda^{-1}\partial_\eta\lambda U', \qquad (2.11b)$$

where γ represents

$$\gamma = \frac{1+\tau f}{1-\tau f} = \left(\frac{1-2s(\sigma+\tau)}{1-2s(\sigma-\tau)}\right)^{1/2}, \quad \text{case (A)}, \qquad (2.12a)$$

$$\gamma = \frac{1-i\tau f}{1+i\tau f} = \left(\frac{1-2s(\sigma-i\tau)}{1-2s(\sigma+i\tau)}\right)^{1/2}, \quad \text{case (B)}. \qquad (2.12b)$$

[If equations are equal in case (A) and (B) up to the substitution $\bar\xi$ for η as Eqs. (2.11) are, we only write down explicitly case (A)].

Eqs. (2.11) imply for normalized $\bar\lambda = \pm\tau^{-2/3}\lambda$ and $\bar U = (-s/f)^{1/3}U'$

$$\partial_\xi\bar U = -\tfrac{1}{2}(1-1/\gamma)\bar\lambda^{-1}\partial_\xi\bar\lambda\bar U, \qquad (2.13a)$$

$$\partial_\eta\bar U = -\tfrac{1}{2}(1-\gamma)\bar\lambda^{-1}\partial_\eta\bar\lambda\bar U. \qquad (2.13b)$$

Because of the normalization

$$\operatorname{Tr}(\bar\lambda^{-1}\partial\bar\lambda) = (\det\bar\lambda)^{-1}\partial\det\bar\lambda = 0 \qquad (2.14)$$

holds and $\det\bar U$ can be chosen equal to 1.

Equations (2.11) and (2.13) are the direct generalizations of the analogous equations in Einstein's theory and of the still more special equations in the O(2,1) nonlinear σ-model.[4] Therefore, it is not surprising that one-parameter families of solutions can also be obtained in a completely analogous manner. We find that

$$\bar{\lambda}^{(s)} = \bar{U}^T(s,\lambda)\bar{\lambda}\bar{U}(s,\lambda), \quad s\in\mathbb{R} \tag{2.15a}$$

and

$$\tau^{(s)} = \left(\frac{\gamma(s)-\gamma(s)^{-1}}{4s\tau}\right)^2 \tau \tag{2.15b}$$

solve

$$\partial_\eta(\tau\bar{\lambda}^{-1}\partial_\xi\bar{\lambda}) + \partial_\xi(\tau\bar{\lambda}^{-1}\partial_\eta\bar{\lambda}) = 0, \tag{2.16a}$$

$$\partial_\eta\partial_\xi\tau = 0. \tag{2.16b}$$

If $\bar{\lambda}$ and τ solve Eqs. (2.16) and $\bar{U}$ solves Eqs. (2.13). Thus, given a solution of Jordan's equations with three commuting Killing vectors, one obtains a one-parameter family of solutions by applying the transformation (2.15).

The symmetry transformation has the further properties

$$\bar{\lambda}^{(s)-1}\partial_\xi\bar{\lambda}^{(s)} = \frac{1}{\gamma}\bar{U}^{-1}\bar{\lambda}^{-1}\partial_\xi\bar{\lambda}\bar{U}, \tag{2.17a}$$

$$\bar{\lambda}^{(s)-1}\partial_\eta\bar{\lambda}^{(s)} = \gamma\bar{U}^{-1}\bar{\lambda}^{-1}\partial_\eta\bar{\lambda}\bar{U} \tag{2.17b}$$

and

$$\tau^{(s)-1}\partial_\xi\tau^{(s)} = \frac{1}{\gamma^2}\tau^{-1}\partial_\xi\tau, \tag{2.18a}$$

$$\tau^{(s)-1}\partial_\eta\tau^{(s)} = \gamma^2\tau^{-1}\partial_\eta\tau, \tag{2.18b}$$

which will be used to construct a linear eigenvalue problem in the next paragraph.

3. FORMULATION OF THE LINEAR PROBLEM IN TERMS OF SL(3,$\mathbb{R}$) INVARIANTS CONSTRUCTED FROM $\bar{\lambda}^{-1}\bar{\lambda}_\xi$ AND $\bar{\lambda}^{-1}\bar{\lambda}_\eta$

Because writing down the eigenvalue problem would be rather lengthy, we only show how to derive it. Starting from $\bar{\lambda}^{-1}\bar{\lambda}_\xi$ and $\bar{\lambda}^{-1}\bar{\lambda}_\eta$ ($\bar{\lambda}_\xi = \partial_\xi\bar{\lambda}$), we construct a basis for the Lie algebra of SL(3,$\mathbb{R}$) at every point of the manifold S (possibly up to a lower dimensional manifold where the system is not linearly independent, which we consider to be part of the boundary of S). For case (A) we take

$$\begin{aligned} &y_1 = \bar{\lambda}^{-1}\bar{\lambda}_\xi, \quad y_2 = \bar{\lambda}^{-1}\bar{\lambda}_\eta, \\ &y_3 = [y_1, y_2], \quad y_4 = [y_1,[y_1,y_2]], \\ &y_5 = [y_2,[y_2,y_1]], \quad y_6 = [y_1,[y_1,[y_1,y_2]]], \\ &y_7 = [y_2,[y_2,[y_2,y_1]]], \quad y_8 = [y_1[y_2,[y_1,y_2]]]. \end{aligned} \tag{3.1}$$

For case (B) we get the real basis by taking real and imaginary parts of (3.1). Because that does not change the essential features of the following discussion, which can be transcribed easily for case (B), we treat only case (A).

From y_1 and y_2 invariants of the form $\mathrm{Tr}(y_1^i y_2^j y_1^k\cdots)$ can be constructed. These invariants are certainly not independent. Because they determine y_1 and y_2 only up to a transformation

$$y_1 \to R^{-1}y_1R, \quad y_2\to R^{-1}y_2R, \tag{3.2}$$

the number of independent invariants is given by the minimal number of independent matrix elements of $R^{-1}y_1R$ and $R^{-1}y_2R$ with arbitrary R. That for $n\times n$ matrices this number is n^2-1 can be seen as follows: In the general case the characteristic polynomial of y_1 has n simple roots. Hence, y_1 can be transformed to diagonal form with, because of $\mathrm{Tr}y_1 = 0$, $n-1$ independent matrix elements. R is determined up to a factor which is irrelevant for $R^{-1}y_1R$ and $R^{-1}y_2R$ and up to the relative normalization of the eigenvectors of y_1 which are the columns of R, with respect to one eigenvector. Therefore, $n-1$ constraints can be imposed on the matrix elements of $R^{-1}y_2R$. Because of $\mathrm{Tr}y_2 = 0$ the minimal number of independent matrix elements of $R^{-1}y_1R$ and $R^{-1}y_2R$ is finally n^2-1.

For 3×3 matrices we can take the following complete set of invariants:

$$A_1 = \mathrm{Tr}y_1^2, \quad A_2 = \mathrm{Tr}y_2^2, \quad A_3 = \mathrm{Tr}y_1y_2, \quad A_4 = \mathrm{Tr}y_1^3, \tag{3.3}$$

$$A_5 = \mathrm{Tr}y_2^3, \quad A_6 = \mathrm{Tr}y_1^2y_2, \quad A_7 = \mathrm{Tr}y_1y_2^2, \quad A_8 = \mathrm{Tr}y_1^2y_2^2.$$

To express the other invariants explicitly in terms of these, one simply uses the fact that

$$y_i^3 = \tfrac{1}{2}(\mathrm{Tr}y_i^2)y_i + \tfrac{1}{3}\mathrm{Tr}y_i^3, \quad (i=1,2) \tag{3.4}$$

holds. [The first three invariants (3.3) are those which were called $2A^2$, $2B^2$ and $2AB\cos\alpha$ in the Einstein case.]

On the linear space sl(3,$\mathbb{R}$) a scalar product invariant under the adjoint representation of sl(3,$\mathbb{R}$) is defined by

$$(X_i,X_k)\equiv\mathrm{Tr}X_iX_k, \quad X_i\in\mathrm{sl}(3,\mathbb{R}). \tag{3.5}$$

Because this scalar product has the signature $(+\,+\,+\,+\,+\,-\,-\,-)$ corresponding to the fact that there are five independent symmetric and three antisymmetric elements of sl(3,$\mathbb{R}$),

$$\eta = \begin{pmatrix} 1 &&&&&&& \\ & 1 &&&&&& \\ && 1 &&&&& \\ &&& 1 &&&& \\ &&&& 1 &&& \\ &&&&& -1 && \\ &&&&&& -1 & \\ &&&&&&& -1 \end{pmatrix} \tag{3.6}$$

is a pseudometric on sl(3,$\mathbb{R}$). According to η the basis y_i can be orthonormalized using Schmidt's orthonormalizing procedure. This yields

$$z_i = b_{ik}y_k \tag{3.7}$$

with

$$(z_i, z_k) = \eta_{ik} \tag{3.8}$$

and b_{ij} which are known functions of the invariants through their dependence on the $a_{ik} = (y_i, y_k)$.

The next step is to express $y_{i,\eta}$ and $y_{i,\xi}$ again through y_k,

$$y_{i,\eta} = \Gamma_{ik}^{(1)}y_k, \quad y_{i,\xi} = \Gamma_{ik}^{(2)}y_k, \tag{3.9}$$

using the equation of motion (2.16a) for $\bar{\lambda}$. That $\Gamma^{(1)}$ and $\Gamma^{(2)}$ depend only on the invariants (3.3) and their first derivatives can be seen as follows: $\bar{\lambda}_{\xi\eta}$, $\bar{\lambda}_{\eta\eta}$, and $\bar{\lambda}_{\xi\xi}$ can be written as a linear combination of the basis elements in the space of symmetric 3×3 matrices

$$x^T \equiv (\bar{\lambda}, \bar{\lambda}_\xi, \bar{\lambda}_\eta, \bar{\lambda}_\xi \bar{\lambda}^{-1} \bar{\lambda}_\xi, \bar{\lambda}_\xi \bar{\lambda}^{-1} \bar{\lambda}_\eta, \bar{\lambda}_\eta \bar{\lambda}^{-1} \bar{\lambda}_\eta). \tag{3.10}$$

The coefficients are functions of the invariants (3.3) and their first derivatives only. For Eq. (2.16a) holds and the coefficients for $\bar{\lambda}_{\eta\eta}$ and $\bar{\lambda}_{\xi\xi}$ can be determined with help of the equations

$$\mathrm{Tr}\tilde{x}_i \lambda_{\eta\eta} = C^1_k \,\mathrm{Tr}\tilde{x}_i x_k, \tag{3.11a}$$

$$\mathrm{Tr}\tilde{x}_i \lambda_{\xi\xi} = C^2_k \,\mathrm{Tr}\tilde{x}_i x_k, \tag{3.11b}$$

with $\tilde{x}_i$ defined by

$$\tilde{x}^T \equiv (\bar{\lambda}^{-1}, \bar{\lambda}^{-1}\bar{\lambda}_\xi \bar{\lambda}^{-1}, \bar{\lambda}^{-1}\bar{\lambda}_\eta \bar{\lambda}^{-1}, \bar{\lambda}^{-1}\bar{\lambda}_\xi \bar{\lambda}^{-1}\bar{\lambda}_\xi \bar{\lambda}^{-1},$$

$$\bar{\lambda}^{-1}\bar{\lambda}_{(\xi} \bar{\lambda}^{-1}\bar{\lambda}_{\eta)} \bar{\lambda}^{-1}, \bar{\lambda}^{-1}\bar{\lambda}_\eta \bar{\lambda}^{-1}\bar{\lambda}_\eta \bar{\lambda}^{-1}) \tag{3.12}$$

and $\mathrm{Tr}\tilde{x}_i \bar{\lambda}_{\eta\eta}$, $\mathrm{Tr}\tilde{x}_i \bar{\lambda}_{\xi\xi}$, $\mathrm{Tr}\tilde{x}_i x_k$ being functions of the invariants and their first derivatives. The same statement concerning the expansion coefficients is true for matrices like $\bar{\lambda}_\xi \bar{\lambda}^{-1}\bar{\lambda}_\eta \bar{\lambda}^{-1}\bar{\lambda}_\xi$ which are built with more than three λ's. This yields that $y_{i,\eta}$ and $y_{i,\xi}$ can be written as linear combinations of matrices $\bar{\lambda}^{-1}\bar{\lambda}_\xi, \ldots, \bar{\lambda}^{-1}\bar{\lambda}_\xi \bar{\lambda}^{-1}\bar{\lambda}_\eta \bar{\lambda}^{-1}\bar{\lambda}_\xi, \ldots$, which can be combined to y_i's because $y_{i,\eta}$ and $y_{i,\xi}$ are elements of $\mathrm{sl}(3,\mathbb{R})$. The coefficients Γ_{ik} are therefore functions of the invariants (3.3) and their first derivatives.

From Eq. (3.7) and (3.9) we get $(b^{kl} b_{lm} = \delta^k{}_m)$

$$z_{i,\eta} = (b_{ik,\eta} b^{km} + b_{ik} \Gamma^{(1)}_{kl} b^{lm}) z_m \equiv C_{1i}{}^m z_m \tag{3.13a}$$

and

$$z_{i,\xi} = (b_{ik,\xi} b^{km} + b_{ik} \Gamma^{(2)}_{kl} b^{lm}) z_m \equiv C_{2i}{}^m z_m. \tag{3.13b}$$

As the $C_{ik}{}^m$ are functions of the invariants and their first derivatives, the compatibility condition of these equations yields the equations of motion for the invariants $A_1, \ldots, A_8$ which turn out to be second order differential equations.

Because of $(z_i, z_k) = \eta_{ik}$ one has

$$C_i \eta = \eta C_i^T = 0. \tag{3.14}$$

Hence the matrices C_i are elements of the vector representation of the lie algebra so(5,3). Because differention is a derivation and $\mathrm{sl}(3,\mathbb{R})$ is semisimple, they are at the same time elements of the adjoint representation of $\mathrm{sl}(3,\mathbb{R})$, i.e.,

$$C_{ik}{}^l = \omega_i{}^j f_{jk}{}^l \tag{3.15}$$

holds where $f_{jk}{}^l$ are the structure constants of $\mathrm{sl}(3,\mathbb{R})$.

We can now put $c_i = \omega_i{}^j Q_j$ with Q_j which are the basis of the vector representation of $\mathrm{sl}(3,\mathbb{R})$ corresponding to the representation matrices $(\tilde{Q}_j)_k{}^l = f_{jk}{}^l$ in the adjont representation. The equations

$$\psi_\eta = c_1 \psi, \quad \psi_\xi = c_2 \psi \quad (\psi \in \mathbb{R}^3) \tag{3.16}$$

are then a possible form of the linear problem to be constructed. The eigenvalue problem can be obtained at once if we construct $c_{1,2}$ using $\bar{\lambda}^{(s)}$ and $\tau^{(s)}$ instead of $\bar{\lambda}$ and τ. Because of Eqs. (2.17) and (2.18) the invariants' τ and the τ derivatives change only by factors depending on γ in which the "eigenvalue" s is hidden. Thus, the matrices c_1 and c_2 depend on s, σ, τ and the invariants constructed from $\bar{\lambda}$. If one could solve the inverse scattering problem, one would get the matrices $c_{1,2}$ and especially τ. We conjecture that each triple ψ_1, ψ_2, ψ_3 of independent solutions which, because of $\mathrm{Tr}c_{1,2} = 0$, can be invariantly normalized to $\det r = 1$ $[r \equiv (\psi_1, \psi_2, \psi_3)]$, yields $\bar{\lambda}$ by the formula $\bar{\lambda} = r^T r$. If this conjecture, which holds in the Einstein case, is not true, we get the z_i's from the O(5,3) representation of r, from z_i we get $\bar{\lambda}^{-1}\bar{\lambda}_\eta$ and $\bar{\lambda}^{-1}\bar{\lambda}_\xi$, and by integration $\bar{\lambda}$.

A different form of the linear eigenvalue problem is provided by the spinor representation of the c_i's considered as elements of SO(5,3). No matter how, a linear eigenvalue problem can be constructed in the case of a five-dimensional theory, too, and we conjecture that this is true even in higher dimensions. However, in higher dimensions the interesting case is the one with noncommuting Killing vectors.[5] Thus, one has to study a theory with this property, and it is not yet clear whether some of the results we obtained survive in spite of the additional difficulties.

ACKNOWLEDGMENT

We are indebted to P. Breitenlohner for clarifying discussions.

[1] P. Jordan, *Schwerkraft und Weltall* (Vieweg, Braunschweig, 1955).
[2] D. Maison, preprint MPI–PAE/PTh 14/78, April 1978.
[3] R. Geroch, J. Math. Phys. **13**, 394 (1972).
[4] M. Lüscher and K. Pohlmeyer, Nucl. Phys. B **137**, 46 (1978).
[5] Y.M. Cho, and P.S. Yang, Phys. Rev. D **12**, 3789 (1975).

Repulsive and attractive timelike singularities in vacuum cosmologies

Bonnie D. Miller

Department of Mathematics and Department of Astronomy and Astrophysics, Michigan State University
(Received 31 October 1978)

Spherically symmetric cosmologies whose big bang is partially spacelike and partially timelike are constrained to occur only in the presence of certain types of matter, and in such cosmologies the timelike part of the big bang is a negative-mass singularity. In this paper examples are given of cylindrically symmetric cosmologies whose big bang is partially spacelike and partially timelike. These cosmologies are vacuum. In some of them, the timelike part of the big bang is clearly a (generalized) negative-mass singularity, while in others it is a (generalized) positive-mass singularity.

I. INTRODUCTION

Among the solutions to Einstein's equations which are generally regarded as "cosmological," most exhibit an initial singularity, i.e., a big bang, which behaves as the source of all matter and information in the spacetime. In the (spatially) homogeneous isotropic Friedmann solutions, as well as in many inhomogeneous cosmologies, the big bang is spacelike, so that none of its points are to the future of any point in spacetime. Inhomogeneous cosmologies, however, also allow the possibility that the big bang is partially spacelike and partially timelike,[1] so that some of its points are subject to influence (as well as observation) from within the spacetime. We will use the term "mixed" to denote such a big bang. When the big bang is mixed, the spacetime has no Cauchy surface, or, equivalently, it is not globally hyperbolic. (Thus, the spacetime does not obey the stronger formulations of the cosmic censorship hypothesis.[2])

When the big bang is mixed in a spherically symmetric cosmology,[3] its timelike segment has a negative mass associated with it. The occurrence of a negative-mass segment of the big bang introduces a number of distinctive features into spherically symmetric cosmologies (even aside from the prospect of being able to interact causally with a singularity), and the primary purpose of this paper is to point out that analogous phenomena occur in cosmologies with other symmetries. In fact, these phenomena occur more readily in cosmologies whose symmetry is weaker than spherical symmetry, in the following sense: They can occur in vacuum, whereas in the spherically symmetric case, not only the presence of matter but also restrictions on the type of matter are required if the big bang is to be mixed.

In Sec. II, we review briefly the known examples of spherically symmetric cosmologies whose big bang is mixed, and we note that there are plane symmetric cosmologies which are closely analogous to them. In particular, that part of the big bang which is timelike, in the plane symmetric cosmologies, is also repulsive, i.e., although null geodesics (both past-directed and future-directed) terminate on it, timelike geodesics do not.

In Sec. III, we exhibit a one-parameter family of cylindrically symmetric vacuum cosmologies in which the big bang is mixed. The timelike part of the big bang is repulsive when the parameter has a negative value, but it is attractive when the parameter has a positive value. The situation in which it is attractive is probably unstable; this is suggested by a simple interpretation of the solutions.

II. SPHERICALLY SYMMETRIC AND PLANE SYMMETRIC COSMOLOGIES

A. The spherically symmetric case

In spherically symmetric spacetimes, the function m, called mass, is defined by $|\nabla R|^2 = 1 - 2m/R$, where R is the geometrically defined areal coordinate. Mass is a constant only in vacuum; if matter is present, m has nonzero derivatives, which are determined by the matter variables.

While Schwarzschild coordinates (R,T) do not adequately describe a hypersurface on which $R = 2m$, on one side of which ∇R is a spacelike vector and on the other side timelike, there may be other coordinate systems which are appropriate for describing the entire spacetime. For instance, in perfect fluid spacetimes, it is convenient to use comoving coordinates. Or if one is discussing Vaidya solutions[4]—spherically symmetric spacetimes whose stress-energy tensor is that of radially outflowing photons—it is convenient to use Eddington–Finkelstein coordinates (R,u), where u is an outgoing null coordinate. In these coordinates, the metric for the Vaidya solutions is

$$ds^2 = -\left[1 - \frac{2m(u)}{R}\right] du^2 - 2du\,dR + R^2(d\theta^2 + \sin^2\theta\,d\varphi^2), \tag{1}$$

where the mass $m(u)$ is an arbitrary decreasing function of u.

Spherically symmetric cosmologies in which the big bang is mixed have been described in detail by Miller[1]; the following is a summary of their properties, in the known examples, which are illustrated schematically in Fig. 1.

(1) Mass changes sign on a hypersurface in spacetime, being negative at points on the side near the timelike segment

 0022-2488/79/071356-06$01.00

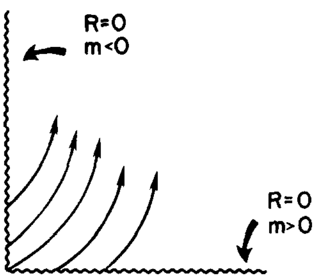

FIG. 1. A spherically symmetric cosmology with a mixed big bang. Matter is being emitted, as the mass of the singularity drops through zero, and the singularity is changing in character from spacelike to timelike.

of the singularity. At these points, spacetime has the local properties of the negative-mass Schwarzschild solution. (For instance, photons are blue shifted as they travel outward.) The function m, defined originally within the spacetime, extends by continuity to the singularity, and it is negative on the timelike points of the singularity.

(2) The negative-mass segment of the singularity is spatially a point. That is, for each of the naturally defined $T=$ constant spacelike hypersurfaces, there is one point on the causal boundary.[5]

(3) The singularity is emitting matter, as it shifts from being spacelike to timelike and its mass drops through zero. The matter cannot be dust; it can be either photons (in the case of a Vaidya solution) or a stiff ($p=\epsilon$) perfect fluid (in the case of the Taub–Cahill[6] self-similar solutions). We note that while there are no known examples in which $0<p<\epsilon$, there is no reason to think that such examples do not occur.

That matter variables should be necessary in order for a positive-mass singularity to (unpredictably)[7] develop into a negative-mass singularity, is a simple consequence of the Birkhoff theorem, which severely constrains the behavior of spherically symmetric vacuum solutions to the field equations. Any such solution, the theorem states, is determined by a single constant, m; the sign of that constant then determines that the singularity at $R=0$ is either spacelike (if $m>0$) or timelike (if $m<0$). In order for the big bang to be mixed in a spherically symmetric cosmology, then, it is necessary that there be more variables in the field equations than occur in vacuum.

B. The plane symmetric case

The term "plane symmetric" has been used in different ways to characterize solutions to Einstein's equations; in this paper we use the term to refer to spacetimes whose isometry group is that of the two-dimensional Euclidean plane. Such spacetimes have been studied extensively by Taub and his co-workers,[8] who emphasize that they bear a strong formal resemblance to spherically symmetric spacetimes and that, in particular, they allow a locally defined masslike function.

The metric on the Euclidean 2-planes may be written $\rho^2(dx^2+dy^2)$, where ρ is a function only of the two coordinates orthogonal to x and y. (If $\nabla\rho\neq 0$, ρ is a natural choice for one of these two coordinates.) The function ρ clearly plays a role similar to that of R in spherically symmetric spacetimes, but there are some notable geometrical and formal differences, most of them in essence restatements of the fact that the flat-space limit of $|\nabla\rho|^2$ is 0 (since $\nabla\rho\equiv 0$), while that of $|\nabla R|^2$ is 1.

In vacuum, the behavior of plane symmetric spacetimes is extremely limited, as it is in spherically symmetric spacetimes, for the same reason: a Birkhoff theorem.[9] The vacuum metric may be written in such a way as to emphasize the resemblance to the Schwarzschild metric[10]:

$$ds^2 = -\left(\frac{2\mu}{\rho}\right)^{-1} d\rho^2 + \left(\frac{2\mu}{\rho}\right) dT^2 + \rho^2(dx^2+dy^2), \quad (2)$$

where μ is a constant. Note that $|\nabla\rho|^2 \equiv -2\mu/\rho$. If $\mu>0$, then $\nabla\rho$ is timelike, and the spacetime is cosmological, with a spacelike singularity at $\rho=0$. If $\mu<0$, then $\nabla\rho$ is spacelike, and the spacetime is static. In this case the singularity at $\rho=0$ is much like the negative-mass Schwarzschild singularity: It is timelike and repulsive. However, rather than being a single point in space, the singularity is, in Taub's terminology, a "big wall": At fixed T, there is a point on the causal boundary for each point (x,y) in the Euclidean plane.

If there is matter present in the spacetime, we can define a function whose vacuum limit is the constant μ. We write

$$|\nabla\rho|^2 \equiv -\frac{2\mu}{\rho}, \quad (3)$$

where μ is now a function of position; Eq. (3) is the analog to the equation $|\nabla R|^2 = 1-2m/R$, which occurs in spherical symmetry. Further considerations of the field equations for plane symmetric spacetimes continue to indicate that μ is reasonably referred to as the analog to m.

As a specific example, consider plane symmetric perfect fluid solutions. The metric in comoving coordinates is

$$ds^2 = -e^{2\varphi}dt^2 + e^{2\psi}dz^2 + \rho^2(dx^2+dy^2). \quad (4)$$

In these coordinates, Eq. (3) takes the form

$$-\frac{2\mu}{\rho} = \Gamma^2 - U^2, \quad (5)$$

where $U \equiv e^{-\varphi}\partial\rho/\partial t$ and $\Gamma \equiv e^{-\psi}\partial\rho/\partial z$. [Taub has stressed that Eq. (5) actually arises as an integral of the field equations, as does its analog in spherical symmetry.] Then the field equations show that

$$\frac{\partial\mu}{\partial t} = -p\rho^2\frac{\partial\rho}{\partial t} \quad (6)$$

and

$$\frac{\partial\mu}{\partial z} = \epsilon\rho^2\frac{\partial\rho}{\partial z}, \quad (7)$$

where p and ϵ are the pressure and energy density, respectively. Equations (6) and (7) are formally identical to their analogs in spherical symmetry, which are the equations for the derivatives of mass.

While the function μ is thus evidently a close formal analog to m, there are significant differences in the conclusions to be drawn from the signs of the two functions. One sees from Eq. (3) that $\nabla\rho$ changes from being spacelike to being timelike on a hypersurface in spacetime on which μ changes sign, whereas the analogous transition in spherical symmetry takes place when $R = 2m > 0$.

In a cosmological spacetime, μ changing sign on a hypersurface indicates that the singularity $\rho = 0$ changes from being spacelike to being timelike, at its intersection with that hypersurface. The timelike part of the singularity then has properties similar to those found in the vacuum "negative mass," i.e., $\mu < 0$) limit: it is repulsive, and is spatially a wall, rather than a point.

Examples of plane symmetric cosmologies in which the big bang is mixed, can be found readily by means entirely analogous to those which yield the spherically symmetric examples. In particular, it may be verified that the procedures of Taub and Cahill for constructing spherically symmetric self-similar spacetimes containing a stiff perfect fluid, lead also in the case of plane symmetry to a solution space which includes cosmologies with a mixed big bang. (As usual, the conditions imposed on these solutions, that the spacetime be self-similar and the perfect fluid be stiff, are made simply to simplify the field equations.)

A second class of plane symmetric cosmologies with mixed big bang is the class analogous to the Vaidya solutions which occur in spherical symmetry. The metric (which may be found by straightforward integration of the field equations) is

$$ds^2 = \frac{2\mu(u)}{\rho}\, du^2 - 2du\, d\rho + \rho^2(dx^2 + dy^2), \tag{8}$$

where again, u is an outgoing null coordinate, and μ is an arbitrary decreasing function of u, which passes through 0. The energy-momentum tensor has a single nonvanishing covariant component, $T_{uu} = -2(d\mu/du)\rho^{-2}$.

We summarize this section by saying that there exist plane symmetric cosmologies whose big bang is mixed, and that the properties of these cosmologies are, for the most part, the same as those seen in the spherically symmetric case. In particular, the timelike part of the singularity is repulsive, and the singularity is necessarily emitting matter—of a type which moves along null geodesics or accelerated timelike curves—as it shifts from being spacelike to being timelike. Figure 1, which illustrates the spherically symmetric case, may also be interpreted as an illustration of the plane symmetric case, if one substitutes μ and ρ for m and R, respectively.

III. CYLINDRICALLY SYMMETRIC COSMOLOGIES

The metric for cylindrically symmetric spacetimes can be written

$$ds^2 = -e^{2(\gamma-\psi)}(dt^2 - dr^2) + e^{2\psi}dz^2 + \alpha^2 e^{-2\psi}d\theta^2, \tag{9}$$

where γ,ψ, and α are functions of r and t. The function $\alpha \equiv |\partial_z||\partial_\theta|$ is closely analogous to R and ρ; in particular, the singularity in a vacuum solution is at $\alpha = 0$. At present, none of the scalar functions which have been defined in cylindrically symmetric spacetimes, appears to be a fully satisfactory analogue to m and μ. (Thorne's "C-energy" scalar[11,12] for instance, which in some respects closely parallels m and μ, does not distinguish the presence of a repulsive singularity from that of an attractive singularity.)

Because of the additional freedom which they give to the field equations (as compared to spherically symmetric or plane symmetric solutions), cylindrically symmetric solutions have been a primary source of information on those general relativistic phenomena which can only occur under sufficiently asymmetrical circumstances—notably phenomena involving gravitational radiation,[13] but also others, such as highly nonspherical gravitational collapse[10,14] and the development of inhomogeneities in cosmology.[15] And while cylindrically symmetric spacetimes are, as such, physically unrealistic, one may suppose that many of the interesting phenomena which they allow do not depend critically for their existence on the existence of an infinite source. Thorne[16] has provided some evidence for this supposition by showing that there are asymptotically flat solutions to the field equations in which the (bounded) source is a ring singularity, and which are locally—as one approaches the singularity—cylindrically symmetric.

When $\nabla\alpha$ is spacelike, then coordinate transformations allow one to choose the coordinate r in Eq. (9) so that $r = \alpha$. A singularity at $r = \alpha = 0$ is then timelike. The vacuum field equations in this case reduce to

$$\frac{\partial^2\psi}{\partial t^2} - \frac{\partial^2\psi}{\partial r^2} - \frac{1}{r}\frac{\partial\psi}{\partial r} = 0, \tag{10a}$$

$$\frac{\partial\gamma}{\partial t} = 2r\frac{\partial\psi}{\partial r}\frac{\partial\psi}{\partial t}, \tag{10b}$$

$$\frac{\partial\gamma}{\partial r} = r\left[\left(\frac{\partial\psi}{\partial r}\right)^2 + \left(\frac{\partial\psi}{\partial t}\right)^2\right]. \tag{10c}$$

Solutions to these equations include the Einstein–Rosen waves, and, in the static limit, the Levi–Civita solutions.

The Levi–Civita solutions, which are described in detail by Thorne,[11] are given in his notation by $\psi = -2k\ln(r/r_0)$ and $\gamma = 4k^2\ln(r/r_0)$, where k and r_0 are constants. There are distinct classes (aside from the locally flat solutions which occur when $k = -\frac{1}{2}$, 0, or ∞) among these solutions. If $k < -\frac{1}{2}$, then the proper circumference $C = 2\pi r e^{-\psi}$ of cylinders varies inversely with r, so the singularity is approached by letting $C \to \infty$. If $-\frac{1}{2} < k < 0$ or $0 < k < \infty$, C varies directly with r, so the singularity lies along the z-axis, $C = 0$; if $-\frac{1}{2} < k < 0$, the singularity is repulsive, while if $0 < k < \infty$, the singularity is attractive.

In the case when $\nabla\alpha$ is timelike,[17] coordinate transformations allow one to choose the coordinate t in Eq. (9) so that $t = \alpha$. The singularity at $t = \alpha = 0$ is then spacelike. The field equations in vacuum are formally the same as Eqs. (10), but with the roles of r and t interchanged:

$$\frac{\partial^2\psi}{\partial r^2} - \frac{\partial^2\psi}{\partial t^2} - \frac{1}{t}\frac{\partial\psi}{\partial t} = 0, \tag{11a}$$

$$\frac{\partial\gamma}{\partial r} = 2t\frac{\partial\psi}{\partial t}\frac{\partial\psi}{\partial r}, \tag{11b}$$

$$\frac{\partial \gamma}{\partial t} = t\left[\left(\frac{\partial \psi}{\partial t}\right)^2 + \left(\frac{\partial \psi}{\partial r}\right)^2\right]. \tag{11c}$$

No general considerations such as a Birkhoff theorem preclude the existence in vacuum of spacetimes in which the singular locus $\alpha = 0$ is partially timelike and partially spacelike. Thus, one might expect to find cylindrically symmetric vacuum cosmologies with a mixed big bang. If this should occur, a convenient global coordinate system would be (α,u,z,θ), where u is an appropriate label for outgoing null rays; the metric would take the form

$$ds^2 = 2g_{u\alpha}du\, d\alpha + g_{uu}\, du^2 + e^{2\psi}dz^2 + \alpha^2 e^{-2\psi}\, d\theta^2. \tag{12}$$

In these coordinates, $|\nabla\alpha|^2$ is given by $|\nabla\alpha|^2 = g^{\alpha\alpha} = -g_{uu}/g^2_{u\alpha}$, so that the critical hypersurface, on which $\nabla\alpha$ is null, is that on which $g_{uu} = 0$.

One way of finding solutions with these properties is to look for solutions to Eqs. (10) which exhibit an appropriate type of incompleteness—namely, which show $|\nabla\alpha|^2 \to 0$ (so that the coordinate system in which $r = \alpha$ is becoming ill-behaved) as one approaches a hypersurface which is apparently well-behaved geometrically—and then extending these solutions to include the hypersurface on which $|\nabla\alpha|^2 = 0$, as well as the region in which $\nabla\alpha$ is timelike. Thus we proceed by looking for a solution to Eq. (10a) of the form $\psi = \psi(w)$, where $w \equiv \tilde{u}/r$ and $\tilde{u} \equiv t - r$. When we assume that ψ depends only on w, Eq. (10a) becomes

$$(w^2 + 2w)\frac{d^2\psi}{dw^2} + (w+1)\frac{d\psi}{dw} = 0. \tag{13}$$

The solution to Eq. (13) is

$$\psi = 2c \ln((w+2)^{1/2} + (w)^{1/2})$$

$$= c \ln\left(\frac{2t + 2(t^2 - r^2)^{1/2}}{r}\right), \tag{14}$$

where c is an arbitrary constant, and an arbitrary additive constant has been set equal to zero. Integration of Eqs. (10b) and (10c) then gives

$$\gamma = c^2 \ln \frac{r}{t^2 - r^2}, \tag{15}$$

if we ignore another additive constant.

In the coordinates $(\alpha,\tilde{u},z,\theta)$, the metric takes the form[18]

$$ds^2 = -e^{2(\gamma-\psi)}(2d\tilde{u}\, d\alpha + d\tilde{u}^2) + e^{2\psi}dz^2 + \alpha^2 e^{-2\psi}d\theta^2$$

$$= \frac{-\alpha^{2c^2+2c}(2d\tilde{u}\, d\alpha + d\tilde{u}^2)}{4^{c}(\tilde{u}^2 + 2\alpha\tilde{u})^{2c^2}[\tilde{u} + \alpha + (\tilde{u}^2 + \alpha\tilde{u})^{1/2}]^{2c}}$$

$$+ \left(\frac{2[\tilde{u} + \alpha + (\tilde{u}^2 + 2\tilde{u}\alpha)^{1/2}]}{\alpha}\right)^{2c} dz^2$$

$$+ \frac{\alpha^{2c+2}\, d\theta^2}{\{2[\tilde{u} + \alpha + (\tilde{u}^2 + 2\tilde{u}\alpha)^{1/2}]\}^{2c}}. \tag{16}$$

As long as $\tilde{u} > 0$, i.e., as long as $t > r$, the metric form (16) is well-behaved. As $\tilde{u} \to 0$, however, $|\nabla\alpha|^2 \to 0$, i.e., $\nabla\alpha$ is becoming null, while $g_{\tilde{u}\tilde{u}} \to -\infty$ and $g_{\tilde{u}\alpha} \to -\infty$.

Two different types of cosmological solutions are described by the family of metrics (16); for purposes of illustrating them simply we restrict attention now to the cases $c = \frac{1}{2}$ and $c = -\frac{1}{2}$. (For other choices of c in the ranges $0 < c < 1/(2)^{1/2}$ or $-1/(2)^{1/2} < c < 0$, the discussion would be essentially the same, depending only on the sign of c, but the extension of the spacetime across the hypersurface on which $\nabla\alpha$ is null would not be as smooth, and the analysis of the geodesic equations not as simple. We ignore the cases $c > 1/(2)^{1/2}$ and $c < -1/(2)^{1/2}$, as well as the case $c = 0$.) In both cases, we make the coordinate transformation $u = \tilde{u}^{1/2}$. In the case $c = \frac{1}{2}$, the metric becomes

$$ds^2 = \frac{-2\alpha^{3/2}(du\, d\alpha + u\, du^2)}{(u^2 + 2\alpha)^{1/2}[u^2 + \alpha + (u^4 + 2\alpha u^2)^{1/2}]}$$

$$+ 2\left(\frac{u^2 + \alpha + (u^4 + 2\alpha u^2)^{1/2}}{\alpha}\right) dz^2$$

$$+ \frac{1}{2}\frac{\alpha^3\, d\theta^2}{[u^2 + \alpha + (u^4 + 2\alpha u^2)^{1/2}]}; \tag{17}$$

in the case $c = -\frac{1}{2}$, the metric becomes

$$ds^2 = \frac{-8[u^2 + \alpha + (u^4 + 2\alpha u^2)^{1/2}]}{\alpha^{1/2}(u^2 + 2\alpha)^{1/2}}(du\, d\alpha + u\, du^2)$$

$$+ \frac{1}{2}\frac{\alpha\, dz^2}{[u^2 + \alpha + (u^4 + 2\alpha u^2)^{1/2}]}$$

$$+ 2\alpha[u^2 + \alpha + (u^4 + 2\alpha\, u^2)^{1/2}]\, d\theta^2. \tag{18}$$

The metrics (17) and (18) are both well-behaved as $|\nabla\alpha|^2$ passes through zero, on the hypersurface $u = 0$, and they are complete. In the region $u > 0$ (i.e., $\tilde{u} > 0$), $\nabla\alpha$ is spacelike, and in the region $u < 0$, $\nabla\alpha$ is timelike. The big bang, at $\alpha = 0$, is thus mixed, in these vacuum cosmologies.

To determine the repulsive or attractive character of the timelike part of the big bang, we consider timelike geodesics in the region $u > 0$. From the metric form (9), with $\alpha = r$, one finds the r-component of the geodesic equations to be

$$\frac{d^2r}{ds^2} = \frac{\partial(\psi - \gamma)}{\partial r}\left[\left(\frac{dr}{ds}\right)^2 + \left(\frac{dt}{ds}\right)^2\right] + 2\frac{\partial(\psi-\gamma)}{\partial t}\frac{dr}{ds}\frac{dt}{ds}$$

$$+ \frac{e^{4\psi - 2\gamma}}{r^2}\left(\frac{1}{r} - \frac{\partial\psi}{\partial r}\right)P_\theta^2 + e^{-2\gamma}\frac{\partial\psi}{\partial r}P_z^2, \tag{19}$$

where $P_\theta \equiv r^2 e^{-2\psi}\, d\theta/ds$ and $P_z \equiv e^{2\psi}\, dz/ds$ are the two conserved momenta. From the condition $u_\beta u^\beta = -1$ obeyed by the tangent vector $\mathbf{u}$ to a geodesic, one has also the relation

$$e^{2(\gamma-\psi)}\left[\left(\frac{dt}{ds}\right)^2 - \left(\frac{dr}{ds}\right)^2\right] - e^{-2\psi}P_z^2 - \frac{e^{2\psi}}{r^2}P_\theta^2 = 1. \tag{20}$$

Our concern here is simply whether or not the geodesics actually reach the singularity at $r=0$, so we look at Eq. (19) under the condition $r \ll t$. When $r \ll t$, ψ and γ may be approximated by $\psi \approx c \ln(4t/r)$ and $\gamma \approx c^2 \ln(r/t^2)$. With the use of these approximations, Eq. (19) becomes (for any c)

$$\frac{d^2r}{ds^2} = -\frac{(c^2+c)}{r}\left[\left(\frac{dr}{ds}\right)^2 + \left(\frac{dt}{ds}\right)^2\right] + 2\frac{(2c^2+c)}{t}\frac{dr}{ds}\frac{dt}{ds} + \frac{4^{4c}t^{4c^2+4c}}{r^{2c^2+4c+3}}(1+c)P_\theta^2 - \frac{ct^{4c^2}}{r^{2c^2+1}}P_z^2. \tag{21}$$

Since $|dr/ds| < dt/ds$ [by Eq. (20)] and $r \ll t$, the term $2(2c^2+c)(dr/ds)(dt/ds)t^{-1}$ is negligible in magnitude compared to the term $-(c^2+c)(dt/ds)^2r^{-1}$, so we ignore the former. When $(dt/ds)^2$ is evaluated from Eq. (20), and then substituted into Eq. (21), the latter equation becomes

$$\frac{d^2r}{ds^2} = -\frac{(c^2+c)}{r}\left[2\left(\frac{dr}{ds}\right)^2 + \frac{4^{2c}t^{2c+4c^2}}{r^{2c^2+2c}}\right], \tag{22}$$

for radial geodesics (i.e., for $P_z = P_\theta = 0$). First, we consider the case $c = \frac{1}{2}$. Equation (22) becomes

$$\frac{d^2r}{ds^2} = -\frac{3}{4r}\left[2\left(\frac{dr}{ds}\right)^2 + \frac{4t^2}{r^{3/2}}\right]. \tag{23}$$

Note that d^2r/ds^2 is negative. Thus, if dr/ds is initially negative along a radial timelike geodesic near $r=0$, the geodesic will reach $r=0$. (Clearly, this conclusion holds for any $c>0$.) That is, the timelike part of the big bang, in the cosmology described by the metric (17), is attractive.

Second, we consider the case $c = -\frac{1}{2}$. In that case, the terms on the right hand side of Eq. (21) which do not involve P_θ or P_z are positive, and any nonzero value of P_θ or P_z makes an additional positive contribution to d^2r/ds^2. To show that no geodesics reach the singularity in this case, then, it suffices to show that radial geodesics do not reach the singularity. For radial geodesics in the case $c = -\frac{1}{2}$, we rewrite Eq. (22), by defining $v \equiv dr/ds$. Then the left-hand side of Eq. (22) is $dv/ds = (dv/dr)\,v = \frac{1}{2}d(v^2)/dr$, and the equation becomes

$$\frac{d(v^2)}{dr} = \frac{v^2}{r} + \frac{1}{8r^{1/2}}. \tag{24}$$

The solution to Eq. (24) is

$$v^2 = -\frac{r^{1/2}}{4} + \kappa r, \tag{25}$$

where the constant κ has the value $\kappa = v_0^2/r_0 + (1/16r_0)^{1/2}$. where v_0 is the value of v at some initial value r_0 of r along the geodesic.

Equation (25) shows that v^2, along an initially ingoing geodesic, must reach zero at a value of r greater than zero; the outward acceleration then directs the geodesic to larger values of r. Thus, the timelike part of the big bang, in the cosmology described by the metric (18), is repulsive.

An explanation of the difference between these two cases, one in which the timelike part of the big bang is attractive and the other in which it is repulsive, is suggested by a rough interpretation of what is occurring in these vacuum cosmologies. The spacelike part of the big bang is emitting gravitational radiation along the two null directions at each of its points (see Fig. 2); the timelike part of the big bang, however, is emitting radiation ("losing mass") outward along only one of the null directions, while radiation is collapsing onto it from the other null direction. Evidently, the singularity is repulsive or attractive, depending on whether the emission or the collapse of radiation, respectively, is the dominant process.

Consider a fully general perturbation of a solution in which both of these processes are present. On the one hand, we see no reason to think that the perturbation will prevent the timelike part of the singularity from emerging or emitting radiation. On the other hand, because of the (generally assumed) instability of collapsing cylinders to the formation of discrete lumps, one would expect that the collapse which in the unperturbed solution is directed, by virtue of the high degree of symmetry, exactly toward the singularity, will in the perturbed solution be redirected, by functions depending on z and θ as well as r and t, toward the lumps which are forming, rather than toward any singularity which may be present. Thus, of the two processes which in the unperturbed solution compete to determine whether the timelike part of the singularity is attractive or repulsive, it is likely that one is eliminated by a perturbation. That is, the cosmologies in which the timelike part of the big bang is attractive are probably unstable, becoming, for instance, ones in which the timelike part is repulsive.

In closing this section we note briefly that there are cylindrically symmetric analogs of the Vaidya metrics; this may be deduced from Liang's[10] discussion of Rao's (unpublished) solutions. As in the spherically symmetric and plane symmetric cases, an arbitrary function of an outgoing null

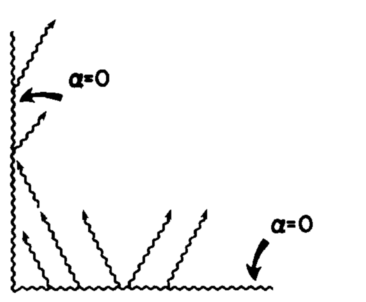

FIG. 2. A cylindrically symmetric vacuum cosmology with a mixed big bang. Gravitational waves leave the spacelike part of the singularity in both null directions; they leave the timelike part of the singularity in one null direction, and enter it along the other.

coordinate allows one to have an arbitrary rate of radiation of photons, and a mixed big bang.

IV. CONCLUDING COMMENTS

A question of considerable interest in theoretical general relativistic cosmology is whether the presence of a Cauchy surface is a generic feature of inhomogeneous cosmologies. (The term "generic" is used in several somewhat different ways, any of which is appropriate in this informal discussion.) Such a surface does not exist if the big bang contains spacelike sections, as well as pieces which are (generalized) negative-mass singularities, which are timelike. While a big bang with this mixed causal structure can occur only under constraints on the type of matter if a cosmology has spherical or plane symmetry, it can occur in vacuum (and thus regardless of the presence or type of matter) if the symmetry is not one whose vacuum limit is governed by a Birkhoff theorem. That is, roughly speaking, the situation depicted in Fig. 1, in which a negative-mass singularity grows out of a positive-mass singularity (by means of a singularity's emitting matter) can occur in vacuum (by means of a singularity's emitting gravitational waves) if the spacetime's symmetry is not one which prohibits the occurrence of gravitational radiation.

It is perhaps the case that such phenomena—spacelike inhomogeneities on the big bang turning into timelike ones—do not occur frequently among solutions to the field equations. That is, it is possible that their occurrence, which has in fact been seen only in highly symmetrical situations, is unstable to perturbation. If their occurrence is stable, however, then the decision to preclude them from the actual big bang is presumably to be made and explained by the (presently unknown) quantum theory of gravity.

We conclude with a comment on the possibility of a naked singularity developing in vacuum. We use the term "naked singularity" as it is most commonly used, to refer to a timelike singularity which develops from the collapse of an initially nonsingular system. (Typically, one has in mind a collapsing star; any example in vacuum would indicate that the formation of a naked singularity need not depend critically on the properties of matter, or, more generally, on the presence of matter.) It is known that the region $m \leqslant 0$ (including the singularity) of the Vaidya solutions discussed in Sec. II A can form as the result of the gravitational collapse of a radiating star,[19] and evidently the field equations allow the formation in vacuum of close analogs to that region, in spacetimes whose symmetry is weaker than spherical symmetry (and thus in which gravitational radiation may play roles which are necessarily taken by matter in spherically symmetric spacetimes). Possibly, then, such analogs (which have so far been seen to develop only from singular initial conditions) can also form from (vacuum) nonsingular initial conditions, in which gravitational radiation is substituting for both the collapsing star and the emitted photons which are essential in the spherically symmetric case.

ACKNOWLEDGMENTS

I thank A. Taub, R. Sachs, and D. Eardley for useful and encouraging discussions regarding this work.

[1]B.D. Miller, Ap. J. **208**, 275 (1976); K. Tomita, Prog. Theor. Phys. **59**, 1150 (1978). In this paper we are precluding from consideration timelike singularities of the "shell-crossing" type.

[2]R. Penrose, in *Theoretical Principles in Astrophysics and Relativity*, edited by N.R. Lebovitz, W.H. Reid, and P.O. Vandervoort (University of Chicago Press, Chicago, 1978).

[3]By a "cosmology" we will mean a spacetime with a big bang: A cosmological spacetime is one in which all past-directed causal curves can be extended for only a finite proper time interval or affine length. The singular structure on which all such curves terminate—the big bang—is discussed in detail in, e.g., S.W. Hawking and G.F.R. Ellis, *The Large-Scale Structure of Spacetime* (Cambridge U.P., Cambridge, 1973).

[4]P.C. Vaidya, Nature **171**, 260 (1953); see also R.W. Lindquist, R.A. Schwartz, and C.W. Misner, Phys. Rev. **137**, B1364 (1965).

[5]R. Geroch, E.H. Kronheimer, and R. Penrose, Proc. R. Soc. London, Ser. A **237**, 545 (1972).

[6]M. Cahill and A.H. Taub, Commun. Math. Phys. **21**, 1 (1971).

[7]The singularity, either prior to or upon reaching zero mass, has the option of simply "shutting off."

[8]A.H. Taub, Ann. Math. **53**, 472 (1951); A.H. Taub, Phys. Rev. **103**, 454 (1956); A.H. Taub, *in General Relativity*, edited by L. O'Raifeartaigh (Clarendon, Oxford, 1972); R. Tabensky and A.H. Taub, Commun. Math. Phys. **29**, 61 (1973).

[9]See Taub (1951), Ref. 8.

[10]E.P. Liang, Phys. Rev. D **10**, 447 (1974).

[11]K.S. Thorne, "Geometrodynamics of Cylindrical Systems," unpublished Ph.D. Thesis, Princeton University (1965).

[12]K.S. Thorne, Phys. Rev. **138**, B251 (1965).

[13]See Ref. 11, and many references therein.

[14]K.S. Thorne, in *Magic Without Magic: John Archibald Wheeler*, edited by J. Klauder (Freeman, San Francisco, 1972).

[15]E.P. Liang, Ap. J. **204**, 235 (1976).

[16]K.S. Thorne, J. Math. Phys. **16**, 1860 (1975).

[17]J. Ehlers, lecture at the Royaumont Conference, published in *Les Théories Relativistes de la Gravitation* (Centre National de la Recherche Scientifique, Paris, 1962); see also the discussion by Thorne in Ref. 11.

[18]D. Eardley has pointed out to me that these metrics contain a (local) conformal Killing vector, whose form is

$$\xi = (c^2 + 1)^{-1}(t\,\partial_t + r\partial_r) + z\partial_z + c^2(c^2 + 1)^{-1}\theta\,\partial_\theta.$$

[19]B. Steinmuller, A.R. King, and J.P. Lasota, Phys. A **51**, 191 (1975); B.D. Miller, Ref. 1, in a brief discussion of solutions by H. Bondi, Proc. R. Soc. London, Ser. A **281**, 39 (1964). Essentially what occurs is the following. A spherically symmetric star collapses, and as it collapses it converts itself to photons, which are then radiated away. No horizon forms during the collapse, since matter is being radiated away sufficiently rapidly, that at any $R \neq 0$, the value of $2m$ is never as large as R. However, as the collapse proceeds, the density at the center of the star (where $m = 0$) becomes infinite, so that a singularity forms, whose mass is initially zero and subsequently (perhaps) negative.

On angular momentum of stationary gravitating systems

Abhay Ashtekar

Département de Physique, Université de Clermont-Fd., 63170 Aubiere, France and Max Planck Institut für Astrophysik, Föhringer Ring 6, 8000 München, Federal Republic of Germany

Michael Streubel

Max Planck Institut für Astrophysik, Föhringer Ring 6, 8000 München, Federal Republic of Germany
(Received 1 January 1979)

A theorem is proved to the effect that, for isolated gravitating systems in equilibrium, the definition of total angular momentum involving fields at null infinity agrees with that involving fields at spatial infinity.

1. INTRODUCTION

Asymptotic properties of the gravitational field of isolated systems have been investigated in two distinct regimes: at large null separations from sources and at large spacelike separations (see, e.g., Refs. 1–6). In each regime, asymptotic symmetry groups have been analyzed, equations satisfied by asymptotic fields have been obtained and conserved quantities have been constructed. These investigations have shed light on a variety of issues concerning isolated gravitating systems. Furthermore, much of one's intuition about the nature of the gravitational interactions in the relativistic regime is based on these analyses.

Unfortunately, however, practically nothing is known about the *relation* between the structure available at null infinity and that available at spatial infinity. Investigation of this relation is of considerable physical interest because one does expect the gravitational field of an isolated system to be asymptotically flat in both regimes. Consider, in particular, the notion of angular momentum. At least in the absence of gravitational radiation, one would expect the sources to give rise to "just one" asymptotic spin vector to which the angular momentum of test particles can couple. The mathematical description, on the other hand, provides one such vector in *each* regime. One is therefore led to ask for the relation between these two vectors.

In order to analyze such issues, a general framework aimed at obtaining a unified description of null and spatial infinity has been recently introduced.[5,7] The purpose of this note is to use this framework to examine the relation between definitions of angular momentum available in the two regimes. We shall show that, in stationary space–times, the definitions do agree; there is, in fact, "just one" spin vector. The investigation of the corresponding question for 4-momenta is greatly simplified due to the availability of the Komar integral for total mass. It is the absence, in general stationary space–times, of similar integrals for angular momentum that makes the present analysis difficult.

Why do we restrict ourselves to stationary space–times? Although a satisfactory definition of angular momentum is available at spatial infinity for a wide class of nonstationary space–times, it is only in the stationary case that the notion is free of supertranslation ambiguities at null infinity.[8–12] That is, the presence of gravitational radiation induces a *qualitative* change in the notion of angular momentum at null infinity. Therefore, in the general nonstationary context, it is difficult to imagine a simple, direct relation between the angular momentumlike quantities that have been introduced so far in the two regimes. It may turn out that if one restricts oneself to nonstationary space–times in which the gravitational radiation falls off in distant past at a suitable rate—and such a restriction may be implicit in the definition[7] of asymptotic flatness to be used in this note—one would be able to remove the supertranslation ambiguities and again compare the two spin vectors. However, even if this turns out to be the case, the present analysis in the stationary context would clearly serve as an essential first step in the required investigation.

2. PRELIMINARIES

In this section, we recall the definition of asymptotic flatness to be used in the main theorem and reexpress the formula for angular momentum at null infinity[8–12] in an intrinsic way, i.e., without reference to spin and conformally weighted functions representing basis vectors in the BMS Lie algebra, which makes its basic properties transparent. This discussion will also be useful for fixing notation.

Definition 1[7]: A space–time $(\hat{M}, \hat{g}_{ab})$ will be said to be asymptotically empty and flat at spatial and null infinity if there exists a space–time (M, g_{ab}) which is C^∞ everywhere except at a point i° where g_{ab} is $C^{>\circ}$, together with an imbedding of $\hat{M}$ into M (with which we identify $\hat{M}$ with its image in M) satisfying the following conditions:

(i) $\bar{J}(i^\circ) = M - \hat{M}$;

(ii) There exists a function Ω on M such that, on $\hat{M}$, $g_{ab} = \Omega^2 \hat{g}_{ab}$; on $[\dot{J}(i^\circ) - i^\circ]$, $\Omega = 0$, $\nabla_a \Omega \neq 0$; and, at i°, $\Omega = 0$, $\nabla_a \Omega = 0$, and $\lim_{\to i^\circ} \nabla_a \nabla_b \Omega = 2 g_{ab}|_{i^\circ}$; and,

(iii) There exists a neighborhood N of $J(i^\circ)$ in M such that (N, g_{ab}) is strongly causal and time orientable, and in $\hat{M} \cap N$, $\hat{g}_{ab}$ satisfies Einstein's vacuum equation.

The point i° represents "spatial infinity" of $(\hat{M}, \hat{g}_{ab})$ while $\mathscr{I} := [\dot{J}(i^\circ) - i^\circ]$ represents "null infinity." [Here, Hawking and Ellis[13] terminology has been used for the causal structure of (M, g_{ab}).] $\mathscr{I}$ is a disjoint union of two sets, $\mathscr{I}^+$

 0022-2488/79/071362-04$01.00

and $\mathscr{I}^-$, which contain points of $\mathscr{I}$ which are, respectively, to the future and to the past of i°. (The $C^{>0}$ differentiability requirement on g_{ab} ensures, in essence, only that g_{ab} is C°, smooth in its "angular dependence" at i°, and that its derivatives—i.e., its connection—suffer only finite "radial" discontinuities at i° (for details, see Ref. 7,5). The relation between the notions of asymptotic flatness expressed in this definition and other definitions available in the literature has been discussed in Ref. 7.

Let $(\widehat{M}, \hat{g}_{ab})$ be also stationary; i.e., let it admit a Killing field $\hat{t}^a$ which is timelike with $\hat{g}_{ab}\hat{t}^a\hat{t}^b$ bounded in $\tilde{N}\cap\widehat{M}$ for some neighborhood $\tilde{N}$ of $\dot{J}(i^\circ)$ in M. Then, a number of simplifications occurs.

Consider first the null regime. The presence of $\hat{t}^a$ enables one to select a preferred four parameter family of cross-sections, the shear-free cuts, of $\mathscr{I}^{\pm}$. Let us, for simplicity, restrict ourselves to $\mathscr{I}^+$; all our remarks will apply equally to $\mathscr{I}^-$. Denote the space of preferred cross sections by m^+. The presence of m^+ enables one to select a preferred Poincaré sub Lie algebra of the BMS Lie algebra associated with $\mathscr{I}^+$. m^+ itself is naturally equipped with the structure of a pseudo-Riemannian manifold, diffeomorphic to $\mathbb{R}^4$, the associated metric being *flat*. Finally, there is a natural isomorphism between the isometry Lie algebra of m^+ and the preferred Poincaré Lie algebra on $\mathscr{I}^+$ (for details, see, e.g., Ref. 14 or 11.) This isomorphism plays an important role in the introduction of conserved quantities. For example, the Bondi–Sachs [1,2] 4-momentum P_α is a constant vector field on m^+ defined by

$$P_\alpha V^\alpha := (-1/32\pi)\int_S {}^*K_{abcd}\, l^a X^b_V\, dS^{cd}, \tag{1}$$

where V^α is a constant vector field on m^+; X^b_V, the corresponding infinitesimal Poincaré transformation on $\mathscr{I}^+$ (which is a BMS translation since V^α is a translational Killing field on m^+); $K_{abcd} = \lim_{\to\mathscr{I}} \Omega^{-1} C_{abcd}$ is the asymptotic Weyl curvature; ${}^*K_{abcd} = \epsilon_{abmn} K^{mn}{}_{cd}$, its dual; S, any 2-sphere cross section of $\mathscr{I}^+$ and 1^a, the null vector field orthonormal to this cross section satifying $1^a\nabla_a\Omega = 1$. (Throughout, Greek indices will refer to m^+ and Latin ones to M.)

Next, we introduce angular momentum, $M_{\alpha\beta}$, at null infinity. The definition of P_α suggests an expression for $M_{\alpha\beta}$. Set

$$M_{\alpha\beta}(p)\, F^{\alpha\beta}(p) = (-1/32\pi)\int_S {}^*K_{abcd}\, l^a X^b{}_{F(p)}\, dS^{cd}, \tag{2}$$

where p is a point of m^+, i.e., a shear-free cross section of $\mathscr{I}^+$; $F^{\alpha\beta}(p)$ is an arbitrary skew tensor at p and $X^b{}_{F(p)}$ the infinitesimal Poincaré transformation on $\mathscr{I}^+$ corresponding to the Lorentz transformation about p generated by $F^{\alpha\beta}(p)$ on m^+. Using (1) and (2), it is straightforward to verify that $M_{\alpha\beta}(p)$ has the following "transformation property"

$$M_{\alpha\beta}(p') = M_{\alpha\beta}(p) + P_{[\alpha}T_{\beta]}, \tag{3}$$

where T_β is the vector connecting p and p' in m^+. Thus, the tensor $M_{\alpha\beta}(p)$ may indeed be regarded as the angular momentum of the given isolated system about the "origin" p.

Introducing suitable tetrad vectors on $\mathscr{I}^+$ and spin and conformally weighted functions (on the space of generators of $\mathscr{I}^+$) corresponding to the various BMS vector fields, one can show that the expression (2) agrees with angular momentum expressions, involving fields on $\mathscr{I}^+$, available in the literature. [8–12] The spin vector S_α is defined as usual by

$$S_\alpha := \epsilon_{\alpha\beta\gamma\delta} M^{\gamma\delta} P^\beta, \tag{4}$$

where $\epsilon_{\alpha\beta\gamma\delta}$ is the natural alternating tensor of m^+. From Eq. (3), it follows that S_α is a constant vector field on m^+ (orthonormal to P_α); unlike $M_{\alpha\beta}$, S_α is origin *independent*.

Finally, the situation at spatial infinity may be summarized as follows. The group of asymptotic symmetries is now the (infinite dimensional) Spi group. [5,7] In its structure, this group is very similar to the BMS group: it is a semidirect product of an infinite dimensional Abelian group (of Spi supertranslations) with the Lorentz group, and, furthermore, admits a preferred four-dimensional normal subgroup (of Spi translations). The ADM 4-momentum $\mathbf{P}_a$ (the generator of Spi translations) may be regarded as a vector in the tangent space at i°. The presence of the Killing field $\hat{t}^a$ again leads to the selection of a preferred Poincaré subgroup of the Spi group.[15] Under certain mild regularity requirements on the asymptotic curvature, one is then led to a definition of angular momentum. Set

$$\mathbf{M}_{ab}\,\mathbf{F}^{ab} = (1/8\pi)\int_{\mathbf{S}} \boldsymbol{\beta}_{ab}\,\mathbf{X}^b_{*\mathbf{F}}\, dS^a, \tag{5}$$

for arbitrary skew tensors $\mathbf{F}^{ab}$ in the tangent space of i°, where, $\mathbf{S}$ is any 2-sphere cross-section of the hyperboloid $\mathscr{D}$ of unit spacelike vectors at i°; $\mathbf{X}^b_{*\mathbf{F}} = {}^*\mathbf{F}^{ba}\,\boldsymbol{\eta}_a$, the restriction to $\mathscr{D}$ of the infinitesimal Lorentz transformation generated by ${}^*\mathbf{F}_{ab}$ in the tangent space of i°; and $\boldsymbol{\beta}_{ab}(\boldsymbol{\eta}) = \lim_{\to i^\circ} {}^*C_{ambn}\,\eta^m\eta^n$, the limit being taken along a spacelike curve with unit tangent η_a, is the tensor field on $\mathscr{D}$ representing the "magnetic part" of the asymptotic Weyl curvature. Thus, for each choice of the conformal metric g_{ab}, Eq. (5) defines a skew tensor $\mathbf{M}_{ab}$ at i°. It turns out that there is a close connection between the permissible conformal rescalings of g_{ab} and Spi translations. Under these rescalings, $\boldsymbol{\beta}_{ab}$ transforms just in the appropriate way for $\mathbf{M}_{ab}$ to satisfy the analog of Eq. (3). (For details, see Ref. 7 or 5.) The four-parameter family of skew tensors $\mathbf{M}_{ab}$ obtained by using all permissible conformal metrics represents the total angular momentum of the given isolated system. Finally, the spin vector $\mathbf{S}_a$ is defined via

$$\mathbf{S}_a := \boldsymbol{\epsilon}_{abcd}\,\mathbf{P}^b\,\mathbf{M}^{cd}, \tag{6}$$

where $\boldsymbol{\epsilon}_{abcd}$ is the alternating tensor at i° defined by g_{ab}. Again, $\mathbf{S}_a$ is a fixed vector at i°; unlike $\mathbf{M}_{ab}$, it is unaffected by conformal rescalings.

3. THE RELATION BETWEEN S_α AND $\mathbf{S}_\alpha$

Consider first the Bondi–Sachs 4-momentum P_α and the ADM 4-momentum $\mathbf{P}_a$. Using Einstein's equation on $\hat{g}_{ab}$, it is not difficult to show that P_α is necessarily colinear to t_α, the time translation on m^+ induced by the Killing field $\hat{t}^a$ on $(\widehat{M}, \hat{g}_{ab})$. Similarly, at spatial infinity one can show that $\mathbf{P}_a$ must be colinear to the vector at i° representing the ator of the infinitesimal Spi translation induced by $\hat{t}^a$.[15] Thus, in each regime the asymptotic rest frame selected by the 4-momentum coincides with that selected by the Killing field. Finally, using the Komar integral associated with $\hat{t}^a$, one can show that the Bondi–Sachs mass equals the ADM mass; $P^\alpha P_\alpha = \mathbf{P}_a \mathbf{P}^a$. In this sense, in stationary space–times the two available 4-momenta do agree.

The corresponding analysis of spin vectors is not as straightforward. Indeed, since the spin vectors S_α and $\mathbf{S}_a$ fail to be colinear to the time translations induced by $\hat{t}^a$ on $\mathscr{I}^+$ and Spi respectively—in fact they are orthogonal to these time translations—we need to introduce additional structure before we can even *formulate* the question of their equality. We therefore begin by introducing this structure. We have:

Lemma 1: There is a natural homomorphism ψ from the BMS Lie algebra associated with $\mathscr{I}^+$ onto the Lie algebra of infinitesimal Lorentz transformations in the tangent space at i°. The kernel of ψ is the supertranslation sub Lie algebra.

Proof: Using the definition of asymptotic flatness, it is easy to establish a natural diffeomorphism between the space of generators of $\mathscr{I}^+$ and the 2-sphere of future pointing null directions in the tangent space of i°. Fix an infinitesimal BMS transformation X^a on $\mathscr{I}^+$. X^a induces, via its action on the space of generators, an infinitesimal motion in the null cone in the tangent space of i°. It is easy to verify that this motion uniquely extends to an infinitesimal Lorentz transformation $\mathbf{X}^a$ in the tangent space of i°. Set $\psi(X^a) = \mathbf{X}^a$. This ψ is clearly a homomorphism. Finally, since the BMS supertranslations can be characterized by their property that they leave each generator of $\mathscr{I}^+$ invariant, it follows that X^a is in the kernel of ψ if and only if it is a BMS supertranslation. □

Fix a point p in $\mathscr{m}^+$ and a skew tensor $F_{\alpha\beta}$ at p. Let $X^a{}_{F(p)}$ be the BMS vector field on $\mathscr{I}^+$ corresponding to the infinitesimal Lorentz transformation about p generated by $F_{\alpha\beta}$ in $\mathscr{m}^+$. Since $\psi(X^a{}_{F(p)}) = \mathbf{X}^a$ is an infinitesimal Lorentz transformation at i°, it defines a unique skew tensor $\mathbf{F}_{ab}$ at i°. Thus, ψ gives rise to an isomorphism $\tilde{\psi}$ between constant (second rank) skew tensor fields on $\mathscr{m}^+$ and (second rank) skew tensors at i°: $\tilde{\psi}(F_{\alpha\beta}) = F_{ab}$. (That $\tilde{\psi}$ is independent of the initial choice of p in $\mathscr{m}^+$ follows from the fact that the kernel of ψ consists of BMS supertranslations.) The mapping $\tilde{\psi}$ in turn gives rise to an isomorphism $\hat{\psi}$ between constant vector fields on $\mathscr{m}^+$ and the tangent space at i°. Using this $\hat{\psi}$ we can now formulate our question concerning the two spin vectors: Does $\hat{\psi}(S_\alpha)$ equal $\mathbf{S}_a$?

In order to answer this question, we first note an important property of the mapping $\tilde{\psi}$. Consider the SO(3) subgroup of the Lorentz group about p (in $\mathscr{m}^+$) which leaves the vector P_α at p invariant, i.e., consider the little group of $P_\alpha|_p$. Denote the corresponding Lie algebra by $\mathscr{L}_p$. Similarly, denote the Lie algebra of the little group of $\mathbf{P}_a|_{i^\circ}$ by $\mathscr{L}_{i^\circ}$. Elements of $\mathscr{L}_p$ are represented by skew tensors $F_{\alpha\beta}(p)$ satisfying $F_{\alpha\beta}(p) P^\beta = 0$ and those of $\mathscr{L}_{i^\circ}$ by skew vectors $\mathbf{F}_{ab}$ at i° with $\mathbf{F}_{ab}\mathbf{P}^a = 0$. We have:

Lemma 2: $\tilde{\psi}$ is an isomorphism between the Lie algebras $\mathscr{L}_p$ and $\mathscr{L}_{i^\circ}$.

Proof: The key idea is to use the fact that on $\mathscr{m}^+$, P_α is parallel to the time translation induced by $\hat{t}^a$, and at i°, $\mathbf{P}_a$ is parallel to the Spi translation induced by $\hat{t}^a$; the two little groups are thus singled out by the presence of the Killing field $\hat{t}^a$. Fix an element $F_{\alpha\beta}(p)$ of $\mathscr{L}_p$ and consider the corresponding BMS vector field $X^a{}_{F(p)}$ on $\mathscr{I}^+$. Then, $\mathscr{L}_{\hat{t}} X^a_{F(p)} = 0$ on $\mathscr{I}^+$. Consider any extension X^a of $X^a_{F(p)}$ to a neighborhood N' of $\dot{J}(i^\circ)$ such that X^a is C^∞ in $N' \cap \hat{M}$, $C^{>0}$ at i°, and satisfies $\mathscr{L}_{\hat{t}} X^a = 0$. Taking the limit of this last equation along a spacelike curve approaching i° and using the fact[15] that $\lim_{\to i^\circ} \Omega^{-1}\hat{t}^a = -1/2\mathbf{K}^a + (\mathbf{K}\cdot\boldsymbol{\eta})\,\boldsymbol{\eta}^a$, where $\mathbf{K}^a$ is a fixed vector at i° and $\boldsymbol{\eta}^a$ the unit tangent at i° to the spacelike curve of approach, one obtains $\mathbf{K}\lim_{\to i^\circ} \nabla_{[a} X_{b]} \equiv \mathbf{K}^a \mathbf{F}_{ab} = 0$. However, $\mathbf{K}^a$ is parallel to the ADM 4-momentum $\mathbf{P}^a$.[15] Hence, $\mathbf{F}_{ab}$ belongs to $\mathscr{L}_{i^\circ}$. By construction, it is clear that $\mathbf{F}_{ab} = \tilde{\psi}(F_{\alpha\beta})$. Since $F_{\alpha\beta}$ is arbitrary, it follows that $\tilde{\psi}(\mathscr{L}_p) \subset \mathscr{L}_{i^\circ}$. Finally, since ψ is a homomorphism from the BMS Lie algebra *onto* the Lorentz Lie algebra at i° and since the kernel of this mapping contains only BMS supertranslations, it follows that $\tilde{\psi}$ is an isomorphism between $\mathscr{L}_p$ and $\mathscr{L}_{i^\circ}$. □

Remark: Lemma 2 implies, in particular, that $\hat{\psi}(P_\alpha)$ is proportional to $\mathbf{P}_a$. Since the existence of the Komar mass integral implies $P_\alpha P^\alpha = \mathbf{P}_a \mathbf{P}^a$, we have $\hat{\psi}(P_\alpha) = \mathbf{P}_a$. Thus, the isomorphism $\hat{\psi}$ enables us to obtain a more satisfactory statement concerning the agreement of the two 4-momenta than the one presented in the beginning of this section.

We can now consider the spin vectors S_α and $\mathbf{S}_a$. We have

Theorem: $\hat{\psi}(S_\alpha) = \mathbf{S}_a$.

Proof: From the definition of the two spin vectors and that of the mappings $\hat{\psi}$ and $\tilde{\psi}$ it follows that $\hat{\psi}(S_a) = \mathbf{S}_a$ if and only if $M_{\alpha\beta}(p) F^{\alpha\beta}(p) = \mathbf{M}_{ab}\mathbf{F}^{ab}$ for *some* p in $\mathscr{m}^+$ and arbitrary skew tensors $F^{\alpha\beta}(p) \in \mathscr{L}_p$, where $\mathbf{F}_{ab} = \tilde{\psi}(F_{\alpha\beta})$. [Because $F_{\alpha\beta}(p)$ belongs to $\mathscr{L}_p$, it follows that $M_{\alpha\beta}(p) F^{\alpha\beta} = M_{\alpha\beta}(p') F^{\alpha\beta}(p') \forall p' \in \mathscr{m}^+$, provided $F^{\alpha\beta}$ is a constant tensor field on $\mathscr{m}^+$, while, since by Lemma 2 $\mathbf{F}_{ab}$ belongs to $\mathscr{L}_{i^\circ}$, $\mathbf{M}_{ab}\mathbf{F}^{ab}$ is invariant under conformal rescalings of g_{ab}.] We shall therefore show that $M_{\alpha\beta} F^{\alpha\beta} = \mathbf{M}_{ab}\mathbf{F}^{ab}$ for all $F_{\alpha\beta}$ in $\mathscr{L}_p$. The key idea is to construct a 2-form in $\hat{M}$ with the property that its integral over a 2-sphere tends to $M_{\alpha\beta} F^{\alpha\beta}$ as the 2-sphere converges to the shear-free cross-section p of $\mathscr{I}^+$ and to $\mathbf{M}_{ab}\mathbf{F}^{ab}$ as the 2-sphere converges to i°. To this end, we introduce certain fields.

Fix an element $F_{\alpha\beta}(p)$ of $\mathscr{L}_p$. Consider the BMS vector field $X^a_{*F(p)}$ on $\mathscr{I}^+$. Fix a C^1, three-dimensional, spacelike submanifold T in (M, g_{ab}) passing through i° which is orthogonal to the ADM 4-momentum $\mathbf{P}_a$ at i°. Consider an extension $X^a_{*F(p)}$ to a neighborhood of $\dot{J}(i^\circ)$ which is $C^{>0}$ at i°, C^∞ elsewhere, and orthogonal to T. [Such an extension is possible because, by Lemma 2, $\psi(X^a_{*F(p)})$ is a "boost" vector field in the tangent space of i°.] Next, introduce a null vector field l^a in a neighborhood of $\dot{J}(i^\circ)$ satisfying the following properties: (i) l^a is orthogonal to the shear-free cut p; (ii) $l^a\nabla_a\Omega = 1$ on $\mathscr{I}$ (so that $\mathscr{L}_{\hat{t}} l^a = (-\lim_{\to\mathscr{I}} \mathscr{L}_{\hat{t}} \log\Omega)\, l^a$ on $\mathscr{I}$); and, (iii) $\mathscr{L}_{\hat{t}} l^a = (-\mathscr{L}_{\hat{t}} \log\Omega)\, l^a$ in the intersection of the this neighborhood with $\hat{M}$. Finally, con-

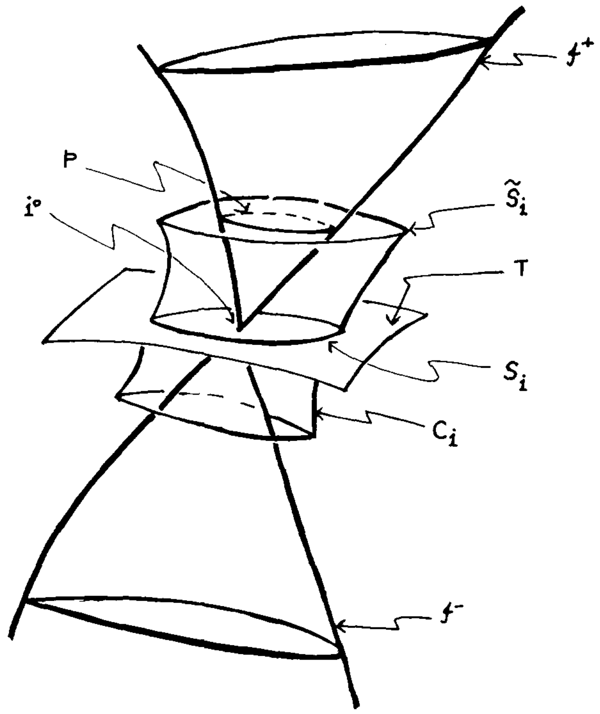

FIG. 1. The physical space–time $(\hat{M}, \hat{g}_{ab})$ consists of points of the completed space–time (M, g_{ab}) which are spacelike related to i°. T is a spacelike hypersurface in (M, g_{ab}) passing through i°, being orthogonal to the ADM 4-momentum $\mathbf{P}_a$ at i°. C_i is a sequence of timelike hypercylinders (defined by, say, $\hat{g}_{ab}\hat{t}^a\hat{t}^b = $ const on each C_i) which converges to $\dot{J}(i^\circ)$. S_i is the sequence of 2-spheres, the intersection of C_i with T, which converges to i°, while $\tilde{S}_i$ is a sequence of 2-sphere cross sections of C_i which converges to a given shear-free cut p of $\mathscr{I}^+$.

sider a family of timelike hypercylinders C_i *in M* (defined by, say, $\hat{g}_{ab}\,\hat{t}^a\hat{t}^b = $ const on each C_i), converging to $\dot{J}(i^\circ)$. Let these C_i intersect the spacelike 3-surface T in a family of 2-spheres S_i which converge to i°. Let $\tilde{S}_i (\subset C_i)$ denote another family of 2-sphere cross sections which converge to the given shear-free cross section p of $\mathscr{I}^+$. (See Fig. 1.)

Now, on the cylinders C_i we have

$$\int_{\Delta C_i} (\nabla_{[m} K_{ab]cd}\, l^c X^d)\, dS^{mab} = \int_{\tilde{S}_i} - \int_{S_i} K_{abcd}\, l^c X^d\, dS^{ab} \tag{7}$$

where ΔC_i is the part of the hypercylinder C_i bounded by $\tilde{S}_i$ and S_i. Using the fact that any 2-form Ω_{ab} satisfies

$$3\hat{t}^m \nabla_{[m}\Omega_{ab]} + 2\nabla_{[b}(\Omega_{a]m}\hat{t}^m) - \mathscr{L}_{\hat{t}}\Omega_{ab} = 0, \tag{8}$$

that the integral of the exact 2-form $\nabla_{[b}(\Omega_{a]m}\hat{t}^m)$ must vanish on a 2-sphere, and that $\mathscr{L}_{\hat{t}} K_{abcd}\, l^c = 0$, Eq.(7) simplifies to

$$\frac{1}{3}\int_{\Delta C_i} K_{abcd}\, l^c (\mathscr{L}_{\hat{t}} X)^d\, dt \wedge dS^{ab} = \int_{\tilde{S}_i} - \int_{S_i} K_{abcd}\, l^c X^d\, dS^{ab}. \tag{9}$$

(Here, the surface element dS^{mab} has been decomposed: $dS^{mab} = \hat{t}^m\, dt \wedge dS^{ab}$.) Now we take the limit as C_i's approach $\dot{J}(i^\circ)$. Using the commutation relations of the Poincaré group, the Bianchi identities and the fact that $F_{\alpha\beta}$ belongs to $\mathscr{L}_p$, we obtain

$$\lim_{\to \dot{J}(i^\circ)} \int_{\Delta C_i} K_{abcd}\, l^c (\mathscr{L}_{\hat{t}} X)^d\, dt \wedge dS^{ab} = 0. \tag{10.a}$$

Using Bianchi identities and the fact that on p, $X^a_{*F(p)} = \epsilon^a{}_b X^b_{F(p)}$ with ϵ^a_b the natural alternating tensor on p, defined by g_{ab}, it follows that

$$\lim_{\to \mathscr{I}} \int_{\tilde{S}_i} K_{abcd}\, l^c X^d\, dS^{ab} = -32\pi M_{\alpha\beta} F^{\alpha\beta}. \tag{10.b}$$

Finally, using the asymptotic field equations [5,7] on the hyperboloid $\mathscr{D}$ at i°, the fact that $\mathbf{F}_{ab}$ belongs to $\mathscr{L}_{i^\circ}$ and that X^a is normal to the spacelike 3-surface T, we have

$$\lim_{\to i^\circ} \int_{S_i} K_{abcd}\, l^c X^d\, dS^{ab} = -32\pi \mathbf{M}_{ab} \mathbf{F}^{ab}. \tag{10.c}$$

From Eqs. (9) and (10) one now has: $M_{\alpha\beta} F^{\alpha\beta} = \mathbf{M}_{ab}\mathbf{F}^{ab}$. Since $F_{\alpha\beta}$ (in $\mathscr{L}_p$) is arbitrary, we have $\hat{\psi}(S_\alpha) = \mathbf{S}_\alpha$. □

ACKNOWLEDGMENT

We wish to thank Jürgen Ehlers for his comments on the first draft of this paper.

[1] H. Bondi, A.W.K. Metzner, and M.J.G. Van der Burg, Proc. R. Soc. London, Ser. A **269**, 21 (1962).
[2] R.K. Sachs, Proc. R. Soc. London Ser. A **270**, 103 (1962); Phys. Rev. **128**, 2851 (1962).
[3] R. Penrose, Proc. R. Soc. London Ser. A **284**, 159 (1965).
[4] R. Arnowitt, S. Deser, and C.W. Misner, in *Gravitation, an Introduction to Current Research*, edited by L. Witten (Wiley, New York, 1962), and references contained therein.
[5] A. Ashtekar and R.O. Hansen, J. Math. Phys. **19**, 1542 (1978).
[6] P.D. Sommers, J. Math. Phys. **19**, 549 (1978).
[7] A. Ashtekar, to appear in the Einstein Centenary Volume, edited by P. Bergmann, J.N. Goldberg, and A. Held (Plenum).
[8] J. Winicour, J. Math. Phys. **9**, 861 (1968).
[9] B. D. Bramson, Proc. R. Soc. London, Ser. A. **341**, 463 (1975).
[10] M. Streubel, J. Gen. Rel. Grav. **9**, 551 (1978).
[11] M. Streubel, *"Conserved" Quantities Related to Asymptotic Symmetries for Isolated systems in General Relativity*, Ph.D. thesis, Max Planck Institut für Astrophysik, München, 1978.
[12] J. Winicour, to appear in the Einstein Centenary Volume, edited by P. Bergmann, J.N. Goldberg, and A. Held (Plenum).
[13] S.W. Hawking and G.F.R. Ellis, *The Large Scale Structure of Space–time* (Cambridge U. P., London, 1973).
[14] E.T. Newman and R. Penrose, J. Math. Phys. **7**, 863 (1966).
[15] A. Ashtekar and A. Magnon–Ashtekar, "On conserved quantities in general relativity" J. Math. Phys. **20**, 793 (1979).

The application of group theory to generate new representability conditions for rotationally invariant density matrices

R. M. Erdahl

Department of Mathematics, Queen's University, Kingston, Ontario, Canada

C. Garrod

Department of Physics, University of California, Davis, California 95616

B. Golli and M. Rosina

Faculty of Natural Sciences and Technology and J. Stefan Institute, University of Ljubljana, Ljubljana, Yugoslavia

(Received 18 July 1978; revised manuscript received 16 November 1978)

New linear conditions are derived which must be satisfied by a two-body density matrix. In the derivation, the ideas of Davidson and McRae are extended so that full use is taken of the symmetries of the system. The coefficients of the linear form are determined by means of reduction of a chosen group in a physically meaningful chain of its subgroups.

1. INTRODUCTION

In the density matrix method the state of an N-fermion system is represented by its two-body density matrix $\widehat{\Gamma}$, $\Gamma_{abcd} = \langle \Psi \,|(\hat{a}_a\hat{a}_b)^\dagger(\hat{a}_c\hat{a}_d)|\,\Psi\rangle$, rather than by its wavefunction Ψ; $\hat{a}_a$, $\hat{a}_b$, $\hat{a}_c$, $\hat{a}_d$ are annihilation operators for some fixed orthonormal basis which is finite. The two-body density matrix for the ground state of the system is determined variationally by minimizing the functional

$$E = \min_{\widehat{\Gamma}\in D_N^2} \operatorname{tr}\widehat{H}\widehat{\Gamma} = \min_{\widehat{\Gamma}\in D_N^2} \tfrac{1}{4}\sum_{abcd} H_{abcd}\Gamma_{abcd}.$$

The variational parameters Γ_{abcd} must satisfy certain subsidiary conditions (representability conditions) in order to correspond to a physical state; the convex set of N-representable density matrices in denoted by D_N^2. The coefficients H_{abcd} are matrix elements of the reduced Hamiltonian

$$H_{abcd} = H^{(0)}(\delta_{ac}\delta_{bd} - \delta_{ad}\delta_{bc})/\tfrac{1}{2}N(N-1)$$
$$+ (H^{(1)}_{ac}\delta_{bd} + H^{(1)}_{bd}\delta_{ac} - H^{(1)}_{ad}\delta_{bc} - H^{(1)}_{bc}\delta_{ad})/$$
$$\times (N-1) + H^{(2)}_{abcd} - H^{(2)}_{abdc}.$$

Since the set of all representability conditions is not known in an explicit form, only subsets of necessary conditions have been used in direct variational calculations.[1] Important conditions are the nonnegativity of the two-body density matrix $\widehat{\Gamma}$, the particle hole matrix $G_{abcd} = \langle \Psi \,|(\hat{a}_a^\dagger\hat{a}_b)^\dagger(\hat{a}_c^\dagger\hat{a}_d)|\,\Psi\rangle$, and the two-hole matrix $Q_{abcd}{}'\langle \Psi \,|(\hat{a}_a\hat{a}_b)(\hat{a}_c\hat{a}_d)^\dagger|\,\Psi\rangle$. This subset gives reasonably good results for some systems with a small number of valence particles. However, calculations with a larger number of particles indicate that other important conditions are still missing.

In the search for new conditions, Davidson and McRae[2] and Erdahl[3] have introduced a finite set of linear inequalities by requiring that the set of mutually commuting operators $\widehat{A}_{ab} = \widehat{N}_a\widehat{N}_b$ (all a and b), $\widehat{N}_a = \hat{a}_a^\dagger\hat{a}_a$, should have physically realizable expectation values.

In this paper we generate new conditions by using a more general set of commuting operators. In this way, the symmetry of the system can be properly taken into account (Sec. 1). The general form of new conditions is presented in Sec. 3. In choosing a more general set of commuting operators it is convenient to demand: (i) easy verification that the operators commute, (ii) an easily computable spectrum of operators, (iii) incorporation of all the symmetries of the system, (iv) some hint that the operators are physically meaningful. We consider here a group and a chain of its subgroups which have been commonly used in the wavefunction method. We choose a commuting subset of generators of the group and the Casimir operators of all subgroups. Group theoretical techniques allow the simultaneous eigenvalues of these operators to be easily obtained. These techniques are developed and illustated in Sec. 4. The use of new conditions is discussed in Sec. 5 and their effectiveness is tested on a simple many-body system in Sec. 6.

2. REPRESENTABILITY CONDITIONS FOR THE SCALAR PART OF THE TWO-BODY DENSITY MATRIX ("SCALAR CONDITIONS")

We first want to introduce the concept of "scalar conditions." In order to express the energy of a system with a rotationally invariant Hamiltonian,

 0022-2488/79/071366-09$01.00

$$E = \mathrm{tr}\hat{H}\hat{\Gamma} = \frac{1}{4}\sum_{\substack{abcd\\LS}} H^{LS}_{abcd}\Gamma^{LS}_{abcd}, \tag{1}$$

it is sufficient to know the scalar part of the density matrix Γ^{LS}_{abcd}. This can be seen as follows. We denote the single-particle basis by $|ams\rangle$, $a\equiv(n_a, j_a)$ denoting the "level" and ms denoting the third component of orbital angular momentum and spin. In nuclear physics j_a and m refer to total angular momentum (half integer) and s to isospin. What we shall call spin should mean spin to an atomic physicist and isospin to a nuclear physicist. A suitable two-body basis is defined by the two-body creation operators

$$(\hat{F}^{LMS\Sigma}_{ab})^\dagger = \sum_{mm'ss'} C^{LM}_{j_a m j_b m'} C^{S\Sigma}_{(s/2)(s'/2)}(\hat{a}_{ams}\hat{a}_{bm's'})^\dagger. \tag{2}$$

Then

$$\hat{\Gamma} = \sum_{\substack{abcd\\LL'SS'\\\lambda\mu\sigma\tau}} \Gamma^{LL'SS'\lambda\mu\sigma\tau}_{abcd} \sum_{MM'\Sigma\Sigma'} (-1)^{L-M} C^{\lambda\mu}_{L-ML'M'}$$

$$\times(-1)^{S-\Sigma} C^{\sigma\tau}_{S-\Sigma S'\Sigma'}(\hat{F}^{LMS\Sigma}_{ab})^\dagger \hat{F}^{L'M'S'\Sigma'}_{cd}. \tag{3}$$

If we ignore spin orbit forces, then H will contain only the scalar part ($\lambda = \mu = \sigma = \tau = 0$):

$$\hat{H} = \sum_{\substack{abcd\\LS}} H^{LS}_{abcd} \sum_{MM'\Sigma\Sigma'} (-1)^{L-M} C^{00}_{L-MLM'}(-1)^{S-\Sigma}$$

$$\times C^{00}_{S-\Sigma S\Sigma'}(\hat{F}^{LMS\Sigma}_{ab})^\dagger \hat{F}^{LM'S\Sigma'}_{cd}, \tag{4}$$

and tr $\hat{H}\hat{\Gamma}$ will also contain only the scalar term of $\hat{\Gamma}$, $\Gamma^{LS}_{abcd} \equiv \Gamma^{LLSS0000}_{abcd}$.

Since we need only the scalar part of $\hat{\Gamma}$, it would be desirable to have representability conditions involving the scalar two-body density matrix alone. The Garrod–Percus theorem[4] that $\mathrm{tr}\hat{A}\hat{\Gamma} \geqslant A_<$ is necessary and sufficient ($\hat{A}$ = any operator, $A_<$ its lowest eigenvalue in N-particle space) is still valid if $\hat{A}$ and $\hat{\Gamma}$ are both restricted to scalar operators (see the proof in Appendix A). We shall use group theory to provide us with a suitable subset of scalar operators $\hat{A}$ for which we know $A_<$. In this way we shall generate a subset of necessary conditions.

3. GENERAL FORM OF THE DIAGONAL CONDITIONS

Davidson and McRae,[2] and Erdahl[3] introduced a set of conditions which are necessary and sufficient for the representability of the diagonal elements of a two-body density matrix in a two-body Slater basis $\Gamma_{\alpha\beta\alpha\beta}$. We shall introduce more general "diagonal conditions" referring to diagonal elements in any two-body basis.

The conditions on $\Gamma_{\alpha\beta\alpha\beta}$ are generated by the set of commuting operators $\hat{A}_{\alpha\beta} = \hat{N}_\alpha \hat{N}_\beta$, $\hat{N}_\alpha = \hat{a}^\dagger_\alpha \hat{a}_\alpha$. General diagonal conditions are generated by a general set $\{\hat{A}^i\}$ of commuting operators $\hat{A}^i$, $i = 1,...,Z$ in the following way.

Take a set of all positive definite operators $\hat{\Gamma}$ with a normalized trace. Define a parameter space $\hat{A}^i = \mathrm{tr}\hat{A}^i\hat{\Gamma}$, $i = 1,...,Z$. If the dimension of the Hilbert space of 0-particle up to N-particle states is K, then there are K simultaneous eigenvalues of $\hat{A}^i$, denoted by A^i_k, $k = 1,...,K$. The vectors A^i_k define K points in the parameter space, spanning a convex polytope. In analogy with Ref. 2, we can state the following

Theorem: A necessary and sufficient condition that a given vector $\hat{A}^i$ is representable is that it lie within the polytope:

$$A^i = \sum_{k=1}^{K} y_k A^i_k, \quad y_k \geqslant 0, \quad \sum_{k=1}^{K} y_k = 1. \tag{5}$$

By representable we mean that there exists a representable $\hat{\Gamma}$ such that

$$A^i = \mathrm{tr}\hat{A}^i\hat{\Gamma}.$$

The proof of this theorem follows the ideas in Ref. 2 and is given in Appendix B.

Let us justify the name "diagonal." If $|\eta\rangle$ are simultaneous eigenstates of $\hat{A}^i$ in the two-particle space with eigenvalues A^i_η, one can write

$$A^i = \mathrm{tr}\hat{A}^i\hat{\Gamma} = \sum_{\eta\eta'} \langle\eta|\hat{A}^i|\eta'\rangle\langle\eta'|\hat{\Gamma}|\eta\rangle$$

$$= \sum_\eta A^i_\eta \Gamma_{\eta\eta}.$$

The conditions on A^i are necessary for $\hat{\Gamma}$. They are necessary and sufficient for the linear combinations $\Sigma_\eta A^i_\eta \Gamma_{\eta\eta}$ of diagonal elements of $\hat{\Gamma}$ in this specific basis. In the two-body Slater basis, the conditions are necessary and sufficient for all diagonal element $\Gamma_{\alpha\beta\alpha\beta}$ (Ref. 2); on the other hand, in an arbitrary basis, the new diagonal conditions are not sufficient for all $\Gamma_{\eta\eta}$ (just for Z linear combinations of $\Gamma_{\eta\eta}$). The reason is that one obtains a complete set of commuting operators only in the two-body Slater basis, while the number Z of a general set of commuting two-body operators $\hat{A}^i$ is smaller, but they are different from the Slater-basis conditions.

In the following we shall be interested in "scalar diagonal conditions," choosing scalar $\hat{A}^i$ only. There are many choices of the set $\{\hat{A}^i\}$, and one should consider physical relevance and numerical feasibility in choosing it. The construction of the polytope would be a prohibitive task if it had to be performed numerically for large Z. It is here that group theory can help.

4. APPLICATION OF GROUPS

First, one chooses a relevant group and a chain of its subgroups. Group theory then offers a convenient set of commuting operators, namely, a few generators plus all Casimir operators. Group theory also offers all their simultaneous eigenvalues ("the polytope"). If the groups in the chain contain O(3) as a subgroup, all Casimir operators are scalars. We shall give three examples.

TABLE I. $j_a = \frac{1}{2}$.

Quantum numbers					Eigenvalues of		
N_a	SU(2) T_a	U(2) $[f]$	Sp(2) (σ)	O(3) L_a	$\hat{N}_a$	$\hat{N}_a^2$	$\frac{1}{2}\hat{C}_a = \hat{L}_a^2 = 2\hat{N}_a - \hat{N}_a^2/2 - \hat{T}_a^2$
0	0	[0]	(0)	0	0	0	0
1	$\frac{1}{2}$	[1]	(1)	$\frac{1}{2}$	1	1	$\frac{3}{4}$
2	0	[2]	(2)	1	2	4	2
	1	[11]	(0)	0	2	4	0
3	$\frac{1}{2}$	[21]	(1)	$\frac{1}{2}$	3	9	$\frac{3}{4}$
4	0	[22]	(0)	0	4	16	0

A. One-level conditions

The single-particle operators acting only on states of a given level (= subshell) are generators of the group $U(2(2j_a + 1))$, so this is the obvious group with which to start. A suitable chain is

$$U(2(2j_a + 1)) \supset SU(2) \otimes U(2j_a + 1),$$

$$\text{or} \begin{cases} U(2j_a + 1) \supset O(2j_a + 1) \supset O(3), & \text{for integer } j_a, \\ U(2j_a + 1) \supset Sp(2j_a + 1) \supset O(3), & \text{for half-integer } j_a. \end{cases}$$

The chain offers the following one- and two-body scalar operators (some care is required to choose only linearly independent operators).

The only scalar generator is $\hat{N}_a$ (the number of particles in the level a).

The Casimir operator of $U(2(2j_a + 1))$ depends only on N_a (in the space of totally antisymmetric wavefunctions), and therefore it can be expressed in terms of $\hat{N}_a$ and $\hat{N}_a^2$.

The Casimir operator of SU(2) is $2\hat{S}_a^2$ (the total spin in level a).

The Casimir operator of $U(2j_a + 1)$ is linearly dependent of $\hat{N}_a$, $\hat{N}_a^2$, and $\hat{S}_a^2$ if restricted to the space of antisymmetric total wavefunctions [the Young tableau of $U(2j_a + 1)$ must be adjoint to the Young tableau of SU(2)].

The Casimir operator of $O(2j_a + 1)$ or $Sp(2j_a + 1)$ is[5]:

$$\hat{C}_a = \frac{1}{2j_a + 1} \sum_{\text{odd } k} (-1)^k (2k+1)^{3/2} \times \sum_q C^{00}_{kqk-q} \hat{U}^k_q \hat{U}^k_{-q}, \tag{6}$$

where

$$\hat{U}^k_q = \sum_{mm's} C^{j_a m}_{kqj_a m'} \hat{a}^\dagger_{ams} \hat{a}_{am's}.$$

The Casimir operator of O(3) is $\{3/j_a(j_a + 1)(2j_a + 1)\} L_a^2$ (the total orbital angular momentum in level a).

Thus we have five commuting operators, $\hat{N}_a$, $\hat{N}_a^2$, $\hat{S}_a^2$,

TABLE II. $j_a = 1$.

Quantum numbers					Eigenvalues of			
N_a	SU(2) S_a	U(3) $[f]$	O(3) (σ)	L_a	$\hat{N}_a$	$\hat{N}_a^2$	$\hat{S}_a^2$	$2\hat{C}_a = \hat{L}_a^2$
0	0	[0]	0	0	0	0	0	0
1	$\frac{1}{2}$	[1]	1	1	1	1	$\frac{3}{4}$	2
2	0	[2]	2	2	2	4	0	6
			0	0	2	4	0	0
	1	[11]	1	1	2	4	2	2
3	$\frac{1}{2}$	[21]	2	2	3	9	$\frac{3}{4}$	6
			1	1	3	9	$\frac{3}{4}$	2
	$\frac{3}{2}$	[111]	0	0	3	9	$\frac{15}{4}$	0
4 5 6	One gets the same sets of quantum numbers as for 6–N, only the columns for $\hat{N}_a$ and $\hat{N}_a^2$ are different.							

TABLE III. $j_a = \frac{3}{2}$.

	Quantum numbers				Eigenvalues of				
N_a	SU(2) T_a	U(4) $[f]$	Sp(4) $(\sigma_1\sigma_2)$	O(3) $L_{a\,\min}, L_{a\,\max}$	$\hat{N}_a$	$\hat{N}_a^2$	$\hat{T}_a^2$	$\hat{C}_a$	$\hat{L}_a^2$
0	0	[0]	(00)	0	0	0	0	0	0
1	$\frac{1}{2}$	[1]	(10)	$\frac{3}{2}$	1	1	$\frac{1}{2}$	$\frac{5}{2}$	$\frac{15}{4}$
2	0	[2]	(20)	1,3	2	4	0	6	2,12
	1	[11]	(00)	0	2	4	2	0	0
			(11)	2	2	4	2	4	6
3	$\frac{1}{2}$	[21]	(10)	$\frac{3}{2}$	3	9	$\frac{3}{4}$	$\frac{5}{2}$	$\frac{15}{4}$
			(21)	$\frac{1}{2},\frac{7}{2}$	3	9	$\frac{3}{4}$	$\frac{15}{2}$	$\frac{3}{4},\frac{63}{4}$
	$\frac{3}{2}$	[111]	(10)	$\frac{3}{2}$	3	9	$\frac{15}{4}$	$\frac{5}{2}$	$\frac{15}{4}$
4	0	[22]	(00)	0	4	16	0	0	0
			(11)	2	4	16	0	4	6
			(22)	2,4	4	16	0	10	6,20
	1	[211]	(20)	1,3	4	16	2	6	2,12
			(11)	2	4	16	2	4	6
	2	[1111]	(00)	0	4	16	6	0	0
5–8	the same values for $\hat{T}_a^2$, $\hat{T}_a$, $\hat{L}_a^2$ as for 8–N								

TABLE IV. $j_a = 2$.

	Quantum numbers				Eigenvalues				
N_a	SU(2) S_a	U(5) $[f]$	O(5) $(\omega_1\omega_2)$	O(3) $L_{a\,\min}, L_{a\,\max}$	$\hat{N}_a$	$\hat{N}_a^2$	$\hat{S}_a^2$	$\hat{C}_a$	$\hat{L}_a^2$
0	0	[0]	(00)	0	0	0	0	0	0
1	$\frac{1}{2}$	[1]	(10)	2	1	1	$\frac{3}{4}$	2	6
2	0	[2]	(00)	0	2	4	0	0	0
			(20)	2,4	2	4	0	5	6,20
	1	[11]	(11)	1,3	2	4	2	3	2,12
3	$\frac{1}{2}$	[21]	(10)	2	3	9	$\frac{3}{4}$	2	6
			(21)	1,5	3	9	$\frac{3}{4}$	6	2,30
	$\frac{3}{2}$	[111]	(11)	1,3	3	9	$\frac{15}{4}$	3	2,12
4	0	[22]	(22)	0,6	4	16	0	8	0,42
			(20)	2,4	4	16	0	5	6,20
			(00)	0	4	16	0	0	0
	1	[211]	(21)	1,5	4	16	2	6	2,30
			(11)	1,3	4	16	2	3	2,12
	2	[1111]	(10)	2	4	16	6	2	6
5	$\frac{1}{2}$	[221]	(22)	0,6	5	25	$\frac{3}{4}$	8	0,42
			(21)	1,5	5	25	$\frac{3}{4}$	6	2,30
			(10)	2	5	25	$\frac{3}{4}$	2	6
	$\frac{3}{2}$	[2111]	(20)	2,4	5	25	$\frac{15}{4}$	5	6,20
			(11)	1,3	5	25	$\frac{15}{4}$	3	2,12
	$\frac{5}{2}$	[11111]	(00)	0	5	25	$\frac{35}{4}$	0	0
6–10 the same values for $\hat{T}_a^2$, $\hat{C}_a$, $\hat{L}_a^2$ as for 10–N									

TABLE V. $j_a = \frac{5}{2}$.

	Quantum numbers				Eigenvalues of				
	SU(2)	U(6)	Sp(6)	O(3)					
N_a	T_a	$[f]$	$(\sigma_1\sigma_2\sigma_3)$	L_a	$\hat{N}_a$	$\hat{N}_a^2$	$\hat{T}_a^2$	$\hat{C}_a$	$\hat{L}_a^2$
0	0	[0]	(000)	0	0	0	0	0	0
1	$\frac{1}{2}$	[1]	(100)	$\frac{5}{2}$	1	1	$\frac{3}{4}$	$\frac{7}{2}$	$\frac{35}{4}$
2	0	[2]	(200)	1,3	2	4	0	8	2,12
	1	[11]	(000)	0	2	4	2	0	0
			(110)	2,4	2	4	2	6	6,20
3	$\frac{1}{2}$	[21]	(100)	$\frac{5}{2}$	3	9	$\frac{3}{4}$	$\frac{7}{2}$	$\frac{35}{4}$
			(210)	$\frac{1}{2},\frac{13}{2}$	3	9	$\frac{3}{4}$	$\frac{21}{2}$	$\frac{3}{4},\frac{195}{4}$
	$\frac{3}{2}$	[111]	(100)	$\frac{5}{2}$	3	9	$\frac{15}{4}$	$\frac{7}{2}$	$\frac{35}{4}$
			(111)	$\frac{3}{2},\frac{9}{2}$	3	9	$\frac{15}{4}$	$\frac{15}{2}$	$\frac{15}{4},\frac{99}{4}$
4	0	[22]	(000)	0	4	16	0	0	0
			(110)	2,4	4	16	0	6	6,20
			(220)	0,8	4	16	0	14	0,72
	1	[221]	(200)	1,5	4	16	2	8	2,30
			(110)	2,4	4	16	2	6	6,20
			(211)	1,7	4	16	2	12	2,56
	2	[1111]	(000)	0	4	16	6	0	0
			(110)	2,4	4	16	6	6	6,20
5	$\frac{1}{2}$	[221]	(100)	$\frac{5}{2}$	5	25	$\frac{3}{4}$	$\frac{7}{2}$	$\frac{35}{4}$
			(210)	$\frac{1}{2},\frac{13}{2}$	5	25	$\frac{3}{4}$	$\frac{21}{2}$	$\frac{3}{4},\frac{195}{4}$
			(111)	$\frac{3}{2},\frac{9}{2}$	5	25	$\frac{3}{4}$	$\frac{15}{2}$	$\frac{15}{4},\frac{99}{4}$
			(221)	$\frac{1}{2},\frac{17}{2}$	5	25	$\frac{3}{4}$	$\frac{31}{2}$	$\frac{3}{4},\frac{323}{4}$
	$\frac{3}{2}$	[2111]	(100)	$\frac{5}{2}$	5	25	$\frac{15}{4}$	$\frac{7}{2}$	$\frac{35}{4}$
			(210)	$\frac{1}{2},\frac{13}{2}$	5	25	$\frac{15}{4}$	$\frac{21}{2}$	$\frac{3}{4},\frac{195}{4}$
			(111)	$\frac{3}{2},\frac{9}{2}$	5	25	$\frac{15}{4}$	$\frac{15}{2}$	$\frac{15}{4},\frac{99}{4}$
	$\frac{5}{2}$	[11111]	(100)	$\frac{5}{2}$	5	25	$\frac{35}{4}$	$\frac{7}{2}$	$\frac{35}{4}$
6	0	[222]	(200)	1,5	6	36	0	8	2,30
			(211)	1,7	6	36	0	12	2,56
			(222)	1,9	6	36	0	18	2,90
	1	[2211]	(000)	0	6	36	2	0	0
			(110)	2,4	6	36	2	6	6,20
			(220)	0,8	6	36	2	14	0,72
			(211)	1,7	6	36	2	12	2,56
	2	[21111]	(200)	1,5	6	36	6	8	2,30
			(110)	2,4	6	36	6	6	6,20
	3	[111111]	(000)	0	6	36	12	0	0

7–12 the same values for $\hat{T}_a^2$, $\hat{C}_a$, $\hat{L}_a^2$ as for 12–N

$\hat{L}_a^2$, and $\hat{C}_a$. Their expectation values can be expressed in terms of the scalar one-body density matrix and the scalar particle hole matrix and are given in Appendix C. For $j_a = 3$ there is an additional group in the chain, G_2. For $j_a > 3$ there may be also additional groups, but they have not yet been worked out. The number of linearly independent commuting operators for low j_a is smaller than five; for j_a equal to $\frac{1}{2}$, 1, $\frac{3}{2}$, 2, $\frac{5}{2}$, 3, the number of linearly independent operators is 3,4,5,5,5,6, respectively.

Examples of the "polytope" are shown in Tables I–V. Simultaneous eigenvalues for five commuting operators are presented for all $j_a \leqslant \frac{5}{2}$. The quantum numbers S_a, $[f]$, $(\omega_1\omega_2)$, and L_a corresponding to the irreducible representations of SU(2), U($2j_a + 1$), O($j_a + 1$) [or Sp($2j_a + 1$)], and O(3) are taken from Ref. 6 for integer j_a and from Ref. 7 for half-integer j_a. The number of vertices of the "polytope" can be reduced by noting that the that differ only in L_a are convex combinations of two vertices, the one that corresponds to the lowest L_a and the one corresponding to the highest L_a. It is therefore enough to take only the lowest and the highest L_a in those cases where more than two L_a correspond to the same $(\omega_1\omega_2)$.

TABLE VI. The chain U(24)⊃SU(4)⊗U(6), SU(4)⊃SU(2)⊗SU(2), U(6)⊃SU(3)⊃O(3).

	Quantum numbers				
N	U(6) $[\tilde{f}]$	SU(4) $[f]$	SU(3) $(\lambda\mu)$	O(3) $L_{\min}\ L_{\max}$	SU(2)⊗SU(2) (TS)
4	[4]	[1111]	(80)	0,8	(00)
			(42)	0,6	
			(04)	0,4	
			(20)	0,2	
4	[31]	[211]	(61)	1,7	(01), (10), (11)
			(42)	0,6	
			(23)	1,5	
			(31)	1,4	
			(12)	1,3	
			(20)	0,2	
4	[22]	[22]	(42)	0,6	(00), (02), (20), (11)
			(31)	1,4	
			(04)	0,4	
			(20)	0,2	
4	[221]	[31]	(50)	1,3	(01), (10), (11), (12), (21)
			(23)	1,5	
			(31)	1,4	
			(12)	1,3	
			(01)	1	
4	[1111]	[4]	(12)	1,3	(00), (11), (22)

[a]The quantum numbers L and TS have to be combined so that in each row any given L is combination with any given pairs TS.

B. Many-level conditions

Having single-particle states in several "levels" $a,b,c,\cdots$ one can choose the following decomposition,

$$\mathrm{U}(\Omega)\supset\mathrm{U}(2(2j_a+1))\oplus\mathrm{U}(2(2j_b+1))$$

$$\oplus\,\mathrm{U}((2j_c+1))\oplus\cdots.$$

There are about five commuting operators per level (see the one-level conditions in the previous section). In addition, there are "mixed" operators $\widehat{N}_a\widehat{N}_b$, $\widehat{S}_a\widehat{S}_b$, $\widehat{L}_a\widehat{L}_b,\cdots$. One can get quite a large parameter space and some care is needed in bookkeeping all the vertices of the polytope.

C. Conditions generated by the SU(3) subgroup

If one uses the orbitals of a harmonic oscillator shell, the following chain of groups is commonly used in nuclear physics:

$$\mathrm{U}(\Omega)\supset\mathrm{SU}(4)\otimes\mathrm{U}(6),$$

$$\mathrm{SU}(4)\supset\mathrm{SU}(2)\otimes\mathrm{SU}(2),$$

$$\mathrm{U}(6)\supset\mathrm{SU}(3)\supset\mathrm{O}(3).$$

In Table VI we present the quantum numbers corresponding to this chain of groups. The vertices of the "polytope" (eigenvalues of operators) are not given in the same way as in Table I. The reader can construct them for himself using the following formulas:

Operator	Eigenvalue
$\widehat{N}$	N,
$\widehat{N}^2$	N^2
$\widehat{C}(\mathrm{SU}(4))$	$f_1(f_1-1)+f_2)(f_2-3)+f_3(f_3-5)$ $+f_4(f_4-7)-N^2/4+4N$,
$\widehat{C}(\mathrm{SU}(3))$	$\lambda^2+\lambda\mu+\mu^2+3\lambda+3\mu$,
L^2	$L(L+1)$,
$\widehat{T}^2$	$T(T+1)$,
$\widehat{S}^2$	$S(S+1)$.

The explicit expressions for the Casimir operators of the SU(3) can be found in Ref. 8, and that of SU(4) in Ref. 9. The table for other particle numbers can be easily obtained by combining the U(6)⊃SU(3)⊃O(3) decomposition[8] and the SU(4)⊃SU(2)⊗SU(2) decomposition.[9] The complete table is not given here because it is rather long and it can be constructed by a computer using subtables from Refs. 8 and 9.

D. The canonical chain of subgroups

Let us finally notice that the set of operators $\widehat{A}_{\alpha\beta}=\widehat{N}_\alpha\widehat{N}_\beta$ corresponds to the decomposition:

$$\mathrm{U}(\Omega)\supset\mathrm{U}(\Omega-1)\supset\mathrm{U}(\Omega-2)\supset\cdots\supset\mathrm{U}(1).$$

This decomposition offers the maximum number of linearly independent commuting operators, but these operators are not scalar.

TABLE VII. Test of violation of one-level conditions for light nuclei in the Zucker model and the $d\,2S$ model.

Nucleus	Model	$N_{\rm val}$	j_a	with $\hat{C}_a$ K	"violation"	with $\hat{Q}\hat{Q}$ operator K	"violation"
15_O	Zucker	3	5/2	13	1.55	20	0.51
16_O	Zucker	4	5/2	27	0	49	0
17_O	Zucker	5	5/2	40	0	85	0
18_O	Zucker	6	5/2	58	0	133	0
19_F	Zucker	7	5/2	71	0	169	0
20_{Ne}	Zucker	8	5/2	85	0	198	0
20_{Ne}	$d\,2S$	4	3/2	18	0	18	0
20_{Ne}	$dS\,2$	4	5/2	27	0.02	49	0
28_{Si}	$d\,2S$	12	3/2	28	0	28	0
28_{Si}	$dS\,2$	12	5/2	98	0	218	1.28

[a] $N_{\rm val}$ is the number of valence nucleons in the model, j_a indicates the level for which the one-level conditions were taken, K is the number of vertices of the polytope. The "violation" is explained in Sec. 5. The conditions for the level $j_a = \frac{1}{2}$ were not violated.

5. USE OF THE NEW CONDITIONS

The new conditions are given in the form of coordinates of a polytope in the parameter space. It is easy to test a given density matrix as to whether or not it violates the new conditions. One needs a computer routine for linear programming to test whether the set of linear equalities and inequalities (5) has a feasible solution. If there is no feasible solution then the program will give information about how far outside the polytope the given point lies (the "violation").

However, it is difficult to incorporate these new conditions into a direct variational calculation in the present form. We now have the vertices of the polytope. It would be more convenient to have a set of linear inequalities (the facets of the polytope) involving the variational parameters directly. The procedure used by Davidson and McRae[2] to generate the facets of a polytope, given the vertices, is too lengthy, and also it would yield an intractable number of facets. We are trying to develop a method of generating only the relevant facets for a given point or a given problem, but we have not yet been successful.

6. TEST OF THE SCALAR DIAGONAL CONDITIONS ON A SMALL MODEL SYSTEM

So far we have tested only the efficiency of the one-level conditions and the conditions generated by the SU(3) group for $N = 4$. We have used two models: (i) the Zuker model of some light nuclei (inert ^{12}C core plus valence nucleons in the $1p_{1/2}$, $1d_{5/2}$, $2s_{1/2}$ levels), and (ii) the $d\,2s$ model of some light nuclei (inert ^{16}O core plus valence nucleons in the $1d_{5/2}$, $1d_{3/2}$, $2s_{1/2}$ levels).

We have tested the two-body density matrix obtained by a direct variational calculation[10] incorporating the nonnegativity of the two-body density matrix, the particle hole matrix and the two-hole matrix ("Γ GQ"). Some of the new conditions were violated and some were not (indicated by "violation" = 0). These results are given in Table VII. To see the sensitivity of the diagonal conditions to the choice of operators, we have also taken the scalar operator $\widehat{Q}\widehat{Q}$,

$$\widehat{Q}\widehat{Q} = \frac{\sqrt{125}}{6} \sum_q C^{00}_{2q2-q}\, \widehat{U}^2_q\, \widehat{U}^2_{-q},$$

instead of the Casimir operator of the Sp(6) group for the level $j_a = \frac{5}{2}$. The operator $\widehat{Q}\widehat{Q}$ commutes with first four one-level operators; however, its eigenvalues have to be computed numerically. The results shown in Table VII indicate that the choice of commuting operators other than Casimir operators leads in some cases to weaker and in some cases to stronger conditions.

The conditions generated by the SU(3) group were tested in the case of ^{20}Ne in the $d\,2s$ model, but they were not violated. However, this may not be too surprising since the constraints requiring the proper $\widehat{N}$, $\widehat{N}^2$, and $\widehat{T}^2$ values were already imposed in the "Γ GQ" calculation. Also, the energy of the "Γ GQ" calculation was reasonably good, 1.5 MeV below the value obtained by complete configuration mixing.

We cannot include the new conditions in direct variational calculations until we derive a tractable algorithm for generating the relevant inequalities. We did calculate in one case, that of ^{15}O in the Zuker model, with the following result for the energy of the ground state:

Complete configuration mixing	− 26.02 MeV,
"Γ GQ"	− 27.09 MeV,
"Γ GQ" plus one-level conditions	− 26.69 MeV.

The improvement is noticable but not complete—the result comes only one third of the way towards the "exact energy." If we include the $\widehat{Q}\widehat{Q}$ operator instead of the Casimir operator, the energy improvement is somewhat smaller, the result being − 26.80 MeV.

From one-level conditions alone one cannot yet draw conclusions about the efficiency of the complete set of new conditions. Work is in progress to test the efficiency of the many level conditions and other alternative choices.

The most difficult remaining problem is how to design a numerical procedure for including the new conditions in the direct variational calculations.

APPENDIX A: THE GARROD–PERCUS THEOREM FOR THE SCALAR TWO-BODY DENSITY MATRIX

For simplicity we shall consider only operators which operate on orbital coordinates and not on spin coordinates. The extension to general operators is straightforward.

A multipole expansion of a two-body operator $\widehat{A}$ can be written:

$$\widehat{A} = \sum_{\substack{abcd\\ JJ'\lambda\mu}} A^{JJ'}_{abcd}(\lambda\mu) \sum_{MM'} (-1)^{J-M} C^{\lambda\mu}_{J-MJ'M'} (\widehat{F}^{JM}_{ab})^\dagger \widehat{F}^{J'M'}_{cd} \equiv \sum_{\lambda\mu} \widehat{A}_{\lambda\mu}. \tag{A1}$$

The following orthogonality relation holds:

$$\mathrm{tr}\widehat{A}_{\lambda\mu}\widehat{B}_{\lambda'\mu'} = \delta_{\lambda\lambda'}\delta_{\mu\mu'}\,\mathrm{tr}\widehat{A}_{\lambda\mu}\widehat{B}_{\lambda\mu}. \tag{A2}$$

Theorem: Let $\widehat{A}$ be a Hermitian operator and $\widehat{A}_0$ be the scalar part ($\lambda = \mu = 0$) in its multipole expansion (A1), then the Garrod–Percus condition on the scalar two-body density matrix $\widehat{\Gamma}$,

$$\mathrm{tr}\widehat{A}\widehat{\Gamma} \geqslant A_<, \tag{A3}$$

is implied by the condition

$$\mathrm{tr}\widehat{A}_0\widehat{\Gamma} \geqslant \widehat{A}_{0<}, \tag{A4}$$

$A_<$ and $A_{0<}$ being the lowest eigenvalues of $\widehat{A}$ and $\widehat{A}_0$, respectively.

Proof: The left-hand sides of (A3) and (A4) are equal because of the orthogonality relation (A2). We have therefore only to prove $A_{0<} \geqslant A_<$. Since $\widehat{A}_0$ is a scalar operator it commutes with the operator $\bar{J}$ and its eigenvectors can be labeled by J and M. Let $|J_0M_0\rangle$ be the eigenvector in the N-body Hilbert space corresponding to $A_{0<}$. Define the ensemble density matrix

$$\hat{\rho}^N = \frac{1}{\sqrt{2J_0+1}}\sum_{M=-J_0}^{J_0}(-1)^{J_0-M}C^{00}_{J_0MJ_0-M} \times |J_0M\rangle\langle J_0M|.$$

Since $\widehat{A}_0$ does not depend on M, it follows that $\mathrm{Tr}\widehat{A}_0\hat{\rho}^N = A_{0<}$.

For $\widehat{A}$, however: $A_< \leqslant \mathrm{Tr}\widehat{A}\hat{\rho}^N$. Applying Eq. (A1), the Wigner–Eckart theorem and the orthogonality relation between Clebsch–Gordon coefficients, one gets

$$\mathrm{Tr}\widehat{A}\hat{\rho}^N = \sum_{M\lambda\mu}\frac{(-1)^{J_0-M}}{\sqrt{2J_0+1}}C^{00}_{J_0MJ_0-M}\langle J_0M|\widehat{A}_{\lambda\mu}|J_0M\rangle$$

$$= \sum_{\lambda\mu}\sum_{M}\frac{(-1)^{J_0-M}}{\sqrt{2J_0+1}}C^{00}_{J_0MJ_0-M}(-1)^{\lambda-J_0+M-\mu} \times \frac{\sqrt{2J_0+1}}{\sqrt{2\lambda+1}}C^{\lambda\mu}_{J_0MJ_0-M}\langle J_0\|\widehat{A}_{\lambda\mu}\|J_0\rangle$$

$$= \sum_{\lambda\mu}\delta_{0\lambda}\delta_{0\mu}\frac{(-1)^{\lambda-\mu}}{\sqrt{2\lambda+1}}\langle J_0\|\widehat{A}_{\lambda\mu}\|J_0\rangle = A_{0<}.$$

Therefore, $A_< \leqslant A_{0<}$.

We have assumed the existence of $\widehat{A}_0$ for every $\widehat{A}$ since for those $\widehat{A}$ with $\widehat{A}_0 = 0$ the condition (A3) is void.

APPENDIX B: PROOF OF THEOREM (5) [EQ. (5)]

Let us rewrite Theorem (5) in the following form:

Theorem: A set of real numbers A^i, $i = 1,\ldots,Z$ represents the expectation values of a set of commuting operators $\{\widehat{A}^i\}$ iff A^i can be written as

$$A^i = \sum_{k=1}^{K} y_k A^i_k, \tag{B1a}$$

where y_k satisfy:

$$y_k \geqslant 0, \quad \sum_{k=1}^{K} y_k = 1 \tag{B1b}$$

(A^i_k and K are defined in Sec. 3).

Proof: Let the simultaneous eigenstates of the set $\{\widehat{A}^i\}$ be denoted by $|k\rangle$, $k = 1,\ldots,K$, then $A^i_k = \langle k|\widehat{A}^i|k\rangle$. The expectation value A^i of $\widehat{A}^i$ in an arbitrary state $|\Psi\rangle$, $|\Psi\rangle = \Sigma_k c^\psi_k|k\rangle$, can be expressed as

$$A^i = \langle\Psi|\widehat{A}^i|\Psi\rangle = \sum_{k=1}^{K}|c^\psi_k|^2A^i_k,$$

where $|c^\psi_k|^2$ are bounded by $|c^\psi_k|^2 \geqslant 0$, $\Sigma^K_{k=1}|c^\psi_k|^2 = 1$.

Ensemble averages of $\widehat{A}^i$ in a system described by an ensemble density matrix, $\hat{\rho}^N$, $\hat{\rho}^N = \Sigma_\phi w_\phi|\Phi\rangle\langle\Phi|$ can be expressed as

$$A^i = \mathrm{Tr}\widehat{A}^i\hat{\rho}^N = \sum_{k=1}^{K}\sum_{\phi}w_\phi|c^\phi_k|^2A^i_k.$$

The coefficients $\Sigma_\phi w_\phi|c^\phi_k|^2$ are again bounded by (B1b). Condition (B1) is therefore necessary.

Condition (B1) is also sufficient because one can always construct a state vector $|y\rangle$, $|y\rangle = \Sigma_k\sqrt{y_k}\,|k\rangle$, where y_k satisfy condition (B1b), such that A^i from (B1a) is equal to $\langle y|\widehat{A}^i|y\rangle$.

APPENDIX C: RELATIONS BETWEEN ONE LEVEL OPERATORS AND THE ONE-BODY AND THE PARTICLE HOLE MATRIX

In accordance with Eq. (3) we define the scalar one-body density matrix, $\hat{\gamma}$, and the scalar particle-hole matrix, $\widehat{G}$:

$$\hat{\gamma} = \sum_{ab}\delta_{j_aj_b}\gamma_{ab}\widehat{B}^{0000}_{ab},$$

and

$$\widehat{G} = \sum_{\substack{abcd\\LS}}G^{LS}_{abcd}\sum_{M\Sigma}(-1)^{L-M}C^{00}_{L-MLM}(-1)^{S-\Sigma} \times C^{00}_{S-\Sigma S\Sigma}(\widehat{B}^{LMS\Sigma}_{ab})^\dagger\widehat{B}^{LMS\Sigma}_{cd},$$

where

$$\widehat{B}^{LMS\Sigma} = \sum_{\substack{mm'\\ss'}}(-1)^{j_b-m'}C^{JM}_{j_amj_bm'}(-1)^{-s'/2} \times C^{S\Sigma}_{(s/2)(-s'/s)}\hat{a}^\dagger_{ams}\hat{a}_{bm's'},$$

and

$$\gamma_{ab} = \frac{1}{\sqrt{2(2j_a+1)}}\langle\Psi|\widehat{B}^{0000}_{ab}|\Psi\rangle,$$

$$G^{LS}_{abcd} = \frac{1}{(2L+1)(2S+1)}\sum_{M}\langle\Psi|(\widehat{B}^{LMS}_{ab})^\dagger\widehat{B}^{LMS}_{cd}|\Psi\rangle.$$

The expectation values of the one-level operators expressed in terms of $\hat{\gamma}$ and $\hat{G}$ are:

$$\langle \Psi | \hat{N}_a | \Psi \rangle = 2(2j_a + 1)\gamma_{aa},$$

$$\langle \Psi | \hat{N}_a^2 | \Psi \rangle = 2(2j_a + 1)G^{00}_{aaaa},$$

$$\langle \Psi | \hat{L}_a^2 | \Psi \rangle = 2j_a(j_a + 1)(2j_a + 1)G^{10}_{aaaaa},$$

$$\langle \Psi | \hat{S}_a^2 | \Psi \rangle = \tfrac{3}{2}(2j_a + 1)G^{01}_{aaaaa},$$

$$\langle \Psi | \hat{C}_a(\mathrm{O}(2j_a + 1)) | \Psi \rangle = 2 \sum_{k \text{ odd}}^{2j_a} (2k + 1)G^{k0}_{aaaa}, \quad \text{integer } j_a,$$

$$\langle \Psi | \hat{C}_a(\mathrm{Sp}(2j_a + 1)) | \Psi \rangle = 2 \sum_{k \text{ odd}}^{2j_a - 1} (2k + 1)G^{k0}_{aaaa}, \quad \text{half-integer } j_a.$$

Relations between G^{LS}_{abcd} and Γ^{LS}_{abcd} can be found in Ref. 1.

The eigenvalues of the Casimir operator $\hat{C}_a(\mathrm{O}(2j_a + 1))$ in terms of the irreducible representation labels of the $O(2j_a + 1)$, $(\omega_1\omega_2\cdots\omega_{j_a})$, and those of the Casimir operator $\hat{C}_a(\mathrm{Sp}(2j_a + 1))$ in terms of $(\sigma_1\sigma_2\cdots\sigma_{j_a + 1/2})$ are[5]

$$\langle \hat{C}_a(\mathrm{O}(2j_a + 1)) \rangle = \tfrac{1}{2}[\omega_1(\omega_1 + 2j_a - 1) + \omega_2(\omega_2 + 2j_a - 3) + \cdots\omega_{j_a}(\omega_{j_a} + 1)],$$

$$\langle \hat{C}_a(\mathrm{Sp}(2j_a + 1)) \rangle = \tfrac{1}{2}[\sigma_1(\sigma_1 + 2j_a + 1) + \sigma_2(\sigma_2 + 2j_a - 1) + \cdots + \sigma_{j_a + 1/2}(\sigma_{j_a + 1/2} + 2)].$$

[1]C. Garrod, M.V. Mihailović, and M. Rosina, J. Math. Phys. **16**, 868 (1975).
[2]W.B. McRae and E.R. Davidoson, J. Math. Phys. **13**, 1527 (1972).
[3]R.M. Erdahl, Queen's Mathematical Preprints No. 1971–48 (Queen's University, Kingston, Ontario, 1971).
[4]C. Garrod and J.K. Percus, J. Math. Phys. **5**, 1765 (1965).
[5]B.F. Bayman, *Groups and their Applications to Spectroscopy*, lecture notes (NOTDITA, Copenhagen 1960).
[6]A. Jahn, Proc. Roy. Soc. A **201**, 516 (1950).
[7]B.H. Flowers, Proc. Roy. Soc. A **212**, 248 (1952).
[8]J.P. Elliott, Proc. Roy. Soc. A **245**, 128 (1958); J.P. Elliott, in *Selected Topics in Nuclear Theory*, edited by F. Janouch (Vienna, IAEA 1963), p. 157.
[9]K.T. Hecht and S.C. Pang, J. Math. Phys. **10**, 1571 (1969); A. Partensky and C. Maquin, J. Math. Phys. **19**, 511 (1978).
[10]M.V. Mihailović and M. Rosina, Nucl. Phys. A **237**, 221 (1975).

Segal quantization of dynamical systems

Franco Gallone[a)]

Department of Physics, Princeton University, Princeton, New Jersey 08540 and Istituto di Scienze Fisiche dell'Università, 20133 Milano, Italy [b)]

Antonio Sparzani

Istituto di Scienze Fisiche dell'Università, 20133 Milano, Italy and Istituto Nazionale di Fisica Nucleare, Sezione di Milano, Milano, Italy

(Received 27 November 1978)

Segal quantization, usually thought of and used as a tool for quantizing kinematical frameworks, is extended to (finite-dimensional) dynamical systems, i.e., to kinematical frameworks plus dynamical motions applied to them. Such a procedure allows us to classify quantum dynamical systems, and to understand how physical inequivalence appears in spite of von Neumann unitary equivalence, for the former is also grounded on the evolution operators of the systems being different functions of the labeled observables of the systems. Second quantization is also examined and is shown to be just one possible procedure of quantization, which can only be used in a particular class of cases.

1. INTRODUCTION

Since the years of Poincaré, the most clear-cut mathematical model for a phase space description of a classical mechanical system has been recognized to be a differentiable manifold. "This manifold always has a special geometric property, pertaining to the occurrence of phase variables in canonically conjugate pairs, called symplectic structure."[1] When the symplectic space on which this classical picture is based features particular properties, a quantization procedure can be used, which is called Segal quantization.

In fact, according to Segal,[2,3] a classical system whose phase space "kinematical" description is based upon a linear symplectic space $(\mathscr{M},B)$ gets quantized through a Weyl system over $(\mathscr{M},B)$; the algebra of observables of the resulting quantum system is an algebra (the Weyl algebra) uniquely determined by the Weyl system. This quantization procedure can be performed at least when $\mathscr{M}$ is finite-dimensional (in which case both B and the Weyl system are essentially unique) or, more generally, when $\mathscr{M}$ is a Hilbert space (the "single particle space" of the quantum field theory), whose inner product has B as its imaginary part. Indeed, in such cases a Weyl system over $(\mathscr{M},B)$ is known to exist; moreover, whenever a Weyl system over a linear symplectic space $(\mathscr{M},B)$ does exist, the resulting Weyl algebra is unique (up to isomorphism), in the sense that it depends just on $(\mathscr{M},B)$ and not on the Weyl system over $(\mathscr{M},B)$ used for its construction.

A nontrivial and meaningful task is to look for a generalization of the Segal quantization procedure expounded above such as to include dynamics along with kinematics. We will call *dynamical system* (both in the classical and in the quantum case) what results from the coupling of a kinematical picture of the system with one of the evolution laws by which the motion of its observables may be ruled. Therefore, the above mentioned task amounts to looking for a general way of quantizing those dynamical systems whose kinematical part may be quantized by the Segal procedure. This quantization procedure of dynamical systems, in which the Segal quantization of the kinematical picture is still used, can be called Segal quantization as well, in a natural way. Of course, a further generalization could be taken up in which one not only adds dynamics to kinematics, but also generalizes kinematics itself, allowing the phase space to be a general differentiable symplectic manifold; naturally, the Segal quantization of kinematical pictures should first be properly generalized in order to cope with this situation.

In the present paper we are tackling the problem of Segal quantization of dynamical systems, even if in a reduced form, as we will impose limitations to the systems we consider. In fact, we will examine here just those classical systems whose (linear) phase spaces are finite-dimensional and whose evolution laws are linear. We believe the thorough analysis of the simple situation we are carrying on here is likely to be suggestive of what a solution to more general problems may look like (as it is often the case), and therefore we hope it will be helpful in the broader analysis we intend to develop in future work. The next (but hopefully not the last) step, which we leave to a forthcoming paper, will be quite naturally to consider systems whose phase spaces are infinite-dimensional Hilbert spaces (namely, field theory models).

As we said before, we are dealing with the Segal quantization of classical dynamical systems in the present paper, considering however just the systems which are linear in both their (finite) phase spaces and their evolution laws. Indeed, this is a sensible choice of a starting point toward more

[a)]A Fulbright–Hays Grantee. Supported in part by the Fondazione A. della Riccia.
[b)]Permanent address.

 0022-2488/79/071375-10$01.00

general analyses, since in a lot of infinite systems linearity ensures the existence and even the uniqueness of the physical vacuum.[4] We will find out what quantum dynamical systems arise—through the Segal quantization—from the classical systems we are considering, how they can be classified, and how they and their classification must be interpreted.

In view of the thorough treatment of the problem of quantization we want to give here for the linear and finite case, we will also investigate what second quantization amounts to for the systems we are dealing with. We will clarify how it is a very special quantization procedure, which can be performed in a limited number of cases only.

Finally, we warn the reader that in order to avoid unnecessarily cumbersome notation we will explicitly deal with and write formulas for the one degree of freedom case only (two-dimensional phase space); the results we will get in this way can easily be generalized to the general finite-dimensional case, for which they remain true, except for a result on the positivity of the energy spectrum, as will be specified below.

2. SEGAL QUANTIZATION

A nondegenerate real skew-symmetric form B on a real vector space $\mathscr{M}$ is called a symplectic form, and the pair $(\mathscr{M},B)$ is called a symplectic space. A linear automorphism of $\mathscr{M}$ which preserves the form B is called a symplectic transformation of $(\mathscr{M},B)$ and the group of such transformations will be denoted by Aut $(\mathscr{M},B)$. These are the basic ingredients occurring in the kinematical pictures of a number of classical systems, once we interpret $\mathscr{M}$ as the phase space of the system, refer B to the occurrence of phase variables in canonically conjugate pairs, and represent the symmetries of the system by elements of Aut$(\mathscr{M},B)$. Notice that more general symmetries could be represented by autodiffeomorphisms of $\mathscr{M}$ which preserve B. However, we are confining our analysis to linear symmetries since the motions we are studying in the present paper are linear, for the reasons touched upon in the introduction. The adjective linear, which should accompany so many words throughout the paper, will be dropped when no confusion may arise; sometimes, however, it will be written in parentheses just to keep the reader aware of the scope of our analysis. Also, as explained in the Introduction, throughout this paper we are dealing explicitly with 1-dimensional linear classical systems only, namely systems whose phase space is $\mathbb{R}+\mathbb{R}$. Therefore, from now on $\mathscr{M}=\mathbb{R}+\mathbb{R}$. As far as the linear structure of $\mathscr{M}$ is concerned, there is just one symplectic form on $\mathscr{M}$; in fact, two symplectic forms B_1 and B_2 are connected by a linear automorphism A of $\mathscr{M}$, namely $A\in$Aut $\mathscr{M}$ exists such that $B_1(m,m')=B_2(Am,Am'),\forall m,m'\in\mathscr{M}$. We recall also that, whatever the form B is, Aut $(\mathscr{M}B)$ is isomorphic to SL$(2,\mathbb{R})$.

If now $\mathscr{U}(\mathscr{H})$ is the group of unitary operators of a separable Hilbert space $\mathscr{H}$, an $\mathscr{H}$-valued Weyl system (WS) over $(\mathscr{M},B)$ is defined as a map $W\colon\mathscr{M}\to\mathscr{U}(\mathscr{H})$ such that

(i) $W(m_1)W(m_2)=\exp[(i/2)B(m_1,m_2)]W(m_1+m_2)$, $\forall m_1,m_2\in\mathscr{M}$,

(ii) $\mathbb{R}\ni t\to W(tm)\in\mathscr{U}(\mathscr{H})$ is weakly continuous, $\forall m\in\mathscr{M}$.

An $\mathscr{H}$-valued WS is called irreducible (IWS) if the only operators in $\mathscr{H}$ which commute with the range of W are the multiples of the unit operator. The problem of existence and uniqueness of IWS's over $(\mathscr{M},B)$ has been solved in a clear-cut way since a long time. As to existence, it is proved by construction: If Q_s and P_s are the operators defined in $L^2(\mathbb{R})$ by $(Q_sf)(x)=xf(x)$, $(P_sf)(x)=-if'(x)$, on suitable domains, and if m_1, m_2 are two elements of $\mathscr{M}$ such that $B(m_1,m_2)=1$ (it is trivial to show that two such elements always exist for any B, and that they are a linear basis in $\mathscr{M}$), the map

$$\mathscr{M}\ni am_1+bm_2\to\exp[-(i/2)ab]\exp[-iaP_s]\times\exp\{-ibQ_s\}\in\mathscr{U}(L^2(\mathbb{R}))$$

is an $L^2(\mathbb{R})$-valued WS over $(\mathscr{M},B)$. It is an IWS which is called the Schrödinger WS over $(\mathscr{M},B)$ and related to (m_1,m_2); also, if m_1',m_2' are any two elements of $\mathscr{M}$ such that $B(m_1',m_2')=1$, the Schrödinger WS over $(\mathscr{M},B)$ and related to (m_1',m_2') is unitarily equivalent to the system related to (m_1,m_2). As for uniqueness, it is settled by a theorem of von Neumann,[5] which establishes that any WS over $(\mathscr{M},B)$ is a Schrödinger WS within unitary equivalence and multiplicity.

Since the main goal of this paper is to discuss the quantization of a classical system as a whole, it is worth recalling here the sense in which a WS over $(\mathscr{M},B)$ quantizes the kinematical picture $(\mathscr{M},B)$ of a classical system. Let m_1 and m_2 be two elements of $\mathscr{M}$ such that $B(m_1,m_2)=1$; such a "canonical" pair $(m_1,m_2)_B$ determines a connection between WS's over $(\mathscr{M},B)$ and localizable one-dimensional quantum systems[6,7]; recall that a localizable one-dimensional quantum system can be defined as a pair (U,Q), where U is a weakly continuous one-parameter group of unitary operators on a separable Hilbert space $\mathscr{H}$, Q is a self-adjoint operator in $\mathscr{H}$, and the following relation is satisfied

$$U(a)QU(-a)=Q-a\mathbb{1}_{\mathscr{H}},\quad\forall a\in\mathbb{R}.$$

To show the connection determined by (m_1,m_2) we first notice that, if W is an $\mathscr{H}$-valued WS over $(\mathscr{M},B)$, then the one-parameter group $U_W^{(m_1)}$ defined by the relation

$$U_W^{(m_1)}(a)=W(am_1),\quad\forall a\in\mathbb{R}$$

and the self-adjoint operator $Q_W^{(m_2)}$ defined by the relation

$$\exp(-iaQ_W^{(m_2)})=W(am_2),\quad\forall a\in\mathbb{R}$$

[which simply means that $Q_W^{(m_2)}$ is the Stone theorem generator of the one-parameter group $\mathbb{R}\ni a\to W(am_2)\in\mathscr{U}(\mathscr{H})$] form a pair $(U_W^{(m_1)},Q_W^{(m_2)})$ which is a localizable one-dimensional quantum system (in $\mathscr{H}$). Second and conversely, we notice that if (U,Q) is a localizable one-dimensional quantum system (in $\mathscr{H}$), then

$$\mathscr{M}\ni am_1+bm_2\to\exp[-(i/2)ab]U(a)\exp(-ibQ)\in\mathscr{U}(\mathscr{H})$$

is a WS W such that $(U_W^{(m_1)},Q_W^{(m_2)})=(U,Q)$. Notice now that a localizable one-dimensional quantum system (U,Q) (in $\mathscr{H}$) admits of the following interpretation: $\mathscr{H}$ is the Hilbert space in which a quantum system is described, Q and the Stone theorem generator P of U are two observables for the system; Q is interpreted as position (along a direction) and P as momentum (along the same direction), as U is interpreted as the group of the displacements along the Q direction. Whether or not all other observables are functions of Q and P depends upon the particular system under consideration: Essentially, they are not if the system is more than one-dimensional, neither are they if the system has nonclassical observables such as the spin, but surely they are if (U,Q) is an irreducible system, namely, if the only operators which commute with Q and the range of U are the multiples of the unit operator (it is clear that irreducible localizable systems are connected with IWS's). Therefore, a WS W over $(\mathscr{M},B)$ quantizes a classical kinematical picture $(\mathscr{M},B)$ and this quantization is determined by a canonical pair $(m_1,m_2)_B$ since W provides both a Hilbert space in which the quantized system is to be described and two labeled observables,[8] namely the momentum $P_W^{(m_1)}$ and the position $Q_W^{(m_2)}$, defined respectively by the relations:

$$\exp(-iP_W^{(m_1)}) = W(m_1), \quad \exp(-iQ_W^{(m_2)}) = W(m_2);$$

m_1 and m_2 can be thought of as those two points of the phase space in which the classic observables position, resp. momentum, assume the values 1, resp. 0, and 0, resp. 1. It is clear that the canonical pair $(m_1,m_2)_B$ that labels the operators to be called momentum and position is not a part of the quantization procedure, nor is it uniquely defined by the property $B(m_1,m_2) = 1$ [in fact, any pair arising from (m_1,m_2) by a symplectic transformation shares this property and is therefore a canonical pair]. Such a pair is, on the contrary, a part of the definition of the classical system itself [in this connection, notice that B *is* uniquely defined by a canonical pair, since if $(m_1,m_2)_B$ and $(m_1,m_2)_{B'}$ are canonical pairs then $B = B'$].

In our discussion of the physical meaning of Segal's quantization procedure, we have used the notion of localizable quantum system. Of course, we could have used the notion of canonical commutation rules as well; as is well known, though, these are equivalent to the Weyl system "exponentiated" formalism, provided that proper care has been taken in the definitions and restrictions on such rules; so we have preferred the more unambiguous formulation in terms of unitary operators. In what follows we will fix a canonical pair, which we will still write simply as $(m_1,m_2)_B$, setting $m_1 = (1,0)$, $m_2 = (0,1)$, whence B results into

$$B((\alpha_1,\beta_1),(\alpha_2,\beta_2)) = \alpha_1\beta_2 - \alpha_2\beta_1, \quad \forall(\alpha_1,\beta_1),(\alpha_2,\beta_2)\in\mathscr{M}.$$

Having fixed $(m_1,m_2)_B$ in this way, we will drop m_1,m_2 and B wherever they occur, and B will always mean the above written symplectic form. So, we will write simply P_W and Q_W and we will speak of WS's over $\mathscr{M}$ and of the Schrödinger WS over $\mathscr{M}$. Notice, by the way, that the Schrödinger WS over $\mathscr{M}$ leads to a quantum description in the Hilbert space $L^2(\mathbb{R})$ such that the momentum and position observables are represented by P_s and Q_s respectively; besides, from elementary properties of the Schrödinger WS, it is easily inferred that an IWS is always an injective map $\mathscr{M} \to \mathscr{U}(\mathscr{H})$.

In what follows we will consider just the WS's over $\mathscr{M}$ which are irreducible; in fact, as we mentioned before, the Weyl algebras defined by the WS's over $\mathscr{M}$ are all isomorphic, and the multiplicity allowed by the von Neumann uniqueness theorem essentially leaves room for the representation (which is not already included in the WS over $\mathscr{M}$) of more than one-dimensional systems and of observables without a classical provenance (such as the spin). And in the present paper we are not interested in any of these possibilities, since we look mainly at the quantization of the motions characterizing one-dimensional classical systems. It is worth noticing explicitly, however, that the multiplicity allowed by the von Neumann theorem does not admit of the following interpretation: The algebra generated by a reducible WS is an algebra of observables with superselection rules; for (in this one-dimensional case) either the algebra generated by a WS is irreducible or it has a non-Abelian commutant.[9,10]

Our sticking to a fixed canonical pair (within a unique phase space) means that we are going to consider classical systems whose kinematical structures are the same and are represented in the same way. The Segal quantization of the kinematical structure of these systems gives a quantum kinematical picture which is unique up to unitary equivalence, as we are considering irreducible WS's only; thus, we obtain an algebra of observables which is unique up to unitary equivalence, and a unique labeling of the kinematical observables momentum and position. In fact, in quantizing all the classical motions which can appear in a one-dimensional classical system, it is good to keep the kinematical picture fixed, in order to compare the various possible motions in the most straightforward way. Obviously, this restricting ourselves to a fixed canonical pair is not an essential limitation, and the results we get are true for any one-dimensional classical system, as it could be shown, for instance, by replacing the canonical pair we have chosen by another arbitrary one, step by step in all what follows. Finally, notice that any linear basis in $\mathscr{M}$ can be made into a canonical pair, by choosing a suitable symplectic form.

3. QUANTIZATION OF CLASSICAL DYNAMICS

As we have already mentioned, a (linear) symmetry of a kinematical picture $(\mathscr{M},B)$ is represented by a symplectic transformation of $(\mathscr{M},B)$. For a quantum kinematical picture set up by means of a $\mathscr{H}$-valued WS, a quantum symmetry is represented by an automorphism of the Weyl algebra; in the particular case we are discussing here (one-dimensional case), it can be shown that a quantum symmetry can be represented by a unitary operator in $\mathscr{H}$ in the following sense: If any two observables, represented by two operators A_1 and A_2 in $\mathscr{H}$, are connected by a given symmetry, then $A_2 = UA_1U^{-1}$ where U is a unitary operator on $\mathscr{H}$, which, by definition, represents the symmetry; moreover U is unique up to a phase factor.[2] Consider a classical kinematical picture $(\mathscr{M},B)$, its quantization constructed by means of an $\mathscr{H}$-valued WS W over $\mathscr{M}$, a classical symmetry represented by a symplectic transformation S and a quantum symmetry represented by a unitary operator U on $\mathscr{H}$. We notice that $W\circ S$ is a WS by the very fact that S is symplectic; therefore the classical symmetry S transforms the labeled observable momentum and position, P_W and Q_W, into the operators $P_{W\circ S}$ and $Q_{W\circ S}$ respectively, since it turns the labeling pairs

(1,0) and (0,1) into S(1,0) and S(0,1), respectively. On the other hand, the quantum symmetry U transforms P_W and Q_W into UP_WU^{-1} and UQ_WU^{-1}, respectively. Moreover, it follows that $P_{W\circ S}$ and $Q_{W\circ S}$ coincide with UP_WU^{-1} and UQ_WU^{-1}, respectively, if and only if the following condition holds,

$$W(S(m)) = UW(m)U^{-1} \quad \forall m \in \mathscr{M}.$$

Therefore, we will call U the quantization of S when such a condition holds. For it follows from the previous discussion that, when this condition holds, U represents S in the quantum kinematical picture which quantizes $(\mathscr{M},B)$. In other words, U being the quantization of S means by definition that the pair (U,S) is compatible with the given labeling. This in turn is equivalent to the fact that U and S represent the same symmetry (U in the quantum picture and S in the classical one). In fact, under such a hypothesis, $P_{W\circ S}$ (for instance) is the momentum observable if we label the observables "after" the action of the symmetry, while UP_WU^{-1} is the momentum observable if we label "before" the symmetry; since the "time" we decide to label must not matter, $P_{W\circ S} = UP_WU^{-1}$ must hold true. If, conversely, U and S are compatible with the labeling, namely, $P_{W\circ S} = UP_WU^{-1}$ and $Q_{W\circ S} = UQ_WU^{-1}$ hold true, then labeling after the action of S amounts to changing the labeled observables with U, and this means precisely that U and S represent the same symmetry. Notice that for any classical symmetry S there is a quantum symmetry which quantizes it; this follows directly from the von Neumann uniqueness theorem. On the contrary, not every quantum symmetry is the quantization of some classical symmetry, as can be easily shown by the construction of unitary operators U for which no symplectic transformation S exists such that $W(S(m)) = UW(m)U^{-1}$ holds for all $m \in \mathscr{M}$.

The discussion of time evolution and its quantization can now be made along very similar lines, since we are dealing with classical dynamical systems whose evolution laws are linear. A (linear) classical motion is represented by a one-parameter group of symplectic transformations

$$\mathbb{R} \ni t \rightarrow S_t \in \mathrm{Aut}(\mathscr{M},B),$$

continuous with respect to the one sensible topology Aut $(\mathscr{M},B)$ can be given (in the finite-dimensional case we are considering). For a quantum system represented by an $\mathscr{H}$-valued WS, a motion is represented by a one-parameter group of automorphisms of the Weyl algebra; in our case, it can be shown[11] that a quantum motion can always be represented by a weakly continuous one-parameter group of unitary operators

$$\mathbb{R} \ni t \rightarrow U_t \in \mathscr{U}(\mathscr{H})$$

in the following sense: if an observable is represented by an operator A in $\mathscr{H}$ at time t_0, it is represented by $U_tAU_t^{-1}$ at time $t_0 + t$ (Heisenberg picture); the one-parameter group U_t is determined within a phase factor which, by the continuity and the group property, can be nothing else than $e^{i\alpha t}, \alpha \in \mathbb{R}$. We point out that, according to the Stone theorem, a quantum motion determines, and is determined by, a unique (up to an additive constant) self-adjoint operator H in $\mathscr{H}$, which is called the Hamiltonian of the quantum system and which satisfies[12]

$$U_t = \exp(itH), \quad \forall t \in \mathbb{R}.$$

According to the previous discussion of symmetries, a quantum motion $t \rightarrow U_t$ will be called the quantization of a classical motion $t \rightarrow S_t$ when the following conditions holds

$$W(S_t(m)) = U_tW(m)U_t^{-1}, \quad \forall m \in \mathscr{M}, \quad \forall t \in \mathbb{R}.$$

We are going to show that, while all the classical motions can be quantized, not every quantum motion is the quantization of a classical one. In fact, we are going to determine the Hamiltonians of the quantum dynamical systems which can be constructed by quantizing the classical ones. First, we need the following technical result:

Lemma 3.1: Let $W: \mathscr{M} \rightarrow \mathscr{U}(\mathscr{H})$ be an IWS over $\mathscr{M}$, $\mathscr{H}$ a separable Hilbert space, and γ,η,ρ any three real numbers. The operator $K_W^{(\gamma,\eta,\rho)} := \gamma P_W^2 + \eta Q_W^2 + \rho\{P_W,Q_W\}_+$ is an essentially self-adjoint operator in $\mathscr{H}$.

Proof: Assume W is the Schrödinger WS and check directly that every Hermite function is an analytic vector for $K_W^{(\gamma,\eta,\rho)}$. Use Nelson's analytic vector theorem (Theorem X.39 of Ref. 13) to get to essential self-adjointness of $K_W^{(\gamma,\eta,\rho)}$. The result is valid for any WS, as essential self-adjointness is conserved by unitary transformations.

From now on we will denote by $H_W^{(\gamma,\eta,\rho)} := \overline{K_W^{(\gamma,\eta,\rho)}}$ the self-adjoint closure of $K_W^{(\gamma,\eta,\rho)}$ (the bar denotes closure here and in the sequel). We can now prove the basic facts about the classical motions and their quantization. To properly understand the scope of the following proposition, it should be borne in mind that we are considering just linear classical motions throughout the present paper.

Proposition 3.2: (a) There is a bijection between $\mathbb{R}^3$ and the family of classical motions; the classical motion related to $(\gamma,\eta,\rho) \in \mathbb{R}^3$ in such a way is—in the matrix representation constructed by means of the "canonical" basis (m_1,m_2) of $\mathscr{M}$—the following one-parameter group of symplectic transformations

$$S^{(\gamma,\eta,\rho)}: t \rightarrow S_t^{(\gamma,\eta,\rho)}:$$

$$= \begin{vmatrix} \cos\omega t - 2\rho(\sin\omega t)/\omega & 2\gamma(\sin\omega t)/\omega \\ -2\eta(\sin\omega t)/\omega & \cos\omega t + 2\rho(\sin\omega t)/\omega \end{vmatrix},$$

where

$$\omega := \begin{cases} 2\sqrt{\eta\gamma - \rho^2}, & \text{if } \eta\gamma - \rho^2 \geqslant 0, \\ 2i\sqrt{\rho^2 - \eta\gamma}, & \text{if } \eta\gamma - \rho^2 < 0, \end{cases}$$

and $(\sin\omega t)/\omega$ means t when $\omega = 0$.

(b) Assume a quantization is given of the classical kinematical picture $(\mathscr{M},B)$ though an IWS W over $\mathscr{M}$; then the classical dynamical system consisting of the kinematical picture $(\mathscr{M},B)$ plus the motion represented by $S^{(\gamma,\eta,\rho)}$ can be quantized and its quantization is essentially unique; it is constructed by applying to the quantum kinematical picture obtained through W the motion characterized by the Hamiltonian $H^{(\gamma,\eta,\rho)}$ (up to an additive constant).

Proof: (a) Observe that the Lie algebra of SL(2,$\mathbb{R}$), which may be considered as the representation of Aut$(\mathscr{M},B)$ through the canonical basis $\{(1,0),(0,1)\}$ of $\mathscr{M}$, is the set

$sl(2,\mathbb{R})$ of all the real 2×2 matrices A such that $\mathrm{Tr}(A) = 0$; observe also that the mapping

$$\mathbb{R}^3 \ni (\gamma,\eta,\rho) \to A^{(\gamma,\eta,\rho)} = \begin{vmatrix} -2\rho & 2\gamma \\ -2\eta & 2\rho \end{vmatrix} \in sl(2,\mathbb{R})$$

is a bijection between $\mathbb{R}^3$ and $sl(2,\mathbb{R})$. To complete the proof, check that the one-parameter subgroup of $SL(2,\mathbb{R})$ generated by $A^{(\gamma,\eta,\rho)}$ is $S^{(\gamma,\eta,\rho)}$.

(b) Setting $U_t^{(\gamma,\eta,\rho)} = \exp(itH^{(\gamma,\eta,\rho)})$ for each $(\gamma,\eta,\rho)\in\mathbb{R}^3$, we can easily show that

$$W(S_t^{(\gamma,\eta,\rho)}(m)) = U_t^{(\gamma,\eta,\rho)} W(m) U_{-t}^{(\gamma,\eta,\rho)}, \quad \forall m \in \mathscr{M}, \quad \forall t\in\mathbb{R}. \qquad (*)$$

For, $W(\alpha,\beta) = \exp(-(i/2)\alpha\beta)\exp(-i\alpha P_W)\exp(-i\beta Q_W)$, which in turn coincides with the exponential of the closure of $-i\alpha P_W - i\beta Q_W$; therefore the equalities (*) hold if and only if the following two equalities hold:

$$[\cos\omega t - 2\rho(\sin\omega t)/\omega] P_W - 2\eta[(\sin\omega t)/\omega] Q_W = U_t^{(\gamma,\eta,\rho)} P_W U_{-t}^{(\gamma,\eta,\rho)},$$

$$2\gamma[(\sin\omega t)/\omega] P_W + [\cos\omega t + 2\rho(\sin\omega t)/\omega] Q_W = U_t^{(\gamma,\eta,\rho)} Q_W U_{-t}^{(\gamma,\eta,\rho)},$$

for each $t\in\mathbb{R}$. There two equalities are in turn equivalent to their infinitesimal forms

$$-2\rho P_W - 2\eta Q_W = i[H^{(\gamma,\eta,\rho)}, P_W],$$

$$2\gamma P_W + 2\rho Q_W = i[H^{(\gamma,\eta,\rho)}, Q_W],$$

which indeed hold true, as can be trivially checked. This shows that every classical dynamical system can be quantized, and that it is quantized to the quantum dynamical system whose Hamiltonian is $H^{(\gamma,\eta,\rho)}$. The uniqueness of the quantization within a phase factor—or equivalently of the Hamiltonian within an additive constant—follows at once from the irreducibility of the Weyl system. If W' is another WS over $\mathscr{M}$, then W and W' are unitarily equivalent (by the von Neumann uniqueness theorem) and it is trivial to show that this equivalence links also P_W with $P_{W'}$, and Q_W with $Q_{W'}$, whence $H_W^{(\gamma,\eta,\rho)}$ and $H_{W'}^{(\gamma,\eta,\rho)}$ are unitarily equivalent.

We observe that, during the proof of the previous proposition, we also showed that a quantum motion is the quantization of a classical one if and only if it is linear in the sense that the following equalities are fulfilled for each $t\in\mathbb{R}$

$$U_t P_W U_{-t} = s_{11}(t) P_W + s_{21}(t) Q_W,$$

$$U_t Q_W U_{-t} = s_{12}(t) P_W + s_{22}(t) Q_W,$$

where $s_{ij}(t)$ are four real functions of the time. In fact, we showed explicitly that this is the case for $U_t = \exp(itH^{(\gamma,\eta,\rho)})$. If conversely this is the case, then the matrix $\{s_{ij}(t)\}$ has to be an element of $SL(2,\mathbb{R})$ because of the commutation relation between P_W and Q_W; moreover, it has to be a continuous function of t because of the continuity of U_t; therefore $\{s_{ij}(t)\}$ has to be $S_t^{(\gamma,\eta,\rho)}$ for a triple (γ,η,ρ), whence $U_t = \exp(itH_W^{(\gamma,\eta,\rho)})$ up to a phase factor since the only operators which commute with Q_W and P_W are the multiples of the identity, by the irreducibility of W.

It is very easy to show that also for a quantum symmetry represented by U to be the quantization of a classical one, it is necessary and sufficient that the linearity condition

$$U\begin{pmatrix} P_W \\ Q_W \end{pmatrix} U^{-1} = \widetilde{S}\begin{pmatrix} P_W \\ Q_W \end{pmatrix}$$

holds, where $\widetilde{S}$ is a 2×2 matrix and an obvious use of the matrix calculus symbolism has been made (incidentally, $\widetilde{S}$ is the transpose of the matrix which represents the classical symmetry). However, this does not imply that a unitary operator which quantizes a classical symmetry has to be the exponential of $H^{(\gamma,\eta,\rho)}$ for some (γ,η,ρ). On the contrary, the previous discussion shows that this is the case if and only if the classical symmetry lies on a classical motion, i.e., on a one-parameter subgroup of $SL(2,\mathbb{R})$ (and this need not be true, because the one-parameter subgroups of $SL(2,\mathbb{R})$ do not fill the whole group).

Proposition 3.2 shows that there are lots of quantum motions which do not have any (linear) classical counterpart, since it shows that the Hamiltonians of the quantum motions which do have a (linear) classical counterpart are quite particular.

4. UNITARY EQUIVALENCE

In Sec. 3 we accounted for the limitations we decided to adopt by choosing a fixed canonical pairs and considering only irreducible WS's. We just recall here that these limitations have the following two precise meanings: First, we are dealing with (one-dimensional linear) classical systems which coincide as far as their kinematical features are concerned (in other words, we are building different motions on equal kinematical pictures to get different classical dynamical systems); second, we quantize the classical kinematical pictures in a minimal way, as we get a quantum picture in which no other degrees of freedom can be represented that those already included in $\mathscr{M}$. Within this frame, we get quantum pictures which coincide (up to unitary equivalence) in both the algebra and the labeling of the observables. In fact, we may surely have two different IWS's W_1 and W_2 over $\mathscr{M}$, but they must be unitarily equivalent by the von Neumann uniqueness theorem; this means that a unitary operator V must exist from the Hilbert space of W_1 onto the Hilbert space of W_2 such that

$$VW_1(m)V^{-1} = W_2(m), \quad \forall m \in \mathscr{M},$$

namely, such as to set up an isomorphism between the algebras generated by W_1 and W_2, which maintains the labeling of momentum and position, and therefore of all other observables, too. Thus the quantum picture we get is unique indeed, as far as kinematics is concerned.

Things are different when dynamics is considered, to the effect that there are many inequivalent quantum dynamical systems. Before discussing this point, however, we have to define equivalence for such systems. Let us denote by (W,H) the system which includes both the quantum kinematical picture obtained from $(\mathscr{M},B)$ through an IWS W and the quantum motion whose Hamiltonian is a self-adjoint operator H. We will call such a pair (W,H) a *a quantum dynamical system*. We will say that two quantum dynamical systems (W_1,H_1) and (W_2,H_2) are equivalent when a unitary operator V exists from the Hilbert space of W_1 onto the Hilbert space of W_2 such that

$$VW_1(m)V^{-1} = W_2(m), \quad \forall m \in \mathscr{M},$$

and

$$VH_1V^{-1} = H_2 + k\mathbb{1}, \quad \text{for some real number } k.$$

Since two IWS's are unitarily equivalent through a unitary operator which is unique up to a factor (because of irreducibility), we can as well say that (W_1,H_1) and (W_2,H_2) are equivalent if H_1 and H_2 are unitarily equivalent up to an additive constant through the unitary operator which sets up the equivalence of W_1 and W_2. This condition is appropriate for the definition of the equivalence of two quantum dynamical systems, since its meaning is that the same unitary operator V through which the kinematical descriptions determined by W_1 and W_2 are brought to coincide, makes the time evolution laws determined by H_1 and H_2 coincide as well. In fact, setting $U_t^1 := \exp(itH_1)$ and $U_t^2 := \exp(itH_2)$, it is obviously true that U_t^1 and U_t^2 define two one-parameter groups of unitary operators which coincide up to a phase factor through V, whenever $VH_1V^{-1} = H_2 + k\mathbb{1}$. Conversely, the necessary condition for the equality up to V of the two quantum motions determined by U_t^1 and U_t^2 is —owing to the irreducibility of W_1—that a real function φ on $\mathbb{R}$ exists such that $VU_t^1V^{-1} = \exp(i\varphi(t))U_t^2$ for each $t \in \mathbb{R}$; moreover, a real number k must exist such that $\varphi(t) = kt$, because of the group laws and continuity of both U_t^1 and U_t^2; in this way we get that $VH_1V^{-1} = H_2 + k\mathbb{1}$ must hold.

It is now clear that, while at the kinematical level all the quantum pictures we are considering here are equivalent by the very fact of being quantizations of the same classical kinematical picture, at the dynamical level two systems (W_1,H_1) and (W_2,H_2) need by no means be equivalent, as it is very trivial to see considering two self-adjoint operators H_1 and H_2 which are not unitarily equivalent through the unique up to a factor unitary operator which sets up the equivalence of two IWS's W_1 and W_2. Indeed, it is worth while pointing out that for two quantum dynamical systems (W_1,H_1) and (W_2,H_2) we may have three different unitary equivalence relations: the equivalence of the underlying kinematical pictures, namely of W_1 and W_2, which in fact we always have; the equivalence of the whole dynamical systems themselves, which does not always hold true and whose validity means that H_1 and H_2 are the same functions—up to an additive constant—of the respective kinematical observables momentum and position; the equivalence of the Hamiltonians H_1 and H_2, which is always valid when the dynamical systems are equivalent and which may also be valid when the equivalence of the dynamical systems is not. Of course, if the Hamiltonians are equivalent while the systems are not, the Hamiltonians are two different labeled observables. We shall see later (in Proposition 4.2) examples of quantum dynamical systems which are not equivalent, while their Hamiltonians are.

We have seen in Proposition 3.2 that each classical dynamical system has its own quantization. It is natural to expect that two different classical dynamical systems give rise to two different, that is nonequivalent, quantum dynamical systems. This is indeed the case, as is shown by the following proposition. We will still denote by $S^{(\gamma,\eta,\rho)}$ the classical motion related to $(\gamma,\eta,\rho) \in \mathbb{R}^3$ in the way exhibited in Proposition 3.2.

Proposition 4.1: Let W and W' be IWS's over $\mathscr{M}$ and let (γ,η,ρ) and (γ',η',ρ') be triples of real numbers. Let us consider, along with W and W', the Hamiltonians $H_W^{(\gamma,\eta,\rho)}$ and $H_{W'}^{(\gamma',\eta',\rho')}$ which define the quantizations—upon the quantum kinematical pictures provided by W and W', respectively—of the classical motions $S^{(\gamma,\eta,\rho)}$ and $S^{(\gamma',\eta',\rho')}$. The quantum dynamical systems $(W,H_W^{(\gamma,\eta,\rho)})$ and $(W',H_{W'}^{(\gamma',\eta',\rho')})$ are equivalent if an only if they quantize the same classical dynamical systems, namely, only if $(\gamma,\eta,\rho) = (\gamma',\eta',\rho')$.

Proof: let V be the (unique up to a factor) unitary operator through which W' and W are equivalent, namely, let the operator V be unitary and such that

$$VW'(m)V^{-1} = W(m), \quad \forall m \in \mathscr{M}.$$

From this relation, the equations

$$VP_{W'}V^{-1} = P_W \quad \text{and} \quad VQ_{W'}V^{-1} = Q_W$$

follow. Therefore, if $(W,H_W^{(\gamma,\eta,\rho)})$ and $(W',H_{W'}^{(\gamma',\eta',\rho')})$ are equivalent, we get

$$\gamma' P_W^2 + \eta' Q_W^2 + \rho' \{P_W,Q_W\}_+ = \gamma P_W^2 + \eta Q_W^2 + \rho \{P_W,Q_W\}_+ + k\mathbb{1}$$

for some real number k, since unitary equivalence and closure "commute" and the two operators which appear in the equation above do have the same domain. Recalling now the way a classical motion is quantized, we get from the last equation that—for each $m \in \mathscr{M}$ and each $t \in \mathbb{R}$—

$$\begin{aligned} W(S_t^{(\gamma,\eta,\rho)}(m)) &= \exp(itH_W^{(\gamma,\eta,\rho)})W(m)\exp(-itH_W^{(\gamma,\eta,\rho)}) \\ &= \exp(itH_{W'}^{(\gamma',\eta',\rho')})W(m)\exp(-itH_{W'}^{(\gamma',\eta',\rho')}) \\ &= W(S_t^{(\gamma',\eta',\rho')}(m)). \end{aligned}$$

By the injectivity of the WS W, this can hold only if $S^{(\gamma,\eta,\rho)} = S^{(\gamma',\eta',\rho')}$. This implies $(\gamma,\eta,\rho) = (\gamma',\eta',\rho')$ by Proposition 3.1.

Notice that the previous proposition in a sense completes part (b) of Proposition 3.2, inasmuch as it amounts to the statement that a quantum dynamical system quantizes a unique, if any, classical dynamical system; in part (b) of Proposition 3.2 we had what is in a way the converse statement, since we proved there that a classical dynamical system determines uniquely its quantization (this could be the "if part" of Proposition 4.1).

Moreover, notice that the previous proposition deals with the equivalence of the quantum dynamical systems which quantize classical ones, but it does not say anything about the equivalence of either the underlying quantum kinematical systems or the Hamiltonians. Indeed, while the former always holds, the latter does not as a general rule, but does hold for particular pairs of Hamiltonians, as can be clearly understood looking at the following proposition.

Proposition 4.2: Let W be an IWS over $\mathscr{M}$ and (γ,η,ρ) an element of $\mathbb{R}^3$. The Hamiltonian $H_W^{(\gamma,\eta,\rho)}$ is unitarily equivalent (up to an additive constant) to the closure of one of the following essentially self-adjoint operators:

$$\pm\sqrt{\eta\gamma-\rho^2}\,(P_W^2+Q_W^2),\quad \text{if}\quad \eta\gamma-\rho^2>0, \tag{1}$$

$$\gamma P_W^2,\quad \text{or}\quad \eta P_W^2,\quad \text{or}\quad \pm 2\rho P_W^2,\quad \text{if}\quad \eta\gamma-\rho^2=0, \tag{2}$$

$$\pm\sqrt{\rho^2-\eta\gamma}\,(P_W^2-Q_W^2),\quad \text{if}\quad \eta\gamma-\rho^2<0. \tag{3}$$

Proof: It is easy to show that a matrix Λ of SL(2,$\mathbb{R}$) exists such that, setting

$$\begin{pmatrix}\pi\\ \varphi\end{pmatrix}=\Lambda\begin{pmatrix}P_W\\ Q_W\end{pmatrix}$$

with an obvious use of the matrix calculus symbolism, the operator $\gamma P_W^2+\eta Q_W^2+\rho(P_WQ_W+Q_WP_W)$ is equal to one of the following operators:

$$k_{(\gamma,\eta,\rho)}(\pi^2+\varphi^2),\quad \text{if}\quad \eta\gamma-\rho^2>0, \tag{4}$$

$$k_{(\gamma,\eta,\rho)}\pi^2,\quad \text{if}\quad \eta\gamma-\rho^2=0, \tag{5}$$

$$k_{(\gamma,\eta,\rho)}(\pi^2-\varphi^2),\quad \text{if}\quad \eta\gamma-\rho^2<0, \tag{6}$$

where $k_{(\gamma,\eta,\rho)}$ is a real number depending on (γ,η,ρ). For, from the solution of an elementary geometry problem, we find that a matrix Λ' of SO(2) exists such that, setting

$$\begin{pmatrix}\pi'\\ \varphi'\end{pmatrix}=\Lambda'\begin{pmatrix}P_W\\ Q_W\end{pmatrix},$$

we have

$$\gamma P_W^2+\eta Q_W^2+\rho(P_WQ_W+Q_WP_W)$$
$$=k^{(1)}_{(\gamma,\eta,\rho)}\pi'^2+k^{(2)}_{(\gamma,\eta,\rho)}\varphi'^2,$$

where $k^{(1)}_{(\gamma,\eta,\rho)}$ and $k^{(2)}_{(\gamma,\eta,\rho)}$ are two real numbers depending on (γ,η,ρ). It is now trivial to construct a diagonal matrix Λ'' of SL(2,$\mathbb{R}$) such that, setting

$$\begin{pmatrix}\pi\\ \varphi\end{pmatrix}=\Lambda''\begin{pmatrix}\pi'\\ \varphi'\end{pmatrix},$$

$k^{(1)}_{(\gamma,\eta,\rho)}\pi'^2+k^{(2)}_{(\gamma,\eta,\rho)}\varphi'^2$ assumes one of the forms (4), (5), (6) listed above. Now $\Lambda:=\Lambda''\Lambda'$ is the matrix whose existence we had to prove.

Should the computations be carried out to find the matrices Λ' and Λ'', we would find that $k_{(\gamma,\eta,\rho)}$ results into one of the following values:

$$\pm\sqrt{\eta\gamma-\rho^2},\quad \text{if}\quad \eta\gamma-\rho^2>0, \tag{7}$$

$$\gamma,\quad \text{or}\quad \eta,\quad \text{or}\quad \pm 2\rho,\quad \text{if}\quad \eta\gamma-\rho^2=0, \tag{8}$$

$$\pm\sqrt{\rho^2-\eta\gamma},\quad \text{if}\quad \eta\gamma-\rho^2<0. \tag{9}$$

Therefore, $\gamma P_W^2+\eta Q_W^2+\rho(P_WQ_W+Q_WP_W)$ equals one of the following operators:

$$\pm\sqrt{\eta\gamma-\rho^2}\,(\pi^2+\varphi^2),\quad \text{if}\quad \eta\gamma-\rho^2>0, \tag{10}$$

$$\gamma\pi^2,\quad \text{or}\quad \eta\pi^2,\quad \text{or}\pm 2\rho\pi^2,\quad \text{if}\quad \eta\gamma-\rho^2=0, \tag{11}$$

$$\pm\sqrt{\rho^2-\eta\gamma}\,(\pi^2-\varphi^2),\quad \text{if}\quad \eta\gamma-\rho^2<0. \tag{12}$$

The next step is to show that the pair $(\bar\pi,\bar\varphi)$ of the closures of the operators π and φ is unitarily equivalent to the pair (P_W,Q_W). In order to do this, we can obviously assume without any lack of generality that W is the Schrödinger WS. We find that the following conditions are satisfied. First, the Schwarz space $\mathscr{S}$—considered as a linear subspace of $L^2(\mathbb{R})$—is contained in the domains of both $\bar\pi$ and $\bar\varphi$; moreover it is invariant with respect to both $\bar\pi$ and $\bar\varphi$. Second the restriction of the operator $\bar\pi^2+\bar\varphi^2$ to $\mathscr{S}$—which coincides with the restriction of $\pi^2+\varphi^2$ to $\mathscr{S}$—is essentially self-adjoint by Nelson's analytic vector theorem, since every Hermite function is an analytic vector for this operator, as can be easily shown. Finally, $\bar\pi\bar\varphi-\bar\varphi\bar\pi$ equals $-i\mathbb{1}$ on $\mathscr{S}$. From this condition and a result of Dixmier[14] it follows that the pair $(\bar\pi,\bar\varphi)$ is unitarily equivalent to a direct sum of copies of the pair (P_S,Q_S) or rather, going back to the general IWS W we are considering in this proposition, to a direct sum of copies of the pair (P_W,Q_W). However, no nontrivial subspace can exist which is invariant with respect to both $\bar\pi$ and $\bar\varphi$, since P_W and Q_W are the closures of the linear combinations of π and φ we get through Λ^{-1}, and W is irreducible. Therefore the pair $(\bar\pi,\bar\varphi)$ has to be unitarily equivalent to the pair (P_W,Q_W).

To conclude the proof, we notice that the closure of $\bar\pi^2+\epsilon\bar\varphi^2$ (with $\epsilon=\pm 1$ or 0) coincides with the closure of $\pi^2+\epsilon\varphi^2$, as can be easily shown using the essential self-adjointness of the restriction of $\pi^2+\epsilon\varphi^2$ to $\mathscr{S}$, which in turn can be checked with the usual argument based on the analytic vector theorem. Therefore $H_W^{(\gamma,\eta,\rho)}$, being the closure of one of the operators listed above in (10), (11) or (12), is in fact the closure of an operator which is unitarily equivalent to the corresponding one written in (1), (2) or (3) in the statement of the proposition. To prove the proposition it is now enough to use the fact that the closure passes through unitary equivalence of operators.

From this proposition it is clear that the Hamiltonians which arise in the quantization of the classical dynamical systems can be grouped not only into unitary equivalence classes, but also according to the kind of the spectrum they have. From Proposition 4.2 and the knowledge of the spectra of $P_S^2, P_S^2+Q_S^2, P_S^2-Q_S^2$ we deduce in fact that the spectrum of $H_W^{(\gamma,\eta,\rho)}$ can be of three possible types according to which situation occurs, among (1), (2), and (3) of Proposition 4.2: in case (1) the spectrum of $H_W^{(\gamma,\eta,\rho)}$ coincides with its discrete spectrum and it is—up to a multiplicative constant and an additive one—the set of the nonnegative integers; in case (2) it coincides with the continuous spectrum and it is either the set of the nonnegative real numbers or the set of the nonpositive ones; in case (3) it coincides with the continuous spectrum and it is the whole real line. This shows, by the way, that the statement of Proposition 4.2 can be strengthened: Hamiltonians belonging to two different classes among the three possible cannot be unitarily equivalent, as they have different spectra. We notice also that in connection with the three possible cases listed above for the type of the spectrum of $H_W^{(\gamma,\eta,\rho)}$, which are nothing else that the three cases listed for $H_W^{(\gamma,\eta,\rho)}$ itself in Proposition 4.2, the classical motion $S^{(\gamma,\eta,\rho)}$ is of one of the three following types: it corresponds either to a periodic motion, to a free particle motion, or to a "hyperbolic" motion, as can be easily seen looking at part (a) of Proposition 3.2. In this sense we can group the classical motions and their quantizations into three groups.

Anyway, it would be very wrong to conclude, on the basis of the considerations above, that by quantizing the classical dynamical systems we get just three types of quantum dynamical systems. In case (1), for instance, we get "harmonic oscillators" only inasmuch as we get Hamiltonians which are up to a factor unitarily equivalent to $P_S^2 + Q_S^2$, but which are not that function of momentum and position which is right to call the Hamiltonian of the harmonic oscillator. In fact, the unitary equivalence which relates one of these Hamiltonians to the harmonic oscillator Hamiltonian does not preserve the labeling so that this mathematical equivalence is not enough to represent a physical one.

It is worth pointing out that neither the statement of Proposition 4.2 nor the results about the spectra established in the discussion which followed the proposition can be directly generalized to a more than one-dimensional finite case. This can be seen quite easily, for instance, by working out the two-dimensional case in a parallel way. Anyway, the analysis above is indicative of what the general situation is and the ultimate meaning of the previous discussion would be the basic meaning of the analogous discussion for any finite-dimensional case. In particular, the last paragraph, which summarizes the sense of the previous analysis, can be straightforwardly generalized to all finite-dimensional (linear) cases.

5. SECOND QUANTIZATION

In this last section we will focus our attention on second quantization, in connection with the one-dimensional linear systems we are studying. After a quick review of what the second quantization procedure means for the systems we are dealing with here, we will prove that not all of them can be dealt with by such a procedure. This shows that second quantization can not be used as a universal quantization procedure for dynamical systems, even in the elementary one-dimensional linear case.

First, let us see how second quantization is a way to get a quantum kinematical picture from the classical one provided by $(\mathscr{M},B)$. To be able to use the second quantization procedure we must enrich the structure of $\mathscr{M}$, defining on it a complex Hilbert space structure which embodies the real vector space structure $\mathscr{M}$ is endowed with; this can be done, as will be specified later. As soon as $\mathscr{M}$ is a complex Hilbert space whose inner product will be denoted by $\langle\ |\ \rangle$, the standard machinery of second quantization applied to $\mathscr{M}$ allows us to get the symmetric Fock space $\mathscr{F}(\mathscr{M})$ over $\mathscr{M}$ and—for each $m\in\mathscr{M}$—the annihilation and creation operators $a(m)$ and $a^\dagger(m)$ in $\mathscr{F}(\mathscr{M})$. It is now possible to obtain the following results (Sec. X.7 of Ref. 13, to which we refer for all the results concerning second quantization we will use): the operator $[1/(2)^{1/2}](a(m)+a^\dagger(m))$ is essentially self-adjoint for each $m\in\mathscr{M}$; if $W^{\rm II}(m)$ denotes the unitary operator $\exp\{[1/(2)^{1/2}]\overline{(a(m)+a^\dagger(m))}\}$, where $m\in\mathscr{M}$ and the bar means closure as usual, the relation

$$W^{\rm II}(m_1)W^{\rm II}(m_2) = \exp[-(i/2)\mathrm{Im}\langle m_1|m_2\rangle] \times W^{\rm II}(m_1+m_2), \quad \forall m_1,m_2\in\mathscr{M}$$

holds true; the mapping $\mathbb{R}\ni t\to W^{\rm II}(tm)\in\mathscr{U}(\mathscr{F}(M))$ is weakly continuous; the only operators of $\mathscr{F}(\mathscr{M})$ which commute with $W^{\rm II}(m)$ for each $m\in\mathscr{M}$ are the multiples of the identity. Therefore, it is clear that we get an $\mathscr{F}(\mathscr{M})$-valued IWS over $\mathscr{M}$ whenever the complex Hilbert space structure on $\mathscr{M}$ can be defined in such a way that

$$B(m_1,m_2) = -\mathrm{Im}\langle m_1|m_2\rangle, \quad \forall m_1,m_2\in\mathscr{M};$$

this to σ can be done, as will be specified below. Thus, second quantization is a straightforward way to construct IWS's over $\mathscr{M}$, namely quantizations of the classical kinematical picture we are considering in this paper. Notice that, as far as the kinematical picture alone is concerned, these quantizations can be thought of as a single one because of the von Neumann uniqueness theorem.

Let us now see how second quantization can be a way to quantize a classical dynamical system. Let us suppose that $t\to S_t$ is a classical motion, i.e., a continuous one-parameter group of symplectic transformations of $(\mathscr{M},B)$; let us further suppose that S satisfies the following condition (SQ): a Hilbert space structure can be given to $(\mathscr{M},B)$, such as to meet the mentioned requirements for the second quantization procedure (in particular B is the imaginary part of the inner product) and such as to make S_t, $\forall t\in\mathbb{R}$, into a unitary operator. If that is the case, by the second quantization procedure we can quantize not only the kinematical picture, but the motion $t\to S_t$ as well. Acutally—if $\Gamma(V)$ denotes the unitary operator on $\mathscr{F}(\mathscr{M})$ which is called the second quantization of a unitary operator V on $\mathscr{M}$—we have, under the hypotheses made above, the following equation: $\Gamma(S_t)W^{\rm II}(m)\Gamma(S_t)^{-1} = W^{\rm II}(S_tm), \quad \forall m\in\mathscr{M}, \quad \forall t\in\mathbb{R}.$ This equation means that the quantum motion $t\to\Gamma(S_t)$ quantizes the classical motion $t\to S_t$, from which it is obtained by second quantization. We notice that if $d\Gamma(A)$ denotes the self-adjoint operator in $\mathscr{F}(\mathscr{M})$ which is called the second quantization of a self-adjoint operator A in $\mathscr{M}$, the last equation can be written also in the following way:

$$\exp(itd\Gamma(h))W^{\rm II}(m)\exp(-itd\Gamma(h)) = W^{\rm II}(S_tm), \quad \forall m\in\mathscr{M}, \quad \forall t\in\mathbb{R},$$

if the self-adjoint operator h in $\mathscr{M}$ is the Stone theorem generator of $t\to S_t$; for, $\Gamma(\exp(itA)) = \exp(itd\Gamma(A))$ holds true for every self-adjoint operator A in $\mathscr{M}$. Therefore, provided $t\to S_t$ fulfills the condition (SQ) a quantization of the classical dynamical system described by $(\mathscr{M},B)$ along with $t\to S_t$ is the quantum dynamical system $(W^{\rm II}, d\Gamma(h))$ in which both the WS $W^{\rm II}$ and the Hamiltonian $d\Gamma(h)$ are constructed by the second quantization procedure; because of the way it is constructed, $(W^{\rm II},d\Gamma(h))$ will be called a second quantized dynamical system. Accordingly, a classical dynamical system will be called second quantizable when its motion meets the condition (SQ). We recall that each classical dynamical system is quantizable and its quantization is essentially unique. The peculiarity of second quantizable systems lies in the way the quantization can be constructed.

Let us now turn our attention to the problem of determining what classical systems are second quantizable. According to the previous results, the first step in this direction is to find all the complex Hilbert space structures which can be defined on $\mathscr{M}$ in a way which is compatible with the real vector space structure of $\mathscr{M}$ and which makes the following relation true:

$$B(m_1,m_2) = -\operatorname{Im}\langle m_1|m_2\rangle, \quad \forall m_1,m_2\in\mathscr{M}.$$

It turns out that this problem amounts to determining all the linear operators J on $\mathscr{M}$ that satisfy the three following conditions: $J^2 = -\mathbb{1}$, $B(Jm_1,Jm_2) = B(m_1,m_2)$, for each $m_1,m_2\in\mathscr{M}$; for $m\in\mathscr{M}$, $B(Jm,m)>0$ if and only if $m\neq 0$. In fact, each operator J with these properties can be used to define a complex Hilbert space structure on $\mathscr{M}$ with the required properties, setting

$$zm := (\operatorname{Re}z\mathbb{1} + \operatorname{Im}zJ)m, \quad \forall z\in\mathbb{C}, \quad \forall m\in\mathscr{M},$$

and

$$\langle m_1|m_2\rangle := B(Jm_1,m_2) - iB(m_1,m_2), \quad \forall m_1,m_2\in\mathscr{M},$$

and conversely, each complex Hilbert space structure on $\mathscr{M}$ with the required properties can be constructed by means of such an operator J in the way written above; also, it is trivial to see that J is unique. A straightforward computation shows that an operator J satisfies the three above-listed conditions if and only if it is represented—with respect to the "canonical" basis (m_1,m_2) of $\mathscr{M}$—by a matrix of the type

$$\begin{pmatrix} -r & g \\ -e & r \end{pmatrix}$$

with $g,e,r\in\mathbb{R}$, $g>0, eg-r^2=1$. It is clear that there is a bijection between the set of such triples of real numbers and the set of the operators J. The symbol $J_{(g,e,r)}$ will denote the operator which corresponds to the triple (g,e,r), and $\mathscr{M}_{(g,e,r)}$ the complex Hilbert space whose structure is defined by $J_{(g,e,r)}$.

The next step is to determine which classical motions $S^{(\gamma,\eta,\rho)}$ become one-parameter groups of unitary operators in some $\mathscr{M}_{(g,e,r)}$. After noticing that a symplectic transformation S is unitary in $\mathscr{M}_{(g,e,r)}$ if and only if $SJ_{(g,e,r)} - J_{(g,e,r)}S = 0$ holds true, it is just a matter of calculations to prove that the motion $S^{(\gamma,\eta,\rho)}$ meets the condition (SQ) if and only if either $\gamma=\eta=\rho=0$ or $\eta\gamma-\rho^2>0$; moreover, if $\eta\gamma-\rho^2>0$, the only space $\mathscr{M}_{(g,e,r)}$ in which $S^{(\gamma,\eta,\rho)}$ is a unitary one-parameter group is the one defined by the triple $g=(\operatorname{sgn}\gamma)(2\gamma/\omega)$, $e=(\operatorname{sgn}\gamma)(2\eta/\omega)$, $r=(\operatorname{sgn}\gamma)(2\rho/\omega)$, where $\omega = 2(\eta\gamma-\rho^2)^{1/2}$ as in Proposition 3.1 a; obviously, in the trivial case $\gamma=\eta=\rho=0$ the motion is unitary in *all* the spaces $\mathscr{M}_{(g,e,r)}$ as $S_t^{(0,0,0)} = \mathbb{1}$ for each $t\in\mathbb{R}$.

Finally, it is worth noticing that the Hamiltonian of a second quantized dynamical system assumes a very convenient form. As a result of the analysis carried out above and of Proposition 3.1 b, it is clear that the Hamiltonians of the second quantized dynamical systems which can be constructed in the Fock space $\mathscr{F}(\mathscr{M}_{(g,e,r)})$ are real multiples of the closure of the operator $\frac{1}{2}(gP_{\mathrm{II}}^2 + eQ_{\mathrm{II}}^2 + r\{P_{\mathrm{II}},Q_{\mathrm{II}}\}_+)$, where P_{II} and Q_{II} are the momentum and position operators defined by the IWS constructed by second quantization in $\mathscr{F}(\mathscr{M}_{(g,e,r)})$. Moreover, up to a phase factor

$$\exp[(it/2)\overline{(gP_{\mathrm{II}}^2 + eQ_{\mathrm{II}}^2 + r\{P_{\mathrm{II}},Q_{\mathrm{II}}\}_+)}] = \Gamma(S_t^{(g/2,e/2,r/2)}), \quad \forall t\in\mathbb{R},$$

by the very definition of Hamiltonian (the bar denotes closure as usual). We notice now that from Proposition 3.1 a it follows that, as operators on $\mathscr{M}_{(g,e,r)}$,

$$S_t^{(g/2,e/2,r/2)} = \cos t\mathbb{1} + \sin tJ_{(g,e,r)} = \exp(it)\mathbb{1} = \exp(it\mathbb{1}), \quad \forall t\in\mathbb{R}.$$

Therefore, the equation

$$\overline{\tfrac{1}{2}(gP_{\mathrm{II}}^2 + eQ_{\mathrm{II}}^2 + r\{P_{\mathrm{II}},Q_{\mathrm{II}}\}_+)} = d\Gamma(\mathbb{1})$$

is true up to an additive constant, since $\Gamma(\exp(itA)) = \exp[itd\Gamma(A)]$ holds true for every self-adjoint operator A.

Summing up, we have proved the following proposition:

Proposition 5.1: (a) A classical dynamical system defined by $(\mathscr{M},B)$ along with a motion $S^{(\gamma,\eta,\rho)}$ is second quantizable if and only if either it is trivial, namely $\gamma=\eta=\rho=0$, or "periodic", namely $\gamma\eta-\rho^2>0$.

(b) Each triple $(g,e,r)\in\mathbb{R}^3$ with $g>0$ and $eg-r^2=1$, determines a second quantization of the one-dimensional linear kinematical picture through the $\mathscr{F}(\mathscr{M}_{(g,e,r)})$-valued IWS $W_{(g,e,r)}^{\mathrm{II}}$ defined by

$$W_{(g,e,r)}^{\mathrm{II}}(m) := \exp[(i/\sqrt{2})\overline{(a_{(g,e,r)}(m) + a_{(g,e,r)}^{\dagger}(m))}], \quad \forall m\in\mathscr{M},$$

where $a_{(g,e,r)}(m)$ and $a_{(g,e,r)}^{\dagger}(m)$ are the annihilation and creation operators in $\mathscr{F}(\mathscr{M}_{(g,e,r)})$. Each second quantizable system is second quantized in the framework of a unique such quantum kinematical picture. Within the quantum kinematical picture determined by $W_{(g,e,r)}^{\mathrm{II}}$, a second quantizable classical dynamical system in second quantized if and only if its motion $S^{(\gamma,\eta,\rho)}$ is determined by a triple (γ,η,ρ) which is a multiple of (g,e,r). If that is the case, the Hamiltonian of the second quantized system is $2\gamma/g$ times the "number operator" $d\Gamma(\mathbb{1})$ in $\mathscr{F}(\mathscr{M}_{(g,e,r)})$, where $\mathbb{1}$ denotes the identity operator in $\mathscr{M}_{(g,e,r)}$.

We notice that the real multiples of $d\Gamma(\mathbb{1})$ in $\mathscr{F}(\mathscr{M}_{(g,e,r)})$ are the second quantizations of the self-adjoint operators in $\mathscr{M}_{(g,e,r)}$; these are in fact, as is easy to see, the real multiples of the identity operator. Therefore the Hamiltonians of the second quantized systems appear as second quantizations of "one particle observables."

Proposition 5.1. shows that not all the classical dynamical systems are quantizable by the second quantization procedure. It is anyway quite clear that each one-dimensional linear quantum dynamical system (W,H) can be described in the framework of the kinematical picture which is obtained by second quantization starting from $\mathscr{M}_{(g,e,r)}$ for any (g,e,r). By the von Neumann uniqueness theorem, in fact, W can be identified with $W_{(g,e,r)}^{\mathrm{II}}$ through a unitary operator, and it is of course possible to identify H with a self-adjoint operator in $\mathscr{F}(\mathscr{M}_{(g,e,r)})$ through the same unitary operator. However, this self-adjoint operator can not be obtained as the second quantization of an observable of the "one particle quantum space" $\mathscr{M}_{(g,e,r)}$ unless H is the Hamiltonian of a quantization of a classical motion $S^{(\gamma,\eta,\rho)}$ for which (γ,η,ρ) is a multiple of (g,e,r); moreover, (γ,η,ρ) is a multiple of a triple (g,e,r) if and only if $\gamma\eta-\rho^2>0$. In Sec. 3 we saw that not all the quantum dynamical systems are quantizations of classical ones. Here we have seen that, besides, not all the quantizations of classical dynamical systems can be constructed by the second quantization procedure. Therefore, the second quantization procedure is not an exhaustive way to construct quantum dynamical systems, even in the elementary one-dimensional case we have been studying in this paper.

Finally, from second quantization we obtain a way to get—for the systems characterized by a triple (γ,η,ρ) such that $\gamma\eta-\rho^2>0$—a deeper insight into the results about equivalence we found in Sec. 4. In fact, we will be able to carry the quantum dynamical systems $(W^{\rm II}_{(g,e,r)}, H^{(\gamma,\eta,\rho)}_{W^{\rm II}_{(g,e,r)}})$ into the same Hilbert space in such a way that their comparison becomes quite straightforward. In order to do this, we consider the symmetric Fock space $\mathscr{F}(\mathbb{C})$ over $\mathbb{C}$, which is nothing else than the Hilbert space l_2 and, for each triple (g,e,r), a unitary operator (which exists and is uniquely determined up to a phase factor, since both $\mathscr{M}_{(g,e,r)}$ and $\mathbb{C}$ are one-dimensional complex Hilbert spaces) $V_{(g,e,r)}$ from $\mathscr{M}_{(g,e,r)}$ onto $\mathbb{C}$. The second quantization $\Gamma(V_{(g,e,r)})$ of $V_{(g,e,r)}$ is a unitary operator from $\mathscr{F}(\mathscr{M}_{(g,e,r)})$ onto $\mathscr{F}(\mathbb{C})$ and the following relations hold[15]:

$$\Gamma(V_{(g,e,r)})d\Gamma(\mathbb{1})\Gamma(V_{(g,e,r)})^{-1}=d\Gamma(\mathbb{1}_{\mathbb{C}}),$$

where in the LHS $d\Gamma(\mathbb{1})$ is the number operator in $\mathscr{F}(\mathscr{M}_{(g,e,r)})$, while in the RHS $d\Gamma$ means second quantization in $\mathscr{F}(\mathbb{C})$ and therefore $d\Gamma(\mathbb{1}_{\mathbb{C}})$ is the number operator in $\mathscr{F}(\mathbb{C})$;

$$\Gamma(V_{(g,e,r)})W^{\rm II}_{(g,e,r)}(m)\Gamma(V_{(g,e,r)})^{-1}$$
$$=W^{\rm II}_{\mathbb{C}}(V_{(g,e,r)}(m)),\quad \forall m\in\mathscr{M},$$

where $W^{\rm II}_{\mathbb{C}}(z)$ is defined for each $z\in\mathbb{C}$ by means of the annihilation and creation operators a and a^+ in l_2 as

$$W^{\rm II}_{\mathbb{C}}(z):=\exp\left[(i/\sqrt{2})\overline{(z^*a+za^\dagger)}\right].$$

We notice that $W^{\rm II}_{\mathbb{C}}\circ V_{(g,e,r)}$ is an IWS over $\mathscr{M}$ by the very way it has been introduced. From the above-written relations it follows that a second quantizable classical dynamical system whose motion is $S^{(\gamma,\eta,\rho)}$ is quantized by the quantum dynamical system defined in the Hilbert space $\mathscr{F}(\mathbb{C})$ by the pair $(W^{\rm II}_{\mathbb{C}}\circ V_{(g,e,r)}, (2\gamma/g)d\Gamma(\mathbb{1}_{\mathbb{C}}))$, where (g,e,r) is either the triple $[2(\operatorname{sgn}\gamma)/\sqrt{\gamma\eta-\rho^2}\,](\gamma,\eta,\rho)$ if the motion is not the trivial one or any triple if $\gamma=\eta=\rho=0$.[16] Thus, it is clear that the Hamiltonians of the nontrivial second quantized systems are oscillatorlike, since they are—up to an additive and a multiplicative constant—nothing else than the number operator in l_2. Besides, it is clear that while the Hamiltonians coincide of the systems for which $(\operatorname{sgn}\gamma)(\gamma\eta-\rho^2)^{1/2}$ has a fixed value, they are not even unitarily equivalent for different values of $(\operatorname{sgn}\gamma)(\gamma\eta-\rho^2)^{1/2}$, since $\gamma/g=(\operatorname{sgn}\gamma)\times(\gamma\eta-\rho^2)^{1/2}$. Moreover, when this value is fixed, (g,e,r) surely varies along with (γ,η,ρ). Whence also the operator $V_{(g,e,r)}$ varies (as is easy to see), and the IWS $W^{\rm II}_{\mathbb{C}}\circ V_{(g,e,r)}$ varies as well (by the injectivity of $W^{\rm II}_{\mathbb{C}}$). Therefore, even if the quantum motions coincide the kinematical pictures do not, and this most clearly shows that in the quantizations of different classical dynamical systems, momentum, position, and energy are different labeled observables.

ACKNOWLEDGMENT

One of the authors (F.G.) would like to thank Prof. A.S. Wightman for the hospitality extended to him at the Department of Physics of Princeton University.

[1] R. Abraham and J.E. Marsden, *Foundations of Mechanics* (Benjamin, New York, 1967).
[2] I.E. Segal, *Mathematical Problems of Relativistic Physics* (Am. Math. Soc., Providence, Rhode Island, 1963).
[3] I.E. Segal, "Representations of Canonical Commutation Relations," in *Applications of Mathematics to Problems in Theoretical Physics*, edited by F. Lurçat (Gordon and Breach, New York, 1967).
[4] P.J.M. Bongaarts, "Linear Fields According to I.E. Segal," in *Mathematics of Contemporary Physics*, edited by R.F. Streater (Academic, London, 1972).
[5] J. von Neumann, Math. Ann. **104**, 570 (1931).
[6] A.S. Wightman, Rev. Mod. Phys. **34**, 845 (1962).
[7] J.M. Jauch, *Foundations of Quantum Mechanics* (Addison–Wesley, Reading, Massachusetts, 1968).
[8] J.E. Roberts, J. Math. Phys. **7**, 1097 (1966).
[9] J.M. Jauch, Helv. Phys. Acta **33**, 711 (1960).
[10] R. Cirelli and F. Gallone, Ann. Inst. H. Poincaré, **A 19**, 297 (1973).
[11] In our case, every authomorphism of the Weyl algebra admits of a unique up to a factor unitary realization. Therefore, a one-parameter group of automorphisms of the Weyl algebra turns into a unitary up to a factor representation of $\mathbb{R}$. It can be shown that the factor can be dropped, thus obtaining a unitary representation of $\mathbb{R}$, if the one-parameter group of automorphisms in continuous in a suitable sense; if that is the case, the unitary representation is continuous as well; see, e.g., V. Bargmann, Ann. Math. **59**, 1 (1954).
[12] It should be realized that these simple definitions are satisfactory enough in the finite-dimensional case only. In the infinite-dimensional case, indeed, it is not *a priori* clear which topology has to be given to $\mathscr{M}$, and the quantum dynamical motion of a system described by a Weyl algebra cannot be discussed without discussing the physical vacuum state as well.
[13] M. Reed and B. Simon, *Methods of Modern Mathematical Physics* (Academic, New York, 1975), Vol. II.
[14] J. Dixmier, Comput. Math. **13**, 263 (1958).
[15] The first relation is trivial to prove, while the second one is proved in Theorem X.41 of Ref. 13.
[16] This result is quite in agreement with what was stated in Proposition 4.2. In fact $\gamma/g=(\operatorname{sgn}\gamma)(\gamma\eta-\rho^2)^{1/2}$; moreover, the unitary operator from l_2 onto $L^2(\mathbb{R})$ which identifies the "natural" orthonormal basis of l_2 with the Hermite orthonormal basis of $L^2(\mathbb{R})$, carries the operator $d\Gamma(\mathbb{1}_{\mathbb{C}})$ into the operator $\frac{1}{2}(P^2_{\mathbb{C}}+Q^2_{\mathbb{C}})$ up to the additive constant $\frac{1}{2}$; therefore, through this unitary operator the Hamiltonian $(2\gamma/g)d\Gamma(\mathbb{1}_{\mathbb{C}})$ goes into the operator $(\operatorname{sgn}\gamma)(\gamma\eta-\rho^2)^{1/2}(P^2_{\mathbb{C}}+Q^2_{\mathbb{C}})$ up to an additive constant.

On some representations of the Poincaré group on phase space[a)]

S. Twareque Ali[b)]

Department of Mathematics, University of Toronto, Toronto, Ontario Canada, M5S 1A1
(Received 31 October 1978)

Some representations of the Poincaré group by functions on phase space are studied both for classical as well as quantum relativistic systems. The classical representations are identified with certain canonically induced representations, and the quantum representations are then obtained on the same Hilbert space. Equations of motion on phase space are also developed.

1. INTRODUCTION

A formulation of nonrelativistic quantum mechanics on phase space has recently been carried out with considerable success [cf. Refs. 1–3, and related papers cited therein]. Mathematically, one distinguishes between the general and the extremal representations of quantum mechanics on phase space. In the case of the extremal representations, one starts with the Hilbert space $L^2(\Gamma)$ of square integrable functions (with respect to the usual Lebesgue measure) on the phase space $\Gamma = \mathbb{R}^6$ of a spinless, noninteracting massive particle. This space carries a reducible unitary representation of the canonical commutation relations, or more precisely, of the Weyl group. One then looks for irreducible subrepresentations by continuous functions on Γ. It turns out that a class of such irreducible subrepresentations is projected out by means of projection operators of the type

$$(\mathbb{P}_e\phi)(\mathbf{q},\mathbf{p}) = \langle e_{\mathbf{q},\mathbf{p}} | \phi \rangle, \tag{1.1}$$

where $\phi \in L^2(\Gamma)$, $(\mathbf{q},\mathbf{p}) \in \Gamma$ and

$$e_{\mathbf{q},\mathbf{p}} = U(\mathbf{q},\mathbf{p})e, \tag{1.2}$$

with e being a fixed vector in $L^2(\Gamma)$ and the $U(\mathbf{q},\mathbf{p})$'s being the unitary operators of the representation of the Weyl group on $L^2(\Gamma)$. We denote by $L^2(\Gamma_e)$ the subspace $\mathbb{P}_e L^2(\Gamma)$ of $L^2(\Gamma)$. For a vector $\phi \in L^2(\Gamma_e)$, with $\|\phi\| = 1$, the quantity $|\phi(\mathbf{q},\mathbf{p})|^2$ then has a very natural interpretation as a probability density on a stochastic phase space

$$\Gamma_e = \{((\mathbf{q},\chi_\mathbf{q}),(\mathbf{p},\chi'_\mathbf{p})) \mid (\mathbf{q},\mathbf{p}) \in \Gamma\}, \tag{1.3}$$

where $\chi_\mathbf{q}$ and $\chi'_\mathbf{p}$ are confidence functions for measurements at the points $\mathbf{q}$ and $\mathbf{p}$, respectively. Under the action of the Weyl group, Γ_e is a Borel space which is Borel isomorphic to Γ. The general, or nonextremal representations, are then obtained as convex combinations of the extremal ones in an appropriate sense.

To carry out a similar analysis in the relativistic context, it is necessary to introduce the dynamics into the picture as well. It then becomes necessary, as pointed out in Ref. 4, to look at representations of the Poincaré group by functions on phase space. As demonstrated in Refs. 4 and 5, a completely consistent relativistic theory can be obtained on stochastic phase spaces. The theory is, moreover, free from many of the problems that plague standard treatments of relativistic wave equations. In particular, a phase space theory gives rise to bona-fide probability densities which satisfy the correct equations of continuity. The physical interpretation of a phase space theory has been given in detail in Ref. 4.

In this paper we wish to dwell on certain mathematical aspects of relativistic quantum mechanics on phase space. To this end we study representations of the proper Poincaré group by functions on phase space. We identify the representations which correspond both to the classical as well as to the quantum descriptions of particles with arbitrary spin and nonzero mass. The quantum mechanical representations are obtained as irreducible subsectors of representations on $L^2(\Gamma)$ and the projection onto these subspaces is done once again as in Eq. (1.1). Finally, we also briefly look at some relativistic wave equations on phase space.

2. CANONICAL INDUCED REPRESENTATION OF $\mathscr{P}_{\uparrow}^{+}$ ON PHASE SPACE

In this section we shall follow the standard Mackey program[6] to construct an induced representation of the Poincaré group on the space of all classical trajectories. A parametrization of each trajectory by its initial conditions will then lead to the identification of the sets of all trajectories with the classical phase space $\Gamma^{(m)}$ of a massive particle. In this way we shall be able to develop a relativistic statistical mechanics of a free classical massive particle. A generalization of the usual free Liouville equation to the relativistic domain will thereby be obtained, as well as an equation of continuity for the phase space density functions $\rho(\mathbf{q},\mathbf{p},t)$. A manifestly covariant version of the theory will then be presented, and an equation of motion in the presence of an electromagnetic field will be developed.

[a)]Supported in part by a research grant from the National Research Council of Canada.

[b)]Present address: Department of Mathematics and Computer Science, University of Prince Edward Island, Charlottetown, P.E.I. Canada, C1A 4P3.

 0022-2488/79/071385-07$01.00

The Poincaré group $\mathscr{P}_{\uparrow}^{+}$ is the semidirect product $\mathbb{R}^4 \circledS L_{\uparrow}^{+}$, of the proper orthocronous Lorentz group $L_{\uparrow}^{+}$ and the Abelian group $\mathbb{R}^4$ of all space–time translations. We shall denote a general element of $\mathscr{P}_{\uparrow}^{+}$ by $\{a,\Lambda\}$, where $a\in\mathbb{R}^4$ and $\Lambda\in L_{\uparrow}^{+}$. The multiplication law in $\mathscr{P}_{\uparrow}^{+}$ is given by

$$\{a_1, \Lambda_1\}\{a_2, \Lambda_2\} = \{a_1 + \Lambda_1 a_2, \Lambda_1\Lambda_2\}. \tag{2.1}$$

Consider, now the subgroup $\mathbb{R}\otimes SO(3)$ of $\mathscr{P}_{\uparrow}^{+}$ consisting of time translations and spatial rotations. When we wish to emphasize this subgroup structure of an element $\{a,\Lambda\}$ of $\mathscr{P}_{\uparrow}^{+}$, we shall write

$$\{a, \Lambda\} = \{(a_0,\mathbf{a}),\Lambda_v R\}, \tag{2.2}$$

where $a_0 = ct$ is a pure time translation (c = velocity of light) and $\mathbf{a}$ is a pure space translation. In Eq. (2.2) we have used the canonical decomposition (cf., for example Ref. 7, chapter 17),

$$\Lambda = \Lambda_v R \tag{2.3}$$

of Λ into a "velocity boost" Λ_v, parametrized by a pure velocity $\mathbf{v}$ and a spatial rotation matrix R. In terms of the three components v_x, v_y, and v_z of $\mathbf{v}$ and the quantity $v^2 = \mathbf{v}\cdot\mathbf{v}$, the explicit form of the matrix Λ_v is

$$\Lambda_v = \left[\begin{array}{c|c|c|c} \gamma & \gamma\dfrac{v_x}{c} & \gamma\dfrac{v_y}{c} & \gamma\dfrac{v_z}{c} \\ \gamma\dfrac{v_x}{c} & 1+(\gamma-1)\dfrac{v_x^2}{v^2} & (\gamma-1)\dfrac{v_xv_y}{v^2} & (\gamma-1)\dfrac{v_xv_z}{v^2} \\ \gamma\dfrac{v_y}{c} & (\gamma-1)\dfrac{v_xv_y}{v^2} & 1+(\gamma-1)\dfrac{v_y^2}{v^2} & (\gamma-1)\dfrac{v_y v_z}{v^2} \\ \gamma\dfrac{v_z}{c} & (\gamma-1)\dfrac{v_z v_x}{v^2} & (\gamma-1)\dfrac{v_y v_z}{v^2} & 1+(\gamma-1)\dfrac{v_z^2}{v^2} \end{array}\right], \tag{2.4}$$

where, of course,

$$\gamma = \left(1 - \frac{v^2}{c^2}\right)^{-1/2}. \tag{2.5}$$

Let us also note, that if u represents the relativistic 4-velocity corresponding to the 3-velocity $\mathbf{v}$, i.e.,

$$u = (u_0,\mathbf{u}) = (\gamma c,\gamma\,\mathbf{v}), \tag{2.6}$$

then

$$\Lambda_v(c,\mathbf{0}) = u. \tag{2.7}$$

Let

$$\mathscr{M} = \mathscr{P}_{\uparrow}^{+}/\mathbb{R}\otimes SO(3) \tag{2.8}$$

denote the left coset space corresponding to the subgroup $\mathbb{R}\otimes SO(3)$. An element of $\mathscr{P}_{\uparrow}^{+}$ then has the (Mackey) decomposition

$$\{a,\Lambda\} = \left\{\left(0,\mathbf{a} - \frac{a_0}{c}\mathbf{v}\right), \Lambda_v\right\}\left\{\left(\frac{a_0}{\gamma},\mathbf{0}\right),R\right\}, \tag{2.9}$$

a relation which may be used to parametrize the elements of $\mathscr{M}$. Indeed, every element of $\mathscr{M}$ is uniquely described by a pair of 3-vectors $(\mathbf{q},\mathbf{v})$, and an element $\{a,\Lambda\}$ in $\mathscr{P}_{\uparrow}^{+}$ belongs to the coset represented by $(\mathbf{q},\mathbf{v})$ if and only if

$$\{a, \Lambda\} = \{(0,\mathbf{q}), \Lambda_v\}\{(a_0/\gamma),\mathbf{0}), R\}, \tag{2.10}$$

i.e., if and only if

$$\Lambda = \Lambda_v R \quad \text{and} \quad a = \left(a_0,\mathbf{q} + \frac{a_0}{c}\mathbf{v}\right). \tag{2.11}$$

Equation (2.11) simply states that the coset representative $(\mathbf{q},\mathbf{v})$ gives just the initial position $\mathbf{q}$ of a classical particle moving with velocity $\mathbf{v}$, if its current space–time configuration is given by $\{a, \Lambda\}$. In other words, $(\mathbf{q},\mathbf{v})$ is the initial condition on a trajectory of a free massive particle. Since for massive particles the initial conditions completely describe the corresponding trajectories, $\mathscr{M}$ can be identified with the set of all classical trajectories of a free massive particle.

The space $\mathscr{M}$ is a homogeneous transformation space, in the sense of Mackey[6] under the action of $\mathscr{P}_{\uparrow}^{+}$. To display explicitly how an element of $\mathscr{M}$ transforms under $\mathscr{P}_{\uparrow}^{+}$, let us first introduce the contravariant and covariant vectors x^μ and x_μ, respectively, in Minkowski space:

$$x^\mu = (x_0,\mathbf{x}), \quad x_\nu = g_{\mu\nu}x^\nu = (x_0, -\mathbf{x}), \tag{2.12}$$

with metric convention

$$g_{\mu\nu} = g^{\mu\nu}, \quad g_{00} = 1 = -g_{11} = -g_{22} = -g_{33}. \tag{2.13}$$

Further, under a Lorentz transformation Λ,

$$x'^\mu = \Lambda^\mu_{\ \nu} x^\nu \tag{2.14}$$

$$(\Lambda^{-1})^\mu_{\ \nu} = (\Lambda^T)^\mu_{\ \nu} = \Lambda_\nu^{\ \ \mu}, \tag{2.15}$$

where Λ^T denotes the transpose of the matrix Λ. It is then easy to verify, using Eqs. (2.1), (2.9), and (2.11) that

$$(\mathbf{q}',\mathbf{v}') = \{a, \Lambda\}^{-1}(\mathbf{q},\mathbf{v}),$$
$$q'^k = -(\Lambda^{-1})^k_\nu a^\nu + (\Lambda^{-1})^k_j q^j + \frac{1}{c}[(\Lambda^{-1})^0_\nu a^\nu - (\Lambda^{-1})^0_j q^j]v'^k, \tag{2.16}$$
$$v'^k = c(\Lambda^{-1})^k_\nu u^\nu/(\Lambda^{-1})^0_\nu u^\nu,$$
is the transformation law for an element $(\mathbf{q},\mathbf{v})$ in $\mathscr{M}$ under the action of the group element
$$\{a, \Lambda\}^{-1} = \{-\Lambda^{-1}a, \Lambda^{-1}\}. \tag{2.17}$$
In Eq. (2.16) we follow standard conventions regarding summation and the use of Greek and Latin indices.

In view of Eq. (2.7) and the fact that elements in $\mathscr{M}$ are parametrizable as $(\mathbf{q},\mathbf{v})$ we see that $\mathscr{M}$ is isomorphic to $\mathbb{R}^3\times\mathscr{V}^+$, where $\mathscr{V}^+$ is the forward "velocity hyperboloid"
$$\mathbf{u}^2 - u_0^2 = -c^2. \tag{2.18}$$

It is well known that $\mathscr{V}^+$ as a homogeneous space under $\mathscr{P}^{\,+}_{\uparrow}$ admits the invariant measure d^3u/u_0. We shall now show that $\mathbb{R}^3\times\mathscr{V}^+$ admits the invariant measure
$$d\mu = d^3\mathbf{q}\, d^3\mathbf{u}. \tag{2.19}$$
Indeed, since both $\mathbb{R}$ and SO(3) are unimodular groups, and so also is $\mathscr{P}^{\,+}_{\uparrow}$, it follows that $\mathscr{M}$ does admit an invariant measure.[7] Let us denote this measure by $d\mu'$. Then if $d(q,\Lambda)$ denotes the invariant Haar measure on $\mathscr{P}^{\,+}_{\uparrow}$ and f is any measurable function on $\mathscr{P}^{\,+}_{\uparrow}$, we must have[7]
$$\int_{\mathscr{P}^{\,+}_{\uparrow}} f(q,\Lambda)d(q,\Lambda) = \int_{\mathbb{R}^3\times\mathscr{V}^+} d\mu'(\mathbf{q},\mathbf{u}) \int_{\mathbb{R}\otimes SO(3)} f(\{q,\Lambda\}\{(a_0\mathbf{0}),R\})da_0 dR, \tag{2.20}$$
where da_0dR is the Haar measure on the subgroup $\mathbb{R}\otimes SO(3)$. Using Eq. (2.9) and the invariance of da_0dR, the right-hand side of Eq. (2.20) can be transformed into
$$\int_{\mathbb{R}^3\times\mathscr{V}^+} d\mu'(\mathbf{q},\mathbf{u}) \times \int_{\mathbb{R}\otimes SO(3)} f\left(\left\{\left(0,\mathbf{q}-\frac{q_0}{c}\mathbf{v}\right),\Lambda_v\right\}\{(a_0,\mathbf{0}),R\}\right)da_0\, dR$$
$$= \int_{\mathbb{R}\times\mathscr{V}^+} d\mu'(\mathbf{q},\mathbf{u}) \times \int f(\{(0,\mathbf{q}),\Lambda_v\}\{(q_0,\mathbf{0}),R\})dq_0\, dR. \tag{2.21}$$
on the other hand, since
$$d(q,\Lambda) = d^4q\frac{d^3\mathbf{u}}{u_0}dR = d^3\mathbf{q}\, dq_0\frac{d^3\mathbf{u}}{u_0}dR, \tag{2.22}$$
the left-hand side of Eq. (2.20) may be rewritten as
$$\int_{\mathscr{P}^{\,+}_{\uparrow}} f(\{(q_0,\mathbf{q}),\Lambda_v R\})\frac{d^3\mathbf{u}}{u_0}d^3\mathbf{q}\, dq_0\, dR$$
$$= \int_{\mathbb{R}^3\times\mathscr{V}^+} d^3\mathbf{q}\frac{d^3\mathbf{u}}{u_0} \times \int_{\mathbb{R}\otimes SO(3)} f\left(\left\{\left(0,\mathbf{q}-\frac{q_0}{c}\cdot\mathbf{v}\right),\Lambda_v\right\}\left\{\left(\frac{q_0}{\gamma},\mathbf{0}\right),R\right\}\right)dq_0\, dR$$
$$= \int_{\mathbb{R}\times\mathscr{V}^+} \gamma d^3\mathbf{q}\frac{d^3\mathbf{u}}{u_0}\int_{\mathbb{R}\otimes SO(3)} f(\{(0,\mathbf{q}),\Lambda_v\} \times\{(q_0,\mathbf{0}),R\})dq_0\, dR. \tag{2.23}$$

Comparing (2.21) and (2.23) we find that
$$d\mu' = \gamma d_3\mathbf{q}\frac{d^3\mathbf{u}}{u_0} = \frac{1}{c}d^3\mathbf{q}\, d^3\mathbf{u}.$$
If we ignore the unimportant multiplicative constant $1/c$, this is the same relationship as (2.19).

In the sequel we shall also parametrize the coset space $\mathscr{M}$ by elements of
$$\Gamma^{(m)} = \mathbb{R}^3\times\mathscr{V}^{\,+}_m, \tag{2.24}$$
where $\mathscr{V}^{\,+}_m$ is the "forward mass hyperboloid" consisting of 4-momentum vectors
$$p = (p_0,\mathbf{p}) = mu = (mc\gamma, m\mathbf{v}\gamma) \tag{2.25}$$
of particles having a fixed mass m. We shall, in this case, denote the matrix in (2.4) by Λ_p, and replace Eq. (2.7) by
$$\Lambda_p(mc,\mathbf{0}) = p. \tag{2.26}$$
The invariant measure on $\Gamma^{(m)}$ will be taken to be
$$d\mu_m = d^3\mathbf{q}\, d^3\mathbf{p}, \tag{2.27}$$
which differs from (2.19) only by the constant factor m^3.

The space $\Gamma^{(m)}$ will be called the relativistic phase space for a particle of mass m. We shall denote by $L^2(\Gamma^{(m)})$ the Hilbert space of all functions on $\Gamma^{(m)}$ which are square integrable with respect to $d\mu_m$. This space will play for us a central role in the building of representations of $\mathscr{P}^{\,+}_{\uparrow}$ on phase space. Let us first construct what we shall call the canonical induced representation on phase space.

Consider the unitary irreducible representation $V^{j,m}$ of $\mathbb{R}\otimes SO(3)$ defined by
$$V^{j,m}(q_0,R)\xi = \exp\left(\frac{imc}{\hbar}q_0\right)L^j(R)\xi. \tag{2.28}$$
Here j is one of the numbers $0, 1/2, 1, 3/2,\cdots$, $m>0$, and ξ is a vector in the Hilbert space $\mathscr{K}^j$, which is a $(2j+1)$-dimensional spinor space. L^j is the usual $(2j+1)$-dimensional unitary irreducible representation of SO(3) [or more properly of its covering group SU(2)]. We proceed to construct the representation of $\mathscr{P}^{\,+}_{\uparrow}$ which is induced from $V^{j,m}$.

The Hilbert space for the induced representation is $\mathscr{H}^{j,m}$, a space of functions $f^{j,m}$ from $\mathscr{P}^{\,+}_{\uparrow}$ to $\mathscr{K}^j$, which satisfy
$$f^{j,m}(\{q,\Lambda\}\{(a_0,\mathbf{0}),R\}) = \exp\left(\frac{-imc}{\hbar}a_0\right)L^j(R)^{-1} \times f^{j,m}(q,\Lambda), \tag{2.29}$$
for all elements $\{(a_0,\mathbf{0}),R\}\in\mathbb{R}\otimes SO(3)$. The scalar product in $\mathscr{H}^{j,m}$ is assumed to be
$$\langle f^{j,m}|g^{j,m}\rangle_{j,m} = \int_{\mathbb{R}^3\times\mathscr{V}^+_m}(f^{j,m}(q,\Lambda), g^{j,m}(q,\Lambda))_j\, d\mu_m, \tag{2.30}$$
where $(\ ,\ )_j$ is the scalar product in $\mathscr{K}^j$. The representation $U^{j,m}$ of $\mathscr{P}^{\,+}_{\uparrow}$ on $\mathscr{H}^{j,m}$ is then defined by
$$(U^{j,m}(q,\Lambda)f^{j,m})(q',\Lambda') = f^{j,m}(\Lambda^{-1}(q'-q),\Lambda^{-1}\Lambda'). \tag{2.31}$$

Equations (2.28)–(2.31) define the canonically induced representation of $\mathscr{P}^{\,+}_{\uparrow}$ from the unitary irreducible representation $V^{j,m}$ of the subgroup $\mathbb{R}\otimes SO(3)$. The representation

$U^{j,m}$ is highly reducible, but it does describe the relativistic statistical mechanics of a free classical particle—a fact we now proceed to analyze.

First let us note that $U^{j,m}$ admits a canonical system of imprimitivity. Let Δ denote a Borel set in $\Gamma^{(m)}$ and $\dot{\Delta}$ its preimage in $\mathscr{P}_{\uparrow}^{+}$, and let us define a projection valued measure $P(\Delta)$ on $\mathscr{M}$, with values in $\mathscr{H}^{j,m}$ as

$$(P(\Delta)f^{j,m})(q,\Lambda)=\chi_{\dot{\Delta}}(q,\Lambda)f^{j,m}(q,\Lambda), \tag{2.32}$$

where $\chi_{\dot{\Delta}}$ is the characteristic function of the set $\dot{\Delta}$,

$$\left.\begin{aligned}\chi_{\dot{\Delta}}(q,\Lambda)&=1, \quad \text{if } \{q,\Lambda\}\in\dot{\Delta}\\ &=0, \quad \text{otherwise}\end{aligned}\right\}. \tag{2.33}$$

One then easily verifies that for all $(q,\Lambda)\in\mathscr{P}_{\uparrow}^{+}$,

$$U(q,\Lambda)P(\Delta)U^{*}(q,\Lambda)=P(\{q,\Lambda\}\Delta), \tag{2.34}$$

$\{q,\Lambda\}\Delta$ denoting the translate of the set Δ by $\{q,\Lambda\}$, i.e., the set in $\mathscr{M}$ whose preimage in $\mathscr{P}_{\uparrow}^{+}$ is $\{q,\Lambda\}\dot{\Delta}$.

Next, let us define a unitary mapping W, between $\mathscr{H}^{j,m}$ and $L^{2}(\Gamma^{(m)})\otimes\mathscr{K}^{j}$ by which an element $f^{j,m}$ in $\mathscr{H}^{j,m}$ is mapping into $f\in L^{2}(\Gamma^{(m)})\otimes\mathscr{K}^{j}$ via

$$f(\mathbf{q},\mathbf{p})=(Wf^{j,m})(\mathbf{q},\mathbf{p})=f^{j,m}(\{(0,\mathbf{q}),\Lambda_{p}). \tag{2.35}$$

The inverse of this map is

$$f^{j,m}(q,\Lambda)=(W^{-1}f)(q,\Lambda)=\exp\left(\frac{-im^{2}c^{2}q_{0}}{p_{0}\hbar}\right)L^{j}(R)^{-1}\times f\left(\mathbf{q}-\frac{q_{0}\mathbf{p}}{p_{0}},p\right). \tag{2.36}$$

Under W the projection operators $P(\Delta)$ transform into

$$E(\Delta)=WP(\Delta)W^{-1}, \tag{2.37}$$

with

$$(E(\Delta)\mathbf{f})(\mathbf{q},\mathbf{p})=\chi_{\Delta}(\mathbf{q},\mathbf{p})f(\mathbf{q},\mathbf{p}). \tag{2.38}$$

Similarly, the operators $U^{j,m}$ transform into

$$U(a,\Lambda)=WU^{j,m}(a,\Lambda)W^{-1} \tag{2.39}$$

with

$$\begin{aligned}(U(a,\Lambda)f)(\mathbf{q},\mathbf{p})&=\exp\left[-\frac{im^{2}c^{2}}{\hbar}(\Lambda^{-1})^{0}_{\nu}(\mathbf{q}-a)^{\nu}/(\Lambda^{-1})^{0}_{\nu}p^{\nu}\right]\\ &\quad\times L^{j}(\Lambda^{-1}_{\Lambda^{-1}p}\Lambda^{-1}\Lambda_{p})^{-1}f(\{a,\Lambda\}^{-1}(\mathbf{q},\mathbf{p})),\end{aligned} \tag{2.40}$$

$\{a,\Lambda\}(\mathbf{q},\mathbf{p})=\{a,\Lambda\}(\mathbf{q},m\gamma\mathbf{v})$ being given as in (2.16), and $\mathbf{q}-a$ being the 4-vector $(-a_{0},\mathbf{q}-\mathbf{a})$.

For $U(a,\Lambda)$, the generator of the subgroup of time translations is easily computed. Let $V(t)$ denote the one-parameter unitary group

$$V(t)=U(\{(-ct,\mathbf{0}),I\}), \tag{2.41}$$

where I is the unit element of SO(3). Now, writing

$$f(\mathbf{q},\mathbf{p},t)\equiv(V(t)f)(\mathbf{q},\mathbf{p})=f^{j,m}(q,\Lambda_{P}), \tag{2.42}$$

for any vector f in $L^{2}(\Gamma^{(m)})\otimes\mathscr{K}^{j}$, where $q=(ct,\mathbf{q})$, we have from Eq. (2.40),

$$\begin{aligned}f(\mathbf{q},\mathbf{p},t)&=\exp\left(\frac{-im^{2}c^{3}t}{\hbar p_{0}}\right)f\left(\mathbf{q}-c\frac{\mathbf{p}t}{p_{0}},\mathbf{p}\right)\\ &=\left(\exp\left\{-\frac{ict}{\hbar p_{0}}[m^{2}c^{2}+\mathbf{p}\cdot\widehat{\mathbf{P}}]\right\}\right)(\mathbf{q},\mathbf{p}).\end{aligned} \tag{2.43}$$

In Eq. (2.43) we have introduced the vector operator

$$\widehat{\mathbf{P}}=(\widehat{P}^{1},\widehat{P}^{2},\widehat{P}^{3})=-i\hbar\left(\frac{\partial}{\partial q^{1}},\frac{\partial}{\partial q^{2}},\frac{\partial}{\partial q^{3}}\right). \tag{2.44}$$

Thus, writing

$$H_{c}=\frac{c}{p_{0}}[m^{2}c^{2}+\mathbf{p}\cdot\widehat{\mathbf{P}}] \tag{2.45}$$

we get

$$V(t)=\exp\left(-\frac{i}{\hbar}H_{c}t\right). \tag{2.46}$$

This generator H_{c} of time evolution is the relativistic analogue of a similar classical evolution operator on nonrelativistic phase space discussed in Refs. 3 and 8.

Further, differentiating Eq. (2.43) with respect to time, we get

$$\frac{\partial}{\partial t}f(\mathbf{q},\mathbf{p},t)=-\frac{ic}{\hbar p_{0}}[m^{2}c^{2}+\mathbf{p}\cdot\widehat{\mathbf{P}}]f(\mathbf{q},\mathbf{p},t), \tag{2.47}$$

for all functions f in $L^{2}(\Gamma^{(m)})\otimes\mathscr{K}^{j}$ which are in the common dense domain of the operators $\widehat{\mathbf{P}}^{j}$. Equation (2.47) can be rewritten in a manifestly covariant form if we introduce the 4-vector operator

$$\widehat{P}^{\mu}=(\widehat{P}^{0},\widehat{\mathbf{P}})=\left(\frac{i\hbar}{c}\frac{\partial}{\partial t},\widehat{\mathbf{P}}\right), \tag{2.48}$$

for then we have

$$(p_{\mu}\widehat{P}^{\mu}-m^{2}c^{2})f=0. \tag{2.49}$$

If on the other hand, we study the time evolution of the square amplitudes

$$\rho(\mathbf{q},\mathbf{p}t)=\|f(\mathbf{q},\mathbf{p},t)\|_{j}^{2} \tag{2.50}$$

where $\|\cdots\|_{j}$ is the norm in $\mathscr{K}^{j}$, it is easily seen from Eq. (2.43) that ρ satisfies

$$p_{\mu}\widehat{P}^{\mu}\rho(\mathbf{q},\mathbf{p},t)=0. \tag{2.51}$$

In fact, if $f_{i}(\mathbf{q},\mathbf{p},t)$ denotes the ith "spin component" of f in $\mathscr{K}^{j}$, then

$$\rho_{i}(\mathbf{q},\mathbf{p},t)=|f_{i}(\mathbf{q},\mathbf{p},t)|^{2} \tag{2.52}$$

also satisfies the equation

$$p_{\mu}\widehat{P}^{\mu}\rho_{i}(q,p,t)=0 \tag{2.53}$$

and, of course,

$$\rho=\sum_{i=1}^{2j+1}\rho_{i}. \tag{2.54}$$

From Eqs. (2.51) and (2.53) we infer that both p and its components ρ_{i} satisfy the relativistic Liouville equation. Furthermore, if f is normalized to unity, then in view of the imprimitivity relations (2.32)–(2.34), $\rho(\mathbf{q},\mathbf{p},t)$ is the probability density of finding a free particle of mass m, at time t, at the phase space point $(\mathbf{q},\mathbf{p})$. In other words, we have here a relativistic description of the classical statistical mechanics of a free particle of mass m and spin j.

If we assume that the particle under consideration has an electric charge e, then in the presence of an electromagnetic field $A^{\mu}(q)$, the equation of motion for ρ is easily obtained by making the replacement

$$p^{\mu}\rightarrow p^{\mu} - \frac{e}{c}A^{\mu}, \tag{2.55}$$

yielding, in place of (2.51), the equation

$$\left(p_{\mu} - \frac{e}{c}A_{\mu}(q)\right)\widehat{P}^{\mu}\rho(\mathbf{q},\mathbf{p},t) = 0. \tag{2.56}$$

The Liouville equation (2.51) can easily be seen to lead to the conservation law

$$\partial_{\mu}j^{\mu} = 0, \tag{2.57}$$

where j^{μ} is a probability 4-current density defined as

$$j^{\mu}(\mathbf{q},t) = \int_{\mathscr{V}_m^+} p^{\mu}\rho(\mathbf{q},\mathbf{p},t)\,\frac{d^3p}{p_0}, \tag{2.58}$$

and

$$\partial_{\mu}\equiv\frac{\partial}{\partial q^{\mu}}. \tag{2.59}$$

also, if J^{μ} is the integrated form of this current density,

$$J^{\mu}(t) = \int_{\mathbb{R}^3} j^{\mu}(\mathbf{q},t)\,d^3q, \tag{2.60}$$

then standard Green's theorem arguments imply

$$\frac{\partial}{\partial t}J^0 = 0. \tag{2.61}$$

3. RELATIVISTIC FREE, MASSIVE PARTICLES ON PHASE SPACE

The representation of $\mathscr{P}_1^+$ considered in Sec. 2 was reducible. In this section we shall construct a class of irreducible representation of $\mathscr{P}_1^+$ on proper subspaces of $L^2(\Gamma^{(m)})\otimes\mathscr{K}^j$. We shall indentify these irreducible sectors of $L^2(\Gamma^{(m)})\otimes\mathscr{K}^j$, with spaces $L^2(\Gamma_e^{(m)})\otimes\mathscr{K}^j$ constructed out of the stochastic phase spaces $\Gamma_e^{(m)}$, and also look at the question of extending these representations to the whole of $L^2(\Gamma^{(m)})\otimes\mathscr{K}^j$. In addition, we shall study some relativistic wave equations on phase space.

Explicity, we shall construct, for each irreducible representation of $\mathscr{P}_1^+$ corresponding to mass m and spin j, an isometric embedding of the underlying Hilbert space into a proper subspace of $L^2(\Gamma^{(m)})\otimes\mathscr{K}^j$. We shall then extend the representation to the whole of $L^2(\Gamma^{(m)})\otimes\mathscr{K}^j$.

Consider, therefore, the irreducible representation of $\mathscr{P}_1^+$ corresponding to mass m and spin j. The Hilbert space is[7]

$$\widetilde{\mathscr{H}}_m = L^2(\mathscr{V}_m^+,\gamma_m)\otimes\mathscr{K}^j, \tag{3.1}$$

where the measure γ_m is given by

$$d\gamma_m = \frac{d^3\mathbf{k}}{k_0}. \tag{3.2}$$

for $(a,\Lambda)\in\mathscr{P}_1^+$ and $\widetilde{U}(a,\Lambda)$ its representative unitary operator on $\widetilde{\mathscr{H}}_m$, we have

$$(\widetilde{U}(a,\Lambda)\widetilde{\psi})(k) = \exp\left(\frac{i}{\hbar}ak\right) \times L^j(\Lambda_{\Lambda^{-1}k}^{-1}\Lambda^{-1}\Lambda_k)^{-1}\widetilde{\psi}(\Lambda^{-1}k), \tag{3.3}$$

for all $\widetilde{\psi}\in\widetilde{\mathscr{H}}_m$, where we have written ak for the scalar product $a_k k^{\mu}$. The L^j is defined as in Sec. 2.

Let $\hat{e}$ be a complex valued function on $\mathbb{R}^3$ which satisfies

$$\int_{\mathbb{R}^3}|\hat{e}(\mathbf{k})|^2\,d^3\mathbf{k} = mc. \tag{3.4}$$

If we consider $\hat{e}$ to be a function on $\mathscr{V}_m^+$, it is easy to see that $\hat{e}\in L^2(\mathscr{V}_m^+,\gamma_m)$. Let us also assume that $\hat{e}$ is rotationally invariant i.e.,

$$\hat{e}(Rk) = \hat{e}(k), \tag{3.5}$$

for all $R\in\mathrm{SO}(3)$ and almost all $k\in\mathscr{V}_m^+$. Using $\hat{e}$, let us define, for each $\widetilde{\psi}\in\widetilde{\mathscr{H}}_m$, a function $\psi_e^{j,m}$ on $\mathscr{P}_1^+$ with values in $\mathscr{K}^j$ in the following way

$$\psi_e^{j,m}(q,\Lambda) = h^{-3/2}\int_{\mathscr{V}_m^+}\overline{\hat{e}(k)}(\widetilde{U}^*(q,\Lambda)\widetilde{\psi})(k)\,d\gamma_m. \tag{3.6}$$

We shall prove below that the mapping $\widetilde{\psi}\rightarrow\psi_e^{j,m}$ so defined embeds $\widetilde{\mathscr{H}}_m$ isometrically onto a proper subspace of $L^2(\Gamma^{(m)})\otimes\mathscr{K}^j$. Assuming this to be true, let us note first that the mapping in (3.6) carries the operators $\widetilde{U}(q,\Lambda)$ into operators $U_e^{j,m}(q,\Lambda)$ defined by

$$(U_e^{j,m}(q,\Lambda)\psi_e^{j,m})(q',\Lambda') = \psi_e^{j,m}(\Lambda^{-1}(q'-q),\Lambda^{-1}\Lambda'). \tag{3.7}$$

As in (2.35) we define next a mapping W from the functions $\psi_e^{j,m}$ to functions ψ_e on $\Gamma^{(m)}$ having values in $\mathscr{K}^j$,

$$\psi_e(\mathbf{q},\mathbf{p}) = (W\psi_e^{j,m})(\mathbf{q},\mathbf{p}) = \psi_e^{j,m}((0,\mathbf{q})\Lambda_p). \tag{3.8}$$

The composite map $\widetilde{\psi}\mapsto\psi_e$ is, therefore,

$$\psi_e(\mathbf{q},\mathbf{p}) = h^{-3/2}\int_{\mathscr{V}_m^+}\overline{\hat{e}(k)}(\widetilde{U}^*((0,\mathbf{q}),\Lambda_p)\widetilde{\psi})(k)\,d\gamma_m. \tag{3.9}$$

We collect below some properties of the functions $\psi_e^{j,m}$, ψ_e and the map W.

Theorem 1: The mapping $\widetilde{\psi}\rightarrow\psi_e$ establishes a linear Hilbert space isometry between $\widetilde{\mathscr{H}}_m$ and a proper subspace of $L^2(\Gamma^{(m)})\otimes\mathscr{K}^j$.

Proof: From the definition of the map in (3.9) it is clear that the functions ψ_e take on values in $\mathscr{K}^j$ and are defined on $\Gamma^{(m)}$. In fact, in view of the continuity of the representation $\widetilde{U}(a,\Lambda)$ it is straightforward to prove that the function ψ_e are also continuous. Linearity follows trivially from the definition. To prove isometry, we see that

$$\|\psi_e\|^2 = \int_{\mathbb{R}^3\times\mathscr{V}_m^+}(\psi_e(\mathbf{q},\mathbf{p}),\psi_e(\mathbf{q},\mathbf{p}))_j\,d\mu_m, \tag{3.10}$$

where, as before, $(\ ,\ )_j$ is the scalar product in $\mathscr{K}^j$. Using (3.9) and (3.3) we then have

$$\|\psi_e\|^2 = h^{-3}\int\exp\left(\frac{-i}{\hbar}(\mathbf{k}'-\mathbf{k})\cdot\mathbf{q}\right)\overline{\hat{e}(\Lambda_p^{-1}k)}\hat{e}(\Lambda_p^{-1}k')$$

$$\times\Big(L^j(\Lambda_k^{-1}\Lambda_p\Lambda_{\Lambda_p^{-1}k}\Lambda_{\Lambda_p^{-1}k'}^{-1}\Lambda_p^{-1}\Lambda_{k'})^{-1}$$

$$\times\,\tilde\psi(k'),\tilde\psi(k)\Big)_j d\mu_m\,d\gamma_m\,d\gamma_m'$$

$$=\int|\hat e(\Lambda_p^{-1}k)|^2\|\tilde\psi(k)\|_j^2\frac{d^3k}{k_0^2}d^3\mathbf{p},\tag{3.11}$$

In view of the unitarity of L^j; also, it follows from Eq. (3.4), and Theorem 3.1 in Ref. 4, that

$$\int|\hat e(\Lambda_p^{-1}k)|^2d^3\mathbf{p}=k_0.\tag{3.12}$$

Thus,

$$\|\psi_e\|^2=\int_{\gamma_m}|\tilde\psi(k)|^2\frac{d^3k}{k_0}=\|\tilde\psi\|^2.\tag{3.13}$$

Let $L^2(\Gamma_e^{(m)})\otimes\mathscr{K}^j$ be the range of the isometry $\tilde\psi\mapsto\psi_e$ in $L^2(\Gamma^{(m)})\otimes\mathscr{K}^j$. It is clearly a closed subspace of $L^2(\Gamma^{(m)})\otimes\mathscr{K}^j$. Q.E.D.

The following results will enable us to study the operator of time evolution

$$\widetilde V(t)=\widetilde U((-ct,\mathbf{0}),I),\tag{3.14}$$

in $\widetilde{\mathscr{H}}_m$ and its image in $L^2(\Gamma_e^{(m)})\otimes\mathscr{K}^j$.

Lemma 1: There exists a dense set of vectors $\tilde\psi_e$ in $\widetilde{\mathscr{H}}_m$ whose images in $L^2(\Gamma_e^{(m)})\otimes\mathscr{K}^j$ form a common dense domain for the operators

$$\hat P_j=i\hbar\frac{\partial}{\partial q^j}.$$

Proof: From the definition in (3.8) it is easily verified that

$$\psi_e(\mathbf{q},\mathbf{p})=h^{-3/2}\int_{\gamma_m}\overline{\hat e(\Lambda_p^{-1}k)}L^j(\Lambda_k^{-1}\Lambda_p\Lambda_{\Lambda_p^{-1}k})^{-1}$$

$$\times\exp\Big(\frac{i}{\hbar}\mathbf{q}\cdot\mathbf{k}\Big)\tilde\psi(k)\,d\gamma_m.\tag{3.15}$$

If $\tilde\psi$ is taken to be of compact support, it follows from the theory of Fourier transforms that ψ_e is infinitely differentiable in $\mathbf{q}$, and hence the result. Q.E.D.

Let $V_e(t)$ be the image of $\widetilde V(t)$, as defined in (3.14), in $L^2(\Gamma_e^{(m)})\otimes\mathscr{K}^j$. By analogy with (2.46) let us write

$$V_e(t)=\exp\Big(-\frac{i}{\hbar}H_et\Big).\tag{3.16}$$

Lemma 2:

$$H_e=c[\hat{\mathbf{P}}^2+m^2c^2]^{1/2}.\tag{3.17}$$

Proof: Consider the expression (3.15) where, as before, $\tilde\psi$ has compact support.

$$(\hat P_j\psi_e)(\mathbf{q},\mathbf{p})=h^{-3/2}\int k_j\overline{\hat e(\Lambda_p^{-1}k)}L^j(\Lambda_k^{-1}\Lambda_p\Lambda_{\Lambda_p^{-1}k})^{-1}$$

$$\times\exp\Big(\frac{i}{\hbar}\mathbf{q}\cdot\mathbf{k}\Big)\tilde\psi(k)\,d\gamma_m.\tag{3.18}$$

On the other hand,

$$(H_e\psi_e)(\mathbf{q},\mathbf{p})=h^{-3/2}\int\overline{\hat e(\Lambda_p^{-1}k)}L^j(\Lambda_k^{-1}\Lambda_p\Lambda_{\Lambda_p^{-1}k})$$

$$\times\exp\Big(\frac{i}{\hbar}\mathbf{q}\cdot\mathbf{k}\Big)(\widetilde H\tilde\psi)(k)\,d\gamma_m,\tag{3.19}$$

where $\widetilde H$ is defined by

$$\widetilde V(t)=\exp\Big(-\frac{i}{\hbar}\widetilde Ht\Big).\tag{3.20}$$

But clearly, for any $\tilde\psi$ with compact support,

$$(\widetilde H\tilde\psi)(k)=ck_0\tilde\psi(k).\tag{3.21}$$

Since $k_0^2-\mathbf{k}=m^2c^2$, it follows from (3.19), (3.18), and Lemma 1 that there exists a dense domain on which

$$H_e^2=c^2[\hat{\mathbf{P}}^2+m^2c^2].$$

Hence, there exists a dense domain in $L^2(\Gamma_e^{(m)})\otimes\mathscr{K}^j$ on which Eq. (3.17) holds. Q.E.D.

By analogy with Eq. (2.42) let us write

$$\psi_e(\mathbf{q},\mathbf{p},t)=(V_e(t)\psi_e)(\mathbf{q},\mathbf{p}).\tag{3.22}$$

Then, it is easily seen that

$$\psi_e(\mathbf{q},\mathbf{p},t)=\psi_e^{j,m}(q,\Lambda_p),\tag{3.23}$$

and that these functions obey the wave equation

$$\Big(\square_q-\frac{m^2c^2}{\hbar^2}\Big)\psi_e(\mathbf{q},\mathbf{p},t)=0,\tag{3.24}$$

with

$$\square_q=-\frac{1}{c^2}\frac{\partial^2}{\partial t^2}+\nabla_{\mathbf{q}}^2.\tag{3.25}$$

Lemma 3: The functions $\psi_e^{j,m}$ and ψ_e satisfy the relation

$$\psi_e^{j,m}(q,\Lambda)=L^j(R)^{-1}(V_e(t)\psi_e)(\mathbf{q},\mathbf{p}),\tag{3.26}$$

where, as before, we have written $\Lambda=\Lambda_pR$.

Proof: From the defining equation (3.6) we have,

$$\psi^{j,m}(q,\Lambda)$$

$$=\psi^{j,m}(\{q,\Lambda_p\}\{0,R\})$$

$$=\int_{\gamma_m}\overline{\hat e(k)}(\widetilde U^*(O,R)\widetilde U^*(q,\Lambda_p)\tilde\psi)(k)d\gamma_m$$

$$=\int_{\gamma_m}\overline{\hat e(k)}L^j(\Lambda_{Rk}^{-1}R\Lambda_k)^{-1}(\widetilde U^*(q,\Lambda_p)\tilde\psi)(Rk)d\gamma_m.\tag{3.27}$$

Next we use the easily verifiable relation

$$\Lambda_{Rk}^{-1}=(R\Lambda_kR^{-1})^{-1}=R\Lambda_k^{-1}R^{-1}\tag{3.28}$$

to rewrite (3.27) in the form

$$\psi^{j,m}(q,\Lambda)=\int_{\gamma_m}\overline{\hat e(k)}L^j(R)^{-1}(\widetilde U^*(q,\Lambda_p)\tilde\psi)(Rk),$$

which upon using the invariance of $\hat e$ under rotations [cf. Eq. (3.5)] and Eqs. (3.22) and (3.23) yields the desired results. Q.E.D.

Equation (3.26) is to be seen as the analogue of Eq. (2.36), in that we have here an expression for the inverse of the map W defined in Eq. (3.8). Indeed,

$$\psi_e^{j,m}(q,\Lambda)=(W^{-1}\psi_e)(\mathbf{q},\mathbf{p})$$

$$=L^j(R)^{-1}\Big[\exp\Big(-\frac{i}{\hbar}H_et\Big)\psi_e\Big](\mathbf{q},\mathbf{p}).\tag{3.29}$$

We are now in a position to write down the form of the operators $U_e(a, \Lambda)$ which arise as the images in $L^2(\Gamma_e^{(m)}) \otimes \mathscr{K}^j$, of the corresponding representation operators $\tilde{U}(a, \Lambda)$ on $\tilde{\mathscr{H}}_m$. Indeed, a straightforward calculation using Eqs. (3.7) and (3.29), establishes the following result.

Theorem 2: The irreducible representation of $\mathscr{P}_{\uparrow}^{+}$ carried by $\Lambda^2(\Gamma_e^{(m)}) \otimes \mathscr{K}^j$, is given by

$$
\begin{aligned}
&[U_e(a, \Lambda)\psi_e](\mathbf{q},\mathbf{p}) \\
&= \exp\left(- \frac{im^2c^2}{\hbar} (\Lambda^{-1})^0_\nu(\mathbf{q} - a)^\nu \Big/ (\Lambda^{-1})^0_\nu p^\nu \right) \\
&\quad \times \left[\exp\left(- \frac{i}{\hbar c} (\Lambda^{-1})^0_\nu(\mathbf{q} - a)^\nu \right) H_e \right. \\
&\quad \times \exp\left(\frac{i}{\hbar c} (\Lambda^{-1})_\nu(\mathbf{q} - a)^\nu \right) H_c \\
&\quad \left. \times L^j (\Lambda^{-1}_{\Lambda^{-1}p} \Lambda^{-1} \Lambda_p)^{-1} \psi_e \right] [\{a, \Lambda\}^{-1}(\mathbf{q},\mathbf{p})]. \qquad (3.30)
\end{aligned}
$$

In writing down Eq. (2.30) we have used the notation introduced in Sec. 2 [cf. Eq. (2.40)].

Having thus obtained irreducible representations of $\mathscr{P}_{\uparrow}^{+}$ on phase space corresponding to any (nonzero) mass and spin, we turn now to the question of extending these representations to the whole of $L^2(\Gamma^{(m)}) \otimes \mathscr{K}^j$. The operator H_e in (3.17) can obviously be extended to an operator H defined on a dense subset of $L^2(\Gamma^{(m)}) \otimes \mathscr{K}^j$, and thus since H_c can also be globally defined on the same dense subset of $L^2(\Gamma^{(m)}) \otimes \mathscr{K}^j$, we may extend Eq. (3.30) to the whole of $L^2(\Gamma^{(m)}) \otimes \mathscr{K}^j$ to obtain

$$
\begin{aligned}
&[U(a, \Lambda)\psi](\mathbf{q},\mathbf{p}) \\
&= \exp\left(- \frac{im^2c^2}{\hbar} (\Lambda^{-1})^0_\nu(\mathbf{q} - a)^\nu \Big/ (\Lambda^{-1})_\nu p^\nu \right) \\
&\quad \times \left[\exp\left(- \frac{i}{\hbar c} (\Lambda^{-1})^0_\nu(\mathbf{q} - a)^\nu \right) H \right. \\
&\quad \times \exp\left(\frac{i}{\hbar c} (\Lambda^{-1})^0_\nu(\mathbf{q} - a)^\nu \right) H_c \\
&\quad \left. \times L^j (\Lambda^{-1}_{\Lambda^{-1}p} \Lambda^{-1} \Lambda_p)^{-1} \psi \right] [\{a, \Lambda\}^{-1}(\mathbf{q},\mathbf{p})]. \qquad (3.31)
\end{aligned}
$$

As is to be expected, this representation does not admit a system of imprimitivity based upon $\Gamma^{(m)}$ [cf. Eqs. (2.32)–(2.38)]. It does, however, admit a generalized system of imprimitivity (based upon $\Gamma^{(m)}$) of the type discussed in Ref. 1. These in turn can be extended to projective systems of imprimitivity on enlarged Hilbert spaces. We propose to report on this matter at greater length elsewhere.

Let e be the vector in $L^2(\Gamma)$ defined as

$$e(\mathbf{q},\mathbf{p}) = h^{-3/2} \overline{\hat{e}(\Lambda_p^{-1}k)} \exp\left(\frac{i}{\hbar} \mathbf{q}\cdot\mathbf{k} \right) \hat{e}(k)\, d\gamma_m. \qquad (3.32)$$

Then, as in Ref. 2, we can show that

$$e(R\mathbf{q},R\mathbf{p}) = e(\mathbf{q},\mathbf{p}), \qquad (3.33)$$

and that the operator $\mathbb{P}_e$ defined as

$$
\begin{aligned}
&(\mathbb{P}_e\psi)(\mathbf{q},\mathbf{p}) \\
&= h^{-3/2} \int_{\mathbb{R} \times \mathscr{V}_m} \overline{e(\mathbf{q}',\mathbf{p}')} [U^*((0,\mathbf{q}), \Lambda_p)\psi](\mathbf{q}',p')\, d\mu'_m \qquad (3.34)
\end{aligned}
$$

is the operator of projection from $L^2(\Gamma^{(m)}) \otimes \mathscr{K}^j$ to $L^2(\Gamma_e^{(m)}) \otimes \mathscr{K}^j$, i.e.,

$$L^2(\Gamma_e^{(m)}) \otimes \mathscr{K}^j = \mathbb{P}_e L^2(\Gamma^{(m)}) \otimes \mathscr{K}^j, \qquad (3.35)$$

and

$$\mathbb{P}_e U(a, \Lambda) \mathbb{P}_e = U_e(a, \Lambda). \qquad (3.36)$$

ACKNOWLEDGMENTS

The author is deeply indebted to Professor E. Prugovečki, for many useful discussions, critical comments and criticisms. He is also indebted to Professor L.T. Gardner for several enlightening comments on the nature of induced representations.

[1]S.T. Ali and E. Prugovečki, J. Math. Phys. **18**, 219 (1977).
[2]S.T. Ali and E. Prugovečki, Physica A **89**, 501 (1977).
[3]E. Prugovečki, Ann. Phys. (N.Y.) **110**, 102 (1978).
[4]E. Prugovečki, J. Math. Phys. **19**, 2260 (1978).
[5]E. Prugovečki, Phys. Rev. D **18**, 3655 (1978).
[6]G.W. Mackey, Ann. Math. **55**, 101 (1952); **58**, 193 (1953).
[7]A.O. Barut and R. Rązka, *Theory of Group Representations and Applications* (PWN-Polish Scientific Publisher, Warszawa, 1977).
[8]E. Prugovečki, Physica A **91**, 202 (1978).

Representations of the duals of gauge field tensors[a)]

Hanno Rund

Department of Mathematics and Optical Sciences Center, University of Arizona, Tucson, Arizona 85721

(Received 21 August 1978; revised 15 November 1978)

Classical electromagnetic fields admit a perfectly symmetrical description in terms of the electromagnetic field tensors and the duals thereof, this phenomenon being a simple consequence of Maxwell's equations for free fields. Since the Yang-Mills equation for free gauge fields possess a somewhat analogous structure, the possibility of a similar symmetrical description of gauge fields is investigated. It is shown that source-free Yang–Mills equations do, in fact, imply the existence of a new gauge in terms of which the duals of the field tensors admit a representation; however, the latter does not in general possess a structure which is identical with that of the field tensors in terms of the original gauge.

I. INTRODUCTION

It is well known that Maxwell's equations for free electromagnetic fields can be written in the form $d(*F) = 0$, where $*F$ denotes the dual of the 2-form F, the latter being defined by $F = dA$, in which $A = A_j dx^j$ is the 1-form determined by the 4-potential A_j $(j = 1,\cdots,4)$. Thus, an immediate consequence of Maxwell's equations is the existence of another 4-potential B_j, such that $*F$ can be expressed in the form $*F = dB$, where $B = B_j dx^j$, the two potentials being related to each other in an obvious manner.

The objective of the present note is the investigation of an analogous complex of ideas for the case of gauge fields. Again, it is well known that the Yang–Mills equations which describe source-free gauge fields may be represented in the form $D(*F^\alpha) = 0$, where D denotes the operator of exterior covariant differentiation as defined by the connection 1-forms A^α in terms of which the field tensors $F_h{}^\alpha{}_j$ are constructed. The question[1] which arises thus is the following: Do the Yang–Mills equations imply the existence of another set of connection 1-forms B^α in terms of which $*F^\alpha$ admits a representation whose structure is similar to that of F^α in terms of A^α? It will be shown below that the answer to this question is generally only partly in the affirmative.

After some analytical preliminaries in Sec. II, a detailed analysis is made in Sec. III of the general consequences of a system of equations of the type $DG^\alpha = 0$, where G^α denotes a set of 2-forms, and D is the operator of exterior covariant differentiation defined for a given set of connection 1-forms A^α. It is shown that these relations do in fact imply the existence of a new set of connection 1-forms B^α, in terms of which G^α admits a representation. Since the latter is not entirely unique, an analysis is given in Sec. IV of the structure of such representations. These conclusions are appplied in Sec. V to the case when $G^\alpha = *F^\alpha$, and it is found that the representation of $*F^\alpha$ in terms of B^α can have the same formal structure as that of F^α in terms of A^α only if a certain $4r \times 4r$ matrix is singular, where r is the dimension of the underlying structural Lie group G. However, when $r = 3$, the determinant of this matrix turns out to be identical with a determinant discussed by Deser and Teitelboim,[2] who have shown that it does not vanish in general. It is therefore concluded, at least for the case when $r = 3$, that—in contrast to electrodynamics— the description of the gauge field theory is not completely symmetrical in terms of F^α and its dual $*F^\alpha$. Clearly this phenomenon is complementary to the fact[3] that duality rotations cannot be consistently implemented in the case of SU(2) gauge fields. Moreover, one may hope to relate the analysis of Secs. III and IV to known results concerning the determination of local gauge fields by their field strengths.[4]

II. ANALYTICAL PRELIMINARIES[5]

Our considerations are based on a flat space–time manifold X_4, referred to local coordinates x^j, with $x^4 = ict$. It is supposed that we are given an r-parameter Lie group G, whose structure constants are denoted by $C_\beta{}^\alpha{}_\gamma$; these are skew symmetric in their subscripts and satisfy the Jacobi identities[6]

$$C_\beta{}^\alpha{}_\gamma C_\epsilon{}^\lambda{}_\gamma + C_\epsilon{}^\alpha{}_\lambda C_\gamma{}^\lambda{}_\beta + C_\gamma{}^\alpha{}_\lambda C_\beta{}^\lambda{}_\epsilon = 0. \tag{2.1}$$

Let $\chi^{\alpha_1\cdots\alpha_p}{}_{\beta_1\cdots\beta_q}$ denote a set of s-forms $(0 \leqslant s \leqslant 4)$. These are said to be a set of type (p,q) forms if the action of G is reflected in the transformation law

$$\begin{aligned}\chi'^{\alpha_1\cdots\alpha_p}{}_{\beta_1\cdots\beta_q} = \chi^{\alpha_1\cdots\alpha_p}{}_{\beta_1\cdots\beta_q} &+ \sum_{t=1}^{p} C_\epsilon{}^{\alpha_t}{}_\gamma \chi^{\alpha_1\cdots\alpha_{t-1}\gamma\alpha_{t+1}\cdots\alpha_p}{}_{\beta_1\cdots\beta_q} u^\epsilon \\ &- \sum_{t=1}^{q} C_\epsilon{}^\gamma{}_{\beta_t} \chi^{\alpha_1\cdots\alpha_p}{}_{\beta_1\cdots\beta_{t-1}\gamma\beta_{t+1}\cdots\beta_q} u^\epsilon,\end{aligned} \tag{2.2}$$

in which $u^\epsilon = u^\epsilon(x)$ denotes the parameters of G in an appropriate representation, the latter being chosen such that the identity element $e \in G$ corresponds to the values $(0,\ldots,0)$.

A connection on the principal bundle (X_4, G) gives rise in a well-known manner to a process of exterior covariant differentiation of type (p,q) forms. For instance, if χ^α denotes a set of type (1,0) forms, one defines

[a)]This research was sponsored in part by NSF GP43070.

$$D\chi^{\alpha} = d\chi^{\alpha} + C_{\epsilon}{}^{\alpha}{}_{\beta} A^{\epsilon} \wedge \chi^{\beta}, \tag{2.3}$$

where the *connection* 1-forms $A^{\alpha}(x)$ represent field variables whose transformation law is given by the gauge transformation

$$A'^{\alpha} = A^{\alpha} - Du^{\alpha} = A^{\alpha} - C_{\epsilon}{}^{\alpha}{}_{\beta} A^{\epsilon} u^{\beta} - du^{\alpha}. \tag{2.4}$$

Moreover,

$$D(D\chi^{\alpha}) = C_{\epsilon}{}^{\alpha}{}_{\beta} F^{\epsilon} \wedge \chi^{\beta}, \tag{2.5}$$

in which F^{α} denotes the *curvature* 2-forms:

$$F^{\alpha} = dA^{\alpha} + \tfrac{1}{2} C_{\epsilon}{}^{\alpha}{}_{\beta} A^{\epsilon} \wedge A^{\beta}, \tag{2.6}$$

these being of type (1,0) by virute of (2.4); also, the structure of (2.6) implies that

$$DF^{\alpha} = 0 \tag{2.7}$$

identically, these relations being generally referred to as the *Bianchi identities*. For the present, no conditions other than the transformation law (2.4) are imposed on the connection 1-forms; the Yang–Mills field equations, which are supposed to govern their behavior, will be introduced in Sec. IV.

Our subsequent analysis will be simplified considerably by the use of certain special bases of the space of s-forms at some point P of X_4, these spaces being denoted as usual by $\Lambda^{s}_{P}(X_4)$. A Minkowskian coordinate system $\{x^j\}$ on X_4 defines the so-called canonical bases $\{dx^j\}$, $\{dx^j \wedge dx^l\}$, $\{dx^j \wedge dx^l \wedge dx^h\}$, $\{dx^j \wedge dx^l \wedge dx^h \wedge dx^k\}$, with $j<h<l<k$, on $\Lambda^{1}_{P}(X_4)$, $\Lambda^{2}_{P}(X_4)$, $\Lambda^{3}_{P}(X_4)$, $\Lambda^{4}_{P}(X_4)$, respectively. A set of *dual bases*[7] is now constructed as follows. For $\Lambda^{3}_{P}(X_4)$ we define basis elements θ_j by putting

$$3!\theta_j = \epsilon_{jlhk} dx^l \wedge dx^h \wedge dx^k, \tag{2.8}$$

which may be inverted to yield

$$dx^l \wedge dx^h \wedge dx^k = \theta_j \epsilon^{jlhk}, \tag{2.9}$$

where $\epsilon_{\dots}$, $\epsilon^{\dots}$ denote the four-dimensional permutation symbols of Levi–Civita. It is easily verified that

$$dx^j \wedge \theta_h = \delta^j_h d(x), \tag{2.10}$$

where $d(x) = dx^1 \wedge dx^2 \wedge dx^3 \wedge dx^4$ is the single basis element of $\Lambda^{4}_{P}(X_4)$. Similarly, for $\Lambda^{2}_{P}(X_4)$ we put

$$2!\theta_{jl} = \epsilon_{jlhk} dx^h \wedge dx^k, \tag{2.11}$$

with inverse

$$2! dx^h \wedge dx^k = \theta_{jl} \epsilon^{jlhk}; \tag{2.12}$$

again, it is readily shown that

$$dx^k \wedge \theta_{jl} = (\delta^k_l \theta_j - \delta^k_j \theta_l). \tag{2.13}$$

Finally, for $\Lambda^{1}_{P}(X_4)$ we put

$$\theta_{jlh} = \epsilon_{jlhk} dx^k, \tag{2.14}$$

for which we have

$$3! dx^k = \epsilon^{jlhk} \theta_{jlh}, \tag{2.15}$$

with

$$dx^k \wedge \theta_{jlh} = (\delta^k_j \theta_{lh} + \delta^k_l \theta_{hj} + \delta^k_h \theta_{jl}). \tag{2.16}$$

It should be observed that $d\theta_{\dots} = 0$ in all three cases.

Let H^{α} denote a set of r arbitrary 2-forms. Relative to the canonical basis of $\Lambda^{2}_{P}(X_4)$ we put

$$H^{\alpha} = -\tfrac{1}{2} H_h{}^{\alpha}{}_k dx^h \wedge dx^k, \tag{2.17}$$

in which, of course, $H_h{}^{\alpha}{}_k = -H_k{}^{\alpha}{}_h$. With the aid of (2.12) we then have

$$H^{\alpha} = -\tfrac{1}{4} H_h{}^{\alpha}{}_k \epsilon^{jlhk} \theta_{jl} = \tfrac{1}{2} i (*H^{\alpha jl}) \theta_{jl}, \tag{2.18}$$

where, as usual, the dual $*H^{\alpha jl}$ of $H_j{}^{\alpha}{}_l$ is defined by

$$*H^{\alpha jl} = \tfrac{1}{2} i \epsilon^{jlhk} H_h{}^{\alpha}{}_k. \tag{2.19}$$

This in turm gives rise to the 2-form

$$*H^{\alpha} = -\tfrac{1}{2} i H^{\alpha jl} \theta_{jl}. \tag{2.20}$$

Moreover, it follows from (2.18) and (2.13) that

$$dH^{\alpha} = \tfrac{1}{2} i \partial_k (*H^{\alpha jl}) dx^k \wedge \theta_{jl}$$

$$= \tfrac{1}{2} i \partial_k (*H^{\alpha jl})(\delta^k_l \theta_j - \delta^k_j \theta_l) = i \partial_h (*H^{\alpha jh}) \theta_j. \tag{2.21}$$

Thus, putting $A^{\alpha} = A^{\alpha}_j dx^j$, it is seen with the aid of (2.3), (2.18), and (2.13) that the exterior covariant derivative of H^{α} may be expressed in the form

$$DH^{\alpha} = dH^{\alpha} + C_{\epsilon}{}^{\alpha}{}_{\beta} A^{\epsilon} \wedge H^{\beta}$$

$$= i[\partial_h (*H^{\alpha jh}) \theta_j + \tfrac{1}{2} C_{\epsilon}{}^{\alpha}{}_{\beta} A^{\epsilon}_l (*H^{\alpha jh}) dx^l \wedge \theta_{jh}]$$

$$= i D_h (*H^{\alpha jh}) \theta_j, \tag{2.22}$$

where the covariant divergence of $*H_j^{\alpha}{}_h$ is defined by

$$D_h (*H^{\alpha jh}) = \partial_h (*H^{\alpha jh}) + C_{\epsilon}{}^{\alpha}{}_{\beta} A^{\epsilon}_h (*H^{\alpha jh}). \tag{2.23}$$

Similarly, it is found that

$$D(*H^{\alpha}) = -i D_h (H^{\alpha jh}) \theta_j. \tag{2.24}$$

For future reference we note the following. First, if

$$\partial_h (*H^{\alpha jh}) = 0, \tag{2.25}$$

it follows directly from (2.21), since $\{\theta_j\}$ is a basis of $\Lambda^{3}_{P}(X_4)$, that $dH^{\alpha} = 0$, which implies the local existence of a 1-form $\phi^{\alpha} = \phi^{\alpha}_h dx^h$ such that $H^{\alpha} = d\phi^{\alpha}$, or in terms of the representation (2.17),

$$H_h{}^{\alpha}{}_k = \partial_k \phi^{\alpha}_h - \partial_h \phi^{\alpha}_k, \tag{2.26}$$

and hence

$$(*H^{\alpha jl}) = -i \epsilon^{jhlk} (\partial_h \phi^{\alpha}_k). \tag{2.27}$$

Clearly the 1-forms ϕ^{α} are determined up to additive exact 1-forms $d\mu^{\alpha}$.

Second, let $f^{jA}(x)$ denote a set $4M$ arbitrary class C^2 functions on $X_4 (A = 1, \cdots, M)$. Then, by (2.10),

$$(\partial_j f^{jA}) d(x) = (\partial_h f^{jA}) \delta^h_j d(x) = \partial_h (f^{jA}) dx^h \wedge \theta_j$$

$$= df^{jA} \wedge \theta_j = d(f^{jA} \theta_j). \tag{2.28}$$

Thus, if

$$\partial_j (f^{jA}) = 0, \tag{2.29}$$

it follows that there exists, at least locally, a set of M 2-forms

Λ^A such that

$$f^{jA}\theta_j = d\Lambda^A. \tag{2.30}$$

If we write

$$\Lambda^A = -\tfrac{1}{2}\Lambda_h{}^A{}_k\, dx^h \wedge dx^k, \tag{2.31}$$

we have, using (2.9)

$$d\Lambda^A = -\tfrac{1}{2}(\partial_l \Lambda_h{}^A{}_k)dx^l \wedge dx^h \wedge dx^k$$
$$= -\tfrac{1}{2}(\partial_l \Lambda_h{}^A{}_k)\theta_j \epsilon^{jlhk},$$

and the relation (2.30) therefore yields

$$f^{jA} = -\tfrac{1}{2}\epsilon^{jlhk}(\partial_l \Lambda_h{}^A{}_k). \tag{2.32}$$

The forms Λ^A which appear in (2.30) are determined up to exact 2-forms $d\lambda^A$, where $\lambda^A = \lambda_j^A dx^j$, in which the coefficients λ_j^A are arbitrary. Thus (2.32) is invariant under the substitution

$$\Lambda^A \to \tilde{\Lambda}^A = \Lambda^A - d\lambda^A, \tag{2.33}$$

the latter being reflected in the coefficients $\Lambda_h{}^A{}_k$ of Λ^A by

$$\Lambda_h{}^A{}_k \to \tilde{\Lambda}_h{}^A{}_k = \Lambda_h{}^A{}_k + \partial_h \lambda_k^A - \partial_k \lambda_h^A. \tag{2.34}$$

III. CONNECTION 1-FORMS GENERATED BY 2-FORMS WITH VANISHING EXTERIOR COVARIANT DERIVATIVES

Let us now suppose that we are given a set of r type (1,0) 2-forms G^α, of which it is assumed that, relative to the given connection 1-forms A^α, their exterior covariant derivatives vanish:

$$DG^\alpha = 0. \tag{3.1}$$

Since $\{\theta_j\}$ is a basis of $\Lambda_P^3(X_4)$, it follows from the counterpart of (2.22) that this condition is equivalent to $D_k(*G^{\alpha lk}) = 0$, or, when written out in full in accordance with (2.23),

$$\partial_k(*G^{\alpha lk}) + C_\epsilon{}^\alpha{}_\beta A_k^\epsilon(*G^{\beta lk}) = 0, \tag{3.2}$$

in which the $*G^{\alpha lk}$ are the coefficients of the 2-forms $*G^\alpha$ relative to the dual bases of $\Lambda_P^2(X_4)$:

$$*G^\alpha = -\tfrac{1}{2}(*G_h{}^\alpha{}_k)dx^h \wedge dx^k. \tag{3.3}$$

Because of the skew symmetry of these coefficients in their superscripts, it is immediately evident that (3.2) implies the relation

$$\partial_l[C_\epsilon{}^\alpha{}_\beta A_k^\epsilon(*G^{\beta lk})] = 0,$$

which has a structure which is identical with that of (2.29). From the concluding remarks of Sec. II one may therefore infer that there exists, at least locally, a set of potentials $\Psi_h{}^\alpha{}_k$, in terms of which we have

$$C_\epsilon{}^\alpha{}_\beta A_k^\epsilon(*G^{\beta lk}) = -\tfrac{1}{2}i\epsilon^{lkhj}(\partial_k \Psi_h{}^\alpha{}_j), \tag{3.4}$$

this representation being invariant under the substitution

$$\Psi_h{}^\alpha{}_j \to \widetilde{\Psi}_h{}^\alpha{}_j = \Psi_h{}^\alpha{}_j + \partial_h \lambda_j^\alpha - \partial_j \lambda_h^\alpha, \tag{3.5}$$

in which the $\lambda_j^\alpha(x)$ are entirely arbitrary class C^2 functions. In terms of the dual $*\Psi^{\alpha hj}$ of $\Psi_h{}^\alpha{}_j$, defined in accordance with (2.19), we may express (3.4) in the form

$$C_\epsilon{}^\alpha{}_\beta A_k^\epsilon(*G^{\beta lk}) = -\partial_k(*\Psi^{\alpha lk}). \tag{3.6}$$

When this is substituted in the original Eq. (3.2), the latter is reduced to the form

$$\partial_k(*G^{\alpha lk} - *\Psi^{\alpha lk}) = 0,$$

which in turn possesses a structure which is identical with that of condition (2.25). By analogy with (2.27) one is therefore led to infer the local existence of a set of r 1-forms $B^\alpha = B_j^\alpha dx^j$, which give rise to the representation

$$(*G^{\alpha lk}) - (*\Psi^{\alpha lk}) = -i\epsilon^{lkhj}(\partial_h B_j^\alpha), \tag{3.7}$$

the latter being invariant under the substitution

$$B_j^\alpha \to \widetilde{B}_j^\alpha = B_j^\alpha - \partial_j \mu^\alpha, \tag{3.8}$$

in which the $\mu^\alpha(x)$ are arbitrary class C^2 functions.

Now, with the aid of (2.11) it is seen that (3.7) is equivalent to

$$\tfrac{1}{2}i(*G^{\alpha lk})\theta_{lk} = \tfrac{1}{2}i(*\Psi^{\alpha lk})\theta_{lk} + (\partial_h B_j^\alpha)dx^h \wedge dx^j,$$

to which we now apply the general formula (2.18) on either side to obtain

$$G^\alpha = dB^\alpha + \Psi^\alpha, \tag{3.9}$$

in which Ψ^α denotes the 2-form defined by

$$\Psi^\alpha = -\tfrac{1}{2}\Psi_h{}^\alpha{}_k\, dx^h \wedge dx^k. \tag{3.10}$$

It is therefore concluded that *any set of type* (1,0) *2-forms* G^α *with vanishing exterior covariant derivatives admits a decomposition such as* (3.9).

From (3.5) it is evident that (3.4) is unaffected by the substitution

$$\Psi^\alpha \to \widetilde{\Psi}^\alpha = \Psi^\alpha - d\lambda^\alpha, \tag{3.11}$$

where $\lambda^\alpha = \lambda_j^\alpha dx^j$. However, this is not the case for (3.7), or the equivalent representation thereof, namely (3.9). Clearly the latter requires that (3.11) be accompanied by the replacement

$$d\widetilde{B}^\alpha = dB^\alpha + d\lambda^\alpha,$$

so that (3.11) must be augmented by the substitution

$$\widetilde{B}^\alpha = B^\alpha + \lambda^\alpha + d\mu^\alpha, \tag{3.12}$$

which in turn also takes account of the lack of uniqueness displayed in (3.8).

The significance of the decomposition (3.9) is best understood in terms of an analysis of the transformation properties of its constituent parts. Since the 2-forms G^α are supposed to be of type (1,0), it follows from (2.2) that they transform as follows under the action of the group G:

$$G'^\alpha = G^\alpha + C_\epsilon{}^\alpha{}_\beta G^\beta u^\epsilon = G^\alpha - C_\epsilon{}^\alpha{}_\beta G^\epsilon u^\beta. \tag{3.13}$$

Hence

$$dG'^\alpha = dG^\alpha - C_\epsilon{}^\alpha{}_\beta dG^\epsilon u^\beta - C_\epsilon{}^\alpha{}_\beta G^\epsilon \wedge du^\beta,$$

and, because of (3.9), this determines the transform $d\Psi'^\alpha$ of $d\Psi^\alpha$. With the aid of (3.9) we may therefore write

$$d\Psi'^\alpha = d\Psi^\alpha - C_\epsilon{}^\alpha{}_\beta d\Psi^\epsilon u^\beta - C_\epsilon{}^\alpha{}_\beta dB^\epsilon \wedge du^\beta$$
$$- C_\epsilon{}^\alpha{}_\beta \Psi^\epsilon \wedge du^\beta$$
$$= d\Psi^\alpha - C_\epsilon{}^\alpha{}_\beta d(\Psi^\epsilon u^\beta) - C_\epsilon{}^\alpha{}_\beta d(B^\epsilon \wedge du^\beta)$$
$$= d[\Psi^\alpha - C_\epsilon{}^\alpha{}_\beta \Psi^\epsilon u^\beta - C_\epsilon{}^\alpha{}_\beta B^\epsilon \wedge du^\beta],$$

which may be integrated to yield

$$\Psi'^{\alpha} = \Psi^{\alpha} - C_{\epsilon}{}^{\alpha}{}_{\beta}\Psi^{\epsilon}u^{\beta} - C_{\epsilon}{}^{\alpha}{}_{\beta}B^{\epsilon}\wedge du^{\beta} + dW^{\alpha}, \quad (3.14)$$

where W^{α} is some 1-form. This relation clearly indicates the structure of the transformation law to be satisfied by the 2-forms Ψ^{α}. Turning now to the 1-forms B^{α}, we observe that the substitution of (3.9) in (3.13) gives

$$\Psi'^{\alpha} - \Psi^{\alpha} + C_{\epsilon}{}^{\alpha}{}_{\beta}\Psi^{\epsilon}u^{\beta} = d(B^{\alpha} - B'^{\alpha}) - C_{\epsilon}{}^{\alpha}{}_{\beta}dB^{\epsilon}u^{\beta}.$$

When this is compared with (3.14) it is seen that

$$d(B^{\alpha} - B'^{\alpha}) - C_{\epsilon}{}^{\alpha}{}_{\beta}dB^{\epsilon}u^{\beta} + C_{\epsilon}{}^{\alpha}{}_{\beta}B^{\epsilon}\wedge du^{\beta} = dW^{\alpha},$$

or

$$d[B^{\alpha} - B'^{\alpha} - C_{\epsilon}{}^{\alpha}{}_{\beta}B^{\epsilon}u^{\beta}] = dW^{\alpha},$$

so that

$$B'^{\alpha} = B^{\alpha} - C_{\epsilon}{}^{\alpha}{}_{\beta}B^{\epsilon}u^{\beta} - W^{\alpha} - dU^{\alpha}, \quad (3.15)$$

where U^{α} denotes a set of r 0-forms. This relation describes the type of transformation law which must be satisfied by the 1-forms B^{α}, Conversely, it is immediately evident that (3.14) and (3.15) together are sufficient to ensure that the combination $dB^{\alpha} + \Psi^{\alpha}$ satisfies the required transformation law (3.13) for arbitrary W^{α} and U^{α}.

The cause of the occurrence of the terms dW^{α} and dU^{α} in (3.14) and (3.15) is obviously related to the invariance of the decomposition (3.9) under the combined substitutions (3.11) and (3.12). In order to reduce the resulting indeterminacy, we shall now require that the 1-forms B^{α} which appear in (3.9) be connection 1-forms in the sense of the gauge transformation (2.14), that is, it is stipulated that under the action of the group G the B^{α} transform as follows

$$B'^{\alpha} = B^{\alpha} - C_{\epsilon}{}^{\alpha}{}_{\beta}B^{\epsilon}u^{\beta} - du^{\alpha}. \quad (3.16)$$

A comparison with (3.15) shows that this is tantamount to the requirement that

$$W^{\alpha} - dU^{\alpha} = du^{\alpha}. \quad (3.17)$$

This in turn implies that $dW^{\alpha} = 0$ and accordingly the transformation law (3.14) of the 2-forms Ψ^{α} is thus also completely specified. Moreover, if it is also required that the 1-forms $\widetilde{B}^{\alpha}$ in the substitution (3.12) satisfy a transformation law which is identical with (3.16), it is immediately evident that the 1-forms λ^{α} are of the type (1,0), while the 0-forms μ^{α} are subject to the extremely restrictive conditions

$$d\mu'^{\alpha} = d\mu^{\alpha} - C_{\epsilon}{}^{\alpha}{}_{\beta}(d\mu^{\epsilon})u^{\beta}. \quad (3.18)$$

It should be remarked that the μ^{α} do not affect the decomposition (3.9) in any way.

We can now summarize our conclusions in the following

Theorem 1: *Let there be given a set of r type* (1,0) 2-*forms* G^{α} *which satisfy the conditions* $DG^{\alpha} = 0$ *relative to a given connection. Then each* 2-*form* G^{α} *admits the decomposition* (3.9), *in which* B^{α} *represents a set of connection* 1-*forms, the latter being determined uniquely up to an additive term* $\lambda^{\alpha} + d\mu^{\alpha}$, *where* λ^{α} *denotes a set of type* (1,0) 1-*forms, while* μ^{α} *is subject to* (3.18).

IV. ANALYSIS OF THE DECOMPOSITION

The decomposition (3.9) of type (1,0) 2-forms with vanishing exterior covariant derivatives can be refined in a manner which significantly facilitates the application of such decompositions to special cases. In order to derive this refinement let us express the condition (3.1) in the form

$$dG^{\alpha} + \omega_{\beta}{}^{\alpha}\wedge G^{\beta} = 0, \quad (4.1)$$

where, as usual, we have put

$$\omega_{\beta}{}^{\alpha} = C_{\epsilon}{}^{\alpha}{}_{\beta}A^{\epsilon}. \quad (4.2)$$

For a specific set of connection 1-forms B^{α}, generated in accordance with the construction of Sec. III from (4.1), we introduce the 1-forms

$$\Pi_{\beta}{}^{\alpha} = \omega_{\beta}{}^{\alpha} - C_{\epsilon}{}^{\alpha}{}_{\beta}B^{\epsilon} = C_{\epsilon}{}^{\alpha}{}_{\beta}E^{\epsilon}, \quad (4.3)$$

where

$$E^{\alpha} = A^{\alpha} - B^{\alpha}. \quad (4.4)$$

From (2.4) and (3.16) it follows directly that the 1-forms E^{α} constitute a type (1,0) set; consequently the 1-forms $\Pi_{\beta}{}^{\alpha}$ are of type (1,1). In terms of the latter and the decomposition (3.9), we can now express (4.1) as

$$-d\Psi^{\alpha} = \Pi_{\beta}{}^{\alpha}\wedge dB^{\beta} + \Pi_{\beta}{}^{\alpha}\wedge\Psi^{\beta} + C_{\epsilon}{}^{\alpha}{}_{\beta}B^{\epsilon}\wedge dB^{\beta} + C_{\epsilon}{}^{\alpha}{}_{\beta}B^{\epsilon}\wedge\Psi^{\beta},$$

which is equivalent to

$$d[-\Psi^{\alpha} + \Pi_{\beta}{}^{\alpha}\wedge B^{\beta} + \tfrac{1}{2}C_{\epsilon}{}^{\alpha}{}_{\beta}B^{\epsilon}\wedge B^{\beta}] = d\Pi_{\beta}{}^{\alpha}\wedge B^{\beta} + \Pi_{\beta}{}^{\alpha}\wedge\Psi^{\beta} + C_{\epsilon}{}^{\alpha}{}_{\beta}B^{\epsilon}\wedge\Psi^{\beta}. \quad (4.5)$$

This suggests the introduction of the 2-forms

$$\Lambda^{\alpha} = \Psi^{\alpha} - \Pi_{\beta}{}^{\alpha}\wedge B^{\beta} - \tfrac{1}{2}C_{\epsilon}{}^{\alpha}{}_{\beta}B^{\epsilon}\wedge B^{\beta}, \quad (4.6)$$

so that (4.5) can be written in the form

$$\begin{aligned} &-d\Lambda^{\alpha} \\ &= \Pi_{\beta}{}^{\alpha}\wedge\Lambda^{\beta} + \tfrac{1}{2}C_{\epsilon}{}^{\beta}{}_{\gamma}\Pi_{\beta}{}^{\alpha}\wedge B^{\epsilon}\wedge B^{\gamma} + C_{\epsilon}{}^{\alpha}{}_{\beta}B^{\epsilon}\wedge\Psi^{\beta} \\ &+ (d\Pi_{\beta}{}^{\alpha} + \Pi_{\epsilon}{}^{\alpha}\wedge\Pi_{\beta}{}^{\epsilon})\wedge B^{\beta}. \end{aligned} \quad (4.7)$$

However, the Jacobi identities (2.1) give rise directly to the relation

$$C_{\epsilon}{}^{\alpha}{}_{\beta}C_{\gamma}{}^{\beta}{}_{\lambda}B^{\epsilon}\wedge B^{\gamma} = -\tfrac{1}{2}C_{\lambda}{}^{\alpha}{}_{\beta}C_{\epsilon}{}^{\beta}{}_{\gamma}B^{\epsilon}\wedge B^{\gamma}, \quad (4.8)$$

which holds identically for any set of 1-forms B^{α}. Thus it follows with the aid of (4.3) that

$$\begin{aligned} \tfrac{1}{2}C_{\epsilon}{}^{\beta}{}_{\gamma}\Pi_{\beta}{}^{\alpha}\wedge B^{\epsilon}\wedge B^{\gamma} &= \tfrac{1}{2}C_{\lambda}{}^{\alpha}{}_{\beta}C_{\epsilon}{}^{\beta}{}_{\gamma}E^{\lambda}\wedge B^{\epsilon}\wedge B^{\gamma} \\ &= -C_{\epsilon}{}^{\alpha}{}_{\beta}C_{\gamma}{}^{\beta}{}_{\lambda}E^{\lambda}\wedge B^{\epsilon}\wedge B^{\gamma} \\ &= -C_{\epsilon}{}^{\alpha}{}_{\beta}B^{\epsilon}\wedge\Pi_{\gamma}{}^{\beta}\wedge B^{\gamma}. \end{aligned} \quad (4.9)$$

This is substituted in (4.7) to give

$$-d\Lambda^{\alpha} = \Pi_{\beta}{}^{\alpha}\wedge\Lambda^{\beta} + C_{\epsilon}{}^{\alpha}{}_{\beta}B^{\epsilon}\wedge(\Psi^{\beta} - \Pi_{\gamma}{}^{\beta}\wedge B^{\gamma}) + (d\Pi_{\beta}{}^{\alpha} + \Pi_{\epsilon}{}^{\alpha}\wedge\Pi_{\beta}{}^{\epsilon})\wedge B^{\beta}. \quad (4.10)$$

But the relation (4.8) also implies the identity

$$C_{\epsilon}{}^{\alpha}{}_{\beta}C_{\gamma}{}^{\beta}{}_{\lambda}B^{\epsilon}\wedge B^{\gamma}\wedge B^{\lambda}=0, \tag{4.11}$$

and accordingly a term of this kind may be associated with the second expression on the right-hand side of (4.10). In view of (4.6) the latter then becomes

$$-dA^{\alpha}=\Pi_{\beta}{}^{\alpha}\wedge A^{\beta}+C_{\epsilon}{}^{\alpha}{}_{\beta}B^{\epsilon}\wedge A^{\beta}$$
$$+(d\Pi_{\beta}{}^{\alpha}+\Pi_{\epsilon}{}^{\alpha}\wedge\Pi_{\beta}{}^{\epsilon})\wedge B^{\epsilon}. \tag{4.12}$$

Moreover, as a direct consequence of (4.3) we have, again by virtue of the identity of (4.8),

$$d\Pi_{\beta}{}^{\alpha}+\Pi_{\epsilon}{}^{\alpha}\wedge\Pi_{\beta}{}^{\epsilon}=C_{\epsilon}{}^{\alpha}{}_{\beta}\mathbf{\Pi}^{\epsilon}, \tag{4.13}$$

where the 2-form $\mathbf{\Pi}^{\epsilon}$ is given by

$$\mathbf{\Pi}^{\alpha}=dE^{\alpha}+\tfrac{1}{2}C_{\epsilon}{}^{\alpha}{}_{\beta}E^{\epsilon}\wedge E^{\beta}. \tag{4.14}$$

This, together with (4.3), enables us to express (4.12) in the form

$$dA^{\alpha}+\omega_{\beta}{}^{\alpha}\wedge A^{\beta}+C_{\epsilon}{}^{\alpha}{}_{\beta}\mathbf{\Pi}^{\epsilon}\wedge B^{\beta}=0, \tag{4.15}$$

which is to be regarded as a system of differential equations for the 2-forms A^{α}, the coefficients of which are completely determined by the connection 1-forms A^{α} and B^{α}. In view of (4.6) this gives rise to the following

Theorem 2: The 2-forms Ψ^{α} which appear in the decomposition (3.9) *of the 2-forms G^{α} with vanishing exterior covariant derivatives can always be represented in the form*

$$\Psi^{\alpha}=\tfrac{1}{2}C_{\epsilon}{}^{\alpha}{}_{\beta}B^{\epsilon}\wedge B^{\beta}+\Pi_{\beta}{}^{\alpha}\wedge B^{\beta}+A^{\alpha}, \tag{4.16}$$

in which $\Pi_{\beta}{}^{\alpha}$ is given by (4.3) *while A^{α} denotes a set of 2-forms which satisfy the differential equations* (4.15).

One would expect that the integrability conditions of (4.15) are compatible with those of the initial hypothesis (4.1). In fact, we shall now verify that *the integrability conditions of* (4.1) *and* (4.15) *are identical.* To this end we take the exterior derivative of (4.15), which gives

$$C_{\epsilon}{}^{\alpha}{}_{\beta}\wedge A^{B}+C_{\epsilon}{}^{\alpha}{}_{\beta}d\mathbf{\Pi}^{\epsilon}\wedge B^{\beta}+C_{\epsilon}{}^{\alpha}{}_{\beta}\mathbf{\Pi}^{\epsilon}\wedge dB^{\beta}$$
$$+C_{\epsilon}{}^{\alpha}{}_{\lambda}\omega_{\beta}{}^{\alpha}\wedge\mathbf{\Pi}^{\epsilon}\wedge B^{\lambda}=0, \tag{4.17}$$

where F^{ϵ} denotes the curvature 2-form (2.6). But from (4.13) it follows that

$$C_{\epsilon}{}^{\alpha}{}_{\beta}d\mathbf{\Pi}^{\epsilon}=C_{\lambda}{}^{\alpha}{}_{\epsilon}\mathbf{\Pi}^{\lambda}\wedge\Pi_{\beta}{}^{\epsilon}-C_{\lambda}{}^{\epsilon}{}_{\beta}\mathbf{\Pi}^{\lambda}\wedge\Pi_{\epsilon}{}^{\alpha}, \tag{4.18}$$

while (4.3) and the identity (5.8) yield

$$C_{\epsilon}{}^{\beta}{}_{\lambda}\omega_{\beta}{}^{\alpha}\wedge\mathbf{\Pi}^{\epsilon}\wedge B^{\lambda}$$
$$=\tfrac{1}{2}C_{\epsilon}{}^{\alpha}{}_{\gamma}(C_{\mu}{}^{\lambda}{}_{\gamma}B^{\mu}\wedge B^{\gamma})\wedge\mathbf{\Pi}^{\epsilon}+C_{\lambda}{}^{\epsilon}{}_{\beta}\mathbf{\Pi}^{\lambda}\wedge\Pi_{\epsilon}{}^{\alpha}\wedge B^{\beta}.$$

This, together with (4.18) is substituted in (4.17), which can then be simplified to the form

$$C_{\epsilon}{}^{\alpha}{}_{\beta}F^{\epsilon}\wedge A^{\beta}+C_{\epsilon}{}^{\alpha}{}_{\beta}\mathbf{\Pi}^{\epsilon}\wedge(dB^{\beta}$$
$$+\tfrac{1}{2}C_{\lambda}{}^{\beta}{}_{\mu}B^{\lambda}\wedge B^{\mu}+\Pi_{\lambda}{}^{\beta}\wedge B^{\lambda})=0. \tag{4.19}$$

Now, with the aid of (4.14), (4.3), (2.6), and (4.4) it is easily verified that

$$\mathbf{\Pi}^{\alpha}=F^{\alpha}-(dB^{\alpha}+\tfrac{1}{2}C_{\epsilon}{}^{\alpha}{}_{\beta}B^{\epsilon}\wedge B^{\beta})-C_{\epsilon}{}^{\alpha}{}_{\beta}E^{\epsilon}\wedge B^{\beta},$$

so that

$$dB^{\alpha}+\tfrac{1}{2}C_{\epsilon}{}^{\alpha}{}_{\beta}B^{\epsilon}\wedge B^{\beta}+\Pi_{\beta}{}^{\alpha}\wedge B^{\beta}=F^{\alpha}-\mathbf{\Pi}^{\alpha}. \tag{4.20}$$

Thus the second term in (4.19) is simply

$$C_{\epsilon}{}^{\alpha}{}_{\beta}\mathbf{\Pi}^{\epsilon}\wedge(F^{\beta}-\mathbf{\Pi}^{\beta})=-C_{\epsilon}{}^{\alpha}{}_{\beta}F^{\epsilon}\wedge\mathbf{\Pi}^{\beta},$$

and accordingly (4.19) is reduced to

$$C_{\epsilon}{}^{\alpha}{}_{\beta}F^{\epsilon}\wedge(A^{\beta}-\mathbf{\Pi}^{\beta})=0, \tag{4.21}$$

which represents the integrability conditions of the system (4.15). But (4.20), taken in conjunction with (4.16), also implies that

$$F^{\alpha}-\mathbf{\Pi}^{\alpha}=dB^{\alpha}+\Psi^{\alpha}-A^{\alpha}=G^{\alpha}-A^{\alpha}, \tag{4.22}$$

where, in the second step, we have used (3.9). Accordingly, (4.21) is equivalent to

$$C_{\epsilon}{}^{\alpha}{}_{\beta}F^{\epsilon}\wedge G^{\beta}=0, \tag{4.23}$$

which is the integrability condition of (4.1), as asserted.

In conclusion, we observe that the differential equation (4.15) can be expressed in slightly more elegant form as follows. Since the 2-forms A^{α} which appear in (4.15) do not constitute a set of type (1,0) forms, we eliminate them from (4.15) in favor of the 2-forms

$$L^{\alpha}=A^{\alpha}+\Pi_{\beta}{}^{\alpha}\wedge B^{\beta}, \tag{4.24}$$

which, by virtue of (4.6), (3.14), (3.17), and (3.16) possess the required transformation properties. We may therefore construct the exterior covariant derivatives

$$DL^{\alpha}=dL^{\alpha}+\omega_{\beta}{}^{\alpha}\wedge L^{\beta}, \tag{4.25}$$

which, because of (4.24) and (4.3), can be expressed as

$$DL^{\alpha}=dA^{\alpha}+\omega_{\beta}{}^{\alpha}\wedge A^{\beta}+(d\Pi_{\beta}{}^{\alpha}+\Pi_{\epsilon}{}^{\alpha}\wedge\Pi_{\beta}{}^{\epsilon})\wedge B^{\beta}$$
$$-\Pi_{\beta}{}^{\alpha}\wedge dB^{\beta}+C_{\lambda}{}^{\alpha}{}_{\beta}B^{\lambda}\wedge\Pi_{\epsilon}{}^{\beta}\wedge B^{\epsilon}. \tag{4.26}$$

By means of (4.3) and (4.8) the last two terms on the right-hand side can be cast into the form

$$-\Pi_{\beta}{}^{\alpha}\wedge(dB^{\beta}+\tfrac{1}{2}C_{\epsilon}{}^{\alpha}{}_{\gamma}B^{\epsilon}\wedge B^{\gamma}),$$

after which (4.26) is substituted in (4.15), the relation (4.13) being taken into account at the same time. This yields

$$DL^{\alpha}+\Pi_{\beta}{}^{\alpha}\wedge(dB^{\beta}+\tfrac{1}{2}C_{\epsilon}{}^{\alpha}{}_{\gamma}B^{\epsilon}\wedge B^{\gamma})=0, \tag{4.27}$$

which is an alternative form of (4.15).

V. SPECIAL CASES AND AN APPLICATION TO THE YANG–MILLS EQUATIONS

We shall consider two special classes of solutions of the differential equation (4.27).

Case I: Because of the Bianchi identity (2.7), it is evident that (4.27) is satisfied if

$$L^{\alpha}=F^{\alpha} \tag{5.1}$$

and

$$dB^{\alpha}+\tfrac{1}{2}C_{\epsilon}{}^{\alpha}{}_{\beta}B^{\epsilon}\wedge B^{\beta}=0. \tag{5.2}$$

In view of (4.6) and (4.24) the identification (5.1) implies that

$$\Psi^{\alpha}=F^{\alpha}+\tfrac{1}{2}C_{\epsilon}{}^{\alpha}{}_{\beta}B^{\epsilon}\wedge B^{\beta},$$

so that, by (3.9) and (5.2),

$$G^\alpha = F^\alpha. \tag{5.3}$$

Condition (5.2) indicates that, under these circumstances, the curvature 2-forms associated with the connection 1-forms B^α vanish. Since this construction does not involve any restriction on the given connection, one may thus infer from (4.4) that *any connection* 1*-form can be expressed as the sum of a type* (1,0) 1*-form and a connection* 1*-form with zero curvature*. In passing we observe that nontrivial solutions of (5.2) can be constructed by means of systems of differential equations of the type

$$b^\alpha_\epsilon da^\epsilon_\beta - C_\epsilon{}^\alpha{}_\beta B^\epsilon = 0, \tag{5.4}$$

in which a^ϵ_β denotes a set r linearly independent type (0,1) 0-forms, while (b^α_ϵ) is the inverse of the matrix (a^ϵ_α). In fact, it is easily verified that the left-hand side of (5.4) represents type (1,1) 1-forms, and that the integrability conditions of (5.4) are satisfied whenever (5.2) holds.

Case II: A rather more stringent stipulation is the requirement that the 2-forms G^α be identical with the curvature 2-forms associated with the 2-forms B^α, that is,

$$G^\alpha = dB^\alpha + \tfrac{1}{2} C_\epsilon{}^\alpha{}_\beta B^\epsilon \wedge B^\beta. \tag{5.5}$$

By virtue of (4.16), (3.9), and (4.24) this is tantamount to

$$L^\alpha = 0,$$

and according to (4.27), (5.5), and (4.3) it is thus necessary that the conditions

$$\Pi_\beta{}^\alpha \wedge G^\beta = C_\epsilon{}^\alpha{}_\beta E^\epsilon \wedge G^\beta = 0 \tag{5.6}$$

be satisfied. If, relative to a canonical basis in $\Lambda^1_P(X_4)$ we put $E^\alpha = E^\alpha_j dx^j$,

these conditions can be expressed as

$$C_\epsilon{}^\alpha{}_\beta G_h{}^\beta{}_j E^\epsilon_l dx^h \wedge dx^j \wedge dx^l = 0,$$

or, if we use (2.9) and the counterpart of (2.19),

$$C_\epsilon{}^\alpha{}_\beta (* G^{\beta lk}) E^\epsilon_l = 0. \tag{5.7}$$

Clearly the left-hand side of (5.7) represents a system of $4r$ linear homogeneous expressions in the $4r$ components E^α_j. Thus, unless the determinant of this system vanishes, it follows that a necessary condition that (5.5) be valid is given by $E^\alpha = 0$. But, under these circumstances, it follows from (4.4) and (5.5) that $B^\alpha = A^\alpha$, and $G^\alpha = F^\alpha$.

This conclusion may be applied directly to the source-free Yang–Mills equations, which, as is well known, may be expressed in the form

$$D(* F^\alpha) = 0. \tag{5.8}$$

Thus, in this case the 2-forms G^α used above are to be identified with $* F^\alpha$; according to the theory of the previous sections the field equations (5.8) give rise to the existence of connection 1-forms B^α and hence to a representation of $* F^\alpha$ by means of (3.9) and (4.16), namely,

$$* F^\alpha = dB^\alpha + \tfrac{1}{2} C_\epsilon{}^\alpha{}_\beta B^\epsilon \wedge B^\beta + \Pi_\beta{}^\alpha \wedge B^\beta + \Lambda^\alpha, \tag{5.9}$$

where Λ^α is any solution of the differential equations (4.15). The condition (5.7) now assumes the form

$$(C_\epsilon{}^\alpha{}_\beta F^{\beta lk}) E^\epsilon_l = 0. \tag{5.10}$$

It is thus inferred that *unless the determinant of this system vanishes, the requirement that* $* F^\alpha$ *be identical with the curvature* 2*-form associated with the connection* 1*-forms* B^α *can be met if and only if* F^α *is self-dual.*

For the case of a 3-parameter Lie group G the equations (5.10) are equivalent to

$$\epsilon_{\alpha\beta\gamma} F^{\beta lk} E^\gamma_l = 0, \tag{5.11}$$

which is identical with a system of 12 linear homogeneous equations studied by Deser and Teitelboim,[8] who proved that, in general, the matrix of the system (5.11) is nonsingular. One may therefore conclude that the source-free Yang–Mills equations do not imply a representation of $* F^\alpha$ such as (5.5), at least for 3-parameter groups.

ACKNOWLEDGMENTS

It is a pleasure to thank Professor C. N. Yang for a valuable discussion. The writer is indebted also to his colleague, Professor W. E. Ferguson, for an algebraic reduction of the determinant of Eq. (5.11).

[1]This problem was raised by C.N. Yang in the course of a personal discussion with the writer.
[2]S. Deser and C. Teitelboim, Phys. Rev. D **13**, 1592 (1976). The writer is indebted to the referee for drawing his attention to this paper.
[3]Ref. 2.
[4]C.H. Gu and C.N. Yang, Scientia Sinica **20**, 47 (1977); D.X. Xia, Scientia Sinica **20**, 145 (1977).
[5]As regards the physical as well as mathematical motivation of the construction discussed here, reference is made to the survey article of C.N. Yang, *Gauge Fields*, Proceedings of the Sixth Hawaii Topical Conference on Particle Physics, edited by P.N. Dobson, Jr., S. Pakvasa, V.Z. Peterson, and S.F. Tuan (University of Hawaii at Manoa, Honolulu, 1975). A self-contained and detailed differential-geometric introduction is presented by W. Drechsler and M.E. Mayer, *Fiber Bundle Techniques in Gauge Theories*, Lecture Notes in Physics, Vol. 67 (Springer–Verlag, New York, 1977).
[6]Latin and Greek indices range from 1–4 and from 1–r, respectively. The summation convention is operative in respect of both sets of suffixes.
[7]These bases have been used with great effect by D.G.B. Edelen, Ann. Phys. **112**, 555 (1978). The construction of such bases is not, however, limited to spaces of four dimensions; H. Rund, Aeq. Math. **13**, 121 (1975).
[8]Ref. 2, Eq. (3.6). In this paper an explicit expression for the determinant of (5.10) due to T.T. Wu and C.N. Yang (private communication) is quoted.

Summation of divergent series by order dependent mappings: Application to the anharmonic oscillator and critical exponents in field theory

R. Seznec and J. Zinn-Justin

CEN—Saclay, Boite Postale No. 2, 91190 Gif-sur-Yvette, France
(Received 27 November 1978)

We study numerically a method of summation of divergent series based on an order dependent mapping. We consider the example of a simple integral, of the ground-state energy of the anharmonic oscillator and of the critical exponents of ϕ_3^4 field theory. In the case of the simple integral convergence can be rigorously proven, while in the other examples we can only give heuristic arguments to explain the properties of the transformed series. For the anharmonic oscillator we have compared our results to an accurate numerical solution (10^{-23}) of the Schrödinger equation. For the critical exponents we have verified the consistency of our results with those obtained before from methods using a Borel transformation.

1. INTRODUCTION

A classical approximation to the energy levels of the anharmonic oscillator whose Hamiltonian H is

$$H = \tfrac{1}{2}p^2 + \tfrac{1}{2}q^2 + \tfrac{1}{2}gq^4$$

is obtained by replacing H by a parameter dependent Hamiltonian H_0:

$$H_0 = \tfrac{1}{2}p^2 + \tfrac{1}{2}\omega^2 q^2.$$

The parameter ω^2 is then fixed by demanding that the first perturbative correction to the energy levels, $\langle (H - H_0)\rangle$, vanish. We have studied the generalization of this method to arbitrary orders. In perturbation theory one expands the energy levels up to order K and fixes the coefficient ω^2 of q^2 by demanding that the K th order term vanish.

This method will be shown to be equivalent to a method of order dependent mapping:

One starts from a Taylor series expansion of a function in powers of a variable g. One then sets

$$g = \rho F(\lambda),$$

where $F(\lambda)$ is a suitable analytic function satisfying

$$F(\lambda) = \lambda + O(\lambda^2).$$

One then transforms the expansion in powers of g in an expansion in powers of λ. The coefficients of the new series will now depend on an adjustable parameter ρ. The K th order approximation to the function we have expanded will now be constructed in the following way. The series expansion in λ is truncated at order K and ρ is taken as the zero of largest module of the K th order term.

The main problem is, of course, the choice of the function $F(\lambda)$, which is constrained by the analytic properties of the function which has been expanded. In the case of the anharmonic oscillator, the previous considerations lead to a natural choice of $F(\lambda)$.

We have applied such a method to the simple integral giving the number of Feynman diagrams (ϕ_0^4), the anharmonic oscillator (ϕ_1^4) and critical exponents in three dimensions (ϕ_3^4). We have analyzed the rate and domain of convergence, giving first heuristic arguments, and then analyzing the numerical results.

Using the same method, we have tried unsuccessfully to sum the perturbation series in a non-Borel summable case, the energy levels of the double well potential. The naïve sum of the series is, of course, complex, and indeed our method is convergent to such a complex result. In the case of a simple integral of similar structure, the real part of the complex result is the correct answer. This is not the case for the double-well potential.

The article is organized as follows: In Sec. 2 we explain the method on a simple integral, and discuss its convergence. In Sec. 3 we apply it to the ground state energy of the anharmonic oscillator. We have used the first 60 terms of the perturbative expansion. In Sec. 4, we analyze the same anharmonic oscillator with a different mapping. In Sec. 5, we sum the perturbative expansion of the critical exponents of the ϕ_3^4 field theory. Here only six or seven terms of the expansion are known. The results are therefore less accurate, but show very good consistency with those obtained by methods based on Borel transformations. In Sec. 6 we apply our method to the series of the ground state energy of the double-well potential. Section 7 contains our conclusions.

2. THE SIMPLE INTEGRAL

(A) Let us apply first this method to a simple integral representing the ϕ^4 theory in "zero" dimension:

$$Z(g) = \frac{1}{\sqrt{\pi}}\int_{-\infty}^{+\infty} e^{-(x^2+gx^4)}dx. \qquad (1)$$

The idea is to rewrite the integral in the following way:

$$Z(g) = \frac{1}{\sqrt{\pi}}\int e^{-\omega^2x^2 - [gx^4 + x^2 - \omega^2x^2]}dx, \qquad (2)$$

 0022-2488/79/071398-11$01.00

where ω is a free parameter. We then expand this integral,

$$Z(g) = \frac{1}{\sqrt{\pi}} \sum_{K=0}^{\infty} \frac{(-1)^K}{K!} \times \int e^{-\omega^2 x^2} (gx^4 + x^2 - \omega^2 x^2)^K dx, \tag{3}$$

and use ω^2 as a variational parameter to minimize the error on the truncated series: At order K, $Z(g)$ will be approximated by

$$Z(g) = \sum_{n=0}^{K} \frac{1}{\sqrt{\pi}} \frac{(-)^n}{n!} \int e^{-\omega^2 x^2} (gx^4 + x^2 - \omega^2 x^2)^n dx. \tag{4}$$

For every ω, these are the first K terms of an asymptotic expansion, and the error $\epsilon_K(g)$ is in magnitude of the order of the last term of the series:

$$\epsilon_K(g) \sim \frac{1}{K!} \frac{1}{\sqrt{\pi}} \int e^{-\omega^2 x^2} (gx^4 + x^2 - \omega^2 x^2)^K dx. \tag{5}$$

We shall choose ω to be the zero of smallest module of the last term of the series.

Let us first rewrite this term:

$$\epsilon_K(g) \sim \frac{1}{K!} \frac{1}{\sqrt{\pi}} \frac{(\omega^2 - 1)^K}{\omega^{2K+1}} \int (\rho x^4 - x^2)^K e^{-x^2} dx, \tag{6}$$

where we have defined

$$\rho = \frac{g}{\omega^2(\omega^2 - 1)}; \tag{7}$$

we shall make the ansatz that ρ is proportional to $1/K$ and evaluate this term by steepest descent. Let us set

$$\rho = 1/\mu K, \quad x^2 = Kt. \tag{8}$$

Then the integral becomes

$$\epsilon_K(g) \sim \frac{1}{K!} \frac{1}{\sqrt{\pi}} \frac{(\omega^2 - 1)^K}{\omega^{2K+1}} \left(\frac{K}{2}\right)^{K+1/2} \times \int \left(\frac{t^2}{\mu} - t\right)^K e^{-Kt} \frac{dt}{\sqrt{t}}. \tag{9}$$

The saddle points are given by the equation:

$$\frac{2t/\mu - 1}{t^2/\mu - t} = 1,$$

or (10)

$$t^2 - t(\mu + 2) + \mu = 0.$$

We see that there exist two saddle points, one with t in the interval $[0,\mu]$ and one with t larger than μ. The first saddle point gives a contribution with a constant sign, whereas the second gives an oscillating contribution. The K th term of the series vanishes when the two contributions cancel.

For one-half of the K's the zero will be real, and for the other half there will two complex conjugate zeroes with an imaginary part decreasing with K, and therefore negligible at larger order. The constant μ is given by

$$\exp[2(1 + \mu^2/4)^{1/2}] = \frac{(1 + \mu^2/4)^{1/2} + 1}{(1 + \mu^2/4)^{1/2} - 1}. \tag{11}$$

Solving this equation numerically yields

$$\mu = 1.325486837\cdots. \tag{12}$$

It is easy to verify that each of the saddle point contributions separately gives now an order of magnitude of the error. So one finds

$$\epsilon_K(g) \sim \left(\frac{\omega^2 - 1}{\omega^2}\right)^K (0.515\ldots)^K. \tag{13}$$

But

$$\omega^2 \sim \sqrt{g/\rho} \sim \sqrt{\mu g K}. \tag{14}$$

We finally have a rough estimate of the error at order K:

$$\epsilon_K(g) \sim (0.515\cdots)^K \exp(-0.896\sqrt{K/g}). \tag{15}$$

Of course, this estimate can be made more precise, and also transformed in a rigorous bound.

What one sees here is that the method converges for all values of g on the whole Riemann surface in a geometrical way.

Now a last remark is in order: Let us set

$$\lambda = \frac{\omega^2 - 1}{\omega^2}. \tag{16}$$

Then g is given by

$$g = \rho \frac{\lambda}{(1 - \lambda)^2}, \tag{17}$$

and the expansion now reads

$$Z(g) = \sqrt{(1 - \lambda)} \sum_{K=0}^{\infty} \lambda^K P_K(\rho). \tag{18}$$

So what we have made is a change of variable, depending on one parameter ρ, which gives a representation for $Z(g)$ which is only singular in the g plane on the interval $[-\rho/4,0]$, which shrinks to zero with K. As the $P_l(\rho)$ up to order K can be calculated from the first K terms of the expansion in powers of g, of $Z(g)$, we have transformed the divergent perturbative expansion, in a set of functions $Z_K(g)$ converging everywhere towards $Z(g)$.

Before going to the problem we are interested in, quantum mechanics, let us consider a slightly different integral.

(B) In the previous case the power series in g was Borel[1] summable, so that the function $Z(g)$ was uniquely defined by the series. Let us now consider the integral:

$$Z(g) = \int_{-\infty}^{+\infty} e^{x^2 - gx^4 - 1/4g} dx. \tag{19}$$

This function admits also an asymptotic expansion in powers of g but with terms all of the same sign, so that the series is not Borel summable, and does not define uniquely the function $Z(g)$:

$$Z(g) = \sum_{0}^{\infty} \sqrt{2} \frac{\Gamma(2K + \frac{1}{2})}{K!} g^K. \tag{20}$$

Of course, this function is the analytic continuation of the function defined in the first part. So it can be calculated by the method exposed there. (One has mainly to change ω^2 in $-\omega^2$.) But we want to use only this non-Borel summable asymptotic expansion.

Let us transform the integral by setting

$$t = (x^2 - 1/2g)\sqrt{g} \quad \text{for} \quad x > 0. \tag{21}$$

The intęgral then becomes

$$Z(g) = \sqrt{2} \int_{-1/2\sqrt{g}}^{+\infty} e^{-t^2} \frac{dt}{(1 + 2\sqrt{g}t)^{1/2}}. \tag{22}$$

For g positive we can then write

$$Z(g) = \mathrm{Re}[F(g + i0)] \tag{23}$$

with

$$F(g) = \sqrt{2} \int_{-\infty}^{+\infty} e^{-t^2} \frac{dt}{(1 + 2t\sqrt{g})^{1/2}}, \tag{24}$$

where $F(g)$ is analytic for $\mathrm{Im} g > 0$. The function $F(g)$ is actually a function which is analytic in the whole complex plane except on the real positive axis, and the expansion (20) is an asymptotic expansion for $F(g)$ in the whole analyticity domain. The Borel sum of this expansion is therefore $F(g)$. It is complex for $g > 0$, and the real part yields $Z(g)$. As it is easy now to verify that the expansion (20) is for g negative, indentical to the expansion for g small, and positive of the function defined by Eq. (1), the same summation method can be used.

As a conclusion, the expansion (20) is not Borel summable for the function $Z(g)$, but the additional information provided by Eq. (23) fixes the ambiguity in its resummation which can then be done by standard methods.

In quantum mechanics, a similar problem occurs with potentials with degenerate minima,[2] and we shall consider it in Sec. 6.

3. THE ANHARMONIC OSCILLATOR

A. The method

The extension of the method to the case of the anharmonic oscillator is straightforward. The Hamiltonian H is

$$H = \tfrac{1}{2}p^2 + \tfrac{1}{2}(q^2 + gq^4). \tag{25}$$

We shall first expand the ground state energy $E(g)$ in powers of g^3:

$$E(g) = \sum_0^\infty E_K g^K. \tag{26}$$

We shall now rewrite the Hamiltonian:

$$H = \tfrac{1}{2}p^2 + \tfrac{1}{2}\omega^2 q^2 + \tfrac{1}{2}[(1 - \omega^2)q^2 + gq^4]\theta \tag{27}$$

and expand the ground state energy in powers in θ, setting at the end $\theta = 1$.

We get now a new expansion, which from simple scaling considerations,

$$E(\omega,g,\theta) = [\omega^2 + \theta(1 - \omega^2)]^{1/2} E\left(\frac{\theta g}{[\omega^2 + \theta(1 - \omega^2)]^{3/2}}\right), \tag{28}$$

can be shown to have the form

$$E(g) = \sum_0^\infty \frac{(\omega^2 - 1)^K}{\omega^{2K-1}} P_K\left(\frac{g}{\omega(\omega^2 - 1)}\right). \tag{29}$$

Equation (29) shows that this expansion is just obtained from expansion (26) by algebraic manipulations. We now truncate this new asymptotic expansion at order K:

$$E_K(g) = \sum_{l=0}^{K} \frac{(\omega^2 - 1)^l}{\omega^{2l-1}} P_l\left(\frac{g}{\omega(\omega^2 - 1)}\right). \tag{30}$$

We fix ω by choosing it be the zero of smallest module of $P_K(g/\omega(\omega^2 - 1))$. At first order this is a well-known approximation for the anharmonic oscillator. One replaces in the Hamiltonian the operator q^4 by $3q^2\langle q^2\rangle$. It is not as easy to study the convergence of this approximation as in the case of the simple integral. We shall do it numerically by summing the first 60 terms of the perturbative expansion, and at the same time give heuristic arguments to explain why and where in the complex g plane it converges. Before doing it, let us define

$$\rho = g/\omega(\omega^2 - 1), \qquad \lambda = (\omega^2 - 1)/\omega^2. \tag{31}$$

Solving these equations yields

$$g = \frac{\rho\lambda}{(1 - \lambda)^{3/2}}, \tag{32}$$

and the expansion (29) becomes

$$E(g) = \frac{1}{\sqrt{1 - \lambda}} \sum_0^\infty \lambda^K P_K(\rho). \tag{33}$$

Again the method can be interpreted as an adjustable change of variable which regularizes the singularity in the g plane at $g = \infty$.

But, and this will be the main difference with the calculation of the previous section, the ground state energy of the anharmonic oscillator has additional singularities on the second sheet. As a result, the convergence will no longer be geometrical with the order. This we shall not prove rigorously, but we shall explain on a simplified example what happens.

B. A simplified example

We shall consider a function, analytic in a cut plane, having an asymptotic expansion in the cut plane, and an integral representation:

$$E(g) = \int_{-\infty}^{0} \frac{D(g')}{g' - g} dg'. \tag{34}$$

We shall assume that $D(g)$ decreases exponentially for g small, and is bounded for all g.

We shall now make the change of variable,

$$g = \frac{\rho\lambda}{(1 - \lambda)^2}, \tag{35}$$

and the corresponding change of variable in the integral,

$$g' = \frac{\rho\lambda'}{(1 - \lambda')^2}. \tag{36}$$

We then get

$$E(g(\lambda)) = -\oint \frac{D((\rho\lambda'/(1 - \lambda')^2)(1 - \lambda)^2}{(\lambda' - \lambda)(1 - \lambda\lambda')}$$

$$\times d\lambda' \frac{(1+\lambda')}{(1-\lambda')}. \tag{37}$$

If we now expand $E(g(\lambda))$ in powers of λ,

$$E(g(\lambda)) = \sum_0^\infty P_K(\rho)\lambda^K, \tag{38}$$

we get for $P_K(\rho)$, $K \geqslant 1$,

$$P_K(\rho) = \oint D\left(\frac{\rho\lambda'}{(1-\lambda')^2}\right)\frac{d\lambda'}{\lambda'}\left[\frac{1}{\lambda'^K} - \lambda'^K\right]. \tag{39}$$

The contour of integration runs on $[0, -1]$ and the half circle $\{|\lambda'| = 1, \operatorname{Im}\lambda' > 0\}$. One sees that $P_K(\rho)$ is the sum of the two contributions which, for K large, have a different behavior. The first one is sensitive to the small values of λ', the second one to the values of D on the circle $|\lambda'| = 1$.

We shall assume that $D(g)$ behaves for g small as:

$$D(g) \underset{g\to 0}{\sim} e^{1/ag}. \tag{40}$$

So the first contribution can be calculated by steepest descent for K large. If again we assume that the product $K\rho$ goes to a constant limit, then the saddle point equation is

$$\frac{\partial}{\partial\lambda'}\left[\frac{(1-\lambda')^2}{a\rho\lambda'} - K\ln\lambda'\right] = 0. \tag{41}$$

It is easy to see that the contribution of the saddle point is an increasing function of λ. If

$$aK\rho < C_1 \quad \text{for} \quad K \text{ large}, \tag{42}$$

the contribution of the saddle point decreases exponentially with K, and if ρ is larger, it increases exponentially.

The second contribution is bounded by a constant, and varies as a power of K if $D(g)$ has singularities, which we shall assume.

For K large on the other hand, at g fixed, λ tends towards one if ρ goes to zero:

$$\lambda \sim 1 - (\rho/g)^{1/2}. \tag{43}$$

So the factor λ^K is a decreasing function of ρ.

It is clear now that we should choose:

$$\rho \sim C_1/aK. \tag{44}$$

This is the region in which the two contributions to $P_K(\rho)$ are comparable, and therefore the domain of the zeroes of $P_K(\rho)$. For such a choice of ρ the error is then at order K of the order of $e^{-\sqrt{K/g}}$.

So if we adopt the following procedure—for each truncated sum

$$E_K(g(\lambda)) = \sum_{l=0}^{K} P_l(\rho)\lambda^l, \tag{45}$$

we take ρ as the zero of largest module of $P_K(\rho)$—then we have transformed the asymptotic series giving $E(g)$ in a convergent expansion in the whole complex plane.

In the case of the simple integral we got in addition an exponentially convergent expansion. The reason can be found in the fact that the corresponding discontinuity $D(g)$ had no singularity except at the origin, so that the second contribution could also be made exponentially small.

C. Application to the anharmonic oscillator

We have seen on the previous analysis that after a mapping made on a asymptotic series of the form

$$g = \rho\frac{\lambda}{(1-\lambda)^\alpha}. \tag{46}$$

The transformed series at large order is the sum of two type of contributions, one related to the large order behavior of the original series, and one related to the singularities of the function after mapping. For the ground state energy we know the location of the singularities of the second sheet qualitatively,[4] and we shall learn more from the actual convergence of the method. Also we know the large order behavior[5,6] of the asymptotic series, or the behavior of the discontinuity of $E(g)$ for g small and negative. We shall use it to explain some facts about the convergenc of the method. It has been shown that

$$E(g) = \frac{g}{\pi}\int_{-\infty}^{0} \frac{\operatorname{Im}[E(g')]}{g'(g'-g)}dg' + E_0, \tag{47}$$

$$\operatorname{Im}[E(g)] \sim \frac{1}{\sqrt{g}}e^{1/ag} \quad \text{and} \quad a = 3/2, \tag{48}$$

$g \to 0$.

We shall now replace g by $\rho\lambda/(1-\lambda)^{3/2}$ and make the change of variable in the integral:

$$g' = \rho\frac{\lambda'}{(1-\lambda')^{3/2}}. \tag{49}$$

Finally we shall consider only the contribution coming from the pole $\lambda = \lambda'$ in the denominator, and expand it in powers of λ. We then get a contribution $P_K^{(1)}(\rho)$,

$$P_K^{(1)}(\rho) = \frac{1}{\pi}\oint_C \operatorname{Im}\left[E\left(\frac{\rho\lambda'}{(1-\lambda')^{3/2}}\right)\right]\frac{d\lambda'}{\lambda'^{K+1}}. \tag{50}$$

The contour C is the image in the λ' plane of the negative real axis of the g plane. We shall calculate this integral by steepest descent, assuming that ρ will behave as $1/K$. The value λ_s' of λ' at the saddle point is given by

$$\frac{\partial}{\partial\lambda'}\left[\frac{(1-\lambda_s')^{3/2}}{a\rho\lambda_s'} - K\ln|\lambda_s'|\right] = 0. \tag{51}$$

The value of λ' at the saddle point depends only of the value of the product $a\rho K$. Depending on this value the contribution will increase or decrease exponentially with K. Again for K large the product $aK\rho$ will tend towards a limit such that

$$\frac{\partial}{\partial\lambda'}\left[\frac{(1-\lambda_s')^{3/2}}{a\rho\lambda_s'} - K\ln|\lambda_s'|\right] = 0,$$
$$\lim_{K\to\infty}\left[\frac{(1-\lambda_s')^{3/2}}{Ka\rho\lambda_s'} - \ln|\lambda_s'|\right] = 0. \tag{52}$$

A numerical calculation yields

$$\lim_{K\to\infty} aK\rho = 4.0312\ldots, \tag{53}$$

or, going back to the parameter ω,

$$\lim_{K\to\infty}\frac{\omega(\omega^2-1)}{aKg} = 0.248063\cdots. \tag{54}$$

As a result the factor λ^K which contains the entire g dependence behaves like

$$\lambda^K \underset{K\to\infty}{\sim} e^{-K(\rho/g)^{2/3}} \sim \exp\left(-C_2 \frac{K^{1/3}}{g^{2/3}}\right), \tag{55}$$

with a constant C_2,

$$C_2 = 1.933\ldots . \tag{56}$$

We shall see that the numerically calculated values of the zeros ρ_K of the polynomials $P_K(\rho)$ can be fitted by

$$(aK\rho_K)^{-1} = 0.248063 + C_3/K^{2/3}, \tag{57}$$

where C_3 can be estimated:

$$C_3 \simeq 1.7.$$

One can then calculate the general behavior of the contribution to the error $\epsilon(K,g)$ coming from the large orders

$$\epsilon(K) \sim \lambda^K \exp\left\{-K\left[\frac{(1-\lambda_s')^{3/2}}{\lambda_s'}\left(\frac{1}{aK\rho_K} - 0.2480\cdots\right)\right]\right\}, \tag{58}$$

where λ_s' is the value at the saddle point calculated before. We get

$$\epsilon(K) \sim \exp\left[-K^{1/3}\left(9.7 + \frac{1.933}{g^{2/3}}\right)\right]. \tag{59}$$

We shall compare later this result with the errors found in the actual calculation.

Assuming that our analysis is correct, we can find the domain of convergence of our method in the complex g plane. The error $\epsilon(K)$ will go to zero for K large as long as

$$\mathrm{Re}\left(1 + \frac{0.20}{g^{2/3}}\right) > 0. \tag{60}$$

Setting

$$g = |g| e^{i\varphi}, \tag{61}$$

we find the union of

$$-\frac{3\pi}{4} < \varphi < \frac{3\pi}{4}(3\pi) \quad \forall |g|, \tag{62}$$

and

$$|g| > 0.09(-\cos 2/3\varphi)^{2/3}. \tag{63}$$

We see therefore that the method will converge on the whole Riemann surface of the ground state energy of the anharmonic oscillator except in a neighborhood of the origin which contains additional singularities corresponding to level crossings.[4]

In particular the line $\varphi = 3\pi/2$ corresponds to the double well potential which has the Hamiltonian H:

$$H = \tfrac{1}{2}p^2 - \tfrac{1}{2}q^2 + \tfrac{1}{8}gq^4. \tag{64}$$

From the asymptotic series of the anharmonic oscillator we can therefore construct a convergent expanion to the ground state energy of the double well potential valid for

$$|g| > 0.09. \tag{65}$$

D. Numerical results[7]

We have used the first 60 terms of the perturbative expansion of the ground state energy of the anharmonic oscillator in our calculation. We have compared the results, with the value given by direct numerical solution of the Schrödinger equation. The Schrödinger equation was solved by transforming it in a finite difference equation. The ground state energy was obtained with a relative accuracy better than 10^{-23} for all values of the coupling constant g.

The first step was to find the zeroes of the polynomials

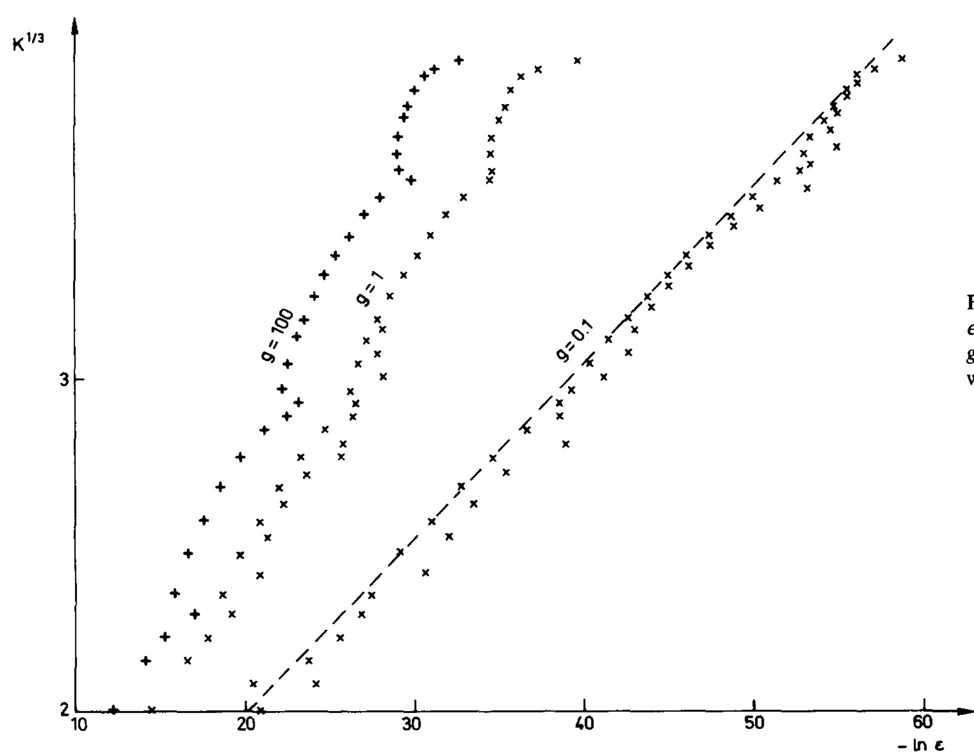

FIG. 1. The logarithm of the relative error $\epsilon(K,g)$ as a function of $K^{1/3}$ for three values of g: $g = 0.1$, $g = 1$, $g = 100$ for the mapping with exponent 3/2.

$P_K(\rho)$. For about half of the values of K, the corresponding zeroes ρ_K were real. As the set of ρ_K was a smooth function of K, at least for K large enough, we decided to fix the missing ρ_K by interpolation rather than by looking for complex zeroes. We just did a few checks to verify that such complex zeroes indeed exist in the neighborhood.

As a result we obtained a matrix of numbers $P_l(\rho_K)$. To calculate the K th order approximation of the energy level for a given value of g, we had then only to solve the equations

$$\omega(\omega^2 - 1) = g/\rho_K', \tag{66}$$

$$\lambda = (\omega^2 - 1)/\omega^2 \tag{67}$$

and to calculate the sum

$$E_K(g) = \frac{1}{\sqrt{1-\lambda}} \sum_0^K P_l(\rho_K)\lambda^l. \tag{68}$$

As mentioned before the ρ_K could be fitted reasonably well by the formula:

$$\rho_K^{-1} = K\left(0.3721 + \frac{2.6}{K^{2/3}}\right). \tag{69}$$

One should remember in addition that as the final result is formally independent of ρ, one does not need to estimate the ρ_K's with great accuracy.

It is not possible to give here the set of all $P_l(\rho_K)$, but we shall give as an example these numbers for $K = 5$. In this case the error on the result is exceptionally small for such a small value of K and a relative accuracy of 10^{-6} is obtained for all values of g, so that this approximation provides a good parametric representation of the energy level.

$$E(g) = \frac{1}{\sqrt{1-\lambda}}[0.5 - 0.1914104\lambda - 0.01992988\lambda^2 - 0.278699\times 10^{-2}\lambda^3 - 0.3246972\times 10^{-3}\lambda^4 + 0.5062714\times 10^{-5}\lambda^5] \tag{70}$$

and $g = \rho\lambda/(1-\lambda)^{3/2}$ with $\rho = 0.15623880$.

The results have been presented as follows: In Fig. 1 we have given the logarithm of the relative error $\epsilon(K)$ as a function of $K^{1/3}$. For K large we should get a straight line with a predicted slope. We have given the results for $g = 0.1, g = 1$, $g = 100$.

One sees that the slopes are in reasonable agreement with the theoretical prediction. Then we have fitted, at fixed K, $\epsilon(K,g)$ by

$$\epsilon(K,g) = \alpha(K) + \beta(K)/g^{2/3}. \tag{71}$$

Figure 2 shows $\alpha(K)$ and $\beta(K)$ as a function of $K^{1/3}$.

One sees on Fig. 1 that about half of the values of K give locally a bigger error than the others, for example 30, 32, 34, 36, 38, 40, 42, 44. On Fig. 2 we have kept only these values of K.

Again we should get, for K large enough, two straight lines with a predicted slope. We see that the results are in good agreement with the theoretical prediction.

To summarize, the relative error can be written approximately

$$-\ln(K,g) = 9.62K^{1/3} - 7.2 + \frac{2.1K^{1/3} - 1.8}{g^{2/3}}. \tag{72}$$

We see that the slopes are within 10% of the predicted values which is probably reasonable. If we use instead the predicted values and the highest values of K, we get instead

$$-\ln(K,g) = 9.7K^{1/3} - 7.6 + \frac{1.93K^{1/3} - 1.16}{g^{2/3}}. \tag{73}$$

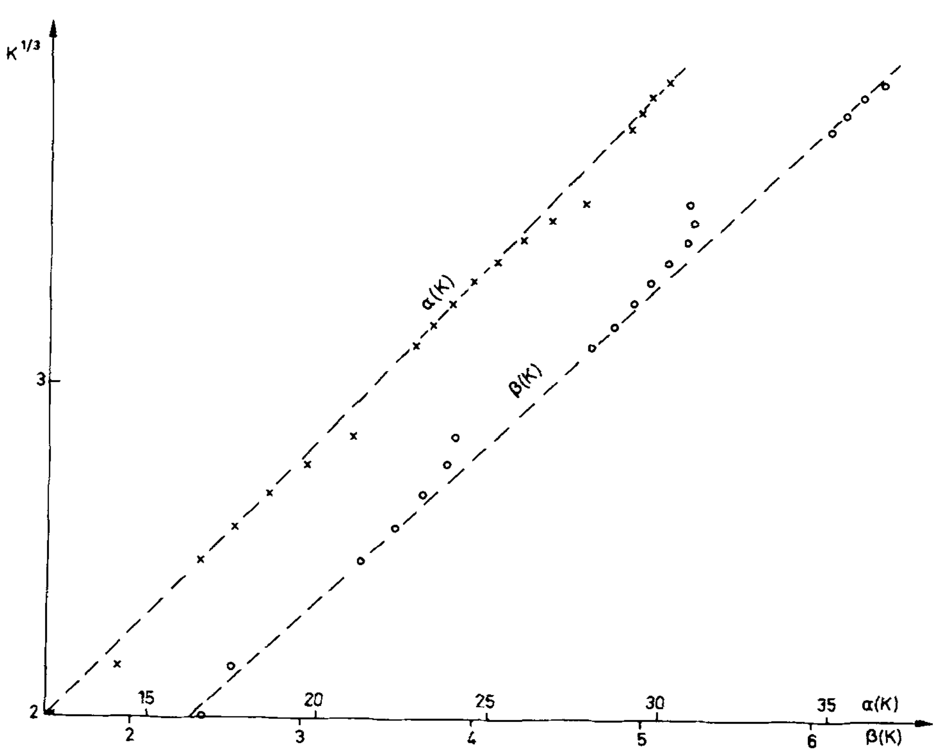

FIG. 2. The coefficients $\alpha(K)$ and $\beta(K)$ as functions of K in a fit: $-\ln[\epsilon(K,g)] = \alpha(K) + \beta(K)/g^{2/3} + \cdots$ for the mapping with exponent 3/2.

4. A MODIFIED SUMMATION METHOD FOR THE ANHARMONIC OSCILLATOR

A. The method

Once one has understood the relevant properties of the summation method, one may wonder if the particular variable mapping we have chosen is the best one. So we have explored another *a priori* reasonable mapping which also suppresses the singularity at $g = \infty$,

$$g = \rho \frac{\lambda}{(1-\lambda)^3}, \tag{74}$$

and we have written

$$E(g(\lambda)) = \frac{1}{(1-\lambda)} \sum_0^\infty \lambda^K Q_K(\rho). \tag{75}$$

From now on the method is completely similar, we look for the zeroes of the polynomials $Q_K(\rho)$, construct the set of quantities $Q_l(\rho_K)$ so that the K th order approximation will be given by

$$E_{(K)}(g) = \frac{1}{1-\lambda} \sum_0^K \lambda^l Q_l(\rho_K), \qquad g = \rho_K \frac{\lambda}{(1-\lambda)^3}. \tag{76}$$

B. Theoretical analysis of the convergence

Following the arguments presented in the previous section, if we assume[4,5] that the discontinuity of $E(g)$ on the negative real axis behaves like

$$\operatorname{Im}[E(g)] \underset{g\to 0}{\sim} e^{1/ag}, \tag{77}$$

then the asymptotic behavior of the ρ_K for K large will be given by the set of the two equations

$$\frac{\partial}{\partial\lambda}\left[\frac{(1-\lambda)^3}{a\rho\lambda} - K\ln|K|\right] = 0, \qquad \frac{(1-\lambda)^3}{a\rho\lambda} - K\ln|\lambda| = 0. \tag{78}$$

The numerical calculation shows that

$$\lim_{K\to\infty} aK\rho_K = 5.31686\cdots. \tag{79}$$

Then for K large, at g fixed, we can calculate λ as a function of g

$$\lambda \sim 1 - \left(\frac{5.316}{aKg}\right)^{1/3}, \tag{80}$$

and λ^K is approximately given by

$$\lambda^K \sim e^{-1.52K^{2/3}/(ag)^{1/3}}. \tag{81}$$

One contribution to the error $\epsilon(K,g)$ has then the form

$$\epsilon(K,g) \sim \epsilon_K e^{-1.52K^{2/3}/g^{1/3}}. \tag{82}$$

The natural domain of convergence of this approximation is therefore

$$\operatorname{Re} \frac{1}{g^{1/3}} > M \tag{83}$$

or, setting

$$g = |g| e^{i\varphi}, \tag{84}$$

we have

$$|g| < \left(\frac{\cos(\varphi/3)}{M}\right)^3, \quad |\varphi| < 3\pi/2. \tag{85}$$

From the known analyticity properties of the anharmonic oscillator,[4] we see that M cannot vanish, because then the analytic continuation of the ground state energy would be analytic up to

$$\arg(g) = 3\pi/2. \tag{86}$$

As a result, if the method converges, the error ϵ should behave as

$$\epsilon(K,g) \sim e^{(C_4 - 1.52/g^{1/3})K^{2/3}}. \tag{87}$$

Figure 3 shows a plot of $-\ln\epsilon$ as a function of $K^{2/3}$ for two values of g: $g = 0.1$ and $g = 10$. A linear fit of the variation in g yields coefficient of $K^{2/3}$ somewhat higher than predicted. To get a more precise result, we have again fitted, at K fixed, ϵ as a function of $g^{-1/3}$. We have then again a subset of values of K. From this we have extracted an expansion

$$\ln[\epsilon(K,g)] = \alpha(K) + \frac{\beta(K)}{g^{1/3}} + \frac{\gamma(K)}{g^{2/3}} + \cdots. \tag{88}$$

The three first coefficients have been plotted against $K^{2/3}$. The coefficients $\beta(K)$ and $\gamma(K)$ are roughly linear and the slope is

$$\beta(K) \sim 1.50K^{2/3}, \quad \gamma(K) \sim 0.08K^{2/3}, \tag{89}$$

while the predictions are

$$\beta(K) \sim 1.52K^{2/3}, \gamma(K) = O(K^{1/3}), \tag{90}$$

so that the agreement is reasonable.

For the coefficient $\alpha(K)$, two remarks can be made: It clearly becomes negative, for K of order 50–60, as expected. But it does not increase linearly, at the order we consider, with $K^{2/3}$. There are two possible reasons, either, as the corrections are of order $K^{1/3}$, we are not yet in the asymptotic regime, or $-\alpha(K)$ increases asymptotically faster than $K^{2/3}$. In the latter case, the method never converges, but is a new kind of asymptotic expansion, with an optimal value of K function of g. In any event, it seems safe to say that $-\alpha(K)$ will grow faster for K large than

$$-\alpha(K) \gtrsim 0.44K^{2/3}, \tag{91}$$

which shows that the method will not converge for

$$1.52 \ \operatorname{Re}(1/g^{1/3}) - 0.44 < 0, \tag{92}$$

which for g real positive yields

$$g \gtrsim 40. \tag{93}$$

For g large there is a best value of K which as for an asymptotic series gives the best estimate of $E(g)$. For example for $g = 700$, the optimal value is of the order of 30, and the relative error is then of the order of 4%.

If $\alpha(K)$ is for instance asymptotically linear in K, then the optimal K is of the form:

$$K_{0p} \sim 2.10^4/g \tag{94}$$

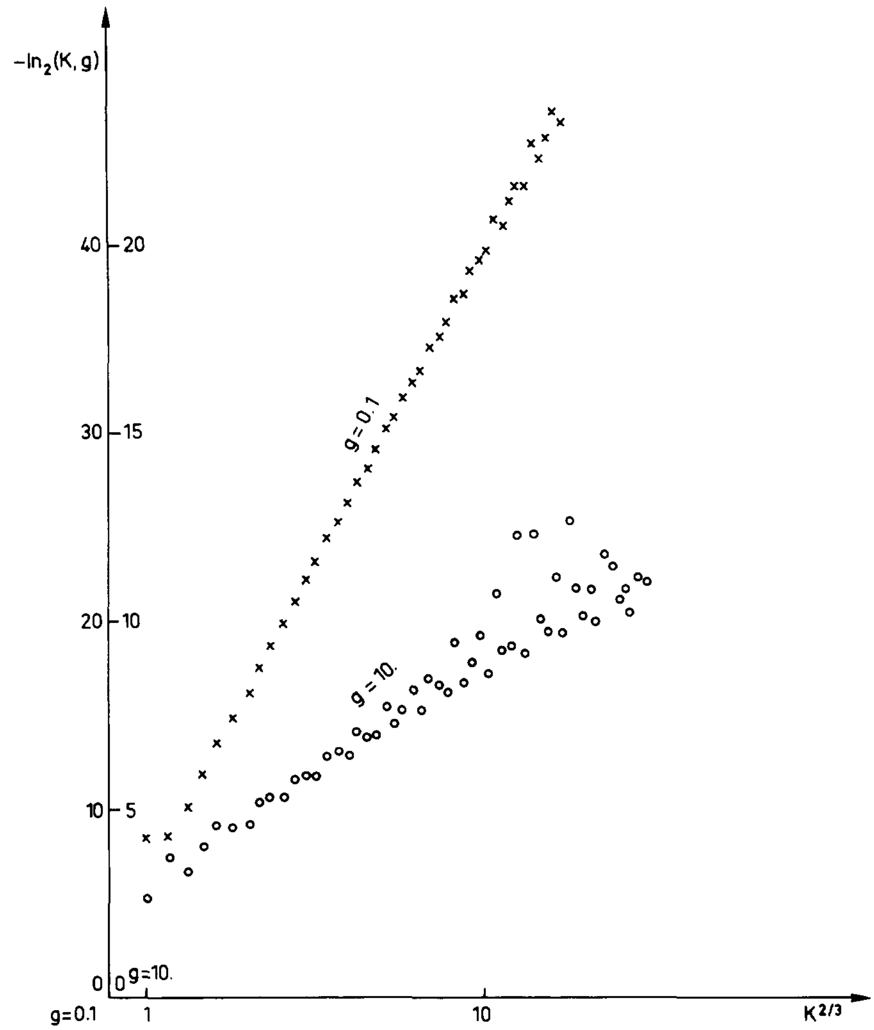

FIG. 3. The logarithm of the relative error $\epsilon(K,g)$ as a function $K^{2/3}$ for two values of g: $g = 0.1$, $g = 10$ for the mapping with exponent 3.

and the error at this value of K would be

$$\epsilon(g) \sim e^{-380/g}. \tag{95}$$

It is difficult from the numerical results to choose and a more detailed theoretical analysis is needed. As a last remark, we want to mention that the following mapping has also been tried:

$$g = \rho \frac{\lambda}{(1-\lambda)^2}. \tag{96}$$

The properties are similar and the convergence is slower for small g as the previous mapping, and better for large g, as expected.

5. FIELD THEORY

It is interesting to try to apply the same method on a field theoretical example. For ϕ^4_3, series expansion up to sixth or seventh order have been calculated[8] for the renormalization group functions. From these functions it is possible to compute critical exponents for ferromagnetic systems and fluid phase transitions.[9,10]

These series have been analyzed by various methods all based on a Borel transformation.[11,7] We shall try again to use a variable mapping method.

A. The method

The Euclidean action $A(\phi)$ in d dimension reads:

$$A(\phi) = \int d^d x [\tfrac{1}{2}(\partial_\mu \phi)^2 + \tfrac{1}{2} m_0^2 \phi^2 + (1/4!) g_0 \phi^4]. \tag{97}$$

The same scaling argument which was appropriate for the simple integral and the anharmonic oscillator would suggest the mapping

$$g_0 = \rho\lambda / (1-\lambda)^{2-d/2}. \tag{98}$$

But for $d = 2$ or 3, the theory needs a mass renormalization. As a result, the large g_0 behavior is not obtained from naïve scaling but from renormalization group arguments.

Defining a renormalized theory by the renormalization conditions

$$\Gamma^{(2)}(0) = 1,$$
$$\frac{\partial}{\partial p^2} \Gamma^{(2)}(0) = 1, \tag{99}$$
$$\Gamma^{(4)}(p_i = 0) = g,$$

where the $\Gamma^{(N)}_{(p_i)}$ are the renormalized 1PI correlation functions, one gets a relation between the bare coupling g_0 and the renormalized one, g:

$$W(g) = \frac{d-4}{d \ln g_0 / dg}, \tag{100}$$

where $W(g)$ is the usual C.S. coupling constant renormalization group function.[12] It is believed that $W(g)$ is a regular function of g in a range $[0, g_{\max}]$ in which it has an infrared stable zero g^*:

$$W(g^*) = 0, \quad \omega = W'(g^*) > 0. \tag{101}$$

The coupling constant $g = g^*$ corresponds to $g_0 = \infty$. If a function is regular in g in the neighborhood of g^*, it will have a regular expansion in powers of $1/(g_0)^{\omega/(4-d)}$ for g_0 large. One generally assumes that the various renormalization group functions are regular at g^*.

As a result for $d = 2$ or 3 one should perform, on the functions expressed in terms of the bare coupling constant g_0, the variable mapping

$$g_0 = \rho\lambda / (1 - \lambda)^{(4-d)/\omega}. \tag{102}$$

This is the natural generalization of the method explained before.

To calculate the critical exponents, we have to set $g_0 = \infty$, which corresponds to $\lambda = 1$. The natural domain of convergence of this method will be given by

$$\mathrm{Re}(g_0^{-\omega/(4-d)}) > M \tag{103}$$

and the rate of convergence will be:

$$\epsilon(K,g) \sim \exp[-K^{1-\omega/(4-d)}(-M + g_0^{-\omega/(4-d)})]. \tag{104}$$

Unfortunately, the numerical values of ω are only known with a poor accuracy[8,11]:

$$d = 2, \quad \omega \sim 1 \text{ to } 1.2, \quad d = 3, \quad \omega \sim 0.79. \tag{105}$$

So we shall have to make a reasonable guess for ω in $d = 3$, and calculate g^* and the critical exponents γ, ν, and η. We shall then compare the results with those given by other methods.[11]

To find a reasonable guess for ω, we have used the known seven terms of the expansion of $W(g)$. In addition the large order behavior calculations tell us the asymptotic behavior of the coefficients. Therefore, what we have done was to extrapolate the known coefficients with the help of their asymptotic behavior:

$$W(g) = \sum W_K g^K,$$
$$W_K \underset{K\to\infty}{\sim} C(-a)^K \Gamma(K+b+1). \tag{106}$$

The quantities W_K for $K \leqslant 7$, and $a,b,c,$[6,13] are known numerically.

We have first calculated ω itself as the function of an input ω, using the relation between g and g_0 derived from Eq. (100)

We have given in Table I the calculated values of ω and the variation of ω as a function of the initial value.

A plausible value of ω seems to be

$$\omega = 0.788. \tag{107}$$

We have calculated the other quantities with this value and indicate the variation of the result as a function of ω. Table II gives the valve for g^* Table III for η,γ, and ν. Again the large order behavior of the corresponding functions $\nu(g)$ and $\gamma(g)$ has been used to extrapolate two terms of the series, which is quite reasonable. (More than one or two terms are, of course, difficult to predict using the large order behavior.)

It is quite satisfactory that these results are in agreement with those obtained by other methods based on a Borel transformation.

TABLE I. The exponent ω at order K, calculated with an input value $\omega = 0.788$. The orders 7 and 8 have been obtained by extrapolation using the large order estimates. The variation of ω calculated when the input ω changes is $d\omega_{\text{cal}}/d\omega_i = -0.6$. The last value has been taken from Ref. 11.

K	ω_K
2	0.5535
3	0.7557
4	0.7121
5	0.7685
6	0.7605
7	0.7812
8	0.7773
	0.782 ± 0.010

TABLE II. The value of g^*, the renormalized coupling constant, as a function of the order K for $\omega_i = 0.788$. The orders 8 and 9 have been obtained by extrapolation, using the large order behavior estimates. $dg^*/d\omega_i = -0.40$. The last value has been taken from Ref. 11.

K	g^*
2	1.2690
3	1.4186
4	1.3902
5	1.4149
6	1.4095
7	1.4152
8	1.4139
9	1.4156
	1.414 ± 0.003

TABLE III. The exponents η, γ, ν as a function of the order K. The orders 7 and 8 are found by extrapolation $d\eta/d\omega_i = -0.04$, $d\gamma/d\omega_i = -0.12$, $d\nu/d\omega_i = -0.07$. The last values have been taken from Ref. 11.

K	η	γ	ν
2	—	1.20926	0.60814
3	0.0192	1.23861	0.62658
4	0.0292	1.23430	0.62409
5	0.0265	1.24002	0.62890
6	0.0306	1.23891	0.62803
7	0.0294	1.24046	0.62956
8	0.0310	1.24008	0.62923
	0.0315 ± 0.0025	1.2402 ± 0.0009	0.6300 ± 0.0008

In addition the rate of convergence is also comparable to other methods, and as can be seen from the tables γ, ν, and η do not vary too rapidly with the input value of ω, so that, for example, an uncertainty of 10^{-2} on ω translate in an uncertainty of order 10^{-3} only on γ.

It is also clear that we have chosen the simplest mapping with the right behavior at $\lambda = 1$. So it is possible that by finding better mappings one can further improve the results.

6. THE DOUBLE WELL POTENTIAL

As we have a simple and rather accurate method to sum divergent series it is tempting to try it on the double well potential whose Hamiltonian is now

$$H = \tfrac{1}{2}p^2 + \tfrac{1}{2}q^2(1 - q\sqrt{g})^2. \tag{108}$$

The problem here is that the series for the energy levels are not Borel summable.[2] As a consequence the levels are not defined without ambiguity by the series. A naïve resummation of the series gives a complex result. In the case of the corresponding simple integral of Sec. 1, we have shown that the real part of the naïve result gives the correct result. We have, of course, no reason to expect that such a simple procedure will also work here. In addition the sum of the series can only give a linear combination of the ground state energy and the first excited state, which for g small is the half sum. Indeed both energy levels $E_{\pm}(g)$ have the same asymptotic expansion:

$$E_+(g) = \sum_0^\infty E_K g^K, \quad E_-(g) = \sum_0^\infty E_K g^K, \tag{109}$$

where the equality here means just that it is an asymptotic expansion. The coefficients E_K can be shown to grow for large K as

$$E_K \underset{K\to\infty}{\sim} 3^K K!; \tag{110}$$

consequently, the series is ambiguous by terms of order the minimum at g fixed of $3^K g^K K!$, i.e., $e^{-1/3g}$. The difference between $E_+(g)$ and $E_-(g)$, $\Delta(g)$, is given by the barrier penetration coefficients between the two minima of the potential and can be shown to behave for g small as:

$$\Delta(g) = \frac{2}{\sqrt{\pi g}} e^{-1/6g}\left[1 - \frac{71}{1!(12)}g - \frac{6299}{2!(12)^2}g^2 - \frac{2691107}{3!(12)^3}g^3 + O(g^4)\right], \tag{111}$$

where the first term is known analytically and the others numerically. Therefore, $E_+(g)$ and $E_-(g)$ receive contributions, in addition to the series, of order $e^{-1/6}$ which are much bigger than the ambiguity of the series and have to be added to it

$$E_+(g) = \sum_0^\infty E_K g^K + \frac{e^{-1/6g}}{\sqrt{g}}\sum A_K g^K,$$
$$E_-(g) = \sum_0^\infty E_K g^K - \frac{e^{-1/6g}}{\sqrt{g}}\sum A_K g^K. \tag{112}$$

The half sum $[E_+(g) + E_-(g)]/2$ is given by the series up to order $e^{-1/3g}$.

We have again calculated the exact result by solving the Schrödinger equation with a 10^{-23} accuracy.

To sum the asymptotic series, we have considered the function

$$F(g) = \sum_0^\infty E_K(-g)^K. \tag{113}$$

This series can be shown to be Borel summable because it is related to the series of the ground state energy of the anharmonic oscillator with O(2) internal symmetry:

$$F(g) = \tfrac{1}{2}E_{O(2)}[2g]. \tag{114}$$

The function $F(g)$ has analytic properties very similar to the ground energy of the ordinary anharmonic oscillator. The function $F(g)f$ has therefore an analytic continuation to $g < 0$, where it is complex, with an imaginary part which for g small is of order $e^{-1/3g}$ and given by the large order behavior calculation[14]:

$$\operatorname{Im}F(g) = \frac{1}{g}e^{-1/3g}\left[1 - \frac{53}{6}g - \frac{1277}{2!(6)^2}g^2 - \frac{336437}{3!(6)^3}g^3 + O(g^4)\right]. \tag{115}$$

We have calculated numerically $F(g)$ using both mappings,

$$g = \rho\frac{\lambda}{(1-\lambda)^{3/2}} \quad\text{and}\quad g = \rho\frac{\lambda}{(1-\lambda)^3}, \tag{116}$$

because the first mapping does not converge for g small.

The calculation of the coefficients $P_l(\rho_K)$ goes exactly as before. Only in the solution of the equation

$$g = \rho\lambda/(1-\lambda)^\alpha \tag{117}$$

is the coupling g taken negative, so that the solution λ is now complex. For g small the mapping with $\alpha = 3$ gives the best result. For g larger the mapping with $\alpha = 3/2$ is preferable. This is expected from the error analysis done in the previous sections. For $\alpha = 3/2$ we have

$$\epsilon(K,g) \sim \exp[-K^{1/3}(11.7 + 1.218/g^{2/3})].$$

But g is now negative, so the method converges only for

$$|g| > 0.012.$$

On the other hand for $\alpha = 3$ the situation is different. For g negative, the real part of $g^{-1/3}$ is still positive, so that the numerical accuracy is best for g small. Nevertheless, the "convergence" is not as rapid as for g positive because

$$\operatorname{Re}(1/g^{1/3}) = 0.5/|g|^{1/3}.$$

We have then studied the function $\theta(g)$:

$$\theta(g) = \frac{E_+(g) + E_-(g)}{2} - \operatorname{Re}|F(-g)|.$$

Notice that the function $[F(-g) + \bar{F}(-\bar{g})]$ has analyticity properties very similar to $E_+(g)$ or $E_-(g)$.[15] Both functions have a set of singularities accumulating at the origin, tangentially to the imaginary axis.

Of course, the first result is that unfortunately $\theta(g)$ does not vanish identically. For g small, it behaves like

$$\theta(g) \underset{g\to 0}{\sim} \frac{A}{g^\alpha} e^{-1/3g}, \quad \alpha = 1.1 \pm 0.2, \quad A \simeq 2.2 \quad \text{for}$$
$\alpha = 1$,

as does the imaginary part of $E(-g)$. Also the ratio

$$\theta(g)/\mathrm{Im}[F(-g)]$$

seems to go to a constant for g small but not exponentially fast. It is a slow varying function, varying betwen 1 and 3, with superficially no remarkable properties. As a result, if we define the sum of the series by $\mathrm{Re}[F(-g)]$, some as yet unknown "instanton" contributions have to be added to it to get the correct result.

7. CONCLUSION

We have studied numerically a summation method for divergent series which does not involve a Borel transformation, and does not use Padé approximants. It is a simple generalization of the method generally used to continue a Taylor series outside of its circle of convergence. It only uses as basic information some known or guessed knowledge of the analytic properties of the function. Its main advantage is its extreme simplicity from the point of view of calculations, as it involves only algebraic manipulations. The calculations can therefore be done with high accurary. At a given order it produces a parametric representation of the function, so that to compute then the function for a given value of the argument involves only the solution of an algebraic equation. In the special case of anharmonic oscillator like problems, these mappings in addition appear naturally as a generalization of standard approximations, which explains why the convergence is so good.

Further analysis of the choice of optimal mappings for a given function, in connection with the theory of the convergence of such approximation would be interesting.

This method can, of course, also be used with functions analytic in a circle, especially when the radius of convergence is unknown, and with entire functions in regions where the argument is large, but the function does not increase maximally, so that there are big compensations between the terms of the Taylor series.

[1]E. Borel, *Leçon sur les séries divergentes* (Paris, 1928); G.H. Hardy, *Divergent Series* (Oxford U.P., Oxford, 1948).

[2]E. Brezin, G. Parisi, and J. Zinn-Justin, Phys. Rev. D **16**, 408 (1977).

[3]C. Bender and T.T. Wu, Phys. Rev. **184**, 1231 (1969).

[4]B. Simon, Ann. Phys. (N.Y.) **58**, 79 (1970); S. Graffi, V. Grecchi, and B. Simon, Phys. Lett. B **32**, 631 (1970)

[5]C. Bender and T.T. Wu, Phys. Rev. Lett. **27**, 461 (1971); Phys. Rev. D **7**, 1620 (1973).

[6]E. Brezin, J.C. Le Guillou, and J. Zinn-Justin, Phys. Rev. D **15**, 1544, 1558 (1977). See also: J. Zinn-Justin, *Salamanca 1977, Lecture Notes in Physics* 77 Number 126, edited by J.A. de Azcerraga (Springer-Verlag, Berlin, 1977); J. Zinn-Justin, *Carġese lectures* 1977 (Plenum, New York, 1977).

[7]For references about numerical calculations see F.T. Hioe and E.W. Montroll, J. Math. Phys. **16**, 1945 (1975).

[8]G. Baker, B.G. Nickel, and D.I. Meiron, Phys. Rev. B **17**, 1365 (1978).

[9]E. Brezin, J.C. LeGuillou, and J. Zinn-Justin, in *Phase Transitions and Critical Phenomena*, edited by C. Domb and M.S. Green (Academic, New York, 1976).

[10]G. Parisi, in Proceedings of the 1973 Cargèse Summer School (unpublished).

[11]J.C. Le Guillou and J. Zinn-Justin, Phys. Rev. Lett. **39**, 95 (1977), and in preparation.

[12]For a review see Ref. 9.

[13]E. Brezin and G. Parisi, Stat. Phys. (to be published).

[14]The value of the first term is derived in Ref. 2.

[15]Crutchfield II, W.Y., Phys. Lett. B **77**, 109 (1978).

On the sound field due to a moving source in a superfluid

J. C. Murray

Department of Mathematical Sciences, University of Petroleum and Minerals, Dhahran, Saudi Arabia
(Received 1 August 1978; revised manuscript received 6 November 1978)

Expressions are obtained for the sound pressure and temperature fields due to the motion of a monopole point source in a superfluid. The acoustic power spectrum is also given for all the Mach number ranges associated with the problem.

INTRODUCTION

The sound field induced by a monopole point source moving in an inviscid gas is well known in the literature of linear acoustics. Expressions can be obtained for any of the dynamic or thermodynamic variables by solving the inhomogeneous classical wave equation which governs the propagation of sound in the fluid. At low temperatures however the gas exhibits superfluidity properties and the wave equation is no longer adequate in describing the acoustic phenomenon. It must be replaced by two linear coupled equations involving two of the thermodynamic variables. Although solutions have been obtained to these equations when various methods of sound excitation are used (for example the sound field produced by a vibrating piston has been investigated by Lifshitz[1]) the solution associated with a moving monopole point source has not been established. It is the purpose of this paper to solve the equations in this case when appropriate boundary conditions are imposed at infinity. An expression is also obtained for the acoustic power for each Mach number range including the limiting classical range.

2. FORMULATION OF THE PROBLEM

The linearized equations governing the propagation of sound in a superfluid have the form,[1]

$$\frac{\partial\rho}{\partial t}+\nabla\cdot\mathbf{j}=Q, \tag{2.1}$$

$$\rho_e\frac{\partial s}{\partial t}+s_e\frac{\partial\rho}{\partial t}+\rho_e s_e\nabla\cdot\mathbf{v}_n=0, \tag{2.2}$$

$$\frac{\partial\mathbf{j}}{\partial t}+\nabla p=0, \tag{2.3}$$

$$\rho_e\frac{\partial\mathbf{v}_s}{\partial t}+\nabla p-\rho_e s_e\nabla T=0, \tag{2.4}$$

where $\mathbf{j}=\rho_{s_e}\mathbf{v}_s+\rho_{n_e}\mathbf{v}_n$. The normal and superfluid velocities are denoted respectively by $\mathbf{v}_n$ and $\mathbf{v}_s$. The quantities ρ, s, p, and T represent the small changes in density, entropy, pressure, and temperature from their constant equilibrium values ρ_e, s_e, p_e, and T_e. Also $\rho_e=\rho_{n_e}+\rho_{s_e}$ where ρ_{n_e} and ρ_{s_e} are the equilibrium values of normal and superfluid density. The quantity Q is a prescribed source term and all dependent variables are functions of position $\mathbf{r}$ $[\equiv(x,y,z)]$ and time t. The three-dimensional gradient and Laplacian operators will be denoted by ∇ and Δ respectively.

After eliminating $\mathbf{v}_n$ and $\mathbf{v}_s$ from Eqs. (2.1)–(2.4) and using the relations

$$\rho=\left(\frac{\partial\rho_e}{\partial p_e}\right)_{T_e}p+\left(\frac{\partial\rho_e}{\partial T_e}\right)_{p_e}T,$$

$$s=\left(\frac{\partial s_e}{\partial p_e}\right)_{T_e}p+\left(\frac{\partial s_e}{\partial T_e}\right)_{p_e}T$$

we obtain the following pair of equations for p and T:

$$\alpha\frac{\partial^2 p}{\partial t^2}-\Delta p-\gamma\frac{\partial^2 T}{\partial t^2}=\frac{\partial Q}{\partial t}, \tag{2.5}$$

$$\beta\frac{\partial^2 T}{\partial t^2}-\Delta T-\mu\frac{\partial^2 p}{\partial t^2}=0. \tag{2.6}$$

The positive coefficients α,β,γ,μ denote the constants:

$$\left(\frac{\partial\rho_e}{\partial p_e}\right)_{T_e},\quad \frac{\rho_{n_e}}{\rho_{s_e}s_e^2}\left(\frac{\partial s_e}{\partial T_e}\right)_{p_e},$$

$$-\left(\frac{\partial\rho_e}{\partial T_e}\right)_{p_e},\quad -\frac{\rho_{n_e}}{\rho_{s_e}s_e^2}\left(\frac{\partial s_e}{\partial p_e}\right)_{T_e}.$$

For a source moving with uniform velocity U along the positive x axis the sound field will be axially symmetric about the x axis and we can write

$$Q(\mathbf{r},t)=q(t)\delta(x-Ut)\frac{\delta(r)}{2\pi r}, \tag{2.7}$$

where $r=(y^2+z^2)^{1/2}$, δ represents the Dirac delta function, and $q(t)$ is a prescribed function of time. We will consider the special case of a simple harmonic source so that

$$q(t)=q_0\cos\omega_0 t. \tag{2.8}$$

To complete the formulation of the problem we require that p and T are sufficiently well behaved at infinity to justify all the mathematical operations employed throughout the analysis.

3. SOLUTION FOR p AND T

To solve for p and T it is convenient to set $\cos\omega_0 t=\frac{1}{2}[\exp(i\omega_0 t)+\exp(-i\omega_0 t)]$ and seek a solution of the form

$$p=p^++p^-,\qquad T=T^++T^-, \tag{3.1}$$

where

$$\alpha\frac{\partial^2 p^+}{\partial t^2}-\Delta p^+-\gamma\frac{\partial^2 T^+}{\partial t^2}=\frac{q_0}{4U\pi}\frac{\partial}{\partial t}[\exp(i\omega_0 t)\,\delta(t-x/U)], \tag{3.2}$$

 0022-2488/79/071409-04$01.00

$$\beta \frac{\partial^2 T^+}{\partial t^2} - \Delta T^+ - \mu \frac{\partial^2 p^+}{\partial t^2} = 0, \tag{3.3}$$

and

$$\alpha \frac{\partial^2 p^-}{\partial t^2} - \Delta p^- - \gamma \frac{\partial^2 T^-}{\partial t^2} = \frac{q_0}{4U\pi} \frac{\partial}{\partial t} [\exp(-i\omega_0 t)\, \delta(t - x/U)], \tag{3.4}$$

$$\beta \frac{\partial^2 T^-}{\partial t^2} - \Delta T^- - \mu \frac{\partial^2 p^-}{\partial t^2} = 0. \tag{3.5}$$

It is only necessary to solve (3.2) and (3.3), as the solution to (3.4) and (3.5) may be found by changing the sign of ω_0.

Introducing the complex Fourier transform pair defined by

$$\bar{f} = \frac{1}{2\pi} \int_{-\infty}^{+\infty} f e^{i\omega t}\, dt, \quad f = \int_{-\infty}^{+\infty} \bar{f} e^{-i\omega t}\, d\omega, \tag{3.6}$$

Eqs. (3.2) and (3.3) become

$$\alpha\omega^2 \bar{p}^+ + \Delta \bar{p}^+ - \gamma\omega^2 \bar{T}^+ = \frac{i\omega q_0}{8\pi^2 U} \exp\left(i \frac{x}{U}(\omega + \omega_0)\right) \frac{\delta(r)}{r}, \tag{3.7}$$

$$\beta\omega^2 \bar{T}^+ + \Delta \bar{T}^+ - \mu\omega^2 \bar{p}^+ = 0. \tag{3.8}$$

Using a Hankel transform pair defined by

$$*\bar{f} = \int_0^\infty \bar{f} r J_0(sr)\, dr, \quad \bar{f} = \int_0^\infty *\bar{f} s J_0(rs)\, ds, \tag{3.9}$$

where J_0 is Bessel's function of the first kind, we obtain

$$\frac{\partial^2 {*\bar{p}^+}}{\partial x^2} + (\alpha\omega^2 - s^2)\, {*\bar{p}^+} - \gamma\omega^2\, {*\bar{T}^+} = \frac{i\omega q_0}{8\pi^2 U} \exp\left(i \frac{x}{U}(\omega + \omega_0)\right), \tag{3.10}$$

$$\frac{\partial^2 {*\bar{T}^+}}{\partial x^2} + (\beta\omega^2 - s^2)\, {*\bar{T}^+} - \mu\omega^2\, {*\bar{p}^+} = 0. \tag{3.11}$$

It can readily be shown that the solution of Eqs. (3.10) and (3.11) has the form

$$*\bar{p}^+ = \frac{\exp[i(x/U)(\omega + \omega_0)]}{J} \cdot \frac{i\omega q_0}{8\pi^2 U} \left[\beta\omega^2 - s^2 - \left(\frac{\omega + \omega_0}{U}\right)^2\right], \tag{3.12}$$

$$*\bar{T}^+ = \frac{\exp[i(x/U)(\omega + \omega_0)]}{J} \cdot \frac{i\omega q_0}{8\pi^2 U} \mu\omega^2, \tag{3.13}$$

where

$$J = \left[\alpha\omega^2 - s^2 - \left(\frac{\omega + \omega_0}{U}\right)^2\right] \times \left[\beta\omega^2 - s^2 - \left(\frac{\omega + \omega_0}{U}\right)^2\right] - \mu\gamma\omega^4. \tag{3.14}$$

The two speeds of sound u_1 and u_2 ($u_1 > u_2$) in a superfluid are such that[1] $u_j^2, j = 1,2$ satisfy the equations

$$(\alpha\beta - \gamma\mu)(u_j^2)^2 - (\alpha + \beta)\, u_j^2 + 1 = 0, \quad j = 1,2, \tag{3.15}$$

so that

$$u_1^2 u_2^2 = \frac{1}{\alpha\beta - \gamma\mu}, \quad u_1^2 + u_2^2 = (\alpha + \beta)\, u_1^2 u_2^2. \tag{3.16}$$

Using the relations in (3.16) we find, after some algebra, that J can be written in the form

$$J = (s^2 - s_1^{+2})(s^2 - s_2^{+2}), \tag{3.17}$$

where

$$s_j^{+2} = \frac{\omega^2}{u_j^2} - \left(\frac{\omega + \omega_0}{U}\right)^2, \quad j = 1,2. \tag{3.18}$$

The right sides of Eqs. (3.12) and (3.13) can then be split into partial fractions so that

$$*\bar{p}^+ = \exp\left(i \frac{x}{U}(\omega + \omega_0)\right) \cdot \frac{i\omega q_0}{8\pi^2 U (u_1^2 - u_2^2)} \times \left(\frac{(1 - \beta u_1^2) u_2^2}{s^2 - s_1^{+2}} - \frac{(1 - \beta u_2^2) u_1^2}{s^2 - s_2^{+2}}\right), \tag{3.19}$$

$$*\bar{T}^+ = \exp\left(i \frac{x}{U}(\omega + \omega_0)\right) \cdot \frac{i\omega q_0}{8\pi^2 U} \mu \cdot \frac{u_1^2 u_2^2}{u_1^2 - u_2^2} \times \left(\frac{1}{s^2 - s_2^{+2}} - \frac{1}{s^2 - s_2^{+2}}\right). \tag{3.20}$$

Using the inversion formula given by (3.9) and the well known identity[2]

$$\int_0^\infty J_0(rs) \frac{s\,ds}{s^2 - b^2} \equiv \frac{\pi i}{2} H_0^{(1)}(br), \quad \mathrm{Im} b > 0, \tag{3.21}$$

where $H_0^{(1)}$ is the Hankel function, we obtain

$$\bar{p}^+ = \exp\left(i \frac{x}{U}(\omega + \omega_0)\right) \cdot \frac{q_0\omega}{16\pi U (u_1^2 - u_2^2)} \times \left((1 - \beta u_2^2) u_1^2 H_0^{(1)}(s_2^+ r) - (1 - \beta u_1^2) u_2^2 H_0^{(1)}(s_1^+ r)\right), \tag{3.22}$$

$$\bar{T}^+ = \exp\left(i \frac{x}{U}(\omega + \omega_0)\right) \cdot \frac{q_0\omega\mu}{16\pi U} \cdot \frac{u_1^2 u_2^2}{u_1^2 - u_2^2} \times [H_0^{(1)}(s_1^+ r) - H_0^{(1)}(s_2^+ r)]. \tag{3.23}$$

We note that $\bar{p}^+$ and $\bar{T}^+$ are defined by analytic continuation as analytic functions of ω provided $s_j^+, j = 1,2$, are defined as single-valued analytic functions.

The solutions for $\bar{p}^-$ and $\bar{T}^-$ are given by equations similar to (3.22) and (3.23) with ω_0 replaced by $-\omega_0$ and s_j^+, $j = 1,2$, replaced by $s_j^-, j = 1,2$, where

$$s_j^{2-} = \frac{\omega^2}{u_j^2} - \left(\frac{\omega - \omega_0}{U}\right)^2, \quad j = 1,2. \tag{3.24}$$

Using the inversion formula given by (3.6) the complete solution may be written in the form

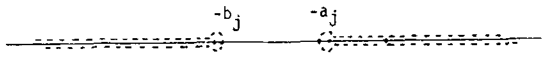

FIG. 1. Complex ω plane for s_j^+.

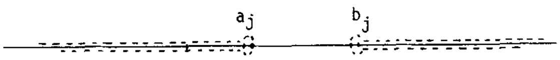

FIG. 2. Complex ω plane for s_j^-.

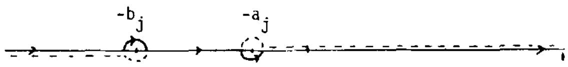

FIG. 3. Integration path for integrals involving s_j^+.

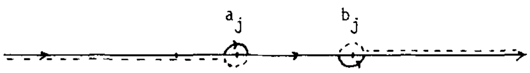

FIG. 4. Integration path for integrals involving s_j^-.

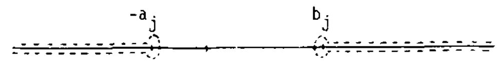

FIG. 5. Complex ω plane for s_j^+.

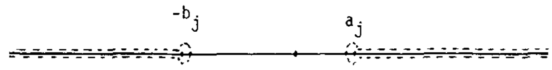

FIG. 6. Complex ω plane for s_j^-.

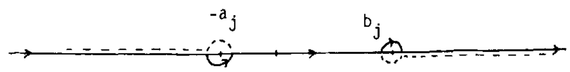

FIG. 7. Integration path for integrals involving s_j^+.

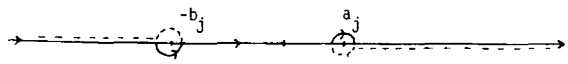

FIG. 8. Integration path for integrals involving s_j^-.

$$p = \frac{1}{(u_1^2 - u_2^2)}[(1-\beta u_2^2)u_1^2 I_2 - (1-\beta u_1^2)u_2^2 I_1], \qquad (3.25)$$

$$T = \frac{\mu u_1^2 u_2^2}{u_1^2 - u_2^2}(I_1 - I_2), \qquad (3.26)$$

where

$$I_j = \frac{q_0}{16\pi U} \times \left\{ \int_{-\infty}^{+\infty} \omega H_0^{(1)}(s_j^+ r) \exp[i(\omega+\omega_0)x/U - i\omega t\,]\, d\omega + \int_{-\infty}^{+\infty} \omega H_0^{(1)}(s_j^- r) \exp[i(\omega-\omega_0)x/U - i\omega t\,]\, d\omega \right\},$$

$$j = 1,2. \qquad (3.27)$$

The integration paths for the integrals in (3.27) depend on the branch of the multiple-valued functions s_j^+ and s_j^-. When M_j $(\equiv U/u_j) < 1$ we can write s_j^+ and s_j^- in the form

$$s_j^{\pm} = -i\frac{(1-M_j)^{1/2}}{U}(\omega \pm a_j)^{1/2}(\omega \pm b_j)^{1/2}, \qquad (3.28)$$

where

$$a_j = \frac{\omega_0}{1+M_j} \quad \text{and} \quad b_j = \frac{\omega_0}{|1-M_j|}, \quad a_j < b_j.$$

For each of the functions s_j^+ and s_j^-, the complex ω plane is cut in the manner depicted in Figs. 1 and 2 and s_j^+, s_j^- are analytic and single-valued outside the branch cuts. The integration paths for the integrals involving s_j^+ and s_j^- are shown in Figs. 3 and 4, respectively. On these integration paths the arguments of the Hankel functions are either real and positive or positive imaginary thus ensuring that the resulting wavefunctions have the correct behavior for large r.

When $M_j > 1$, then we can write

$$s_j^{\pm} = \frac{(M_j^2 - 1)^{1/2}}{U}(\omega \pm a_j)^{1/2}(\omega \mp b_j)^{1/2} \qquad (3.29)$$

and these functions are analytic and single-valued in the cut complex planes shown in Figs. 5 and 6. In this case the integration paths for the integrals involving s_j^+ and s_j^- are those shown in Figs. 7 and 8, respectively. Again, on these integration paths, the arguments of the Hankel functions are either real or positive or positive imaginary so that the resulting wavefunctions will have the correct behavior for large r.

The quantity I_j in Eqs,. (3.27) is identical [see, e.g., (3)] in form to that representing the pressure field due to a source moving in a normal fluid with velocity U and in which the speed of sound is u_j. In both the cases $M_j < 1$ and $M_j > 1$ the closed form equivalent is well known[3] and in the interests of brevity need not be written down. In any event the form of the solution for p and T given by Eqs. (3.25) and (3.26) is best suited to the calculation of the acoustic power spectrum which follows in the next section.

4. THE POWER SPECTRUM

The acoustic power emitted by the moving source is denoted by Π and is the time average over one period, $2\pi/\omega_0$, of $\Pi(t)$ where

$$\Pi(t) = \int_{\Sigma} p\mathbf{v}\cdot\mathbf{v}\, d\sigma, \qquad (4.1)$$

with Σ an infinitely large surface enclosing the line of motion of the source and $\mathbf{v}$ its outward unit normal. In what follows we will take Σ to be the infinitely long circular cylinder of radius r which is parallel to the x axis so that

$$\Pi(t) = \int_{-\infty}^{+\infty} 2\pi r p v_r\, dx, \qquad (4.2)$$

where v_r is the radial component of fluid velocity.

From Eqs. (2.3) and (2.4) and the relations $\mathbf{j} = \rho_{s_e}\mathbf{v}_s + \rho_{n_e}\mathbf{v}_n$, $\mathbf{v} = \mathbf{v}_n + \mathbf{v}_s$ we obtain

$$\frac{\partial}{\partial t}\mathbf{v} = s_e\left(\frac{\rho_{n_e} - \rho_{s_e}}{\rho_{n_e}}\right)\nabla T - \frac{2}{\rho_e}\nabla p. \qquad (4.3)$$

Using (3.25) and (3.26) together with (4.3) we obtain the radial component of velocity in the form

$$v_r = \frac{s_e(\rho_{n_e} - \rho_{s_e})}{\rho_{n_e}} \cdot \frac{\mu u_1^2 u_2^2}{u_1^2 - u_2^2}(K_1 - K_2) - \frac{2}{\rho_e(u_1^2 - u_2^2)}\left[(1 - \beta u_2^2)u_1^2 K_2 - (1 - \beta u_1^2)u_2^2 K_1\right] \tag{4.4}$$

$$K_j = \frac{q_0}{16\pi U i}\Big\{\int_{-\infty}^{+\infty} s_j'^{+} H_1^{(1)}(s_j'^{+} r) \times \exp[i(\omega' - \omega_0)x/U - i\omega' t]\, d\omega' + \int_{-\infty}^{+\infty} s_j'^{-} H_1^{(1)}(s_j'^{-} r) \times \exp[i(\omega' - \omega_0)x/U - i\omega' t]\, d\omega'\Big\}, \quad j = 1, 2, \tag{4.5}$$

where the functions $s_j'^{+}$, $s_j'^{-}$ are identical to s_j^{+}, s_j^{-} with ω replaced by ω'. Equation (4.2) may then be written

$$\Pi(t) = \frac{2\pi}{(u_1^2 - u_2^2)^2}\int_{-\infty}^{+\infty} r[(1 - \beta u_2^2)u_1^2 I_2 - (1 - \beta u_1^2) \times u_2^2 I_1]\Big(\frac{s_e(\rho_{n_e} - \rho_{s_e})}{\rho_{n_e}}\mu u_1^2 u_2^2(K_1 - K_2) - \frac{2}{\rho_e}[(1 - \beta u_2^2)u_1^2 K_2 - (1 - \beta u_1^2)u_2^2 K_1]\Big)\, dx. \tag{4.6}$$

The integrals in (4.6) which have integrands containing one of the factors $\exp[i(\omega + \omega' \pm 2\omega_0)x/U - i(\omega + \omega')t]$ will, after integrating with respect to x, reduce to integrals containing the factor $2\pi U\delta(\omega' + \omega \pm 2\omega_0)\exp[-i(\omega + \omega')t]$. After a further integration with respect to ω' (or ω) we obtain integrals which have the factor $\exp(\pm 2i\omega_0 t)$. The time average over one period of these integrals is zero. The remaining integrals have the form

$$\int_{-\infty}^{+\infty} r dx \int_{-\infty}^{+\infty}\int \omega H_0^{(1)}(s_j^{-} r)\, H_1^{(1)}(s_k'^{+} r) \times s_k'^{+} \exp[i(\omega + \omega')x/U - i(\omega + \omega')t]\, d\omega d\omega', \quad k, j = 1, 2, \tag{4.7}$$

or

$$\int_{-\infty}^{+\infty} r dx \int_{-\infty}^{+\infty}\int \omega H_0^{(1)}(s_j^{+} r)\, H_1^{(1)}(s_k'^{-} r) \times s_k'^{-} \exp[i(\omega + \omega')x/U - i(\omega + \omega')t]\, d\omega d\omega', \quad k, j = 1, 2, \tag{4.8}$$

For large r the Hankel functions in the integrands of (4.7) and (4.8) become exponentially small when the arguments of these functions are positive imaginary. We need integrate only over the ranges of ω and ω' for which the arguments are real and positive. In (4.7) we integrate first with respect to x to obtain $2\pi U\delta(\omega + \omega')$ in the integrand of the remaining double integral. After a subsequent integration with respect to ω' we obtain

$$2\pi U\int r\omega\, H_0^{(1)}(s_j^{-} r)\, H_1^{(1)}(s_k^{+} r)\, s_k'^{+}\, d\omega, \tag{4.9}$$

where ω' is replaced by $-\omega$ in $s_k'^{+}$. In fact $s_k'^{+} = -s_k^{-}$. Also the range of integration in (4.9) is such that s_j^{-} is positive. Using the asymptotic form of the Hankel functions in (4.9) the integral becomes

$$-4iU\int \omega\left(\frac{s_k^{-}}{s_j^{-}}\right)^{1/2}\exp[i(s_j^{-} - s_k^{-})r]\, d\omega. \tag{4.10}$$

Unless $j = k$ the integral in (4.10) tends to zero as $r \to \infty$. When $j = k$ its value is

$$-4iU\int \omega d\omega. \tag{4.11}$$

The integrals in (4.8) can be treated in a similar manner. The range of integration in (4.11) is determined by the Mach number range. When $M_j < 1$ the ranges for which the arguments of the appropriate Hankel functions are positive are $(-b_j, -a_j)$ and (a_j, b_j). In the case $M_j > 1$ the ranges are $(-\infty, -a_j)$, (b_j, ∞) and $(-\infty, -b_j)$, (a_j, ∞) and the integral in (4.11) has to be interpreted as

$$\lim_{\Omega \to \infty}\left(\int_{-\Omega}^{-a_j} + \int_{b_j}^{\Omega}\right) \quad \text{or} \quad \lim_{\Omega \to \infty}\left(\int_{-\Omega}^{-b_j} + \int_{a_j}^{\Omega}\right).$$

Evaluating the integrals in (4.6) in the manner shown above we obtain (for each of the Mach number ranges $M_1 < M_2 < 1$; $M_1 < 1, M_2 > 1$; $1 < M_1 < M_2$) an expression for Π of the form

$$\Pi = \frac{q_0^2\omega_0^2}{8\pi(u_1^2 - u_2^2)^2}\Big[\frac{s_e(\rho_{n_e} - \rho_{s_e})}{\rho_{n_e}}\mu\, u_1^2 u_2^2 \times\left(\frac{u_2^2(\beta u_1^2 - 1)}{u_1(1 - M_1^2)^2} + \frac{u_1^2(\beta u_2^2 - 1)}{u_2(1 - M_2^2)^2}\right) - \frac{2}{\rho_e}\left(\frac{u_2^4}{u_1}\frac{(\beta u_1^2 - 1)^2}{(1 - M_1^2)^2} + \frac{u_1^4}{u_2}\frac{(\beta u_2^2 - 1)^2}{(1 - M_2^2)^2}\right)\Big]. \tag{4.12}$$

We note that the expression for Π is formally the same for all Mach number ranges. Moreover, it is consistent with the classical result in the limit $\rho_{s_e} = 0$. When $\rho_{s_e} \to 0$ it can readily be shown that u_2^2, u_1^2, βu_2^2, μu_2^2 assume the values 0, c^2, 1, $-(\partial s_e/\partial p_e)_{T_e}/(\partial s_e/\partial T_e)_{p_e}$, respectively, where $c^2 = (\partial p_e/\partial \rho_e)_{s_e}$. We consider the Mach number ranges $M_1 < 1, M_2 > 1$; $1 < M_1 < M_2$ and let $M_2 \to \infty$ so that in the limit $\rho_{s_e} = 0$ Eq. (4.12) reduces to

$$\Pi = \frac{q_0^2\omega_0^2}{8\pi\rho_e c}\frac{1}{(1 - M^2)^2}, \tag{4.13}$$

where $M = U/c$. Equation (4.13) is the classical result and is valid for $M < 1$ or $M > 1$.

ACKNOWLEDGMENT

The author wishes to thank the referee for helpful comments on the first version of this paper.

[1]L.D. Landau and E.M. Lifshitz, *Fluid Mechanics* (Pergamon, New York, 1959), pp. 517–22.

[2]G.N. Watson, *Theory of Bessel Functions* (Cambridge U.P., Cambridge, 1966), p. 429.

[3]P.M. Morse and K.O. Ingard, *Theoretical Acoustics* (McGraw-Hill, New York, 1968), pp. 717–32.

Static conformally flat solution in a scalar-tensor theory of gravitation

D. R. K. Reddy

Department of Applied Mathematics, Andhra University, Waltair, India
(Received 14 September 1978)

Vacuum field equations for the static spherically symmetric conformally flat metric are obtained in a scalar-tensor theory of gravitation proposed by Ross. It is observed that unlike in Brans–Dicke scalar-tensor theory the only spherically symmetric static conformally flat vacuum solution in this theory is simply the empty flat space–time of Einstein's theory.

1. INTRODUCTION

Recently Ross[1] has constructed a scalar–tensor theory of gravitation using the Weyl formulation of Riemannian geometry. The scalar field, in this theory, is given an important geometrical role to play and is related to the integrable change in length of a vector as it is transported from point to point in space–time. This results in a Riemannian geometry, which appears quite different from general relativity, with modified covariant derivative and a metric tensor, which is not covariantly constant. It is pointed out that this theory is an alternative to the usual Brans–Dicke[2] formalism in regions of space free of mass and charge densities. The scalar field enters the two theories very differently.

The field equations given by Ross[1] for the combined scalar and tensor fields in the regions free of mass and charge densities are

$$S_{\pi\beta} - \tfrac{1}{2} g_{\pi\beta} S = 0$$

and

$$\phi^{;\alpha}_{;\alpha} = g^{\alpha\beta}\phi_{,\alpha,\beta} - g^{\alpha\beta}\phi_{,\pi}\begin{Bmatrix}\pi\\ \alpha\beta\end{Bmatrix}, - 4\phi_{,\alpha}\phi^{,\alpha} = 0, \tag{1}$$

where

$$\begin{aligned} S_{\pi\beta} &= R_{\pi\beta} - 2\phi_{,\pi,\beta} - 2\phi_{,\pi}\phi_{,\beta} + 2g_{\pi\beta}\phi^{,\alpha}\phi_{,\alpha} - g_{\pi\beta}g^{\gamma\alpha}\phi_{,\alpha,\gamma} \\ &\quad + 2\phi_{,\alpha}\begin{Bmatrix}\alpha\\ \pi\beta\end{Bmatrix} + g_{\pi\beta}g^{\gamma\alpha}\phi_{,\delta}\begin{Bmatrix}\delta\\ \alpha\gamma\end{Bmatrix} \end{aligned}$$

and

$$\phi^{,\alpha} = g^{\alpha\beta}\phi_{,\beta}.$$

A semicolon here denotes covariant differentiation, ϕ is the fundamental scalar field in the theory. $R_{\pi\beta}$ is the usual contracted Riemann curvature tensor and

$$\begin{Bmatrix}\delta\\ \alpha\gamma\end{Bmatrix}$$

is the Christoffel symbol of the second kind. By using an exact solution of the coupled nonlinear field equations for a static point mass, Ross[3] has shown that this theory, which is conformally equivalent to the empty space Einstein equations, predicts the same results for experiments as the usual theory of Brans and Dicke,[2] which has a nonzero energy–momentum tensor. Krori and Nandy[4] have shown that a sufficient condition for Birkhoff's theorem to be valid in this theory is that the scalar field is time invariant. Penney[5] and Gurses[6] have obtained exact solutions for massless scalar meson fields with a conformally flat metric while Ray[7] has given a complete set of exact solutions for both massive and massless scalar mesons in a conformally flat space–time. Recently, Reddy[8] has obtained exact solutions for both Brans–Dicke and Sen–Dunn Scalar–tensor fields in a static conformally flat space. The present note is an attempt to study the scalar–tensor theory of gravitation, proposed by Ross, in a static conformally flat space.

2. FIELD EQUATIONS AND SOLUTION

We consider the spherically symmetric conformally flat line element

$$ds^2 = e^u(dr^2 + r^2 d\theta^2 + r^2\sin^2\theta d\Phi^2 - dt^2), \tag{2}$$

where u is a function of r alone. For this space–time the nonvanishing components of the Einstein tensor are given by

$$\begin{aligned} G^1_1 &= e^{-u}\left(\frac{3}{4}u'^2 + \frac{2u'}{r}\right), \\ G^2_2 = G^3_3 &= e^{-u}\left(u'' + \frac{u'^2}{4} + \frac{u'}{r}\right), \\ G^4_4 &= e^{-u}\left(u'' + \frac{u'^2}{4} + \frac{2u'}{r}\right). \end{aligned} \tag{3}$$

Here a superscript prime indicates differentiation with respect to r.

Taking ϕ as a function of r only and using (3) and (2) in (1), the filed equations of Ross scalar–tensor theory of gravitation, in vacuum, can be written as

$$\begin{aligned} &\frac{3}{4}u'^2 + \frac{2u'}{r} - 3\phi'^2 + \phi'\left(3u' + \frac{4}{r}\right) = 0, \\ &u'' + \frac{u'^2}{4} + \frac{u'}{r} + 2\phi'' - \phi'^2 + \phi'\left(u' + \frac{2}{r}\right) = 0, \\ &u'' + \frac{u'^2}{4} + \frac{2u'}{r} + 2\phi'' - \phi'^2 + \phi'\left(u' + \frac{4}{r}\right) \\ &\qquad \phi'' + \phi'\left(u' + \frac{2}{r}\right) - 4\phi'^2 = 0. \end{aligned} \tag{4}$$

It can be easily verified that when the scalar field ϕ is a constant, the field equations yield a solution which describes an empty flat space time in Einstein's theory. Also, in fact, it

 0022-2488/79/071413-02$01.00

can be seen, from (4) that $u = \text{const}$ implies $\phi = \text{const}$, which is the trivial solution. This is easily verified by letting $u' = u'' = 0$ in (4) and solving for ϕ. More generally, the only solution of (4) is the trivial $u = \text{const}$ and $\phi = \text{const}$ solution which describes nothing but the flat space–time of general relativity. That is, a static conformally flat vacuum metric, in this theory, describes simply the empty flat space–time of general relativity. Thus, unlike in Brans–Dicke scalar–tensor theory,[8] the only spherically symmetric static conformally flat solution, in this theory, is simply the empty flat space–time of Einstein's theory. It is to be noted, in this context, that Einstein's equations have no conformally flat solution space–times other than the trivial one of the Minkowski metric.

3. CONCLUSIONS

Vacuum field equations in a scalar–tensor theory, formulated by Ross, are obtained for a static spherically symmetric conformally flat metric. It is observed that unlike in Brans–Dicke theory, the only spherically symmetric static conformally flat solution, in this theory, is simply the empty flat space–time of general relativity. This agrees with the result obtained by Reddy[8] in the scalar–tensor theory of gravitation proposed by Sen and Dunn.[9]

[1] D.K. Ross, Phys. Rev. D **5**, 284 (1972).
[2] C. Brans and R.H. Dicke, Phys. Rev. **124**, 925 (1961).
[3] D.K. Ross, Gen. Rel. Grav. **6**, 157 (1975).
[4] K.D. Krori and D. Nandy, J. Phys. A. Math. Gen. **10**, 993 (1977).
[5] R.V. Penney, Phys. Rev. D **14**, 910 (1975).
[6] M. Gurses, Phys. Rev. D **15**, 2731 (1977).
[7] Dipankar Ray, J. Math. Phys. **18**, 1899 (1977).
[8] D.R.K. Reddy, J. Math. Phys. (to be published).
[9] D.K. Sen and K.A. Dunn, J. Math. Phys. **12**, 578 (1971).

The algebra generated by Hodge's star and external differential

Jerzy F. Plebański[a)]

Centro de Investigación y de Estudios Avanzados del I.P.N., México, México D.F., 14—740

(Received 17 April 1978)

The algebra **A** of all secondary operators generated by Hodge's star and the external differential is investigated. It is established that the elements of **A** are spanned by **1**, $*d$, $*\delta$, $\nabla\!\!\!\!\nabla$, $*$, d, δ, $*\nabla\!\!\!\!\nabla$, where $\nabla\!\!\!\!\nabla := -d\delta + \delta d$, with the coefficients being formal power series in $\Delta = d\delta + \delta d$. The multiplication law of **A** and its representations are studied.

1. INTRODUCTION

We are working in this text basically on a finite dimensional real manifold M_n endowed with the nonsingular metric of the signature (n_+,n_-), $n_+ + n_- = n$, although all constructions and results will apply also *mutatis mutandi* for complex M_n's with nonsingular metric. The differential forms with which we are going to work are to be understood in both cases as complex-valued. Acting on the (complex) multiforms $\Lambda = \oplus \Lambda^P$, over a finite-dimensional manifold M, the external differential d and the normalized Hodge star, $\star$, have the two basic properties:

$$\sim : d^2 = 0, \quad \star^2 = \mathbb{1}, \tag{1.1}$$

where $\mathbb{1} \equiv$ identity over Λ. Differential geometry often employs the secondary concepts of the codifferential,

$$\delta := -i\star d\star \quad \rightarrow \quad \delta^2 = 0, \tag{1.2}$$

and of the Laplacian (Laplace–Beltrami operator),

$$\Delta := d\delta + \delta d. \tag{1.3}$$

(For the comparison of $\star$ and δ with the standard definitions see the Appendix.)

Apart from these, we have found recently[1] that it is useful to introduce the notion of the "anti-Laplacian" defined by

$$\not\Delta := -d\delta + \delta d. \tag{1.4}$$

Indeed, while Δ computes with d and $\star$:

$$[\Delta, d]^{(-)} = 0 = [\Delta, \star]^{(-)}, \tag{1.5}$$

the operator $\not\Delta$ anticommutes with d and $\star$:

$$[\not\Delta, d]^{(+)} = 0 = [\not\Delta, \star]^{(+)}. \tag{1.6}$$

(We use the notation

$$[a,b]^{(\pm)} := ab \pm ba.) \tag{1.7}$$

These simple facts were shown to be useful in applications in the theory of the Hertz potentials.[1]

It seems, therefore, of some interest to investigate more systematically *all* secondary operators generated by d and $\star$. That is the basic motivation of this paper.

For this purpose, we shall consider the algebra generated by the associative operations d and $\star$, i.e., the set of the formal series $\mathbb{A}$, with the coefficients in $\mathbb{C}$, of those products formed from d and $\star$ which should be considered independent modulo the equivalence class (1.1). Of course, the identity over Λ, $\mathbb{1}$, and the unity of $\mathbb{C}$, have to be identified within $\mathbb{A}$. The elements $\mathbb{1}$, d, $\star$, Δ, δ, $\not\Delta$ are elements of $\mathbb{A}$, and the equalities (1.5) and (1.6) are the equalities in the sense of this algebra.

A few comments about the basic consequences of (1.5) and (1.6). The operator Δ, commuting with the generators of $\mathbb{A}$, commutes also with any product of these operators and hence:

$$\mathbb{A} \ni a \rightarrow [f(\Delta), a]^{(-)} = 0, \tag{1.8}$$

where $f(\Delta)$ is any formal series in Δ. Similarly, if a^{odd} and a^{even} are those elements of $\mathbb{A}$ spanned respectively by products of the generators odd and even in the number of factors, the rule (1.6) implies

$$\mathbb{A} \ni a^{\text{odd}}, a^{\text{even}} \rightarrow [\not\Delta, a^{\text{odd}}]^{(+)} = 0 = [\not\Delta, a^{\text{even}}]^{(-)}. \tag{1.9}$$

Next, we observe the important fact that

$$\not\Delta^2 = (d\delta)^2 + (\delta d)^2 = \Delta^2. \tag{1.10}$$

Knowing this, a useful lemma immediately follows:

$$\mathscr{L}^+ : (\star d)^{2n} = (i\Delta)^{n-1}(\star d)^2, \quad n = 1,2,\cdots. \tag{1.11}$$

Indeed, from the definitions of Δ and $\not\Delta$

$$(\star d)^2 = \tfrac{1}{2}i(\Delta + \not\Delta), \quad (d\star)^2 = \tfrac{1}{2}i(\Delta - \not\Delta). \tag{1.12}$$

Consequently, employing $[\Delta, \not\Delta]^{(-)} = 0$, $\not\Delta^2 = \Delta^2$, we have

$$(\star d)^4 = \tfrac{1}{4}i^2(\Delta^2 + 2\Delta\not\Delta + \not\Delta^2) = \tfrac{1}{2}i^2\Delta(\Delta + \not\Delta) = i\Delta(\star d)^2, \tag{1.13}$$

and (1.11) follows thus by a trivial induction.

We observe now that according to the definitions of Δ and $\not\Delta$

$$\Delta + \not\Delta = 2\delta d, \quad \Delta - \not\Delta = 2d\delta. \tag{1.14}$$

Therefore, acting on these equalities from the right with the nilpotent d and δ respectively, we obtain at once

$$\not\Delta d = -\Delta d, \quad \not\Delta\delta = \Delta\delta. \tag{1.15}$$

It follows in particular that:

$$(\star d)^3 = \tfrac{1}{2}i(\Delta + \not\Delta)\star d = \tfrac{1}{2}i \star(\Delta - \not\Delta)d = i\Delta \star d. \tag{1.16}$$

Consequently, $\mathscr{L}^+$ implies also a general rule:

$$\mathscr{L}^- : (\star d)^{2n+1} = (i\Delta)^n(\star d), \quad n = 0,1,2,\cdots. \tag{1.17}$$

 0022-2488/79/071415-08$01.00

2. THE REDUCTION OF $a \in \mathbb{A}$ TO ITS CANONICAL FORM

The most general element of $\mathbb{A}$ can be written in the form of:

$$\mathbb{A} \ni a = \sum_{n=0} A_n(\star d)^n + \sum_{n=1} B_n \star(\star d)^n\star + \sum_{n=0} C_n(\star d)^n\star + \sum_{n=1} D_n \star(\star d)^n,$$

$$A_n, B_n, C_n, D_n \in \mathbb{C}. \tag{2.1}$$

Indeed, the terms $A_0\cdot\mathbb{1} + C_0\cdot\star + D_1\cdot d$, provide the entries of the basic generators; in the remaining terms, A-terms provide the products beginning with $\star$ and ending with d, B-terms describe the products beginning with d and ending with $\star$, C-terms include all products beginning and ending with $\star$, and finally, the D-terms secure the entries of all products beginning and ending with d. It is clear that in this way we exhaust all possibilities for the products among d's and stars restricted by the equivalence class (1.1).

We can now reduce (2.1) to its canonical form by employing our lemmas $\mathscr{L}^{\pm}$. Indeed, as the consequence of $\mathscr{L}^{\pm}$ we easily obtain:

$$\begin{aligned}
&(\star d)^{2n} = (i\Delta)^{n-1}\tfrac{1}{2}i(\Delta + \not\Delta), && (\star d)^{2n+1} = (i\Delta)^n(\star d),\\
&\star(\star d)^{2n}\star = (i\Delta)^{n-1}\tfrac{1}{2}i(\Delta - \not\Delta), && \star(\star d)^{2n+1}\star = (i\Delta)^n(d\star),\\
&(\star d)^{2n}\star = (i\Delta)^{n-1}\tfrac{1}{2}i\ \star(\Delta - \not\Delta), && (\star d)^{2n+1}\star = (i\Delta)^n(\star d\star),\\
&\star(\star d)^{2n} = (i\Delta)^{n-1}\tfrac{1}{2}i\ \star(\Delta + \not\Delta), && \star(\star d)^{2n+1} = (i\Delta)^n d.
\end{aligned} \tag{2.2}$$

The formulas from the first column are valid for $n = 1,2,\cdots$, and from the second column apply for $n = 0,1,\cdots$.

Distinguishing now in the series (2.1) the terms even and odd in n and substituting from (2.2), remembering that Δ commutes with all elements of $\mathbb{A}$, after ordering, we can represent the result in the form of

$$\mathbb{A} \ni a = \mathscr{P}^-\cdot d + \mathscr{Q}^-\cdot\delta + \mathscr{R}^-\cdot\star + \mathscr{S}^-\cdot\star\not\Delta + \mathscr{P}^+\cdot\star d + \mathscr{Q}^+\cdot\star\delta + \mathscr{R}^+\cdot\mathbb{1} + \mathscr{S}^+\cdot\not\Delta, \tag{2.3}$$

where $\mathscr{P}^{\pm}$, $\mathscr{Q}^{\pm}$, $\mathscr{R}^{\pm}$, $\mathscr{S}^{\pm}$ are all formal power series in Δ, all eight of them in general with nontrivial numeric leading term.

It is thus natural to categorize the possible elements of $\mathbb{A}$ as follows: First, we have the class of the algebraically neutral elements, which commute with all elements of $\mathbb{A}$, consisting of the formal power series in Δ:

$$\mathscr{N} \ni N = \sum_{k=0} n_k \Delta^k, \quad n_k \in \mathbb{C}. \tag{2.4}$$

We include, of course, in $\mathscr{N}$ the subclass of numbers, $\mathscr{N}_0 \ni n_0\cdot\mathbb{1}$, $n_0 \in \mathbb{C}$. Therefore, $\mathscr{N}$ forms the center of the algebra $\mathbb{A}$.

Thus, in the structure of $a \in \mathbb{A}$ only the seven elements d, δ, $\star$, $\star\not\Delta$, $\star d$, $\star\delta$, and $\not\Delta$ can participate as the algebraically active elements. Adjoining to them the element $\mathbb{1}$, we can say therefore that the most general element of $\mathbb{A}$ is spanned by the eight elements from the list:

$$\mathscr{L}': \mathbb{1}, \star d, \star\delta, \not\Delta, \star, d, \delta, \star\not\Delta, \tag{2.5}$$

with the coefficients in the algebraically neutral $\mathscr{N}$.

Notice that in the first line of (2.3) the entries have only the "odd" operators, d, $\delta = -i\star d\star$, $\star$, and $\star\not\Delta = -id\star d + i\star d\star d\star$; the second line contains only the "even" operators, $\star d$, $\star\delta = -id\star$, $\mathbb{1}$, and $\not\Delta = i[(d\star)^2 - (\star d)^2]$. The construction of all these operators involves at most no more then two d's and three stars.

The question now naturally arises concerning the structure of the multiplication table of the elements from our list $\mathscr{L}'$. In order to obtain this table in an esthetical form, it is convenient to take as the independent entries in the list some linear combinations of our eight operators.

Suppose that we define as the secondary operators:

$$\pmb{\not\Delta} := i\not\Delta, \quad \xi := \delta + d, \quad \eta := i(\delta - d). \tag{2.6}$$

In the terms of these, the rules (1.15) take now the form of

$$\pmb{\not\Delta}\xi = \Delta\eta, \quad \pmb{\not\Delta}\eta = -\Delta\xi, \tag{2.7}$$

and we easily see that

$$\begin{aligned}
\star\xi &= -id\star + \star d = \eta\star,\\
\star\eta &= d\star - i\star d = \xi\star.
\end{aligned} \tag{2.8}$$

It is thus obvious that $\star\xi$ and $\star\eta$ can replace in our list $\star d$ and $\star\delta = -id\star$. It will be convenient to introduce for these operators the symbols

$$\rho := \star\xi, \quad \sigma := \star\eta. \tag{2.9}$$

Clearly, then, $\star\rho = \xi = \delta + d$, $\star\sigma = \eta = i(\delta - d)$ can replace in our list d and δ. We arrive in this manner at the list of operators equivalent to (2.5):

$$\mathscr{L}: \begin{cases} \mathscr{L}^+: \mathbb{1}, \rho, \sigma, \pmb{\not\Delta},\\ \mathscr{L}^-: \star, \star\rho, \star\sigma, \star\pmb{\not\Delta}. \end{cases} \tag{2.10}$$

The sublist $\mathscr{L}^+$ contains "even" operators and the operators from $\mathscr{L}^-$ are all "odd."

The general element of the studied algebra can be now written in its final "canonical" form:

$$\mathbb{A} \ni a = (N + A\rho + B\sigma + C\pmb{\not\Delta}) + \star(\check{N} + \check{A}\rho + \check{B}\sigma + \check{C}\pmb{\not\Delta}), \tag{2.11}$$

where all coefficients $N,\check{N}$, $A,\check{A}$, $B,\check{B}$, and $C,\check{C}$ belong to $\mathscr{N}$.

The multiplication table for the eight elements from the list $\mathscr{L}$ is given in Table I. This table is to be understood, of course, in the sense of the products of the elements listed in the first line by the elements from the first column (from the left).

The structure of the table shows that our choice for the base of $a \in \mathbb{A}$ is indeed advantageous. First of all, we notice that all elements from the list $\mathscr{L}$ have simple squares: Six elements have algebraically neutral squares,

TABLE I.

$\mathbb{1}$	ρ	σ	$\not\Delta$	$\star$	$\star\rho$	$\star\sigma$	$\star\not\Delta$
ρ	$\not\Delta$	Δ	$-\Delta\cdot\sigma$	$\star\sigma$	$\Delta\cdot\star$	$-\star\not\Delta$	$\Delta\cdot\star\rho$
σ	Δ	$-\not\Delta$	$\Delta\cdot\rho$	$\star\rho$	$\star\not\Delta$	$\Delta\cdot\star$	$-\Delta\cdot\star\sigma$
$\not\Delta$	$-\Delta\cdot\sigma$	$\Delta\cdot\rho$	$-\Delta^2$	$-\star\not\Delta$	$\Delta\cdot\star\sigma$	$-\Delta\cdot\star\rho$	$\Delta^{2}\cdot\star$
$\star$	$\star\rho$	$\star\sigma$	$\star\not\Delta$	$\mathbb{1}$	ρ	σ	$\not\Delta$
$\star\rho$	$\star\not\Delta$	$\Delta\cdot\star$	$-\Delta\cdot\star\sigma$	σ	Δ	$-\not\Delta$	$\Delta\cdot\rho$
$\star\sigma$	$\Delta\cdot\star$	$-\star\not\Delta$	$\Delta\cdot\star\rho$	ρ	$\not\Delta$	Δ	$-\Delta\cdot\sigma$
$\star\not\Delta$	$-\Delta\star\sigma$	$\Delta\cdot\star\rho$	$-\Delta^{2}\cdot\star$	$-\not\Delta$	$\Delta\cdot\sigma$	$-\Delta\rho$	Δ^2

$$\mathbb{1}^2=\mathbb{1},\quad \not\Delta^2=-\Delta^2,\quad \star^2=\mathbb{1},\quad (\star\rho)^2=\Delta,$$
$$(\star\sigma)^2=\Delta,\quad (\star\not\Delta)^2=\Delta^2, \tag{2.12}$$

while the squares of the two remaining elements are still algebraically active,

$$\rho^2=\not\Delta,\quad \sigma^2=-\not\Delta. \tag{2.13}$$

It follows, however, that the fourth powers of these elements are already algebraically neutral:

$$\rho^4=-\Delta^2,\quad \sigma^4=-\Delta^2. \tag{2.14}$$

[Notice also $\rho^3=-\Delta\sigma, \sigma^3=-\Delta\rho$.] Another advantage of the arrangement of the elements in the list $\mathscr{L}$, is that they all reappear in the lines and columns of the multiplication table with algebraically neutral coefficients of six types only: ± 1, $\pm\Delta$, $\pm\Delta^2$. Our normalization of the elements in the list $\mathscr{L}$ assured the absence of complex coefficients in the table.

We can also observe the important fact: The upper left square of the table consisting of the products of the even elements with the even elements consists of the same even elements with the algebraically neutral coefficients. Moreover, the matrix of these products is symmetric. Thus, the even elements are spanning a conmutative subalgebra.

3. THE EXTENDED ALGEBRA $\mathbb{A}^{\text{ext}}$

In our multiplication table there appear as coefficients the algebraically neutral $\pm\Delta$ and $\pm\Delta^2$. A natural idea thus arises to "normalize" properly our operators from the list $\mathscr{L}$ in such a manner that these coefficients shall disappear. In order to be able, however, to proceed with such a normalization, we must assume additionally that there exists over $\Lambda\cdot$ the new operation

$$\Delta^{-1/2}, \tag{3.1}$$

which is endowed with the two basic properties: (1) it is algebraically passive, $a\in\mathbb{A}\rightarrow[\Delta^{-1/2},a]^{(-)}=0$, and (2) has the property that

$$\Delta\cdot(\Delta^{-1/2})^2=\mathbb{1}. \tag{3.2}$$

How to construct effectively such an operation for the usual applications in differential geometry is a problem which we leave open in this paper. It can be pointed out, however, that nowadays, by employing the theory of distributions, one can proceed with all rigor to divide operators by the square root of the d'Alembertain in quantum field theory. The standard trick consists in the use of Fourier transforms and properly interpreted principal values. Of course, equation of the type $\Delta(\Delta^{-1}\alpha)=\alpha$, $\alpha\in\Lambda$ remains true if we replace $\Delta^{-1}\alpha$ by $\Delta^{-1}\alpha+\alpha^{\text{harm}}$, $\Delta\alpha^{\text{harm}}=0$, and when working with the inverse powers of Δ one must remember the possible ambiguities related to the existence of harmonic (multi-harmonic) forms. In this paper dedicated to purely algebraic considerations, realizing the difficulties related to giving a rigorous meaning to the postulate of the existence of the operator $\Delta^{-1/2}$, we shall only explore the algebraic consequences of this postulate.

Assuming thus the existence of such an algebraically neutral $\Delta^{-1/2}$ that (3.2) is true, we extend the set of the neutral elements to:

$$\mathscr{N}^{\text{ext}}\ni N=\sum_{k=-\infty}\Delta^{k/2}n_k,\quad n_k\in\mathbb{C}, \tag{3.3}$$

with the series understood as a formal power series. Then we can understand, as the extension of our algebra $\mathbb{A}$, the algebra $\mathbb{A}^{\text{ext}}$, whose general element has the form of (2.11), with the algebraically neutral coefficients in $\mathscr{N}^{\text{ext}}$. Thus, we construct the natural base for $\mathbb{A}^{\text{ext}}$ by defining:

$$\alpha:=\Delta^{-1/2}\rho,\quad \beta:=\Delta^{-1/2}\sigma,\quad \gamma:=\Delta^{-1}\not\Delta, \tag{3.4}$$

and we have the typical element of $\mathbb{A}^{\text{ext}}$ in the form of

$$\mathbb{A}^{\text{ext}}\ni a=(N+A\alpha+B\beta+C\gamma)+\star(\check{N}+\check{A}\alpha+\check{B}\beta+\check{C}\gamma),$$
$$N,\check{N},A,\check{A},B,\check{B},C,\check{C},\in\mathscr{N}^{\text{ext}}. \tag{3.5}$$

The multiplication table of the base elements is now the properly reduced form of Table I, which we give in Table II.

We immediately recognize from the table that the elements from the "Abelian corner", $(\mathbb{1},\alpha,\beta,\gamma)$, normalized according to $(\mathbb{1},\alpha e^{-i\pi/4},\beta e^{i\pi/4},-i\gamma)$, amount to the cyclic group C_4, i.e., to the group of the roots of the fourth order of unity. [The products among $(\mathbb{1},\alpha e^{-i\pi/4},\beta e^{i\pi/4},-i\gamma)$, agree with the products among $(1,i,-i,-1)$.] The whole struc-

TABLE II.

$\mathbb{1}$	α	β	γ	$\star$	$\star\alpha$	$\star\beta$	$\star\gamma$
α	γ	$\mathbb{1}$	$-\beta$	$\star\beta$	$\star$	$-\star\gamma$	$\star\alpha$
β	$\mathbb{1}$	$-\gamma$	α	$\star\alpha$	$\star\gamma$	$\star$	$-\star\beta$
γ	$-\beta$	α	$-\mathbb{1}$	$-\star\gamma$	$\star\beta$	$-\star\alpha$	$\star$
$\star$	$\star\alpha$	$\star\beta$	$\star\gamma$	$\mathbb{1}$	α	β	γ
$\star\alpha$	$\star\gamma$	$\star$	$-\star\beta$	β	$\mathbb{1}$	$-\gamma$	α
$\star\beta$	$\star$	$-\star\gamma$	$\star\alpha$	α	γ	$\mathbb{1}$	$-\beta$
$\star\gamma$	$-\star\beta$	$\star\alpha$	$-\star$	$-\gamma$	β	$-\alpha$	$\mathbb{1}$

ture of the table exhibits the fact that all eight elements constitute a projective representation of a group.[2] Indeed, by adjoining to the list of elements $(\mathbb{1},...,\star\gamma)$ the list $(-\mathbb{1},...,-\star\gamma)$ it is evident that the 16×16 multiplication table forms a group. The quotient of this group by its center spanned by $\mathbb{1}$, $-\mathbb{1}$ is obviously the semidirect product of the two cyclic groups:

$$G = C_4 \circledS C_2. \tag{3.6}$$

From the point of view of a physicist, the interesting aspect of the algebra $\mathrm{A}^{\mathrm{ext}}$ consists in its matrix representations. From the fact that our multiplication table is the projective representation of the group (3.6), with C_4 having one-dimensional representations, one easily infers that the lowest-dimensional *faithful* representation of the algebra $\mathrm{A}^{\mathrm{ext}}$ must be four-dimensional.

In order to construct such a representation, consider the two copies of the ring of the Pauli operators $(\mathbb{1},\sigma_k)$ and $(\mathbb{1},\rho_k)$, $k = 1,2,3$ with the familiar multiplication rules,

$$\sigma_k\cdot\sigma_l = \delta_{kl}\cdot\mathbb{1} + i\epsilon_{klm}\sigma_m, \tag{3.7}$$

$$\rho_k\cdot\rho_l = \delta_{kl}\cdot\mathbb{1} + i\epsilon_{klm}\rho_m, \tag{3.8}$$

(of course, $\epsilon_{123} = 1$) which possess the standard matrix realization:

$$(\mathbb{1},\sigma_k) = \begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}, \begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}, \begin{pmatrix}0 & -i\\ i & 0\end{pmatrix},$$

$$\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix} = (\mathbb{1},\rho_k). \tag{3.9}$$

We shall seek the representation of $\mathrm{A}^{\mathrm{ext}}$ in the cartesian product $(\mathbb{1},\sigma_k)\times(\mathbb{1},\rho_l)$, understanding $\mathbb{1}$ as the product of unities of both rings and proceeding with $[\sigma_k,\rho_l]^{(-)} = 0$; the four-dimensional matrix realizations of this Cartesian product one constructs in the standard manner, following Dirac.[3]

One then easily finds that the objects

$$\mathbb{1} = \mathbb{1}, \quad \alpha := (1/\sqrt{2})(\mathbb{1} - i\sigma_3)\rho_3,$$
$$\beta := (1/\sqrt{2})(\mathbb{1} + i\sigma_3)\rho_3, \quad \gamma := -i\sigma_3 \tag{3.10}$$

fulfill the multiplication rules of the Abelian corner of Table II. Moreover, realizing the star simply in the form of $\star = \sigma_1\cdot\rho_3$, i.e., postulating for the "odd" elements the representation

$$\star := \sigma_1\cdot\rho_3, \quad \star\alpha := (1/\sqrt{2})(\sigma_1 - \sigma_2),$$
$$\star\beta := (1/\sqrt{2})(\sigma_1 + \sigma_2), \quad \star\gamma = -\sigma_2\rho_3, \tag{3.11}$$

one easily proves that the so defined eight elements $(\mathbb{1},...,\star\gamma)$ do fulfill *all* multiplication rules from Table II.

In the 4×4 matrix realization, the representation described by (3.40) and (3.11) amounts to

$$\mathbb{1} = \begin{pmatrix}1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&1\end{pmatrix}, \quad \alpha = e^{-i\pi/4}\begin{pmatrix}1&0&0&0\\0&i&0&0\\0&0&-1&0\\0&0&0&-i\end{pmatrix},$$

$$\beta = e^{i\pi/4}\begin{pmatrix}1&0&0&0\\0&-i&0&0\\0&0&-1&0\\0&0&0&i\end{pmatrix},$$

$$\gamma = -i\begin{pmatrix}1&0&0&0\\0&-1&0&0\\0&0&1&0\\0&0&0&-1\end{pmatrix} \tag{3.12}$$

and

$$\star = \begin{pmatrix}0&1&0&0\\1&0&0&0\\0&0&0&-1\\0&0&-1&0\end{pmatrix}, \quad \star\alpha = e^{-i\pi/4}\begin{pmatrix}0&i&0&0\\1&0&0&0\\0&0&0&i\\0&0&1&0\end{pmatrix},$$

$$\star\beta = e^{i\pi/4}\begin{pmatrix}0&-i&0&0\\1&0&0&0\\0&0&0&-i\\0&0&1&0\end{pmatrix},$$

$$\star\gamma = -i\begin{pmatrix}0&-1&0&0\\1&0&0&0\\0&0&0&1\\0&0&-1&0\end{pmatrix}. \tag{3.13}$$

It is now of interest to work out in this representation our original operators from the list $\mathcal{L}'$, (2.5); one easily finds that they amount to

$$\mathbb{1}=\begin{pmatrix}1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&1\end{pmatrix},$$

$$\star d=\sqrt{\Delta}\,e^{i\pi/4}\begin{pmatrix}0&0&0&0\\0&1&0&0\\0&0&0&0\\0&0&0&-1\end{pmatrix},$$

$$\star\delta=\sqrt{\Delta}\,e^{-i\pi/4}\begin{pmatrix}1&0&0&0\\0&0&0&0\\0&0&-1&0\\0&0&0&0\end{pmatrix},$$

$$\not\Delta=\Delta\begin{pmatrix}-1&0&0&0\\0&1&0&0\\0&0&-1&0\\0&0&0&1\end{pmatrix}, \tag{3.14}$$

and

$$\star=\begin{pmatrix}0&1&0&0\\1&0&0&0\\0&0&0&-1\\0&0&-1&0\end{pmatrix},$$

$$d=\sqrt{\Delta}\,e^{i\pi/4}\begin{pmatrix}0&1&0&0\\0&0&0&0\\0&0&0&1\\0&0&0&0\end{pmatrix},$$

$$\delta=\sqrt{\Delta}\,e^{-i\pi/4}\begin{pmatrix}0&0&0&0\\1&0&0&0\\0&0&0&0\\0&0&1&0\end{pmatrix},$$

$$\not\Delta=\Delta\begin{pmatrix}0&1&0&0\\-1&0&0&0\\0&0&0&-1\\0&0&1&0\end{pmatrix}. \tag{3.15}$$

We can now substitute (3.14) and (3.15) into (2.3), which leads to the representation of $a\in\mathbb{A}$ in the form of a 4×4 matrix with the commuting entries:

$$\mathbb{A}^{\text{ext}}\ni a=\begin{pmatrix}K&L&0&0\\M&N&0&0\\0&0&\bar K&\bar L\\0&0&\bar M&\bar N\end{pmatrix}, \tag{3.16}$$

where we have denoted

$$\begin{aligned}K&:=\mathcal{R}^+-\Delta\mathcal{S}^++\sqrt{\Delta}\,e^{-i\pi/4}\mathcal{Q}^+,\\ L&:=\mathcal{R}^-+\Delta\mathcal{S}^-+\sqrt{\Delta}\,e^{i\pi/4}\mathcal{P}^-,\\ M&:=\mathcal{R}^--\Delta\mathcal{S}^-+\sqrt{\Delta}\,e^{-i\pi/4}\mathcal{Q}^-,\\ N&:=\mathcal{R}^++\Delta\mathcal{S}^++\sqrt{\Delta}\,e^{i\pi/4}\mathcal{P}^+,\end{aligned} \tag{3.17}$$

and

$$\begin{aligned}\bar K&:=\mathcal{R}^+-\Delta\mathcal{S}^+-\sqrt{\Delta}\,e^{-i\pi/4}\mathcal{Q}^+,\\ \bar L&:=-(\mathcal{R}^-+\Delta\mathcal{S}^-)+\sqrt{\Delta}\,e^{i\pi/4}\mathcal{P}^-,\\ \bar M&:=-(\mathcal{R}^--\Delta\mathcal{S}^-)+\sqrt{\Delta}\,e^{-i\pi/4}\mathcal{Q}^-,\\ \bar N&:=\mathcal{R}^++\Delta\mathcal{S}^+-\sqrt{\Delta}\,e^{i\pi/4}\mathcal{P}^+.\end{aligned} \tag{3.18}$$

It is now clear that the commuting objects from $\mathcal{N}^{\text{ext}}$, $(K,\ldots,\bar N)$, can replace conveniently the coefficients $(\mathcal{P}^+,\ldots,\mathcal{S}^-)$ in (2.3); from (3.16), which provides a faithful representation of our algebra, it is obvious that, working with $(K,\ldots,\bar N)$, the multiplication law of the elements from $\mathbb{A}^{\text{ext}}$ amounts to two copies of the (independent) multiplication of the general 2×2 matrices.

Solving (3.17) and (3.18) for $(\mathcal{P}^+,\ldots,\mathcal{S}^-)$ and substituting into (2.3), we can organize the result in the form of

$$\mathbb{A}^{\text{ext}}\ni a=K\kappa+L\lambda+M\mu+N\nu+\bar K\bar\kappa+\bar L\bar\lambda+\bar M\bar\mu+\bar N\bar\nu, \tag{3.19}$$

where $(K,\ldots,\bar N)\in\mathcal{N}^{\text{ext}}$ and the algebraically active operators are defined according to

$$\begin{aligned}\kappa&:=\frac14\Big(\mathbb{1}+\frac{2}{\sqrt{\Delta}}e^{i\pi/4}\star\delta-\frac{1}{\Delta}\not\Delta\Big),\\ \bar\kappa&:=\frac14\Big(\mathbb{1}-\frac{2}{\sqrt{\Delta}}e^{i\pi/4}\star\delta-\frac{1}{\Delta}\not\Delta\Big),\\ \lambda&:=\frac14\star\Big(\mathbb{1}+\frac{2}{\sqrt{\Delta}}e^{-i\pi/4}\star d+\frac{1}{\Delta}\not\Delta\Big),\\ \bar\lambda&:=\frac14\star\Big(-\mathbb{1}+\frac{2}{\sqrt{\Delta}}e^{-i\pi/4}\star d-\frac{1}{\Delta}\not\Delta\Big),\\ \mu&:=\frac14\star\Big(\mathbb{1}+\frac{2}{\sqrt{\Delta}}e^{i\pi/4}\star\delta-\frac{1}{\Delta}\not\Delta\Big),\\ \bar\mu&:=\frac14\star\Big(-\mathbb{1}+\frac{2}{\sqrt{\Delta}}e^{i\pi/4}\star\delta+\frac{1}{\Delta}\not\Delta\Big),\\ \nu&:=\frac14\Big(\mathbb{1}+\frac{2}{\sqrt{\Delta}}e^{-i\pi/4}\star d+\frac{1}{\Delta}\not\Delta\Big),\\ \bar\nu&:=\frac14\Big(\mathbb{1}-\frac{2}{\sqrt{\Delta}}e^{-i\pi/4}\star d+\frac{1}{\Delta}\not\Delta\Big).\end{aligned} \tag{3.20}$$

Clearly, these operators all belong to $\mathbb{A}^{\text{ext}}$. For various purposes they represent perhaps the most advantageous linear combinations of the basic eight operators from the list $\mathcal{L}'$—because of the very simple multiplication table between them which is given in Table III.

This table can be computed either directly by using the definitions (3.20) and the basic multiplication rules among the elements from the list $\mathcal{L}'$, or by employing the remark that with (3.16) being the faithful representation of $\mathbb{A}^{\text{ext}}$, the multiplication rule among the elements of the form (3.19) must amount to the matrix multiplication of their coefficients from $\mathcal{N}^{\text{ext}}$:

TABLE III.

	κ	λ	μ	ν	$\bar\kappa$	$\bar\lambda$	$\bar\mu$	$\bar\nu$
κ	κ	λ	0	0	0	0	0	0
λ	0	0	κ	λ	0	0	0	0
μ	μ	ν	0	0	0	0	0	0
ν	0	0	μ	ν	0	0	0	0
$\bar\kappa$	0	0	0	0	$\bar\kappa$	$\bar\lambda$	0	0
$\bar\lambda$	0	0	0	0	0	0	$\bar\kappa$	$\bar\lambda$
$\bar\mu$	0	0	0	0	$\bar\mu$	$\bar\nu$	0	0
$\bar\nu$	0	0	0	0	0	0	$\bar\mu$	$\bar\nu$

$$a = a_1 \cdot a_2 \rightarrow \begin{pmatrix} K & L \\ M & N \end{pmatrix} = \begin{pmatrix} K_1 & L_1 \\ M_1 & N_1 \end{pmatrix} \begin{pmatrix} K_2 & L_2 \\ M_2 & N_2 \end{pmatrix},$$

$$\begin{pmatrix} \bar K & \bar L \\ \bar M & \bar N \end{pmatrix} = \begin{pmatrix} \bar K_1 & \bar L_1 \\ \bar M_1 & \bar N_1 \end{pmatrix} \begin{pmatrix} \bar K_2 & \bar L_2 \\ \bar M_2 & \bar N_2 \end{pmatrix}. \tag{3.21}$$

We can now interpret this result by saying that the algebra $\mathbb{A}^{\text{ext}}$ splits into the direct product of the subalgebras A and $\bar A$ spanned respectively by $(\kappa\lambda\mu\nu)$ and $(\bar\kappa\bar\lambda\bar\mu\bar\nu)$. These subalgebras do not contain unity; however,

$$\kappa + \bar\kappa + \nu + \bar\nu = \mathbb{1}. \tag{3.22}$$

Let $\omega \in \mathscr{N}^{\text{ext}}$; we then easily find the secular determinant of the matrix (3.16), $\mathscr{P}(\omega) := \det(a - \mathbb{1}\omega)$ in the form of

$$\mathscr{P}(\omega) = [\omega^2 - (K+N)\omega + (KN - ML)] \times [\omega^2 - (\bar K + \bar N)\omega + (\bar K\bar N - \bar M\bar L)]. \tag{3.23}$$

According to the Hamilton–Cayley theorem, the matrix a from (3.16) fulfills the equation

$$\mathscr{P}(a) = 0, \tag{3.24}$$

with $a^\circ \equiv \mathbb{1}$. The same equation must be fulfilled by $a \in \mathbb{A}^{\text{ext}}$ from (3.19) because of the multiplication rules (3.21) and (3.23). We can, however, by feeding (3.17) and (3.18) into (3.24) work out the coefficients of $\mathscr{P}(\omega)$ in the terms of the original formal power series in Δ, $(\mathscr{P}^{\pm},\dots,\mathscr{S}^{\pm})$ which parametrize according to (2.3) $a \in \mathbb{A}^{\text{ext}}$. Writing the polynomial in the form of

$$\mathscr{P}(\omega) = \omega^4 - C_1\omega^3 + C_2\omega^2 - C_3\omega + C_4, \tag{3.25}$$

we find the coefficients in the form of

$$C_1 := 4\mathscr{R}^+,$$

$$C_2 := \{6(\mathscr{R}^+)^2 - 2(\mathscr{R}^-)^2\} + \{i(\mathscr{Q}^+)^2 - i(\mathscr{P}^+)^2 - 2\mathscr{P}^-\mathscr{Q}^-\}\Delta + 2[(\mathscr{S}^-)^2 - (\mathscr{S}^+)^2\}\Delta^2,$$

$$\begin{aligned} C_3 := {} & 4\mathscr{R}^+((\mathscr{R}^+)^2 - (\mathscr{R}^-)^2) \\ & + 2[\mathscr{R}^+(i(\mathscr{Q}^+)^2 - \mathrm{i}(\mathscr{P}^+)^2 - 2\mathscr{P}^-\mathscr{Q}^-) \\ & + \mathscr{R}^-[\mathrm{i}\mathscr{P}^+\mathscr{P}^- - i\mathscr{Q}^+\mathscr{Q}^- + \mathscr{Q}^+\mathscr{P}^- + \mathscr{P}^+\mathscr{Q}^-)]\cdot\Delta \\ & + 2[\mathscr{R}^+((\mathscr{S}^-)^2 - (\mathscr{S}^+)^2) + \mathscr{S}^+(i(\mathscr{Q}^+)^2 \\ & + i(\mathscr{P}^+)^2) + \mathscr{S}^-(-i\mathscr{Q}^+\mathscr{Q}^- - \mathrm{i}\mathscr{P}^+\mathscr{P}^- \\ & + \mathscr{P}^+\mathscr{Q}^- - \mathscr{Q}^+\mathscr{P}^-)]\Delta^2, \end{aligned} \tag{3.26}$$

and

$$\begin{aligned} C_4 = {} & [(\mathscr{R}^+)^2 - (\mathscr{R}^-)^2]^2 \\ & + [i(\mathscr{R}^+\mathscr{Q}^+ - \mathscr{R}^-\mathscr{Q}^-)^2 - i(\mathscr{R}^+\mathscr{P}^+ - \mathscr{R}^-\mathscr{P}^-)^2 \\ & - 2(\mathscr{R}^+\mathscr{P}^- - \mathscr{R}^-\mathscr{P}^+)(\mathscr{R}^+\mathscr{Q}^- - \mathscr{R}^-\mathscr{Q}^+)]\cdot\Delta \\ & + \{2i(\mathscr{R}^+\mathscr{Q}^+ - \mathscr{R}^-\mathscr{Q}^-)(\mathscr{S}^+\mathscr{Q}^+ - \mathscr{S}^-\mathscr{Q}^-) \\ & + 2i(\mathscr{R}^+\mathscr{P}^+ - \mathscr{R}^-\mathscr{P}^-)(\mathscr{S}^+\mathscr{P}^+ - \mathscr{S}^-\mathscr{P}^-) \\ & + (\mathscr{P}^+\mathscr{Q}^+ - \mathscr{P}^-\mathscr{Q}^-)^2 \\ & + 2[(\mathscr{R}^+)^2 - (\mathscr{R}^-)^2][(\mathscr{S}^-)^2 - (\mathscr{S}^+)^2] \\ & + 2(\mathscr{R}^+\mathscr{S}^- - \mathscr{R}^-\mathscr{S}^+)(\mathscr{P}\mathscr{Q}^- - \mathscr{P}^-\mathscr{Q}^+)\}\cdot\Delta^2 \\ & + [i(\mathscr{S}^+\mathscr{Q}^+ - \mathscr{S}^-\mathscr{Q}^-)^2 - i(\mathscr{S}^+\mathscr{P}^+ - \mathscr{S}^-\mathscr{P}^-)^2 \\ & + 2(\mathscr{S}^+\mathscr{P}^- - \mathscr{S}^-\mathscr{P}^+)(\mathscr{S}^+\mathscr{Q}^- - \mathscr{S}^-\mathscr{Q}^+)]\cdot\Delta^3 \\ & + [(\mathscr{S}^+)^2 - (\mathscr{S}^-)^2]^2\cdot\Delta^4. \end{aligned} \tag{3.27}$$

The structure of the coefficients $C_1,\dots,C_4$ is such that with $(\mathscr{P}^{\pm},\dots,\mathscr{S}^{\pm}) \in \mathscr{N}$, these coefficients are also all in $\mathscr{N}$; all roots of Δ cancel out. This has an important consequence: with (3.25) applying for an $a \in \mathbb{A}^{\text{ext}}$, we can infer that this relation must also apply on the level of the original algebra *without* the neccesity of introducing the operation $\Delta^{-1/2}$. In other words, given some a in $\mathbb{A}$, with all coefficients in (2.3) being in $\mathscr{N}$, we can assert that this element fulfills the equation $\mathscr{P}(a) = 0$, with the coefficients $C_1,\dots,C_4$ given explicitly by (3.27) and (3.28) also belonging to $\mathscr{N}$, i.e., consisting in a formal power series in Δ. The device of $\mathbb{A}^{\text{ext}}$ was needed only as an intermediate step in order to obtain this result and a simple manner of determining the coefficients $C_1,\dots,C_4$.

4. THE INVERSE ELEMENT $a^{-1} \in \mathbb{A}$ AND SOME APPLICATIONS

When working with $\mathbb{A}^{\text{ext}}$, one easily sees that its general element, given in the form of (3.19) if only

$$D := KN - ML \neq 0 \neq \overline{KN} - \overline{ML} =: \bar D, \tag{4.1}$$

posseses the inverse

$$\mathbb{A}^{\text{ext}} \ni a^{-1} := D^{-1}\{N\kappa - L\lambda - M\mu + K\nu\} + \bar D^{-1}\{\overline{N\kappa} - \overline{L\lambda} - \overline{M\mu} + \overline{K\nu}\}, \tag{4.2}$$

such that

$$a \cdot a^{-1} = \mathbb{1} = a^{-1} \cdot a. \tag{4.3}$$

(The objects D and $\bar D$ in $\mathscr{N}^{\text{ext}}$, if different from zero, do possess inverses.) Consequently, the algebra of the nonsingular elements of $\mathbb{A}^{\text{ext}}$, such that $D\bar D \neq 0$, is isomorphic to the group $\mathrm{GL}(2,\mathbb{C}) \times \mathrm{GL}(2,\mathbb{C})$.

Now, we shall call $a \in \mathbb{A}$ strongly nonsingular, if this element given in the form of (2.3) has the property that C_4 from

(3.27) has a nontrivial numerical leading term, i.e., if $(\mathscr{R}^{+})^2 - (\mathscr{R}^{-})^2 = N_0(\mathbb{1} + \Delta\,[\text{something}])$, $N_0 \neq 0$; in this way one secures the existence of a formal series in Δ, C_4^{-1}, such that $C_4^{-1} \cdot C_4 = 1$.

Because a strongly nonsingular element must fulfill its Hamilton–Cayley equation, (3.24), it follows that it possesses an inverse element given explicitly in the form of

$$a^{-1} = C_4^{-1}(C_3 \cdot \mathbb{1} - C_2 \cdot a + C_1 \cdot a^2 - a^3). \tag{4.4}$$

When $C_4 \neq 0$, but the formal power series for C_4 begins with some nontrivial power of Δ, say $C_4 = \Delta^n \cdot \tilde{C}_4$, $n \geqslant 1$, $\tilde{C}_4^{-1}$ exists, we can then only construct such an element

$$\tilde{a} = \tilde{C}_4^{-1}(C_3 \cdot \mathbb{1} - C_2 \cdot a + C_1 \cdot a^2 - a^3) \tag{4.5}$$

that

$$a \cdot \tilde{a} = \Delta^n + \tilde{a} \cdot a \tag{4.6}$$

An element of $\mathbb{A}$ with $C_4 \neq 0$ will be called nonsingular. In various applications of $\mathbb{A}$ (or $\mathbb{A}^{\text{ext}}$) one will encounter the problem of solving for x the equation

$$a \cdot x = y, \quad a \in \mathbb{A}, \; x, y \in \Lambda \tag{4.7}$$

with a and y known, or the similar system of equations:

$$a_{ij} x^j = y_i, \quad a_{ij} \in \mathbb{A}, \; x^j, y_i \in \Lambda. \tag{4.8}$$

It is clear that when a is strongly nonsingular, (4.7) has an unique formal solution $x = a^{-1}y$; when it is just nonsingular, (4.7) has a solution in the form of

$$x = \tilde{a}y + x_0, \quad \Delta^n x_0 = 0, \tag{4.9}$$

with the multiharmonic $x_0 \in \Lambda$ being arbitrary.

In the case of a system of equations of the type (4.8), the chief idea in trying to devise their formal solutions would consist in the construction of such elements $\tilde{a}_{ij} \in \mathbb{A}$ that $\tilde{a}_{ij} \cdot a_{jk} = N_{ik} \in \mathscr{N}$. In principle, the algebraic information contained in this paper is sufficient for this purpose, but we will not elaborate this point in the present publication.

We will describe now briefly a typical example consisting in the search for solutions of the Maxwell equations in a curved space–time.

Let $\dim(M) = 4$ and the signature of the real metric induced over the manifold M by $\star$ through $i\alpha \lrcorner \alpha : = \star(\alpha \wedge \star\alpha)$, $\alpha \in \Lambda^1$, be $(+\;+\;+\;-)$. Then the real electromagnetic field given in local components by

$$f = \tfrac{1}{2} f_{\mu\nu}\, dx^\mu \wedge dx^\nu \in \Lambda^2 \tag{4.10}$$

is described entirely by the complex self-dual 2-form

$$\omega : = f + \star f \in \Lambda^2, \tag{4.11}$$

and the Maxwell equations without currents amount to

$$d\omega = 0, \quad (\star - \mathbb{1})\omega = 0. \tag{4.12}$$

The solution of (4.12) in the terms of the Hertz potentials was described in some detail in Ref. 1; see also the excellent paper,[4] where the theory of the Hertz potentials is outlined from the point of view of many formalisms appearing in the literature.

In the present text, we will consider conditions (4.12) under more general assumptions: instead of working with $\dim(M) = 4$ and specific signature, we will only assume that $\dim(M) = 2n$ and instead of confining ω to a 2-form, we will consider it as a general linear combination of multiforms from Λ over M.

It presents no difficulty to show that the general solution of (4.12) can be represented in the form of

$$\omega = \not{\Delta}\bar{\Pi}, \quad \bar{\Pi} : = \tfrac{1}{2}(\mathbb{1} - \star)(\bar{\Pi}_0 + \bar{\Pi}_1 + \cdots + \bar{\Pi}_n) \to \star\bar{\Pi} = -\bar{\Pi}, \tag{4.13}$$

where the forms $\bar{\Pi}_p \in \Lambda^p$ are all harmonic,

$$\Delta\bar{\Pi}_p = 0, \quad p = 0, 1, \ldots, n, \tag{4.14}$$

and apart from these conditions arbitrary. In fact, $d\omega = d\not{\Delta}\bar{\Pi} = -\not{\Delta}d\bar{\Pi} = \Delta d\bar{\Pi} = d\Delta\bar{\Pi} = d(\mathbb{1} - \star)\Delta \times (\bar{\Pi}_0 + \cdots + \bar{\Pi}_n) = 0$ and $(\mathbb{1} - \star)\omega = \not{\Delta}(\mathbb{1} + \star)\bar{\Pi} = 0$, according to the basic properties of d, $\star$, and $\not{\Delta}$. The proof that for ω which fulfills (4.12) one can always find the representation in the form of (4.13), requires the study of the gauges for the forms $\bar{\Pi}_p$, along similar lines as in Ref. 1. If among the $\bar{\Pi}_p$'s only $\bar{\Pi}_n \neq 0$, then $\omega \in \Lambda^n$ and thus this case corresponds for $n = 2$ to the solution of the Maxwell equations via the Hertz potential.

ACKNOWLEDGMENTS

I would like to express thanks to Dr. T. Selingmann and Dr. S. Gitler for helpful discussion. My best thanks are also due to Mr. S. Alarcón Gutiérrez for the help in working out the multiplication tables.

APPENDIX

This text employs $\star$ and δ renormalized conveniently comparing with the standardly used $\star_0$ and δ_0. If a real M_n is endowed with a nonsingular metric of the signature with n_+ pluses and n_- minuses $(n = n_+ + n_-)$, then the usual definition of $\star_0 : \Lambda^p \to \Lambda^{p'}$; $0 \leqslant p \leqslant n$, $p' = n - p$, stated in a most elementary form—in a local chart $\{x^\mu\}$—is:

$$\begin{aligned} &\text{if } \alpha : = \frac{1}{p!} \alpha_{\mu_1 \cdots \mu_p}\, dx^{\mu_1} \wedge \cdots \wedge dx^{\mu_p} \in \Lambda^p, \\ &\text{then } \star_0\alpha : = \frac{|\det(g_{\alpha\beta})|^{1/2}}{p!p'!} \epsilon^{\lambda_1 \cdots \lambda_p}{}_{\mu_1 \cdots \mu_{p'}} \\ &\qquad \times \alpha_{\lambda_1 \cdots \lambda_p} dx^{\mu_1} \wedge \cdots \wedge dx^{\mu_{p'}}, \end{aligned} \tag{A1}$$

where $|\;|$ means the absolute value and $\epsilon^{\lambda_1 \cdots \lambda_p}{}_{\mu_1 \cdots \mu_{p'}}$ is the Levi-Civita symbol $\epsilon_{\alpha_1 \cdots \alpha_n}$ $(\epsilon_{1 \cdots n} = +1)$ with a part of indices raised by $g^{\alpha\beta}$ as indicated. This leads to

$$\alpha \in \Lambda^p \to \star_0 \cdot \star_0 \alpha = (-1)^{p'p + n_-} \alpha. \tag{A2}$$

At the same time, the usual definition of δ_0 is

$$\alpha \in \Lambda^p \to \delta_0 \alpha : = (-1)^{np - n_- + 1} \star_0 d \star_0 \alpha, \tag{A3}$$

which assures the validity of

$$d(\alpha \wedge \star_0 \beta) = d\alpha \star_0 \beta - \alpha \wedge \star_0 \delta\beta \quad (\alpha \in \Lambda^p, \beta \in \Lambda^{p+1}) \tag{A4}$$

and, therefore, secures that, for compact M_n's with $n_- = 0$, δ_0 is conjugated to d in the sense of $\langle \alpha, \beta \rangle = \int_{M_n} \alpha \wedge \star_0 \beta$, $\alpha, \beta \in \Lambda^p$. When working with the complex valued forms (in general) on manifolds with indefinite metric it is convenient to employ a differently normalized $\star$:

$$\alpha \in \Lambda^{p} \rightarrow \star\alpha := e^{i\pi(pp' + n_-)/2} \star_0 \alpha, \tag{A5}$$

which has a manifest property $\star\star\alpha = \alpha$ for $\alpha \in \Lambda^{p}$ $(p = 0,1,...,n)$, and hence, $\star\star = \mathbb{1}$ over the whole $\oplus_p \Lambda^{p}$. One then easily sees that the old δ_0 has a simplified form in terms of the new $\star$:

$$\alpha \in \Lambda^{p} \rightarrow \delta_0 \alpha = e^{-i\pi(n+1)/2} \star d \star \cdot \alpha, \tag{A6}$$

where the normalization factor is independent of p and of signature. The codifferential $\delta = -\mathrm{i}\star\mathrm{d}\star$ used in this paper is related to δ_0 by: $\delta = e^{i\pi n/2}\delta_0$, the factor $e^{i\pi n/2}$ being of no importance in general, and reducing to unity for the important in relativity $n = 4$. The operator $\Delta_0 := d\delta_0 + \delta_0 d$ (i.e., for the pseudo-Riemannian geometry the deRahm–Lichnerowicz operator) is thus related to Δ from our text again by $\Delta = e^{+i\pi n/2}\Delta_0$, the unimportant factor again being unity for $n = 4m$. The renormalized $\star$ and δ with $\star\star = \mathbb{1}$, $\delta = -i\star d\star$ have self-evident algebraic advantages, simplifying simultaneously all numerical coefficients. The complex factors implies by this normalization which occur for some n's and p's when working on real forms, amount to ± 1 and $\pm i$ and can be easily traced in practical applications. In fact, the renormalized $\star$ and δ are already conveniently used in standard applications in the relativity theory.

[1]J.F. Plebański and Ivor Robinson, "Electromagnetic and Gravitational Hertz Potentials," J. Math. Phys. **19**, 2350 (1978).

[2]T. Kahan, *Theory of Groups in Classical and Quantum Physics* (Oliver & Boyd, London, 1965), p. 220.

[3]P.A.M. Dirac, *The Principles of Quantum Mechanics* (Oxford at the Clarendon Press, Oxford 1958), 4th ed., p. 257.

[4]J.N. Cohen and L.S. Kegels, Phys. Rev. D **10**, 1070 (1974); Phys. Lett. A **47**, 261 (1974).

Distributional geometry

Phillip E. Parker

Department of Mathematics, State University of New York at Buffalo, Buffalo, New York 14214 and Syracuse University, Syracuse, New York 13210 [a)]

(Received 22 September 1978; revised manuscript received 7 December 1978)

Recently there has been considerable interest in geometry with nonsmooth coefficients. The general theory of such *distributional geometry* is begun and an application in general relativity to black holes is given.

INTRODUCTION

In mathematics the study of classical mechanics in any depth is usually identified with the study of symplectic geometry. The constraints[1] in a problem are used to define a pseudo-Riemannian manifold such that motion in the absence of a potential occurs along the geodesics. The manifold is called the *configuration space* and its cotangent bundle the *phase space*. Usually one assumes that the potential is smooth, but some interesting problems require discontinuous potentials: e.g., hard balls bouncing around inside a box. Since the theory employed in the smooth case is *differential* geometry, one is interested in *distributional* geometry.

This theory has recently entered mathematical physics in several places independently. Isham showed in Ref. 2 that distributional Lorentzian structures would have to be considered in order to obtain a consistent quantization of general relativity. Glimm and Jaffe[3] have constructed *multiple meron* gauge fields; these are distributional connections in principal SU(2) bundles over $\mathbb{R}^4$. The author in his thesis[4] used distributional Lorentzian structures to formalize general relativity in such a way that it does not break down at spacetime singularities. There is thus ample motivation for a general study of distributional geometry.

This paper contains the elements of the theory. In Sec. 1 flows for distributional vector fields are constructed. This extends the work of Marsden[5] in which the special case of measurable vector fields was extensively studied. These are used in Secs. 2 and 3 to discuss symplectic and pseudo-Riemannian geometry, respectively. The last section describes the application to general relativity mentioned above.

1. FLOWS

Let X be a smooth (C^∞) paracompact manifold. Denote by X^α the *bundle of α-densities* over X and by Ω_α its space of smooth sections, the space of *α-densitites* on X.[6] If E is any vector bundle over X, then $\Gamma(E)$ denotes the space of smooth sections and $\mathrm{B}(E)$ those with compact support. The space of *α-density sections* of E is $\Gamma_\alpha(E) := \Gamma(E \otimes X^\alpha)$, the *geometric α-dual* is $E'_{1-\alpha} := E^* \otimes X^{1-\alpha}$ and the space of *α-distributional sections* is $\mathrm{B}'_\alpha(E) := (\mathrm{B}(E'_{1-\alpha}))'$. To avoid confusion, fiberwise contraction will be denoted (,) and the action of distributions $\langle\, , \rangle$. When $\alpha = 0$, it will be suppressed. The distributions on X are thus denoted as usual by $\mathscr{D}'(X)$. If $P: \mathrm{B}_\alpha(E) \to \mathrm{B}_\delta(F)$ is a differential operator of order m, then there is a unique extension $P: \mathrm{B}'_\alpha(E) \to \mathrm{B}'_\delta(F)$. Thus there is an exterior derivative d for distributional forms and a Lie derivative £ for distributional tensors. One can extend contraction and the usual relations among d, £, and continue to hold. The exterior product $\wedge$ also extends provided one can multiply the distributional coefficients. Finally, the Lie derivative can also be extended to distributional vector fields if only smooth tensors are differentiated. For more details see Marsden[5] and Parker.[4]

The following definition of generalized paths is based on Young[7] and avoids the delicate measure theory of Marsden[5]. It will be clear that these paths are more general than currents.[8]

Let $C^\infty_{p\,\ell oc}(\mathbb{R}, X)$ be the set of all proper smooth maps $p: I_p \to X$ where $0 \in I_p$, the domain of p. Define an action of this space on $\mathscr{D}(X)$, the smooth functions with compact support, by

$$\langle p, \varphi \rangle := \int_{I_p} (\varphi \circ p)(t)\, dt, \quad \varphi \in \mathscr{D}(X). \tag{1.1}$$

This yields an injection $C^\infty_{p\,\ell oc}(\mathbb{R}, X) \to \mathscr{D}'_1(X)$ and one defines $P(X)$, the space of *generalized proper paths* in X, to be the sequential closure. More generally, one defines $P_K(X)$ to be the sequential closure of the set of all p such that $I_p \supseteq K$, replacing I_p in (1.1) with K, where K is a compact subset of $\mathbb{R}$. The space of *generalized paths* in X is $\widehat{P}(X)$:proj-$\lim_K P_K(X)$.[9] Since $\mathbb{R}$ is σ-compact, this is a countable inverse limit, and one may think of elements of $\widehat{P}(X)$ as sequences in $\mathscr{D}'_1(X)$ which may not converge. The sequence of compact sets $[-i, i]$, $i \geqslant 1$, will be used as the standard sequence so that $p \in \widehat{P}(X)$ can be written as $p = (p_i)$ with $p_i \in P_{[-i,i]}(X)$. The p_i are called the *representatives* of p.

In $\widehat{P}(TX)$ one finds the *generalized velocity field* $\dot{p}$ of a generalized path p [replace p with $\dot{p}$ and take $\varphi \in \mathscr{D}(TX)$ in (1.1)]. If $p = (p_i) \in \widehat{P}(X)$ define $\operatorname{supp} p := \cup_i \operatorname{supp} p_i$ so that $\operatorname{supp} p = \pi(\operatorname{supp} \dot{p})$. Observe that one may consider $\dot{p}_i$ as a 1-distributional vector field on X with support $\operatorname{supp} p_i$. If one defines a generalized integral curve p for a distributional vector field ξ in the obvious way and if X has a distinguished 1-density, then $\xi \,|\operatorname{supp} p_i = \dot{p}_i$ whenever the restriction is defined (see Hörmander[6] for restrictions). The obvious way does work.

[a)]Current address.

0022-2488/79/071423-04$01.00

Theorem 1.1: *Let $\xi \in \mathfrak{X}' := \mathrm{B}'(TX)$ and let $\{u_n\}$ and $\{v_n\}$ be two sequences of smooth vector fields converging to ξ in $\mathfrak{X}'$. As sets of paths, the flows of $\{u_n\}$ and $\{v_n\}$ converge to the generalized flow of ξ.*

Proof: Choose $x \in X$, and let c_n and e_n be maximal integral curves through x for u_n and v_n, respectively. We may assume that each c_n and e_n is proper. To show that $c_n \to p \in P(X)$, let $\varphi \in \mathscr{D}(X)$ and consider the sequence of integrals.

$$\int_{I_n} (\varphi \circ c_n)(t)\, dt, \tag{1.2}$$

where I_n is the domain of c_n. In order to see that this sequence converges, recall that $\dot{c}_n = u_n \circ c_n$ and let $\widetilde{\varphi} = \varphi \circ \pi$ so that (1.2) becomes

$$\int_{I_n} (\widetilde{\varphi} \circ c_n)(t)\, dt.$$

Now $\widetilde{\varphi} \circ u_n = (\omega, u_n)$ for some $\omega \in \mathrm{B}(T^*X)$, hence $\widetilde{\varphi} \circ u_n \circ c_n = (\omega, u_n \circ c_n)$. By assumption the sequence $\int_X (\omega, u_n) v$ converges for every $v \in \Omega_1(X)$, and replacing u_n with $u_n - u_m$ one sees it suffices to show that for some v,

$$\left| \int_X (\omega, u_n)\, v \right| \geqslant \left| \int_{I_n} (\omega, u_n \circ c_n)(t)\, dt \right| \tag{1.3}$$

in order to establish convergence of (1.2). Using a partition of unity argument, it will be enough to prove (1.3) when X is an open subset of $\mathbb{R}^n$. In this case it is clear that one can find $v = f\,dx$, where dx is Lebesque measure and f is smooth, such that

$$\int_X |(\omega, u_n) f|\, dx \geqslant \int_{I_n} |(\omega, u_n \circ c_n)(t)|\, dt;$$

moreover, v can be further chosen so that

$$\int_X |(\omega, u_n) f|\, dx = \left| \int_X (\omega, u_n) f\, dx \right|.$$

Thus $c_n \to p$ and $e_n \to q \in P(X)$. If $x \in X \setminus \operatorname{sing\,supp} \xi$, then clearly $p = q$. However, if $x \in \operatorname{sing\,supp} \xi$, then although q is still a generalized integral curve it need no longer be p. Thus there is a unique flow, but it only foliates $X \setminus \operatorname{sing\,supp} \xi$. □

As an example take $X = \mathbb{R}^2$ and consider the energy functions $\frac{1}{2}(x^2 + y^2) + \delta(x)$ and $\frac{1}{2}(x^2 + y^2) + h(x)$, where $\delta(x)$ is the Dirac delta in the x variable and $h(x)$ is the Heaviside function. Approximating δ with the Poisson kernel as $\delta(x) = \lim_{\epsilon \searrow 0} \epsilon / \pi(\epsilon^2 + x^2)^{-1}$ and looking at the flow for fixed ϵ, a straighforward calculation yields a flow of reflection off the y axis in the first case and refraction by Snell's law across the y axis in the second.

2. SYMPLECTIC GEOMETRY

Let σ be a symplectic form on the manifold X. $u \in \mathfrak{X}'$ is *symplectic* iff $\pounds_u \sigma = 0$. Since $\pounds_u \sigma = d(u \;\; \sigma)$, it follows from the local Poincaŕe lemma that u is symplectic iff $u \;\; \sigma = df$ locally for some $f \in \mathscr{D}'(X)$. If $f \in \mathscr{D}'(X)$ define the *Hamiltonian vector field* H^f by $df = H^f \;\; \sigma$. In this case f is sometimes called the *Hamiltonian distribution* or *energy* of H^f and is unique modulo constants. If $g \in C^\infty(X)$, then the *Poisson bracket* of f and g is $\{f,g\} := \sigma(H^f, H^g)$. The basic formulas continue to hold: e.g., $H^{\{f,g\}} = [H^g, H^f]$, $\pounds_{H^f} \sigma = 0$, $\{f,g\} = \pounds_{H^g} f$.

The first problem one encounters is defining pullbacks along generalized flows. Let $\{u_n\}$ be a sequence in $\mathfrak{X}$ converging to u in $\mathfrak{X}'$, and let c^n be the flow of u_n. Denote by p the generalized flow of u. If $p = (p_i)$, let $\{c_i^n\}$ be a sequence of proper restrictions such that $c_i^n \to p_i$. For $\omega \in \Omega'$ define $p_t^* \omega := (\lim_n (c_t^n)^* \omega)$.

Proposition 2.1.

$$\frac{d}{dt} p_t^* \omega = p_t^* \pounds_u \omega.$$

Proof: As above,

$$\frac{d}{dt} p_t^* \omega = \left(\lim_n \frac{d}{dt} (c_t^n)^* \omega \right) = (\lim_n (c_t^n)^* \, \pounds_u \omega)$$

$$= ((p_t)^* \, \pounds_u \omega) = p_t^* \, \pounds_u \omega. \qquad \square.$$

Proposition 2.2. *If p is the generalized flow of a symplectic vector field, than $p_t^* \sigma = \sigma$.*

Proof: Check that $p_t^* \sigma$ exists. □

Conservation of Energy. If $f \in \mathscr{D}'(X)$ and p is the generlized flow of H^f, then $p_t^* f = f$.

Proof: Apply Proposition 1. □

Recall that the *phase volume* of (X, σ) is $\mu := (1/n!)(-1)^{n(n-1)/2} \sigma^n$, where $\dim X = 2n$.

Liouville's Theorem: *If p is the generalized flow of a symplectic vector field, then $p_t^* \mu = \mu$.*

Proof: Apply Proposition 2 to the definition of μ. □

These are the basic results needed to do mechanics. As pointed out to the author by Brian Hassard, one can recover classical scattering theory, as indicated by the two examples at the end of the preceding section, using them. The main application here will be to pseudo-Riemannian geometry.

3. PSEUDORIEMANNIAN GEOMETRY

Henceforth the summation convention is in effect.

A distributional pseudo-Riemannian (PR) structure β on X is a distributional symmetric (2,0)-tensor on X such that $\beta \,|\, X \setminus \operatorname{sing\,supp} \beta$ is a smooth PR structure. In the smooth case β is usually defined as a (0,2)-tensor, but the most natural environment for a PR structure is the cotangent bundle. Locally one can represent β by a matrix $[\beta^{ij}]$ of distributions on X. One may consider $\beta \in \mathscr{D}'(T^*X)$ via $\langle \beta, \varphi \rangle := \langle \pi^* \beta, \xi \otimes \xi \varphi \rangle := \langle \pi^* \beta^{ij}, \xi_i \xi_j \varphi \rangle$ locally, where ξ_i is the ith component function. Define the *geodesics* of β to be the projections of the generalized integral curves of $H^{\beta/2}$. This is well known in the smooth case.

To define conjugate points first consider smooth β. Regard the exponential map of β as defined on the cotangent bundle and let g be geodesic flow. Observe that $\exp_* = \pi_* g_{1*}$ drops rank at $t\xi \in T^*X$ iff $\pi_* g_{t*}$ drops rank at ξ. Thus, if $\dim X = n$, conjugate points lie below precisely those points where $\text{rank}\pi_* g_{t*} < n$. Since g_t is a symplectomorphism and $T^*_{x_0}X$ is a conic Lagrangian submanifold[6] of T^*X, $g_t(T^*_{x_0}X)$ when defined is a conic Lagrangian immersion[6] in T^*X. In any case one can think of $g_{t*}T_\xi T^*_{x_0}X$ as an infinitesimal conic Lagrangian submanifold through $g_t(\xi)$. Thus there is associated with each integral curve g_ξ a lifting λ_ξ to the Grassmann bundle of Lagrangian planes ΛTT^*X given by $\lambda_\xi(t) := g_{t*}T_\xi T^*_{x_0}X$. Therefore, $x = \pi g_t(\xi)$ is conjuage to x_0 iff $\text{rank}(\pi_* |\lambda_\xi(t)) < n$.

More generally, the *caustic set* of a Lagrangian immersion $\iota: L \to T^*X$ is $\{\pi\iota(\ell); \ell \in L \text{ and } \text{rank}(\pi_* |\iota_* T_l L) < n\}$. If c is the flow of a symplectic vector field on T^*X, then the *caustic set of the flow c out of* $\iota(L)$ is $\{\pi c_t\iota(\ell); \ell \in L$ and $\text{rank}(\pi_* |c_{t*}\iota_* T_\ell L) < n\}$.

This proves the following:

Theorem 3.1: *The conjugate locus* $C(x_0)$ *of* x_0 *is the caustic set of the geodesic flow out of* $T^*_{x_0}X$. □

The set $\{c_t\iota(\ell); \ell \in L \text{ and } \text{rank}(\pi_* |c_{t*}\iota_* T_\ell L) < n\}$ is an unfolding in the sense of Thom, the *unfolding of the caustic*. For more details on this see Duistermaat[10] and Jänich.[11] The name is motivated by caustics in geometric optics.

Recall that ΛTT^*X is closed in $G_n TT^*X$, the Grassmann bundle of n-planes. Following the procedure of Sec. 1, if p is a generalized integral curve of $H^{\beta/2}$, where β is a distributional PR structure, through $\xi \in T^*_{x_0}X$, then one can associate to it $\lambda \in \widehat{P}(\Lambda TT^*X)$. Define the *conjugate locus* of x_0 to be $C(x_0) := \{\pi(\ell); \ell \in \text{supp}\lambda$ for some λ as above and $\text{rank}(\pi_* |\ell) < n\}$. Thus conjugate points in distributional PR structures are limits of sequences of conjugate points in approximating smooth structures. One defines the *order* or *multiplicity* of a conjugate point as usual to be $\dim \text{Ker}(\pi_* |\ell)$.

In the smooth case it is well known that the horizontal bundle is Lagrangian. (The vertical bundle is trivially Lagrangian). Recall that the Lagrangian subspaces of a symplectic vector space transverse to a given Lagrangian subspace can be coordinatized with the symmetric $n \times n$ matrices.[12] Thus each choice of a Lagrangian horizontal bundle gives rise to a vector bundle structure on the set of all Lagrangian subspaces transverse to the vertical subspaces, $\Lambda_0 TT^*X$. In general there is no natural vector bundle structure on this bundle.

To define the horizontal bundle of a distributional PR structure some general constructions on fiber bundles are necessary. Let $\pi: E \to B$ be a smooth fiber bundle and denote by $\Gamma(E)$ the smooth sections. Observe that any smooth section is a proper map $B \to E$. If B has a preferred density μ, one defines an action of $\Gamma(E)$ on $\mathscr{D}(E)$ by

$$\langle s, \varphi \rangle = \int_B (\varphi \circ s)\mu,$$

where $s \in \Gamma(E)$ and $\varphi \in \mathscr{D}(E)$. This yields as before an injection $\Gamma(E) \to \mathscr{D}'_1(E)$ and one defines the *distributional sections* of E to be the sequential closure of $\Gamma(E)$ in $\mathscr{D}'_1(E)$, denoted $\text{B}'(E)$. If E is a vector bundle and one replaces $\mathscr{D}(E)$ by $\mathscr{F}(E)$, the smooth functions with compact support which are linear on the fibers, one recovers the usual distributional sections of E. Clearly these constructions can be extended to obtain α-distributional sections of E.

Continuing now with a distributional PR structure β on X, let $\{\beta_n\}$ be a sequence of smooth structures converging to β and let $\mathscr{H}_n$ be the field of horizontal spaces Levi-Civita associated to β_n. Using the phase volume on T^*X as the preferred density, define $\mathscr{H} \in \text{B}'(\Lambda TT^*X)$ to be the *field of horizontal spaces Levi-Civita associated to* β iff $\mathscr{H}_n \to \mathscr{H}$ in $\text{B}'(\Lambda TT^*X)$. Thus $\mathscr{H}$ provides $\Lambda_0 TT^*X$ with a distributional vector bundle structure. Let ∇_n be the Levi-Civita covariant derivative of β_n. Consider $\nabla_n: \Gamma(T^*X) \to \Gamma(L(TX, T^*X))$ as a first order differential operator. Define ∇ to be the *Levi-Civita covariant derivative* of β iff $\nabla_n \to \Delta: \Gamma(T^*X) \to \text{B}'(L(TX, T^*X))$. Note that neither $\mathscr{H}$ nor ∇ depends on the sequence $\{\beta_n\}$ since convergence takes place in a distribution topology. ∇ is a differential operator with distributional coefficients.

By now it should be clear how to define any geometric object usually associated to β. The existence of such objects can be proved by methods analogous to those employed in Theorem 1.1. The only real problem that arises is that resulting from the infamous "multiplication of distributions" syndrome: There is no general theory of products available. In any concrete case one can usually find a way. For example one can use Hörmander's definition of products.[6] One can also define products when the distributions are obtained via the Łojasiewicz division theorem.

4. AN APPLICATION IN GENERAL RELATIVITY

The discovery of stars in various stages of gravitational collapse and the urge to explain the beginning (and possible end) of the universe has caused a renaissance in Einstein's general theory of relativity (GR). It is still the best candidate for a general explanation of gravitation. One frustrating aspect of GR at present is that it seems to predict that we can no longer predict. The theory apparently breaks down when it attempts to explain a completely collapsed star (black hole) or the big bang (a naked singularity). These breakdowns are examples of "spacetime singularities." Thus the first problem that arises is to find a precise mathematical definition of a singularity.

As first it was thought that the known exact solutions had too much symmetry built into them and that generically singularities should not occur. Using geodesic incompleteness as a means of identifying smooth spacetimes from which singular points had been excised, Hawking and Penrose proceeded to prove some rather deep theorems to the effect that any physically reasonable spacetime must be geodesically incomplete.[13] Thus the geometric objects should carry the singular information.

If singularities are truly inevitable, then allowance must be made for them from the beginning. Thus define a *spacetime* (X, β) to be a connected manifold X provided with a distributional Lorentzian structure β. $\Sigma := \text{sing supp}\beta$ will be called the set of *spacetime singularities*. Let p be a generalized geodesic. The following proposition is an easy exercise in distribution theory.

Proposition 4.1: *If $\Sigma \neq \phi$, then the smooth part $p|X \setminus \Sigma$ of a generalized geodesic p that meets Σ is incomplete.* □

This gives the precise relationship of the present definition of singularities to the geodesic incompleteness one. Notice that in this theory Taub–NUT space is nonsingular by definition, without the necessity of ad hoc arguments.

The extension of GR to distributional Lorentzian structures is largely purely formal; for more details see Parker.[4] For example, the Ricci tensor and the scalar curvature exist as distributional tensors. If the product βR is definable, then one can require that β satisfy the Einstein equation

$$\mathrm{Ric}(\beta) - \tfrac{1}{2}\beta R = 8\pi T,$$

where T is also a distributional tensor now. In particular the Einstein equations hold at space time singularities.[14] Also, all the solutions in Carter's four-parameter family[15] are naturally distributional structures on $\mathbb{R}^4$.[14] This family includes the Schwarzschild, Reissner–Nordström, and Kerr structures. One must keep in mind that distributions are not functions.

As an example let us look at the Schwarzschild structure.

On $\mathbb{R}^4$ consider Cartesian coordinates t, x, y, z. The Schwarzschild structure is given by the line element

$$ds^2 = dt^2 - dx^2 - dy^2 - dz^2 - \frac{2m}{r}\left[\frac{xdx + ydy + zdz}{r} + dt\right]^2.$$

This can be regarded as defining a distributional Lorentzian structure β on $\mathbb{R}^4$ given by

$$[\beta^{ij}] = \begin{bmatrix} 1 + \frac{2m}{r} & -2m\frac{x}{r^2} & -2m\frac{y}{r^2} & -2m\frac{z}{r^2} \\ -2m\frac{x}{r^2} & -1 + 2m\frac{x^2}{r^3} & 2m\frac{xy}{r^3} & 2m\frac{xz}{r^3} \\ -2m\frac{y}{r^2} & 2m\frac{xy}{r^3} & -1 + 2m\frac{y^2}{r^3} & 2m\frac{yz}{r^3} \\ -2m\frac{z}{r^2} & 2m\frac{xz}{r^3} & 2m\frac{yz}{r^3} & -1 + 2m\frac{z^2}{r^3} \end{bmatrix}. \tag{4.1}$$

Observe that β is a *Kerr–Schild* structure: $\beta = \eta + f\lambda \otimes \lambda$, where η is a flat Lorentzian structure, $f \in \mathscr{D}'(\mathbb{R}^4)$ and $\lambda \in \mathfrak{X}'$ is lightlike. According to Gürses and Gürsey[16] the full Einstein equations are

$$\partial_l \partial_i [\eta^{ij}\beta^{kl} - \eta^{ik}\beta^{lj} - \eta^{il}\beta^{kj} + \eta^{kl}\beta^{ij}] = 16\pi\eta^{kh}T^{l}_{h}, \tag{4.2}$$

where T is the stress–energy tensor of everything except the gravitational field.

Since each entry in (4.1) is locally integrable and homogeneous of degree -1, it follows from (4.2) that each entry in T is homogeneous of degree -3 and thus some constant multiple of δ_t, where

$$\langle \delta_t, \varphi \rangle := \int_{-\infty}^{\infty} \varphi(t,0,0,0)\, dt, \quad \varphi \in \mathscr{D}(\mathbb{R}^4).$$

The easiest way to compute the constants is by Fourier transforms. Note that since β is independent of t it suffices to compute in x, y, z. The computation is straightforward and details are omitted. One obtains $T^{t}_{t} = m\delta_t$ and all other components vanish. For more details see Parker.[17,4]

ACKNOWLEDGMENTS

I would like to thank Frank Flaherty, my major professor, and Bent Peterson for helpful conversation. This paper is based on Chaps. II and IV of my thesis at Oregon State University.[4]

[1] H. Goldstein, *Classical Mechanics* (Addison-Wesley, Reading, Massachusetts, 1950).
[2] C.J. Isham, Proc. R. Soc. Lond. Ser. A **351**, 209 (1976).
[3] J. Glimm and A. Jaffe, preprint (1977).
[4] P.E. Parker, thesis, Oregon State University (1977).
[5] J.E. Marsden, Arch. Rat. Mech. Anal. **28**, 323 (1968).
[6] L. Hörmander, Acta Math. **127**, 79 (1971).
[7] L.C. Young, Sprawozdania z posieden Towarzystwa Nankwego Warzawskiego, Wydzial III **30**, 211 (1937).
[8] G. DeRham, *Variétés Différentiables* (Hermann, Paris, 1953).
[9] See F. Trèves, *Topological Vector Spaces, Distributions and Kernels* (Academic, New York, 1967), for inverse or projective limits.
[10] J.J. Duistermaat, Commun. Pure Appl. Math. **27**, 207 (1974).
[11] K. Janich, Mat. Ann. **209**, 161 (1974).
[12] V.I. Arnol'd, Funct. Anal. Appl. **1**, 1 (1967).
[13] S.W. Hawking and G.F.R. Ellis, *The Large Scale Structure of Spacetime*, (Cambridge, U.P., Cambridge, 1973).
[14] Cf. p. 287 in Ref. 13.
[15] B. Carter, "Black Hole Equilibrium States," *Black Holes: Les Houches* 1972 (Gordon and Breach, New York, 1973), p. 57.
[16] M. Gürses and F. Gürsey, J. Math. Phys. **16**, 2385 (1975).
[17] P.E. Parker, preprint (1977).

Center of inertia and coordinate transformations in the post-Newtonian charged n-body problem in gravitation

B. M. Barker

Department of Physics and Astronomy, The University of Alabama, University, Alabama 35486

R. F. O'Connell

Department of Physics and Astronomy, Louisiana State University, Baton Rouge, Louisiana 70803

(Received 8 June 1978; revised manuscript received 7 November 1978)

We generalize the field theory propagator by finding a way to make it a function of some additional arbitrary parameters. Thus, it is now possible to obtain Lagrangians (which contain the propagator parameters) from field theory in a more general coordinate system than had previously been possible. We find the n-body (classical) Bażański Lagrangian in this more general coordinate system and we give the relationship between the various coordinate systems by an n-body coordinate transformation involving the propagator parameters. We find the center of inertia for the case of the n-body Baszański Lagrangian in the general coordinate system and find that the potential energy terms $-Gm_im_j/r_{ij}$ and e_ie_j/r_{ij} do *not* in general split equally between particles i and j as they *do* in the case of Bażański coordinates. We also find the center of inertia for the case of the n-body (unchanged) post-Newtonian Lagrangian with parameterized post-Newtonian (PPN) parameters γ and β in standard coordinates, and show that the potential energy terms *do* split equally between a pair of particles.

INTRODUCTION

In a recent paper[1] we started with the two-body Bażański Lagrangian[1,2] (the post-Newtonian Lagrangian for two charged bodies in general relativity) in Bażański coordinates. We then made a coordinate transformation[3] involving two arbitrary dimensionless parameters α_g and α_p to obtain the Lagrangian $\mathscr{L}(\alpha_g,\alpha_p)$ in the new coordinate system. The form of the coordinate transformation[3] was chosen so that the Lagrangian $\mathscr{L}(\alpha_g,\alpha_p)$ would be consistent with the Hamiltonian $\mathscr{H}(\alpha_g,\alpha_p)$ derived from quantum field theory. In deriving $\mathscr{H}(\alpha_g,\alpha_p)$ a graviton propagator[1,4] involving α_g and a photon propagator[1] involving α_p were used. The form of these propagators (first given by Hiida and Okamura[4] for the graviton case) determined the form of the coordinate transformation.

It is now clear to us that the form of the graviton and photon propagators can be generalized by adding additional dimensionless parameters. This will lead to the same additional parameters in the corresponding coordinate transformation.

In Sec. I we shall derive the one-graviton-exchange potential energy term (i.e., term of order G) and the one-photon-exchange potential energy term (i.e., term of order e^2) for two scalar particles where parameters α_g, a_{12g}, and a_{21g} are used in the graviton propagator and parameters α_p, a_{12p}, and a_{21p} are used in the photon propagator; and then in Sec. II we shall give the corresponding two-body coordinate transformations.

In Sec. III we shall give the n-body coordinate transformations and the n-body Lagrangian. They will contain parameters α_{ijg}, a_{ijg} and α_{ijp}, a_{ijp}.

In Sec. IV we shall turn our attention to the topic of center of inertia.[5,6] It is well known[5-7] [for n-body problems involving Darwin (D), Einstein, Infeld, Hoffmann (EIH) or Bażański (B) Lagrangians] that in finding the center of inertia the potential energy terms $-Gm_im_j/r_{ij}$ and e_ie_j/r_{ij} must be split equally between the particles i and j. However, while this $\frac{1}{2}$, $\frac{1}{2}$ split, as we shall call it, holds in Bażański coordinates (same as used by Darwin and EIH) there are certain coordinate systems in which it does *not* hold. We shall show that for the coordinate system introduced in Sec. III the $\frac{1}{2}$, $\frac{1}{2}$ split does not hold in general. We shall also explicitly show what split does occur for this coordinate system. Finally, in Sec. IV we consider the n-body (uncharged) post-Newtonian Lagrangian with parameterized post-Newtonian (PPN) parameters γ and β in the usually given coordinate system and find that the $\frac{1}{2}$, $\frac{1}{2}$ split holds in this case also. In Sec. V we present our conclusions.

I. DERIVATION OF $V_1(r)$ FROM FIELD THEORY

Let us review the most important points in the derivation of the one-graviton exchange interaction as well as the one-photon exchange interaction for the case of two scalar particles. Let e_1, m_1, $\mathbf{P}_1$, E_1, and e_2, m_2, $\mathbf{P}_2$, E_2 denote the charge, rest mass, momentum, and energy of particles 1 and 2 respectively. We also have

$$\mathbf{P}_1=\hbar\mathbf{p},\quad E_1=c\hbar p_0,\quad \lambda_1=m_1c/\hbar,\quad p_\mu^2=-\lambda_1^2, \tag{1}$$

$$\mathbf{P}_2=\hbar\mathbf{q},\quad E_2=c\hbar q_0,\quad \lambda_2=m_2c/\hbar,\quad q_\mu^2=-\lambda_2^2, \tag{2}$$

where p_μ and q_μ are the propagation 4-vectors for particle 1 and 2, respectively. The graviton coupling constant κ is related to Newton's constant of gravitation G and the speed of

 0022-2488/79/071427-08$01.00

light c by the relation

$$\kappa^2 = 16\pi G/c^4. \tag{3}$$

We shall also use Gaussian electromagnetic units.

Consider the gravitational (or electromagnetic) scattering of two particles of spin 0. Let the initial and final propagation four-vectors for particle 1 be p and p', respectively, and those of particle 2 be q and q', respectively. The quantity $V_1(\mathbf{k})$ is defined in terms of the S matrix as

$$S_2 = (-i/c\hbar V^2)(2\pi)^4\delta(p+q-p'-q') \times a_1^*(\mathbf{p}')a_2^*(\mathbf{q}')V_1(\mathbf{k})a_2(\mathbf{q})a_1(\mathbf{p}), \tag{4}$$

where a_1, a_1^* and a_2, a_2^* denote the annihilation and creation operators for particles 1 and 2, respectively, and the factor V is a volume factor. The results for $V_1(\mathbf{k})$ for the one-graviton exchange and the one-photon exchange interactions are, respectively (see the Appendix),

$$V_{1g}(\mathbf{k}) = -\frac{c^2\hbar^2\kappa^2}{4(p_0'p_0q_0'q_0)^{1/2}}\frac{1}{\mathbf{k}^2-k_0^2}[(p'q')(pq)+(p'q)(pq') - (p'p)(q'q) - \lambda_1^2(q'q) - \lambda_2^2(p'p) - 2\lambda_1^2\lambda_2^2], \tag{5}$$

$$V_{1p}(\mathbf{k}) = \frac{-\pi e_1e_2}{(p_0'p_0q_0'q_0)^{1/2}}\frac{1}{\mathbf{k}^2-k_0^2}[(p'+p)(q'+q)], \tag{6}$$

where $k=p'-p=q-q'$. The quantity $V_1(\mathbf{k})$ can also be defined in terms of the potential energy $V_1(\mathbf{r})$ as

$$V_1(\mathbf{k}) = \int d\mathbf{r}\,\exp(-i\mathbf{p}'\cdot\mathbf{r}_1)\exp(-i\mathbf{q}'\cdot\mathbf{r}_2)V_1(\mathbf{r})\exp(i\mathbf{p}\cdot\mathbf{r}_1) \times\exp(i\mathbf{q}\cdot\mathbf{r}_2), \tag{7}$$

where $\mathbf{r}\equiv\mathbf{r}_1-\mathbf{r}_2$ and $\mathbf{r}_1$ and $\mathbf{r}_2$ are the position vectors of particles 1 and 2, respectively. The potential energy $V_1(\mathbf{r})$ is a Hermitian operator and is also momentum dependent [i.e., $V_1(\mathbf{r})\equiv V_1(\mathbf{r},\mathbf{p}_{\rm op},\mathbf{q}_{\rm op})$ where $\mathbf{p}_{\rm op}$ and $\mathbf{q}_{\rm op}$ are operators].

If we are only interested in the classical result, as we are in this paper, the ordering of the factors in $V_1(\mathbf{r},\mathbf{p}_{\rm op},\mathbf{q}_{\rm op})$ makes no difference (i.e., we can neglect delta function terms) and we can thus write Eqs. (5)–(7) as

$$V_{1g}(\mathbf{k}) = -\frac{c^2\hbar^2\kappa^2}{4p_0q_0}\frac{1}{\mathbf{k}^2-k_0^2}[2(pq)^2-\lambda_1^2\lambda_2^2], \tag{8}$$

$$V_{1p}(\mathbf{k}) = -\frac{\pi e_1e_2}{p_0q_0}\frac{1}{\mathbf{k}^2-k_0^2}[4(pq)], \tag{9}$$

$$V_1(\mathbf{k}) = \int d\mathbf{r}\,e^{-i\mathbf{k}\cdot\mathbf{r}}V_1(\mathbf{r}). \tag{10}$$

The inverse of Eq. (10) is

$$V_1(\mathbf{r}) = \frac{1}{(2\pi)^3}\int d\mathbf{k}\,e^{i\mathbf{k}\cdot\mathbf{r}}V_1(\mathbf{k}). \tag{11}$$

In this paper we shall be interested only in the post-Newtonian approximation in which case Eqs. (8) and (9) can be approximated as

$$V_{1g}(\mathbf{k}) = -\frac{c^2\hbar^2\kappa^2\lambda_1\lambda_2}{4\mathbf{k}^2}\left(1+\frac{3}{2}\frac{\mathbf{p}^2}{\lambda_1^2}+\frac{3}{2}\frac{\mathbf{q}^2}{\lambda_2^2} - 4\frac{\mathbf{p}\cdot\mathbf{q}}{\lambda_1\lambda_2}+\frac{k_0^2}{\mathbf{k}^2}\right), \tag{12}$$

$$V_{1p}(\mathbf{k}) = \frac{4\pi e_1e_2}{\mathbf{k}^2}\left(1-\frac{\mathbf{p}\cdot\mathbf{q}}{\lambda_1\lambda_2}+\frac{k_0^2}{\mathbf{k}^2}\right). \tag{13}$$

A. The propagator term

We shall call $1/(\mathbf{k}^2-k_0^2)$ the propagator term and in Eq. (12) and (13) it was expanded as

$$\frac{1}{\mathbf{k}^2-k_0^2} = \frac{1}{\mathbf{k}^2}+\frac{k_0^2}{\mathbf{k}^4}, \tag{14}$$

where

$$\mathbf{k} = \mathbf{p}'-\mathbf{p} = \mathbf{q}-\mathbf{q}', \tag{15}$$

$$k_0 = p_0'-p_0 = q_0-q_0', \tag{16}$$

and

$$p_0'-p_0 = \frac{p_0'^2-p_0^2}{p_0'+p_0} = \frac{\mathbf{p}'^2-\mathbf{p}^2}{p_0'+p_0} = \frac{\mathbf{k}\cdot(\mathbf{p}'+\mathbf{p})}{p_0'+p_0}, \tag{17}$$

$$q_0-q_0' = \frac{q_0^2-q_0'^2}{q_0+q_0'} = \frac{\mathbf{q}^2-\mathbf{q}'^2}{q_0+q_0'} = \frac{\mathbf{k}\cdot(\mathbf{q}+\mathbf{q}')}{q_0+q_0'}. \tag{18}$$

Since we are interested in the classical post-Newtonian result, Eqs. (17) and (18) may be expressed as

$$k_0 = \frac{\mathbf{k}\cdot\mathbf{p}}{\lambda_1} \quad\text{or}\quad k_0 = \frac{\mathbf{k}\cdot\mathbf{q}}{\lambda_2}. \tag{19}$$

The factor k_0^2 which appears in Eqs. (12) and (13) may be written in a symmetrical way as[1,4,8]

$$k_0^2 = (1+4\alpha)\left(\frac{(\mathbf{k}\cdot\mathbf{p})(\mathbf{k}\cdot\mathbf{q})}{\lambda_1\lambda_2}\right) - 2\alpha\left(\frac{(\mathbf{k}\cdot\mathbf{p})^2}{\lambda_1^2}+\frac{(\mathbf{k}\cdot\mathbf{q})^2}{\lambda_2^2}\right). \tag{20}$$

At this stage we must consider $\mathbf{k}$ as an independent variable [i.e., we no longer use Eq. (15)] since we will be using Eq. (11) to obtain $V_1(\mathbf{r})$. For the special case $\alpha=0$ the resulting Hamiltonian (or Lagrangian) will be in the coordinate system[1,4] of B–EIH–D.

If we go to center-of-mass coordinates[9] (i.e., total momentum equals zero) where $\mathbf{p}=-\mathbf{q}$ we then have, from Eq. (20),

$$k_0^2 = -(1+4\alpha)\left(\frac{(\mathbf{k}\cdot\mathbf{p})^2}{\lambda_1\lambda_2}\right) - 2\alpha\left(\frac{(\mathbf{k}\cdot\mathbf{p})^2}{\lambda_1^2}+\frac{(\mathbf{k}\cdot\mathbf{p})^2}{\lambda_2^2}\right). \tag{21}$$

Thus, $k_0^2=0$ in the special case where

$$\frac{1+4\alpha}{\lambda_1\lambda_2}+2\alpha\left(\frac{1}{\lambda_1^2}+\frac{1}{\lambda_2^2}\right)=0, \tag{22}$$

which implies that[1,10]

$$\alpha = -\frac{\lambda_1\lambda_2}{2(\lambda_1+\lambda_2)^2} = -\frac{m_1m_2}{2(m_1+m_2)^2}. \tag{23}$$

The one-graviton-exchange interaction for particles of various spins in center-of-mass coordinates (with $k_0^2=0$) was first given by Barker, Gupta, and Haracz[11] using Gupta's[12] quantum theory of gravitation.

In Ref. 1 we first suggested that there could be *two* independent α's, an α_p in the photon propagator as well as an α_g in the graviton propagator.

Let us now generalize the form of the propagator term so that k_0^2 will have the form

$$k_0^2 = [1+2\alpha(a_{12}+a_{21})]\left(\frac{(\mathbf{k}\cdot\mathbf{p})(\mathbf{k}\cdot\mathbf{q})}{\lambda_1\lambda_2}\right) - 2\alpha\left(a_{12}\frac{(\mathbf{k}\cdot\mathbf{p})^2}{\lambda_1^2} + a_{21}\frac{(\mathbf{k}\cdot\mathbf{q})^2}{\lambda_2^2}\right), \tag{24}$$

where a_{12} and a_{21} are new dimensionless parameters. It is to be understood that the α's and a's are to have a subscript g when used in the graviton propagator and a subscript p when used in the photon propagator. As we would like our final Hamiltonian or Lagrangian to be in a symmetrical form with respect to particles 1 and 2 we must impose symmetry conditions on $\alpha\equiv\alpha(m_1,m_2)$ and on $a_{12}\equiv a_{12}(m_1,m_2)$ and[13] $a_{21}\equiv a_{21}(m_2,m_1)$ of the form

$$\alpha(m_2,m_1) = \alpha(m_1,m_2), \tag{25}$$

$$a_{21}(m_2,m_1) = a_{12}(m_2,m_1). \tag{26}$$

Note that α of Eq. (23) satisfies Eq. (25).

Let us again go to center-of-mass coordinates in which case

$$k_0^2 = \left[-\frac{1}{\lambda_1\lambda_2} - 2\alpha\left(\frac{a_{12}+a_{21}}{\lambda_1\lambda_2} + \frac{a_{12}}{\lambda_1^2} + \frac{a_{21}}{\lambda_2^2}\right)\right](\mathbf{k}\cdot\mathbf{p})^2. \tag{27}$$

If we want the above result to be the same as for the case when $a_{12}=a_{21}=1$ we must have [from Eqs. (1) and (2) we note that $m_1/m_2=\lambda_1/\lambda_2$]

$$a_{12}m_2^2 + (a_{12}+a_{21})m_1m_2 + a_{21}m_1^2 = (m_1+m_2)^2. \tag{28}$$

Dividing Eq. (28) by m_1+m_2 gives us

$$a_{12}m_2 + a_{21}m_1 = m_1+m_2. \tag{29}$$

The solution to Eq. (29) which also satisfies the symmetry condition of Eq. (26) is

$$a_{12} = [(1-a_0)m_1 + a_0m_2]/m_2, \tag{30}$$

$$a_{21} = [(1-a_0)m_2 + a_0m_1]/m_1, \tag{31}$$

where

$$a_0 \equiv a_0(m_1,m_2) = a_0(m_2,m_1). \tag{32}$$

It can easily be shown that Eq. (29) does not restrict the result of Eq. (24) since it is always possible to remove a factor from the a's and absorb it in α in such a way as the resulting a's will satisfy Eq. (29). We shall thus require a_{12} and a_{21} to satisfy Eq. (29). Let us also note that if a_{12} and a_{21} are not mass dependent then $a_{12}=a_{21}=1$ and there will be no generalization of the propagator.

B. Results for $V_{1g}(\mathbf{r})$ and $V_{1p}(\mathbf{r})$

Using the form of k_0^2 given by Eq. (24) in Eqs. (12) and (13) together with Eq. (11) gives us

$$\begin{aligned}V_{1g}(\mathbf{r}) = &-\frac{Gm_1m_2}{r}\Bigg[1+\left(\frac{3}{2}-\alpha_g a_{12g}\right)\frac{\mathbf{P}_1^2}{m_1^2c^2} + \left(\frac{3}{2}-\alpha_g a_{21g}\right)\frac{\mathbf{P}_2^2}{m_2^2c^2}\\ &+ [\alpha_g(a_{12g}+a_{21g}) - \tfrac{7}{2}]\frac{\mathbf{P}_1\cdot\mathbf{P}_2}{m_1m_2c^2}\\ &- [\tfrac{1}{2}+\alpha_g(a_{12g}+a_{21g})]\frac{(\mathbf{P}_1\cdot\mathbf{r})(\mathbf{P}_2\cdot\mathbf{r})}{m_1m_2c^2r^2}\\ &+ \alpha_g\left(a_{12g}\frac{(\mathbf{P}_1\cdot\mathbf{r})^2}{m_1^2c^2r^2} + a_{21g}\frac{(\mathbf{P}_2\cdot\mathbf{r})^2}{m_2^2c^2r^2}\right)\Bigg],\end{aligned} \tag{33}$$

$$\begin{aligned}V_{1p}(\mathbf{r}) = &\frac{e_1e_2}{r}\Bigg[1-\alpha_p\left(a_{12p}\frac{\mathbf{P}_1^2}{m_1^2c^2} + a_{21p}\frac{\mathbf{P}_2^2}{m_2^2c^2}\right)\\ &+ [\alpha_p(a_{12p}+a_{21p}) - \tfrac{1}{2}]\frac{\mathbf{P}_1\cdot\mathbf{P}_2}{m_1m_2c^2}\\ &- [\tfrac{1}{2}+\alpha_p(a_{12p}+a_{21p})]\frac{(\mathbf{P}_1\cdot\mathbf{r})(\mathbf{P}_2\cdot\mathbf{r})}{m_1m_2c^2r^2}\\ &+ \alpha_p\left(a_{12p}\frac{(\mathbf{P}_1\cdot\mathbf{r})^2}{m_1^2c^2r^2} + a_{21p}\frac{(\mathbf{P}_2\cdot\mathbf{r})^2}{m_2^2c^2r^2}\right)\Bigg],\end{aligned} \tag{34}$$

where a_{12g}, a_{21g} and a_{12p}, a_{21p} must satisfy Eqs. (29)–(32). Of course subscripts g and p must be added to the a's of these equations for use in Eqs. (33) and (34), respectively.

II. TWO-BODY COORDINATE TRANSFORMATIONS

Let us now consider the two-body coordinate transformations

$$\mathbf{r}_{1B} = \mathbf{r}_1 - \mathbf{r}\left(\alpha_g a_{12g}\frac{Gm_2}{c^2r} - \alpha_p a_{12p}\frac{e_1e_2}{m_1c^2r}\right), \tag{35}$$

$$\mathbf{r}_{2B} = \mathbf{r}_2 + \mathbf{r}\left(\alpha_g a_{21g}\frac{Gm_1}{c^2r} - \alpha_p a_{21p}\frac{e_1e_2}{m_2c^2r}\right), \tag{36}$$

relating the Bażański coordinates $\mathbf{r}_{1B}$, $\mathbf{r}_{2B}$, to the new coordinates $\mathbf{r}_1$, $\mathbf{r}_2$. It can readily be shown that if one starts with the two-body Bażański Lagrangian in Bażański coordinates, and then makes the transformations of Eqs. (35)–(36) to obtain the Lagrangian in the new coordinates, then the corresponding Hamiltonian in the new coordinates will be in agreement with Eqs. (33) and (34). Note that the Hamiltonian will also contain G^2, e^4, and Ge^2 terms which, if derived from field theory, must come from fourth-order S matrix calculations.

From Eqs. (35)–(36) we obtain

$$\mathbf{r}_B = \mathbf{r}\left(1 - \alpha_g \frac{GM}{c^2 r} + \alpha_p \frac{e_1 e_2}{\mu c^2 r}\right), \tag{37}$$

where $\mathbf{r}_B = \mathbf{r}_{1B} - \mathbf{r}_{2B}$ and the reduced mass and total mass are given, respectively, by $\mu = m_1 m_2/(m_1 + m_2)$, and $M = m_1 + m_2$. In obtaining Eq. (37) we have made use of Eq. (29), and for this reason Eq. (37) does not depend on the a's.

Let us now consider some special cases of Eqs. (35)–(36).

A. Result with $a_{0g} = a_{0p} = 1$

From Eqs. (30) and (31) we obtain $a_{12g} = a_{12p} = 1$, and $a_{21g} = a_{21p} = 1$, which gives us the transformations[3] corresponding to the propagator with k_0^2 in the form of Eq. (20).

B. Result with $a_{0g} = a_{0p} = \frac{1}{2}$

From Eqs. (30) and (31) we obtain $a_{12g} = a_{12p} = M/2m_2$, and $a_{21g} = a_{21p} = M/2m_1$, which gives us the transformations[14] that were considered in Ref. 1. We rejected these transformations in Ref. 1 since they did *not* correspond to the propagator with k_0^2 in the form of Eq. (20). However, in the light of this present paper they *are* consistent with the field theory results.

C. Result with $a_{0g} = a_{0p} = 0$

From Eqs. (30) and (31) we obtain $a_{12g} = a_{12p} = m_1/m_2$, and $a_{21g} = a_{21p} = m_2/m_1$, which is the simplest form (with mass dependence) that the a's can have.

In Ref. 1 we also considered the transformations[15]

$$\mathbf{r}_{1B} = \mathbf{r}_1\left(1 - \alpha_g \frac{GM}{c^2 r} + \alpha_p \frac{e_1 e_2}{\mu c^2 r}\right), \tag{38}$$

$$\mathbf{r}_{2B} = \mathbf{r}_2\left(1 - \alpha_g \frac{GM}{c^2 r} + \alpha_p \frac{e_1 e_2}{\mu c^2 r}\right), \tag{39}$$

from which we can also obtain Eq. (37). These transformations have some unusual properties[15] and are *not* a special case of Eqs. (35)–(36). We have not found a propagator where Eqs. (38)–(39) would be the corresponding transformations and we are not sure that such a propagator exists.

III. n-BODY RESULTS

We shall now generalize our results to the case of n bodies.

A. n-body coordinate transformations

The n-body generalization of Eqs. (35)–(36) is

$$\mathbf{r}_{iB} = \mathbf{r}_i - \sum_{\substack{j=1\\ j\neq i}}^{n} \mathbf{r}_{ij}\left(A^g_{ij} \frac{Gm_j}{c^2 r_{ij}} - A^p_{ij} \frac{e_i e_j}{m_i c^2 r_{ij}}\right), \tag{40}$$

where $\mathbf{r}_{ij} \equiv \mathbf{r}_i - \mathbf{r}_j$ and (for the sake of brevity) we have set $A^g_{ij} \equiv \alpha_{ijg} a_{ijg}$ and $A^p_{ij} \equiv \alpha_{ijp} a_{ijp}$. Note that only the α_{ij}'s and a_{ij}'s for which $i\neq j$ are used. The symmetry conditions on $\alpha_{ijg} \equiv \alpha_g(m_i, m_j)$ and on $a_{ijg} \equiv a_{ijg}(m_i, m_j)$ are

$$\alpha_g(m_j, m_i) = \alpha_g(m_i, m_j), \tag{41}$$

$$a_{jig}(m_j, m_i) = a_{ijg}(m_j, m_i), \tag{42}$$

and thus $\alpha_{jig} = \alpha_{ijg}$. Note for the two-body problem there would be only one of the quantities $\alpha_g \equiv \alpha_{12g} \equiv \alpha_{21g}$.

We shall also require that

$$a_{ijg} m_j + a_{jig} m_i = m_i + m_j, \tag{43}$$

which is the n-body generalization of Eq. (29). The solution of Eq. (43) which satisfies the symmetry condition of Eq. (42) is

$$a_{ijg} = [(1 - a_{0ijg}) m_i + a_{0ijg} m_j]/m_j, \tag{44}$$

where[16]

$$a_{0ijg} \equiv a_{0g}(m_i, m_j) = a_{0g}(m_j, m_i), \tag{45}$$

and thus $a_{0jig} = a_{0ijg}$. For the two-body problem we have $a_{0g} \equiv a_{012g} = a_{021g}$. For α_{ijp} and a_{ijp} we will have a set of equations similar to Eqs. (41)–(45), where the subscript g is replaced by p. Note also that all the α's and a's must be dimensionless.

From Eqs. (40) and (43) we obtain

$$\mathbf{r}_{ijB} = \mathbf{r}_{ij}\left(1 - \alpha_{ijg} \frac{GM_{ij}}{c^2 r_{ij}} + \alpha_{ijp} \frac{e_i e_j}{\mu_{ij} c^2 r_{ij}}\right) - \sum_{\substack{k=1\\ k\neq i\\ k\neq j}}^{n} \left[\mathbf{r}_{ik}\left(A^g_{ik} \frac{Gm_k}{c^2 r_{ik}} - A^p_{ik} \frac{e_i e_k}{m_i c^2 r_{ik}}\right) - \mathbf{r}_{jk}\left(A^g_{jk} \frac{Gm_k}{c^2 r_{jk}} - A^p_{jk} \frac{e_j e_k}{m_j c^2 r_{jk}}\right)\right], \tag{46}$$

where

$$\mu_{ij} = m_i m_j/(m_i + m_j) \quad \text{and} \quad M_{ij} = m_i + m_j. \tag{47}$$

B. n-body Lagrangian

Starting with the n-body Bażański Lagrangian[1,2] in Bażański coordinates and using Eq. (40) gives us the n-body Lagrangian in the new coordinates, as

$$\mathscr{L} = \sum_{i=1}^{n} (-m_i c^2 + \tfrac{1}{2} m_i v_i^2 + \tfrac{1}{8} m_i v_i^4/c^2) + \tfrac{1}{2} \sum_{i,j=1}^{n}{}' \left[\frac{Gm_i m_j}{r_{ij}}\left(1 + (3 - 2A^g_{ij}) \frac{v_i^2}{c^2} + (2A^g_{ij} - \tfrac{7}{2}) \times \frac{\mathbf{v}_i\cdot\mathbf{v}_j}{c^2} - (\tfrac{1}{2} + 2A^g_{ij}) \frac{(\mathbf{v}_i\cdot\mathbf{r}_{ij})(\mathbf{v}_j\cdot\mathbf{r}_{ij})}{c^2 r_{ij}^2} + 2A^g_{ij} \frac{(\mathbf{v}_i\cdot\mathbf{r}_{ij})^2}{c^2 r_{ij}^2}\right)\right.$$

$$-\frac{e_i e_j}{r_{ij}}\Big(1 - 2A^p_{ij}\frac{v_i^2}{c^2} + (2A^p_{ij} - \tfrac{1}{2})\frac{\mathbf{v}_i\cdot\mathbf{v}_j}{c^2}$$

$$- (\tfrac{1}{2} + 2A^p_{ij})\frac{(\mathbf{v}_i\cdot\mathbf{r}_{ij})(\mathbf{v}_j\cdot\mathbf{r}_{ij})}{c^2 r_{ij}^2} + 2A^p_{ij}\frac{(\mathbf{v}_i\cdot\mathbf{r}_{ij})^2}{c^2 r_{ij}^2}\Big)$$

$$+ (\alpha_{ijg} - \tfrac{1}{2})\frac{G^2 m_i m_j M_{ij}}{c^2 r_{ij}^2} + \alpha_{ijp}\frac{e_i^2 e_j^2}{\mu_{ij} c^2 r_{ij}^2}$$

$$+ (1 - \alpha_{ijg} - \alpha_{ijp})\frac{G e_i e_j M_{ij}}{c^2 r_{ij}^2} - \frac{G(e_i^2 m_j + e_j^2 m_i)}{2c^2 r_{ij}^2}\Big]$$

$$+ \sum_{i,j,k=1}^{n}{}' \Big\{ -\frac{G^2 m_i m_j m_k}{2c^2 r_{ij} r_{ik}} + \frac{G e_i e_j m_k}{2c^2}$$

$$\times\Big(\frac{1}{r_{ij} r_{ik}} + \frac{1}{r_{ji} r_{jk}} - \frac{1}{r_{ki} r_{kj}}\Big)$$

$$+ \Big[A^g_{ik}(G^2 m_i m_j m_k - G e_i e_j e_k)$$

$$+ A^p_{ik}\Big(\frac{e_i^2 e_j e_k}{m_i} - G e_i m_j e_k\Big)\Big]\frac{\mathbf{r}_{ij}\cdot\mathbf{r}_{ik}}{c^2 r_{ij}^3 r_{ik}}\Big\}, \tag{48}$$

where Σ' means that no two summation indices are the same. If we set $\alpha_{ijg} = \alpha_{ijp} = 0$ (which implies that $A^g_{ij} = A^p_{ij} = 0$) in Eq. (48) we will obtain the n-body Bażański Lagrangian in *Bażański coordinates*. The rest-energy terms (which do not effect the equations of motion) have been included in the above Lagrangian and are needed for our treatment of the center of inertia in the following section.

IV. CENTER OF INERTIA

Let us start with an n-body Lagrangian of the form $\mathscr{L} = \mathscr{L}(\mathbf{r}_{ij}, \mathbf{v}_k)$, where $\mathscr{L}$ is a scalar in three dimensions. We define, as usual, $\mathbf{P}_i \equiv \partial\mathscr{L}/\partial\mathbf{v}_i$ and $\mathbf{L}_i \equiv \mathbf{r}_i \times \mathbf{P}_i$. The total energy[17]

$$\mathscr{E} = \sum_{i=1}^{n} \mathbf{P}_i\cdot\mathbf{v}_i - \mathscr{L} \tag{49}$$

is conserved since $\mathscr{L}$ is not an explicit function of time. The total momentum

$$\mathbf{P} = \sum_{i=1}^{n} \mathbf{P}_i \tag{50}$$

is conserved since $\mathscr{L}$ is a function of the *differences* in coordinates $\mathbf{r}_{ij}$. The total angular momentum[17]

$$\mathbf{L} = \sum_{i=1}^{n} \mathbf{L}_i \tag{51}$$

is conserved since $\mathscr{L}$ is a scalar in three dimensions. Since $\mathbf{P}_i$ and $\mathbf{L}_i$ are explicitly defined there is no problem in finding $\mathscr{E}$, $\mathbf{P}$, or $\mathbf{L}$. Let us note that, so far, an $\mathscr{E}_i$ where

$$\mathscr{E} = \sum_{i=1}^{n} \mathscr{E}_i \tag{52}$$

has not been defined. If we did have such an $\mathscr{E}_i$ we could then define the center of inertia $\mathbf{r}_{\rm CI}$ by the definition[18]

$$\mathscr{E}\mathbf{r}_{\rm CI} \equiv \sum_{i=1}^{n} \mathscr{E}_i \mathbf{r}_i. \tag{53}$$

We also wish to require that

$$(\mathscr{E}/c^2)\mathbf{v}_{\rm CI} = \mathbf{P}, \tag{54}$$

which will hold if

$$\frac{d}{dt}\Big(\sum_{i=1}^{n} \frac{\mathscr{E}_i}{c^2}\mathbf{r}_i\Big) = \sum_{i=1}^{n} \mathbf{P}_i. \tag{55}$$

We thus must find $\mathscr{E}_i$ such that Eqs. (52) *and* (55) *are satisfied*. In order to show that Eq. (55) is satisfied we will always have to use the equations of motion

$$\dot{\mathbf{P}}_i - \partial\mathscr{L}/\partial\mathbf{r}_i = 0. \tag{56}$$

In order to have $\mathscr{E}_i$ which satisfies Eqs. (52) and (55) (even in the one-body case where $\mathscr{E}_1 = \mathscr{E}$) the Lagrangian must satisfy additional symmetry conditions.[6] We *will not* be concerned with such symmetry conditions in this paper. We *will* be interested in finding (and do find) $\mathscr{E}_i$ which satisfy Eqs. (52) and (55) for some particular Lagrangians.

Let us first look at a very simple case[6] that works out exactly to all orders (in c).

A. Free particle case

The Lagrangian is

$$\mathscr{L} = \sum_{i=1}^{n} -m_i c^2 (1 - v_i^2/c^2)^{1/2}, \tag{57}$$

from which it follows that

$$\mathbf{P}_i = m_i \mathbf{v}_i/(1 - v_i^2/c^2)^{1/2}, \tag{58}$$

$$\mathscr{E} = \sum_{i=1}^{n} m_i c^2/(1 - v_i^2/c^2)^{1/2}. \tag{59}$$

We shall choose

$$\mathscr{E}_i = m_i c^2/(1 - v_i^2/c^2)^{1/2}, \tag{60}$$

and note that Eq. (52) is satisfied. We also have

$$\frac{d}{dt}\Big(\frac{\mathscr{E}_i}{c^2}\mathbf{r}_i\Big) = \frac{m_i \mathbf{v}_i}{(1 - v_i^2/c^2)^{1/2}} + \frac{m_i \mathbf{r}_i(\mathbf{v}_i\cdot\mathbf{a}_i)/c^2}{(1 - v_i^2/c^2)^{3/2}}. \tag{61}$$

Since $\mathbf{a}_i = 0$ from the equations of motion, Eq. (55) follows from Eq. (61) after summing.

B. Post-Newtonian case for general relativity (with charge)

For this case we shall use the Lagrangian of Eq. (48). We will have to check that both sides of Eq. (55) are in agreement to order c^{-2}. Thus, for the left-hand side we will need the rest energy and Newtonian kinetic and potential energy terms (but *not* post-Newtonian energy terms), while for the right-hand side we will need the Newtonian and post-Newtonian momentum terms.

Let us try $\mathscr{E}_i$ of the form

$$\mathscr{E}_i = m_i c^2 + \tfrac{1}{2} m_i v_i^2 + \sum_{\substack{j=1\\ j\neq i}}^{n} \left(-(\tfrac{1}{2} + b_{ijg}) \frac{G m_i m_j}{r_{ij}} + (\tfrac{1}{2} + b_{ijp}) \frac{e_i e_j}{r_{ij}} \right), \quad (62)$$

where b_{ijg} and b_{ijp} are to be determined. It follows that

$$\mathscr{E} = \sum_{i=1}^{n} \left(m_i c^2 + \frac{1}{2} m_i v_i^2 \right) + \sum_{i,j=1}^{n}{}' \left(-\frac{1}{2} \frac{G m_i m_j}{r_{ij}} + \frac{1}{2} \frac{e_i e_j}{r_{ij}} \right), \quad (63)$$

if we require that

$$b_{ijg} = -b_{jig}, \quad b_{ijp} = -b_{jip}. \quad (64)$$

Clearly, Eq. (63) is the energy corresponding to the Lagrangian of Eq. (48) to the order that we need.

Using Eq. (62) in the left-hand side of Eq. (55) and eliminating the acceleration term by using the equations of motion

$$m_i \mathbf{a}_i = \sum_{\substack{j=1\\ j\neq i}}^{n} \left(-\frac{G m_i m_j}{r_{ij}^2} + \frac{e_i e_j}{r_{ij}^2} \right) \frac{\mathbf{r}_{ij}}{r_{ij}}, \quad (65)$$

we obtain

$$\frac{d}{dt}\left(\sum_{i=1}^{n} \frac{\mathscr{E}_i}{c^2} \mathbf{r}_i \right) = \sum_{i=1}^{n} (m_i \mathbf{v}_i + \tfrac{1}{2} m_i v_i^2 \mathbf{v}_i / c^2)$$
$$+ \sum_{i,j=1}^{n}{}' \left[-\frac{G m_i m_j}{r_{ij}} \left((\tfrac{1}{2} + b_{ijg}) \frac{\mathbf{v}_i}{c^2} + (\tfrac{1}{2} - b_{ijg}) \frac{(\mathbf{v}_i \cdot \mathbf{r}_{ij})}{c^2 r_{ij}^2} \mathbf{r}_{ij} \right) + \frac{e_i e_j}{r_{ij}} \left((\tfrac{1}{2} + b_{ijp}) \frac{\mathbf{v}_i}{c^2} + (\tfrac{1}{2} - b_{ijp}) \frac{(\mathbf{v}_i \cdot \mathbf{r}_{ij})}{c^2 r_{ij}^2} \mathbf{r}_{ij} \right) \right]. \quad (66)$$

We next find that $\Sigma_{i=1}^{n} \mathbf{P}_i$ is exactly the same as the right-hand side of Eq. (66) if we put

$$b_{ijg} = \alpha_{ijg}(a_{ijg} - a_{jig}), \quad b_{ijp} = \alpha_{ijp}(a_{ijp} - a_{jip}), \quad (67)$$

which is consistent with Eq. (64).

Thus in the special case of Bażański coordinates (α's $= 0$) or coordinates systems where α's$\neq 0$ but a's $= 1$, the b's $= 0$ and Eq. (62) gives the $\frac{1}{2},\frac{1}{2}$ split (i.e., the potential energy terms $-Gm_i m_j/r_{ij}$ and $e_i e_j/r_{ij}$ are split equally between the particles i and j).

On the other hand, for coordinate systems where α's$\neq 0$ and a's$\neq 1$ we do *not* get the $\frac{1}{2},\frac{1}{2}$ split.

C. Post-Newtonian case with γ and β (without charge)

We shall next consider the n-body (uncharged) post-Newtonian Lagrangian with PPN parameters γ and β which can be written as[19,20]

$$\mathscr{L} = \sum_{i=1}^{n} (-m_i c^2 + \tfrac{1}{2} m_i v_i^2 + \tfrac{1}{8} m_i v_i^4 / c^2)$$
$$+ \tfrac{1}{2} \sum_{i,j=1}^{n}{}' \left[\frac{G m_i m_j}{r_{ij}} \left(1 + (1 + 2\gamma) \frac{v_i^2}{c^2} - (\tfrac{3}{2} + 2\gamma) \frac{\mathbf{v}_i \cdot \mathbf{v}_j}{c^2} - \frac{1}{2} \frac{(\mathbf{v}_i \cdot \mathbf{r}_{ij})(\mathbf{v}_j \cdot \mathbf{r}_{ij})}{c^2 r_{ij}^2} \right) + (\tfrac{1}{2} - \beta) \frac{G^2 m_i m_j M_{ij}}{c^2 r_{ij}^2} \right]$$
$$+ \sum_{i,j,k=1}^{n}{}' \left((\tfrac{1}{2} - \beta) \frac{G^2 m_i m_j m_k}{c^2 r_{ij} r_{ik}} \right), \quad (68)$$

in the standard coordinate system (i.e., the Lagrangian becomes the same as EIH Lagrangian in EIH coordinates when $\gamma = \beta = 1$).

If we use $\mathscr{E}_i$ in the form of the $\frac{1}{2},\frac{1}{2}$ split, that is

$$\mathscr{E}_i = m_i c^2 + \tfrac{1}{2} m_i v_i^2 + \sum_{\substack{j=1\\ j\neq i}}^{n} \left(-\frac{1}{2} \frac{G m_i m_j}{r_{ij}} \right), \quad (69)$$

we find that Eq. (55) is satisfied. It should be noted that while $\mathbf{P}_i$ contains γ, $\Sigma_{i=1}^{n} \mathbf{P}_i$ does not. We find that

$$\mathbf{P}_i = m_i \mathbf{v}_i + \tfrac{1}{2} m_i v_i^2 \mathbf{v}_i / c^2 + \sum_{\substack{j=1\\ j\neq i}}^{n} \left[-\frac{G m_i m_j}{r_{ij}} \left(-(1 + 2\gamma) \frac{\mathbf{v}_i}{c^2} + (\tfrac{3}{2} + 2\gamma) \frac{\mathbf{v}_j}{c^2} + \frac{1}{2} \frac{(\mathbf{v}_j \cdot \mathbf{r}_{ij})}{c^2 r_{ij}^2} \mathbf{r}_{ij} \right) \right], \quad (70)$$

$$\sum_{i=1}^{n} \mathbf{P}_i = \sum_{i=1}^{n} (m_i \mathbf{v}_i + \tfrac{1}{2} m_i v_i^2 \mathbf{v}_i / c^2) + \sum_{i,j=1}^{n}{}' \left[-\frac{G m_i m_j}{r_{ij}} \left(\frac{1}{2} \frac{\mathbf{v}_i}{c^2} + \frac{1}{2} \frac{(\mathbf{v}_i \cdot \mathbf{r}_{ij})}{c^2 r_{ij}^2} \mathbf{r}_{ij} \right) \right]. \quad (71)$$

Will[21] has shown, for a perfect fluid system, that the $\frac{1}{2},\frac{1}{2}$ split holds in any fully-conservative theory of gravity (these theories contain only the PPN parameters γ and β). However, this result as well as our result depends on the fact that a particular coordinate system has been used.

V. CONCLUSION

We have found a way to generalize the propagator [see Eqs. (14) and (24)] by making it a function of the parameters a_{12} and a_{21} as well as the parameter α which had been included before.[1,4] Thus, it is now possible to directly obtain Hamiltonians from field theory in a wider variety of coordinate systems than had previously been possible. The relationship between the various coordinate systems has been given by our n-body coordinate transformations of Eq. (40).

We have also found the center of inertia for two cases involving post-Newtonian n-body Lagrangians. For the case of the Lagrangian with PPN parameters γ and β in standard coordinates [see Eq. (68)] we found that the $\frac{1}{2},\frac{1}{2}$ split of the potential energy was correct as had previously been found[5,6,7] to be the case for the Bażański, EIH, and Darwin Lagrangians in standard coordinates. However, for the Bażański Lagrangian of Eq. (48), which is in a more general coordinate system, the split was found *not* to be $\frac{1}{2},\frac{1}{2}$ in general, but was found instead to be in accordance with Eqs. (62) and (67).

APPENDIX

In this Appendix we shall fill in some of the intermediate steps leading to Eqs. (5) and (6). Our field theory notation is similar to that of Gupta,[22] except that we are using Gaussian units instead of rationalized Gaussian units. Let U and U be complex scalar fields with mass m_1, charge e_1, and mass m_2, charge e_2, respectively. The interaction terms (using ordered products[22]) with the photon field, A_μ, are[22]

$$:L': = \frac{ie_1}{c\hbar}:\left(\frac{\partial U_1^*}{\partial x_\mu}U_1 - \frac{\partial U_1}{\partial x_\mu}U_1^*\right)A_\mu: - \frac{e_1^2}{c^2\hbar^2}:U_1^*U_1A_\mu^2: + \frac{ie_2}{c\hbar}:\left(\frac{\partial U_2^*}{\partial x_\mu}U_2 - \frac{\partial U_2}{\partial x_\mu}U_2^*\right)A_\mu: - \frac{e_2^2}{c^2\hbar^2}:U_2^*U_2A_\mu^2:, \tag{A1}$$

and the interaction terms with the graviton field, $h_{\mu\nu}$, are[11,12] (to order κ)

$$:L': = -\tfrac{1}{2}\kappa:\left(\frac{\partial U_1^*}{\partial x_\mu}\frac{\partial U_1}{\partial x_\nu} + \frac{\partial U_1^*}{\partial x_\nu}\frac{\partial U_1}{\partial x_\mu} - \delta_{\mu\nu}\frac{\partial U_1^*}{\partial x_\rho}\frac{\partial U_1}{\partial x_\rho} - \delta_{\mu\nu}\lambda_1^2U_1^*U_1\right)h_{\mu\nu}:$$

$$-\tfrac{1}{2}\kappa:\left(\frac{\partial U_2^*}{\partial x_\mu}\frac{\partial U_2}{\partial x_\nu} + \frac{\partial U_2^*}{\partial x_\nu}\frac{\partial U_2}{\partial x_\mu} - \delta_{\mu\nu}\frac{\partial U_2^*}{\partial x_\rho}\frac{\partial U_2}{\partial x_\rho} - \delta_{\mu\nu}\lambda_2^2U_2^*U_2\right)h_{\mu\nu}:. \tag{A2}$$

The contractions are[12,22,23]

$$A_\mu^\cdot(x)A_\nu^\cdot(x') = -4\pi ic\hbar\delta_{\mu\nu}D_F(x-x'), \tag{A3}$$

$$h_{\mu\nu}^\cdot(x)h_{\lambda\rho}^\cdot(x') = -ic\hbar(\delta_{\mu\lambda}\delta_{\nu\rho} + \delta_{\mu\rho}\delta_{\nu\lambda} - \delta_{\mu\nu}\delta_{\lambda\rho})D_F(x-x'), \tag{A4}$$

where

$$D_F(x-x') = \lim_{\epsilon\to+0}\frac{1}{(2\pi)^4}\int dk\, e^{ik(x-x')}\frac{1}{k^2-i\epsilon}. \tag{A5}$$

We also have[22]

$$S_2 = \frac{-1}{2c^2\hbar^2}\int dx\int dx'\, T[:H'(x)::H'(x'):], \tag{A6}$$

where $:H':$ may be replaced[22,24] by $-:L':$. Using Eqs. (A1) and (A2) in Eq. (A6) we get, respectively, the one-photon and one-graviton exchange results

$$S_{2p} = \frac{e_1e_2}{c^4\hbar^4}\int dx\int dx':\left(\frac{\partial U_1^*(x)}{\partial x_\mu}U_1(x) - \frac{\partial U_1(x)}{\partial x_\mu}U_1^*(x)\right)A_\mu^\cdot(x)\left(\frac{\partial U_2^*(x')}{\partial x'_\nu}U_2(x') - \frac{\partial U_2(x')}{\partial x'_\nu}U_2^*(x')\right)A_\nu^\cdot(x'):, \tag{A7}$$

$$S_{2g} = -\frac{\kappa^2}{4c^2\hbar^2}\int dx\int dx':\left(\frac{\partial U_1^*(x)}{\partial x_\mu}\frac{\partial U_1(x)}{\partial x_\nu} + \frac{\partial U_1^*(x)}{\partial x_\nu}\frac{\partial U_1(x)}{\partial x_\mu} - \delta_{\mu\nu}\frac{\partial U_1^*(x)}{\partial x_\rho}\frac{\partial U_1(x)}{\partial x_\rho} - \delta_{\mu\nu}\lambda_1^2U_1^*(x)U_1(x)\right)h_{\mu\nu}^\cdot(x)$$

$$\times\left(\frac{\partial U_2^*(x')}{\partial x'_\alpha}\frac{\partial U_2(x')}{\partial x'_\beta} + \frac{\partial U_2^*(x')}{\partial x'_\beta}\frac{\partial U_2(x')}{\partial x'_\alpha} - \delta_{\alpha\beta}\frac{\partial U_2^*(x')}{\partial x'_\lambda}\frac{\partial U_2(x')}{\partial x'_\lambda} - \delta_{\alpha\beta}\lambda_2^2U_2^*(x')U_2(x')\right)h_{\alpha\beta}^\cdot(x'):. \tag{A8}$$

Using Eqs. (A3), (A4), and (A5) in Eq. (A7) and (A8) along with[22,25]

$$U_1(x) = (c\hbar/2p_0V)^{1/2}a_1(\mathbf{p})e^{ipx}, \tag{A9}$$

$$U_1^*(x) = (c\hbar/2p'_0V)^{1/2}a_1^*(\mathbf{p}')e^{-ip'x}, \tag{A10}$$

$$U_2(x') = (c\hbar/2q_0V)^{1/2}a_2(\mathbf{q})e^{iqx'}, \tag{A11}$$

$$U_2^*(x') = (c\hbar/2q'_0V)^{1/2}a_2^*(\mathbf{q}')e^{-iq'x'}, \tag{A12}$$

and then integrating gives us [after comparing with Eq. (4)] the results of Eqs. (6) and (5).

[1]B.M. Barker and R.F. O'Connell, J. Math. Phys. **18**, 1818 (1977); **19**, 1231(E) (1978).
[2]S. Bażański, Acta. Phys. Pol. **15**, 363 (1956); **16**, 423 (1957); in *Recent Developments in General Relativity* (Pergamon, New York, 1962), p. 137, see n-body Lagrangian on p. 149.
[3]See Eqs. (3) and (4) of Ref. 1.
[4]K. Hiida and H. Okamura, Prog. Theor. Phys. **47**, 1743 (1972).
[5]L.D. Landau and E.M. Lifshitz, *The Classical Theory of Fields* (Pergamon, New York, 1975), 4th revised English ed. see Sec. 14, and p. 168–9, and p. 342.
[6]P. Havas and J. Stachel, Phys. Rev. **185**, 1637 (1969).
[7]O. Costa de Beauregard, Phys. Lett. A **28**, 365 (1968); S.C. Coleman and J.H. Van Vleck, Phys. Rev. **171**, 1370 (1968).
[8]Sometimes it is convenient to set $x = 4\alpha + 1$ as is done in Refs. 1 and 4.
[9]See Appendix of Ref. 1. Note also that center of mass is *not* the same thing as center of inertia according to the way we define these quantities. The position $\mathbf{r}_{CM}$ is defined for a two-body problem ($\mathbf{r} = \mathbf{r}_1 - \mathbf{r}_2$, $\mathbf{r}_{CM} = \nu_1\mathbf{r}_1 + \nu_2\mathbf{r}_2$ with $\nu_1 + \nu_2 = 1$; which gives us $\mathbf{P} = \nu_2\mathbf{P}_1 - \nu_1\mathbf{P}_2$, $\mathbf{P}_{CM} = \mathbf{P}_1 + \mathbf{P}_2$) in which case $\dot{\mathbf{v}}_{CM} \neq 0$, $\dot{\mathbf{P}}_{CM} = 0$. The position $\mathbf{r}_{CI}$ is defined for the n-body problem (see Sec. IV of this paper) in which case $\dot{\mathbf{v}}_{CI} = 0$.
[10]B.M. Barker and R.F. O'Connell, Phys. Rev. D **12**, 329 (1975), see Sec. II.

[11]B.M. Barker, S.N. Gupta, and R.D. Haracz, Phys. Rev. **149**, 1027 (1966).
[12]S.N. Gupta, Proc. Phys. Soc. London Ser. A **65**, 161, 608 (1952); Phys. Rev. **96**, 1683 (1954); Rev. Mod. Phys. **29**, 334 (1957); in *Recent Developments in General Relativity* (Pergamon, New York, 1962), p. 251; Phys. Rev. **172**, 1302 (1968); Phys. Rev. D **14**, 2596 (1976).
[13]Note that $a_{21} \not\equiv a_{21}(m_1, m_2)$.
[14]See Eqs. (3b) and (4b) in Appendix of Ref. 1.
[15]See Eqs. (3a) and (4a) and discussion in Appendix of Ref. 1.
[16]In Sec. IV of Ref. 1 we had an α_g without *ij* indices and had thought of this quantity as being a totally symmetric function of *all* the masses. We prefer the procedure of this paper which is *more* general in the sense of the *ij* indices, but *less* general in that it involves only the *two* masses m_i and m_j. In this paper one could make the following generalization: Let α_{ijg} and a_{0ijg} be replaced, respectively, by $\alpha_{ijg}\bar{\alpha}_g$ and $a_{0ijg}\bar{a}_g$ where $\bar{\alpha}_g$ and $\bar{a}_g$ are totally symmetric functions of *all* the masses. Everything said here also holds for p replacing g.
[17]L.D. Landau and E.M. Lifshitz, *Mechanics* (Pergamon, New York, 1969) 2nd English ed., Chap. II.
[18]This definition assumes that (a) the total energy can be partitioned among the n particles as given by Eq. (52), and (b) that all energy associated with the ith particle is at $\mathbf{r}_i$. This is not at all obvious, but it turns out to be the case for the Lagrangians considered in this paper.
[19]B.M. Barker and R.F. O'Connell, Phys. Rev. D **14**, 861 (1976).
[20]K. Nordtvedt, Jr., Phys Rev. **14**, 1511 (1976).
[21]C.M. Will, in *Experimental Gravitation*, Proceedings of the International School of Physics "Enrico Fermi," Course 56, edited by B. Bertotti (Academic, New York, 1974), p. 41, Eq. (4.121).
[22]S.N. Gupta, *Quantum Electrodynamics* (Gordon and Breach, New York, 1977).
[23]The factor 4π in the photon contraction is due to our use of Gaussian units.
[24]P.T. Matthews, Phys. Rev. **76**, 684 (1949).
[25]We have given the appropriate terms in the Fourier expansion of the U's for the process under consideration.

Mass dependence of Schrödinger wavefunctions

C. N. Leung and J. Rosner[a)]

School of Physics and Astronomy, University of Minnesota, Minneapolis, Minnesota 55455
(Received 19 July 1978)

It is shown that for nonrelativistic central potentials of specific forms, the bound state wavefunctions $u(r)$ have the property that $G(r) \equiv \int_0^r dx\, u(x)(\partial u(x)/\partial\mu) \geq 0$ for all r. Here μ is the reduced mass. Thus, for such potentials, the probability that a particle lies within a spherical shell of radius r is a monotonically increasing function of μ. The forms for which this property has been established are (a) $V(r) = Cr^\epsilon$, $-2<\epsilon<\infty$; (b) $V(r) = C\ln r$, and (c) S waves in the potential $V(r) = -V_0$ $(0 \leq r < a)$, $V(r) = 0$ $(a \leq r < \infty)$. For *all* monotonically increasing potentials, $G(r)>0$ for nodeless states. It is shown explicitly that $G(r)$ cannot be nonnegative for *all* bound states in an *arbitrary* monotonically increasing potential. The question remains open regarding the widest class of potentials for which $G(r)$ is nonnegative for all bound states.

I. INTRODUCTION

Consider a particle of mass μ bound to a center of force by a potential $V(r)$. What happens when μ is increased?

One very general result of increasing μ is that all the levels become more deeply bound[1,2]:

$$\frac{\partial E}{\partial \mu} = -\frac{\langle T \rangle}{\mu}, \tag{1}$$

where $\langle T \rangle$ is the expectation value of the kinetic energy. In qualitative terms, for a monotonically increasing potential $V(r)$,

$$V'(r) > 0. \tag{2}$$

This could be taken as meaning that the bound states "fall deeper into the well." The present note explores the degree to which this qualitative notion can be expected to be true. We examine the probability $P(r)$ that a particle is found within a spherical shell of radius r. In terms of the radial wavefunction $u(r)$ for any specific bound state, normalized according to

$$\int_0^\infty dr\, [u(r)]^2 = 1, \tag{3}$$

we have

$$P(r) \equiv \int_0^r dx\, [u(x)]^2, \tag{4}$$

so that

$$0 \leqslant P(r) \leqslant 1, \tag{5}$$

and

$$\frac{dP(r)}{dr} \geqslant 0. \tag{6}$$

We will ask: When is $P(r)$ a monotonically increasing function of the mass μ, *for all r*? Equivalently, we seek the widest class of potentials $V(r)$ [satisfying (2)] such that the function

$$G(r) \equiv \tfrac{1}{2}\frac{\partial P(r)}{\partial \mu} = \int_0^r dx\, u(x)\frac{\partial u(x)}{\partial \mu} \tag{7}$$

satisfies

$$G(r) \geqslant 0, \quad 0 \leqslant r < \infty, \tag{8}$$

for all bound states.[3]

Equation (8) is a quantitative statement that as μ increases, the bound particle "falls deeper into the well."

We shall show that Eq. (8) is true for

$$\text{(a) } V(r) = Cr^\epsilon, \quad -2 < \epsilon < \infty, \tag{9}$$

$$\text{(b) } V(r) = C\ln r, \tag{10}$$

$$\text{(c) } V(r) = \begin{cases} -V_0, & 0 \leqslant r < a \\ 0, & a \leqslant r < \infty \end{cases} \quad (S \text{ waves}) \tag{11}$$

and

(d) *not* for *every* potential satisfying Eq. (2). Specifically, we shall show that $G(r)$ can become negative for certain states in a finite square well nested inside an infinite square well.[4] The widest class of monotonic potentials with $G(r) \geqslant 0$ for all bound states thus is still in question. It has been shown previously[5] that Eq. (8) holds for *ground states* as long as Eq. (2) is true.

In Sec. II we demonstrate Eq. (8) for the power law and logarithmic potentials (9) and (10). Section III is devoted to a proof of Eq. (8) for the finite square well (11). Section IV contains the proof that $G(r) \geqslant 0$ for nodeless states in a monotonic potential. Section V discusses the example of the nested wells, in which Eq. (8) is found not to hold for higher-lying S wave bound states. Section VI discusses the results and summarizes open questions.

II. POWER-LAW AND LOGARITHMIC POTENTIALS

The radial Schrödinger equation is

$$-u''(r)/2\mu + [V(r) + l(l+1)/2\mu r^2 - E]u(r) = 0. \tag{12}$$

Here and elsewhere primes denote derivatives with respect to r. Eigenfunctions $u(r)$ obey $\lim_{r\to 0}[u(r)/r^{l+1}] = \text{const}$. For $V(r)$ of the power law form (9), the mass may be absorbed into the coordinate by rescaling.[6] Thus, if we define

a)Work supported in part by U.S. Department of Energy under contract no. E(11-1)-1764.

 0022-2488/79/071435-04$01.00

$$\rho \equiv (2\mu C)^{1/(2+\epsilon)} r, \tag{13}$$

$$\mathscr{E} \equiv (2\mu)^{\epsilon/(2+\epsilon)} (C)^{-2/(2+\epsilon)} E, \tag{14}$$

$$w(\rho) = (2\mu C)^{-1/2(2+\epsilon)} u(r), \tag{15}$$

then

$$-w''(\rho) + (\rho^\epsilon + l(l+1)/\rho^2 - \mathscr{E}) w(\rho) = 0, \tag{16}$$

and

$$\int_0^\infty d\rho\, [w(\rho)]^2 = 1. \tag{17}$$

For the logarithmic potential (10), Eqs. (13)–(15) and (17) remain valid if we set $\epsilon = 0$, and (16) is replaced by

$$-w''(\rho) + \{\ln\rho + l(l+1)/\rho^2 - [\mathscr{E} + \tfrac{1}{2}\ln(2\mu C)]\} w(\rho) = 0. \tag{18}$$

Note that $w(\rho)$ is a universal function of ρ, independent ot μ. Hence

$$\frac{\partial u}{\partial \mu} = w(\rho) \frac{\partial}{\partial \mu}((2\mu C)^{1/2(2+\epsilon)})$$

$$+ (2\mu C)^{1/2(2+\epsilon)} w'(\rho) \frac{\partial \rho}{\partial \mu}$$

$$= \frac{1}{\mu(2+\epsilon)} [u(r)/2 + r u'(r)] \tag{19}$$

and

$$G(r) = \frac{r[u(r)]^2}{2\mu(2+\epsilon)} \geqslant 0. \tag{20}$$

As noted above, Eq. (20) with $\epsilon = 0$ also applies to the logarithmic potential (10). Note that $G(r) = 0$ for $\epsilon = \infty$, which may be regarded as the infinite square well.

Equation (20) shows that for bound states in power law and logarithmic potentials the probability that a particle is located within a shell of radius r never decreases if the reduced mass μ increases. The derivative of this probability with respect to μ *vanishes* at the nodes of the radial wavefunction, however. This last property is not valid for the example to be discussed in the next section.

III. FINITE SQUARE WELL

In the potential (11), the S wave eigenfunctions are

$$0 \leqslant r \leqslant a; \quad u(r) = A \sin kr, \tag{21}$$

$$a \leqslant r \leqslant \infty, \quad u(r) = B e^{-Kr}, \tag{22}$$

where, for $E < 0$,

$$k^2 = 2\mu(E + V_0), \quad K^2 = -2\mu E. \tag{23}$$

The eigenvalue condition is

$$k \cot ka = -K, \tag{24}$$

and (3) implies

$$A^2 = \frac{2K}{1 + aK}. \tag{25}$$

By virtue of (23),

$$\frac{dk}{d\mu} = \frac{1}{k}\left(V_0 + E + \mu \frac{dE}{d\mu}\right)$$

$$= \frac{V_0 + E - \langle T \rangle}{k} = \frac{V_0 + \langle V \rangle}{k} \geqslant 0. \tag{26}$$

Now, for $r \leqslant a$,

$$\frac{\partial u}{\partial \mu} = \left(\frac{r u'(r)}{k} + \frac{u}{A}\frac{dA}{dk}\right)\frac{dk}{d\mu}, \tag{27}$$

so

$$G(r) = \frac{dk}{d\mu}\left\{\frac{r u^2}{2k} + \frac{1}{2}\frac{d}{dk}\left[\ln\left(\frac{A^2}{k}\right)\right] \int_0^r dx\, u^2(x)\right\}. \tag{28}$$

A short calculation shows

$$\frac{d}{dk}\ln\left(\frac{A^2}{k}\right) = \frac{ka}{K(1 + aK)} > 0, \tag{29}$$

so that we have shown

$$G(r) \geqslant 0 \quad \text{for } r \leqslant a. \tag{30}$$

To show $G(r) \geqslant 0$ for $r \geqslant a$ we use a general expression that will be useful later. Solve Eq. (12) for u'', differentiate with respect to μ, use the Feynman–Hellmann theorem (1), and multiply by $u(r)$. The result is

$$u\frac{\partial u''}{\partial \mu} = 2(V - \langle V \rangle) u^2 + [2\mu(V - E)$$

$$+ l(l+1)/r^2] u \frac{\partial u}{\partial \mu}. \tag{31}$$

Add to this Eq. (12)$\times 2\mu(\partial u/\partial \mu)$, and integrate by parts. The result is

$$u\frac{\partial u'}{\partial \mu} - u'\frac{\partial u}{\partial \mu}$$

$$= 2\int_0^r dx\, u^2(x)(V(x) - \langle V \rangle)$$

$$= 2\int_r^\infty dx\, u^2(x)(\langle V \rangle - V(x)). \tag{32}$$

For a monotonic potential (2), the right-hand side of (32) is negative. Now $G'(r) = u(\partial u/\partial \mu)$ vanishes when and only when $u = 0$ or $\partial u/\partial \mu = 0$. To see whether these points are maxima or minima we examine $G''(r) = u'(\partial u/\partial \mu) + u(\partial u'/\partial \mu)$. When $u = 0$, $G'' = u'(\partial u/\partial \mu)$, while when $\partial u/\partial \mu = 0$, $G'' = u(\partial u'/\partial \mu)$. Equation (32) then permits one to conclude, for a monotonic potential (2), that for $0 < r < \infty$: When

$$u = 0, \quad G''(r) > 0 \quad \text{(a minimum)}, \tag{33}$$

when

$$\frac{\partial u}{\partial \mu} = 0, \quad G''(r) < 0 \quad \text{(a maximum)}. \tag{34}$$

Clearly $G(\infty) = 0$. If G becomes negative in $a < r < \infty$, it must have a minimum at some point r_0, $a < r_0 < \infty$. Then

$u(r_0) = 0$, which is incompatible with the form (22). Hence $G(r)$ must stay nonnegative [cf. (30)] in $a < r < \infty$, and hence $G(r) \geqslant 0$ everywhere.

The form (28) shows that, in contrast to the power law and logarithmic situations, the finite square well gives rise to a function $G(r)$ possessing minima, but *not nodes*, at the zeroes of $u(r)$.

IV. THEOREM FOR NODELESS STATES

For the lowest state for fixed $l, u(r) \neq 0$ for $0 < r < \infty$. But then Eq. (33) shows that $G(r)$ cannot have a minimum in $0 < r < \infty$. Since $G(0) = G(\infty) = 0$,

$G(r) \geqslant 0$ for the lowest state for fixed l in any monotonic potential. (35)

This discussion leaves open the possibility that for *higher* excited states in certain monotonically increasing potentials it may be possible to find wavefunctions for which $G(r)$ can be negative for some r. An example in which this behavior occurs very strikingly is shown in the next section.

V. NESTED WELLS

The potential is assumed to be

$$V(r) = \begin{cases} -V_0, & 0 \leqslant r < a, \\ 0, & a \leqslant r < b, \\ \infty, & b \leqslant r < \infty. \end{cases} \tag{36}$$

The S wave eigenfunctions are

$$\begin{aligned} u(r) &= A \sin kr, \quad 0 \leqslant r < a, \\ &= A \frac{\sin ka}{\sinh K(b-a)} \sinh K(b-r), \quad a \leqslant r < b, \\ &= 0, \quad b \leqslant r < \infty, \end{aligned} \tag{37}$$

with k and K defined as in Eq. (23). (In the present example, E may be negative or positive.)

The eigenvalue condition is

$$\tan ka = -\frac{k}{K} \tanh K(b-a), \tag{38}$$

and the normalization condition (3) implies

$$1 = \frac{A^2}{2}\left(b + \frac{V_0}{2kE}\sin 2ka + (b-a)\frac{V_0}{E}\cos^2 ka\right). \tag{39}$$

For $r \leqslant a$ the expression (28) still holds for $G(r)$. Let us look at the values of $G(r)$ at its minima, which are nodes of $u(r)$ according to Eq. (33):

$$kr = m\pi, \quad m = 1,2,\cdots. \tag{40}$$

At these points, $G(r)$ will be negative if $d[\ln(A^2/k)]/dk < 0$. Equivalently, $G(r)$ will be negative if

$$\frac{\partial}{\partial k}\left(\frac{k}{A^2}\right) > 0. \tag{41}$$

The left hand side of Eq. (41) is easily computed for values of k satisfying

$$ka = n\pi, \quad n = 1,2,\cdots, \quad n \geqslant m, \tag{42}$$

where the condition on n guarantees $r \leqslant a$. In order that Eq. (42) be consistent with Eq. (38), one must have $E > 0$, $K = i|K|$, and $\tan(|K|(b-a)) = 0$. This last condition means, as does (42), that $r = a$ is a node of the wavefunction. Then, when (42) holds, we find that

$$\frac{\partial}{\partial k}\left(\frac{k}{A^2}\right) = (V_0 + E)\left[\frac{b}{2E} - \frac{V_0(b-a)}{E^2}\right]. \tag{43}$$

The term $V_0 + E$ is positive. If we can demonstrate that the term in square brackets can be made positive in a manner consistent with the eigenvalue condition, then we have constructed an example with $G(r) < 0$.

Thus, let

$$b = ha \quad (h > 1), \tag{44}$$

$$k^2 = 2\mu(E + V_0) = \left(\frac{n\pi}{a}\right)^2 \quad (n = \text{integer} \geqslant 1), \tag{45}$$

$$|K|^2 = 2\mu E = \left(\frac{p\pi}{b-a}\right)^2 \quad (p = \text{integer} \geqslant 1), \tag{46}$$

and look for solutions with

$$\frac{V_0}{E} < \frac{b}{2(b-a)} = \frac{h}{2(h-1)}. \tag{47}$$

These are possible when

$$1 < 1 + \frac{V_0}{E} = \frac{n^2(h-1)^2}{p^2} < 1 + \frac{h}{2(h-1)}. \tag{48}$$

Thus, for example, with $h = 2$ we seek n,p with $1 < n^2/p^2 < 2$, and the lowest such pair is $(n,p) = (4,3)$. For these values

$$E = (2\mu a^2)^{-1}(9\pi^2), \tag{49}$$

$$E + V_0 = (2\mu a^2)^{-1}(16\pi^2), \tag{50}$$

so the depth of the well must be chosen to be

$$V_0 = (2\mu a^2)^{-1}(7\pi^2). \tag{51}$$

Equation (47) says that in order to obtain negative values of $G(r)$ *in the above fashion*, one must discuss excitations above a certain energy. It is still an open question whether there is a certain minimum excitation energy below which one could prove $G(r) \geqslant 0$ for all r in the nested-well potential.

The reason one can exhibit solutions with $G(r) < 0$ for some r in the nested-well potential is that the smaller well gives rise to resonancelike behavior for $E > 0$. A shift in μ can shift the system on or off "resonance," and hence the larger value of μ need not always be associated with a spatially more compact wavefunction.

VI. DISCUSSION

We have given only a partial answer to the question: "When does an increased mass μ lead to a spatially more compact bound state wavefunction?" We have shown that this occurs *locally*, in the sense that $G(r) \geqslant 0$ [see Eq. (7)], for power law and logarithmic potentials and for S waves in the finite square-well potential, and that the monotonicity condition is *not* sufficient to guarantee $G(r) \geqslant 0$ for excited states. It is still unknown whether $G(r) \geqslant 0$ for all states in the following well-known monotonic potentials:

(a) Superpositions of powers, $V(r) = Ar^{\epsilon} + Br^{\eta}$;

(b) Yukawa potential, $V(r) = \lambda e^{-br}/r$;

(c) Exponential potential, $V(r) = \lambda e^{-br}$;

(d) The potential $V(r) = A/\cosh^2 br$.

It would be interesting to turn the question about: If $G(r) \geqslant 0$ for all bound states in a potential, for all values of μ and for all r, what does this imply about the potential?

The interest of this problem is particularly current in light of the discovery of bound states of heavy quarks and antiquarks like the ψ/J[7] and Υ[8] families. Some idea of the behavior of the spatial properties of the wavefunctions with increasing quark mass can be particularly helpful in anticipating the decay widths and other properties of these and heavier states.[5]

ACKNOWLEDGMENTS

One of us (J.L.R.) thanks J.F. Donoghue, C. Quigg, and H. Thacker for helpful discussions. Part of this work was performed at Fermi National Accelerator Laboratory and at the Aspen Center for Physics.

[1]H. Hellmann, *Einführung in die Quantenchemie* (Franz Deuticke, Leipsig and Vienna, 1937), p. 286.
[2]R.P. Feynman, Phys. Rev. **56**, 340 (1939).
[3]We thank C. Quigg for the crucial observations that most of our results are not limited to S waves, and that Sec. II and Eqs. (31)–(35) apply, *mutatis mutandis*, to the one-dimensional case.
[4]One of us (J.L.R.) is grateful to R.P. Feynman for suggesting this example.
[5]Jonathan L. Rosner, C. Quigg, and H.B. Thacker, Phys. Lett. B**74**, 350 (1978).
[6]See, e.g., C. Quigg and Jonathan L. Rosner, "Scaling the Schrödinger equations," in Comments on Nuclear and Particle Physics **8** A, 11 (1978).
[7]J.J. Aubert *et al.*, Phys. Rev. Lett. **33**, 1404 (1974); J.-E. Augustin *et al.*, Phys. Rev. Lett. **34**, 1406 (1974).
[8]S.W. Herb *et al.*, Phys. Rev. Lett. **39**, 252 (1977); W.R. Innes *et al.*, Phys. Rev. Lett. **39**, 1240, 1640 (E) (1977).

Equations for the propagation of finite local disturbances in vector fields of physics and mechanics

Robert C. Costen

NASA Langley Research Center, Hampton, Virginia
(Received 25 May 1978)

The purpose of this theoretical investigation is to determine the propagation velocities of local disturbances that may occur in the curl and divergence of a vector field **v** that satisfies the equation $\partial\mathbf{v}/\partial t + \mathbf{f} = 0$, where **f** is a general vector field. Equations of this form frequently occur in physics and mechanics. A *disturbance* in the curl **v** (div **v**) field is defined by $\partial\text{curl }\mathbf{v}/\partial t \neq 0$ ($\partial\text{div }\mathbf{v}/\partial t \neq 0$). Such a disturbance is termed a *local* disturbance if it is surrounded by a quiescent layer of finite thickness. Equations are derived for the velocity $\mathbf{U}(\mathbf{r},t)$ of curl **v** lines and the velocity $\mathbf{W}(\mathbf{r},t)$ of div **v** elements. These velocities conserve, respectively, the circulation of **v** about arbitrary circuits and the efflux of **v** through arbitrary closed surfaces. These conservation conditions are not sufficient to completely specify the two velocity fields, and each contains an arbitrary term. More convenient for three-dimensional applications than the circulation is the gyration **G** (spatial integral of curl **v**); and **U** is found to also conserve **G** in regions T where the curl **v** lines retain their individual shapes and orientations as they move, and our consideration of curl **v** disturbances is confined to such regions. The centroid of efflux is defined for arbitrary volumes, and the centroid of gyration for volumes in T. The movement of these centroids is linked to the velocities **U** and **W** by the conservation relations. Contributions from the arbitrary term in **U** (**W**) are suppressed by integration over a *local* curl **v** (div **v**) disturbance and by construction of a scalar potential field that satisfies Laplace's equation in the interior of the disturbance with Dirichlet (Neumann) conditions related to **f** on the closed boundary. Formulas for the velocities of the centroids are thus obtained for local curl **v** and div **v** disturbances as functions of **f**, and these velocities are interpreted as propagation velocities of the disturbances. Only one type of disturbance need be localized for its formula to apply. These formulas enable us to calculate how such local disturbances will propagate without integrating forward in time; and in certain cases, only the fields outside a disturbance are required for this calculation. Applications to Maxwell's electrodynamic equations are presented as examples.

I. INTRODUCTION

Dynamical vector equations of the form $\partial\mathbf{v}/\partial t + \mathbf{f} = 0$, where **v** and **f** are vector functions of space and time occur frequently in physics and mechanics, and they are characteristically difficult to solve, especially when **f** is a nonlinear combination of dependent variables. However, for applications where local fluctuations in curl**v** and in div**v** tend to occur and persist, it is useful to know how such would propagate through space. Examples of such local fluctuations include electric charge density $\rho = \text{div}\mathscr{D}$, magnetic field strength $\mathscr{B} = \text{curl}\mathscr{A}$, fluid vorticity $\boldsymbol{\omega} = \text{curl}\mathbf{u}$, and atmospheric convection characterized by div**u**. The purpose of this article is to derive general formulas for the propagation velocities of such local disturbances.

Rudiments of this theory lie in the classical concepts for the movement of magnetic field lines[1] and for the movement of fluid vortex lines (Kelvin's circulation theorem[2]). However, these classical concepts are precise only for perfect conductors and idealized fluids. In the context of the equation in the preceding paragraph, they treat the propagation of curl**v** disturbances with severe restrictions imposed on **f**. We shall generalize these concepts, removing most of the restrictions, and shall extend the treatment to div**v** disturbances, for which no prior theorems appear to exist.

II. GENERAL CONSIDERATIONS

As indicated in the Introduction, our analysis starts with a dynamical vector equation of the form

$$\frac{\partial\mathbf{v}}{\partial t} + \mathbf{f} = 0. \tag{1}$$

The vector field $\mathbf{v}(\mathbf{r},t)$ may be characterized by its curl **c** and its divergence D, where **c** and D satisfy the equations

$$\frac{\partial\mathbf{c}}{\partial t} + \text{curl}\mathbf{f} = 0, \tag{2}$$

$$\frac{\partial D}{\partial t} + \text{div}\mathbf{f} = 0. \tag{3}$$

As used throughout this article, a *disturbance* in any vector or scalar field is defined by $\partial(\text{field})/\partial t \neq 0$. A distur-

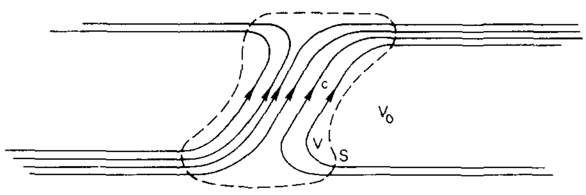

FIG. 1. Illustration of a local **c**-disturbance: A bundle of **c**-lines is moving in the horizontal direction; the disturbance ($\partial\mathbf{c}/\partial t\neq 0$) is contained within volume V bounded by closed surface S; on S and in V_0 (which in this illustration extends to infinity) $\partial\mathbf{c}/\partial t = 0$.

bance in $\mathbf{v}$ ($\partial\mathbf{v}/\partial t\neq 0$) may be characterized by a **c**-disturbance ($\partial\mathbf{c}/\partial t\neq 0$) or a D-disturbance ($\partial D/\partial t\neq 0$), or both. By Eqs. (2) and (3), **c**-disturbances and D-disturbances are distinct, although they may intersect in $(\mathbf{r},t)$. A **c**-disturbance is termed a *local* **c**-disturbance if there exists a surrounding shell of finite thickness wherein $\partial\mathbf{c}/\partial t = -\operatorname{curl}\mathbf{f} = 0$.

This definition of a local **c**-disturbance has physical relevance. For example, suppose that a moving bundle of **c**-lines is flared out at both ends, as shown in Fig. 1. Although the **c**-lines extend to infinity, as shown, the disturbance ($\partial\mathbf{c}/\partial t\neq 0$) is contained within the local volume V. Similarly, a D-disturbance is a local D-disturbance if there exists a surrounding shell of finite thickness wherein $\partial D/\partial t = -\operatorname{div}\mathbf{f} = 0$. Such local disturbances may propagate, and as mentioned earlier, our objective is to derive formulas for their propagation velocities.

The approach is (a) to derive the velocity fields of three-dimensional **c**- and *D-elements*, although these velocity fields may not be defined uniquely, and (b) to integrate these velocity fields over *local* **c**- and D-disturbances in order to determine unique formulas for the velocities of the centroids of these disturbances. These centroidal velocities are interpreted as propagation velocities. The velocity of **c**-*elements* is obtained indirectly by first determining the velocity of **c**-*lines* and then determining the condition under which this velocity also applies to **c**-*elements*.

III. EQUATIONS FOR THE VELOCITY OF c-LINES AND c-ELEMENTS

Take σ to be an arbitrary open surface with unit normal $\hat{n}$ and bounding circuit line γ, as shown in Fig. 2, where the senses of the unit vectors $\hat{n}$ and $\hat{\gamma}$ follow the right-hand screw convention. Integrating the normal component of Eq. (2) over σ and using Stokes' theorem gives

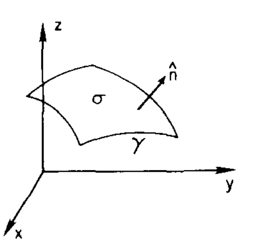

FIG. 2. Arbitrary open surface σ with unit normal $\hat{n}$ and boundary circuit line γ, where the senses of $\hat{n}$ and $\hat{\gamma}$ conform to the right-hand screw convention.

$$\int\!\!\int \frac{\partial\mathbf{c}}{\partial t}\cdot\hat{n}\,d\sigma + \oint \mathbf{f}\cdot d\boldsymbol{\gamma} = 0. \tag{4}$$

The first term in Eq. (4) may be modified by using Helmholtz' vector-flux theorem[2]

$$\frac{d}{dt}\oint \mathbf{v}\cdot d\boldsymbol{\gamma} + \oint (\mathbf{U}'\times\mathbf{c} + \mathbf{f})\cdot d\boldsymbol{\gamma} = 0, \tag{5}$$

where γ is taken to be moving with the mathematical velocity field $\mathbf{U}'(\mathbf{r},t)$, which is distinct from any physical field. Since the circulation $\Gamma(\mathbf{v},\gamma)$ is purely a function of t, it is appropriate to take its time derivative.

We now stipulate that the mathematical velocity field, henceforth denoted by $\mathbf{U}$, shall conserve $\Gamma(\mathbf{v},\gamma)$:

$$\frac{d}{dt}\oint \mathbf{v}\cdot d\boldsymbol{\gamma} = 0, \tag{6}$$

$$\oint (\mathbf{U}\times\mathbf{c} + \mathbf{f})\cdot d\boldsymbol{\gamma} = 0. \tag{7}$$

Applying Eq. (7) to arbitrary circuits and assuming that the integrand is continuous gives

$$\mathbf{U}\times\mathbf{c} + \mathbf{f} = \operatorname{grad}\phi, \tag{8}$$

where $\phi(\mathbf{r},t)$ is an arbitrary scalar field. The field $\mathbf{U}$ is the velocity of the **c**-lines; this velocity is not unique because of the arbitrarines of ϕ. (Fortunately, contributions from ϕ may be made to vanish upon subsequent integration.) It is clear from Eq. (8) that if **f** contains any explicit gradient terms, such terms may be suppressed by absorption into $\operatorname{grad}\phi$ without affecting $\mathbf{U}\times\mathbf{c}$. A special case of Eq. (8) for fluid vortex lines has been derived and solved for the movement of idealized buoyant-core vortices subject to gravity.[3,4]

Substituting Eq. (8) back into Eq. (2) gives

$$\frac{\partial\mathbf{c}}{\partial t} - \operatorname{curl}(\mathbf{U}\times\mathbf{c}) = 0. \tag{9}$$

Equation (9) is the differential form of Eq. (6) for the conservation of circulation $\Gamma(\mathbf{v},\gamma)$, and it is a generalization of an equation derived by Helmholtz[2] for vorticity in an idealized fluid. Another integral form of Eq. (9) is obtained through integration over volume τ of Fig. 3 and use of a kinematic theorem[5]

$$\frac{d}{dt}\oint\!\!\oint \hat{n}\times\mathbf{v}\,d\sigma - \int\!\!\int\!\!\int (\mathbf{c}\cdot\nabla)\mathbf{U}\,d\tau = \mathbf{0}, \tag{10}$$

where σ and τ move with $\mathbf{U}$, and where the last integrand is

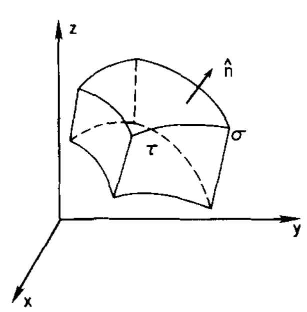

FIG. 3. Arbitrary volume τ bounded by closed surface σ with outward unit normal $\hat{n}$.

expressed, for brevity, in Cartesian notation. In general, therefore, while $\mathbf{U}$ conserves the circulation $\Gamma(\mathbf{v},\gamma)$, it may not conserve the gyration $\mathbf{G}(\mathbf{v},\sigma)$.

However, for spatial regions T wherein

$$(\mathbf{c}\cdot\nabla)\mathbf{U}=0 \quad (\text{in } T); \tag{11}$$

i.e., wherein $\mathbf{U}$ is constant on each $\mathbf{c}$-line but may vary from line to line, the gyration $\mathbf{G}(\mathbf{v},\sigma)$ is also conserved, and the integral of the first term in Eq. (8) can be expressed as follows:

$$\begin{aligned}\iiint \mathbf{U}\times\mathbf{c}\,d\tau &= \iiint\left(\left(\frac{\partial}{\partial t}+\mathbf{U}\cdot\nabla\right)\mathbf{r}\right)\times\mathbf{c}\,d\tau\\ &= \iiint \frac{\partial}{\partial t}(\mathbf{r}\times\mathbf{c})d\tau + \oiint \mathbf{r}\times\mathbf{c}(\mathbf{U}\cdot\hat{n})d\sigma\\ &= \frac{d}{dt}\iiint \mathbf{r}\times\mathbf{c}\,d\tau \quad (\tau\in T),\end{aligned} \tag{12}$$

where τ and σ move with $\mathbf{U}$. The last integral in Eq. (12) may be written

$$\iiint \mathbf{r}\times\mathbf{c}\,d\tau = \mathbf{R}_{\mathbf{G}}\times\mathbf{G} + (\mathbf{R}_{\mathbf{G}}\cdot\widehat{G})\mathbf{G}, \tag{13}$$

where $\mathbf{G}$ denotes $\mathbf{G}(\mathbf{v},\sigma)$, and where $\mathbf{R}_{\mathbf{G}}$ (denoting $\mathbf{R}_{\mathbf{G}(\mathbf{v},\sigma)}$) is a generalized *centroid* of $\mathbf{G}(\mathbf{v},\sigma)$. By Eqs. (12) and (13), we have

$$\iiint \mathbf{U}\times\mathbf{c}\,d\tau = \dot{\mathbf{R}}_{\mathbf{G}}\times\mathbf{G} + (\dot{\mathbf{R}}_{\mathbf{G}}\cdot\widehat{G})\mathbf{G} \quad (\tau\in T). \tag{14}$$

Integrals of the remaining terms in Eq. (8) will be treated in a subsequent section.

IV. EQUATION FOR THE VELOCITY OF *D*-ELEMENTS

To derive an equation for the velocity of D-elements, we integrate Eq. (3) over the volume τ bounded by a closed surface σ, as shown in Fig. 3, and apply Gauss' divergence theorem:

$$\iiint \frac{\partial D}{\partial t}d\tau + \oiint \mathbf{f}\cdot\hat{n}d\sigma = 0. \tag{15}$$

A kinematic theorem[5] then gives

$$\frac{d}{dt}\oiint \mathbf{v}\cdot\hat{n}d\sigma + \oiint(-D\,\mathbf{W}'+\mathbf{f})\cdot\hat{n}d\sigma = 0, \tag{16}$$

where σ is moving with the mathematical velocity field $\mathbf{W}'(\mathbf{r},t)$, which is distinct from any physical field and is also distinct from $\mathbf{U}$.

We now stipulate that this mathematical velocity field, hereafter denoted by $\mathbf{W}$, shall conserve the efflux $E(\mathbf{v},\sigma)$:

$$\frac{d}{dt}\oiint \mathbf{v}\cdot\hat{n}d\sigma = 0, \tag{17}$$

$$\oiint(-D\,\mathbf{W}+\mathbf{f})\cdot\hat{n}d\sigma = 0. \tag{18}$$

Applying Eq. (18) to arbitrary closed surfaces σ and assuming that the integrand is continuous gives

$$-D\,\mathbf{W}+\mathbf{f} = \mathrm{curl}\mathbf{A}, \tag{19}$$

where $\mathbf{A}(\mathbf{r},t)$ is an arbitrary vector field. The field $\mathbf{W}(\mathbf{r},t)$ is the velocity of D-elements, although, as with $\mathbf{U}$, this velocity is not uniquely defined. Again contributions from the arbitrary field shall be made to vanish upon integration. It is clear from Eq. (19) that if $\mathbf{f}$ contains any explicit curl terms, such terms may be suppressed by absorption into curl$\mathbf{A}$ without affecting $D\,\mathbf{W}$.

Substituting Eq. (19) back into Eq. (3) gives

$$\frac{\partial D}{\partial t} + \mathrm{div}D\,\mathbf{W} = 0. \tag{20}$$

Equation (20) is the differential form of Eq. (17) for the conservation of efflux $E(\mathbf{v},\sigma)$. It is analogous to Maxwell's equation for the conservation of electric charge, with $D\,\mathbf{W}$ as the current. This current may be definite at points where D vanishes, since the magnitude of $\mathbf{W}$ is not bounded by physical considerations. [Similarly comments apply to $\mathbf{U}$ and $\mathbf{c}$ in Eq. (9).]

Because $E(\mathbf{v},\sigma)$ is conserved by $\mathbf{W}$, the integral of the first term in Eq. (19) can be expressed as follows:

$$\begin{aligned}\iiint D\,\mathbf{W}d\tau &= \iiint D\left(\left(\frac{\partial}{\partial t}+\mathbf{W}\cdot\nabla\right)\mathbf{r}\right)d\tau\\ &= \iiint \frac{\partial}{\partial t}(\mathbf{r}D)d\tau + \oiint \mathbf{r}D\,(\mathbf{W}\cdot\hat{n})d\sigma\\ &= \frac{d}{dt}\iiint \mathbf{r}D\,d\tau,\end{aligned} \tag{21}$$

where τ and σ move with $\mathbf{W}$. The last integral in Eq. (21) may be written

$$\iiint \mathbf{r}D\,d\tau = \mathbf{R}_{\mathbf{E}(\mathbf{v},\sigma)}E(\mathbf{v},\sigma), \tag{22}$$

where $\mathbf{R}_{\mathbf{E}(\mathbf{v},\sigma)}$ is the centroid of $E(\mathbf{v},\sigma)$. Hence, by substituting Eq. (22) into (21), we obtain

$$\iiint D\,\mathbf{W}d\tau = \dot{\mathbf{R}}_{\mathbf{E}(\mathbf{v},\sigma)}E(\mathbf{v},\sigma). \tag{23}$$

Integrals of the remaining terms in Eq. (19) will be treated in a subsequent section.

V. CONSTRAINED FORMS OF THE EQUATIONS

The equations that apply where $\mathbf{c}$ is undisturbed are

$$\frac{\partial\mathbf{c}}{\partial t}=0, \tag{24a}$$

$$\mathrm{curl}\mathbf{f}=0, \tag{24b}$$

$$\mathbf{f}=\mathrm{grad}\xi, \tag{24c}$$

$$\frac{\partial\mathbf{v}}{\partial t} = -\mathrm{grad}\xi, \tag{24d}$$

$$\frac{\partial D}{\partial t} = -\mathrm{div}\,\mathrm{grad}\xi, \tag{24e}$$

$$\mathbf{U}\times\mathbf{c}=\mathrm{grad}(\phi-\xi), \tag{24f}$$

$$-D\,\mathbf{W}=\mathrm{curl}\mathbf{A}-\mathrm{grad}\xi. \tag{24g}$$

The movement of $\mathbf{c}$-lines, as constrained by Eq. (24f), is consistent with the static nature of the $\mathbf{c}$-field; however, Eq. (24g) confirms that the movement of D-elements remains unconstrained. Equation (24f) also shows that a static **c**-field can make arbitrary contributions to the integral in Eq. (14); we shall suppress these contributions in the next section by restricting ϕ.

The equations that apply where D is undisturbed are

$$\frac{\partial D}{\partial t}=0, \tag{25a}$$

$$\text{div}\mathbf{f}=0, \tag{25b}$$

$$\mathbf{f}=\text{curl}\mathbf{M}, \tag{25c}$$

$$\frac{\partial \mathbf{v}}{\partial t}=-\text{curl}\mathbf{M}, \tag{25d}$$

$$\frac{\partial \mathbf{c}}{\partial t}=-\text{curl curl}\mathbf{M}, \tag{25e}$$

$$-D\mathbf{W}=\text{curl}(\mathbf{A}-\mathbf{M}), \tag{25f}$$

$$\mathbf{U}\times\mathbf{c}=-\text{curl}\mathbf{M}+\text{grad}\phi. \tag{25g}$$

The movement of D-elements, as constrained by Eq. (25f), is consistent with the static nature of the D-field; however, Eq. (25g) confirms that the motion of **c**-lines remains unconstrained. Equation (25f) also shows that a static D-field can make arbitrary contributions to the integral in Eq. (23); we shall suppress these contributions in a subsequent section by restricting **A**.

VI. PROPAGATION OF A LOCAL c-DISTURBANCE

In this section the D-field may be arbitrarily disturbed; however, the disturbance in the **c**-field is taken to be a *local* **c**-disturbance that is contained within volume V, bounded by closed surface S, as shown in Fig. 4. Equations (24) apply on S and in an external layer of finite thickness termed V_0; therefore, the interface S between V and V_0 is considered to be included in V_0.

In order to determine the movement of a local **c**-disturbance, we may integrate Eq. (8) over V; however, as mentioned earlier, contributions from the arbitrary field ϕ must be made to vanish. The first steps towards accomplishing this is to restrict ϕ to that subset θ of arbitrary functions that satisfy the condition

$$\text{grad}(\theta-\xi)=0 \quad (\text{in } V_0), \tag{26}$$

and Eq. (8) becomes

$$\mathbf{U}\times\mathbf{c}+\mathbf{f}=\text{grad}\theta. \tag{27}$$

A consequence of this restriction is that

$$\mathbf{U}\times\mathbf{c}=0 \quad (\text{in } V_0); \tag{28}$$

therefore, this restriction suppresses the static **c**-field contributions discussed in the preceding section.

The second step in suppressing arbitrary contributions is to construct a vector field **g** that satisfies the conditions

$$\mathbf{g}=\text{grad}\psi \quad (\text{in V and V}_0), \tag{29a}$$

$$\mathbf{g}=\mathbf{f} \quad (\text{in } V_0), \tag{29b}$$

and to subtract Eq. (29a) from Eq. (27):

$$\mathbf{U}\times\mathbf{c}+\mathbf{f}-\mathbf{g}=\text{grad}(\theta-\psi). \tag{30}$$

Since both sides of Eq. (30) vanish in V_0, we obtain upon integration over V

$$\iiint \mathbf{U}\times\mathbf{c}dV+\iiint(\mathbf{f}-\mathbf{g})dV=0. \tag{31}$$

The first integral may be expressed as in Eq. (14) provided the volume V is contained in T where Eq. (11) applies; proceeding under this provision, we find

$$\dot{\mathbf{R}}_G\times\mathbf{G}+(\dot{\mathbf{R}}_G\cdot\hat{G})\mathbf{G}=-\iiint(\mathbf{f}-\mathbf{g})dV \quad (V\in T), \tag{32}$$

where **G** now denotes $\mathbf{G}(\mathbf{v},S)$ and where by Eq. (13)

$$\mathbf{R}_G\times\mathbf{G}+(\mathbf{R}_G\cdot\hat{G})\mathbf{G}=\iiint\mathbf{r}\times\mathbf{c}dV. \tag{33}$$

Taking the vector and scalar products of Eqs. (32) and (33) with $\mathbf{G}(\mathbf{v},S)$, we obtain

$$\dot{\mathbf{R}}_{G\perp}=-\frac{\mathbf{G}}{G^2}\times\iiint(\mathbf{f}-\mathbf{g})dV \quad (V\in T) \tag{34a}$$

$$\dot{\mathbf{R}}_{G\parallel}=-\frac{\mathbf{G}}{G^2}\cdot\iiint(\mathbf{f}-\mathbf{g})dV \quad (V\in T), \tag{34b}$$

where

$$\mathbf{R}_{G\perp}=\frac{\mathbf{G}}{G^2}\times\iiint\mathbf{r}\times\mathbf{c}dV, \tag{34c}$$

$$\mathbf{R}_{G\parallel}=\frac{\mathbf{G}}{G^2}\cdot\iiint\mathbf{r}\times\mathbf{c}dV, \tag{34d}$$

and

$$\mathbf{G}=\oiint\hat{n}\times\mathbf{v}dS=\iiint\mathbf{c}dV. \tag{34e}$$

The task remaining is to construct the vector field **g** in V. Conditions (29) will be satisfied provided

$$\mathbf{g}=\mathbf{f} \quad (\text{in } V_0), \tag{35a}$$

$$\hat{n}\times[\mathbf{g}]=0 \quad (\text{on } S), \tag{35b}$$

$$\text{curl}\mathbf{g}=0 \quad (\text{in } V). \tag{35c}$$

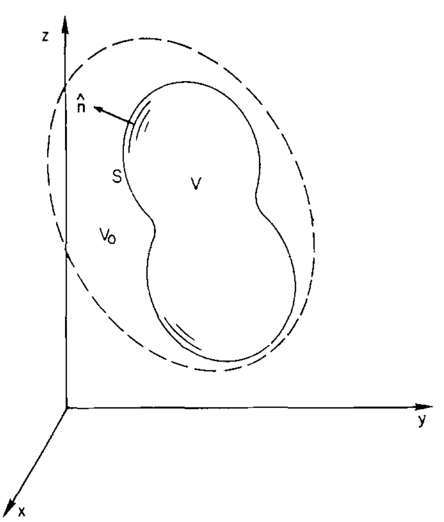

FIG. 4. Volume V containing a local curlv disturbance, with bounding surface S and outward unit normal $\hat{n}$. Equations (24) apply on S and in the external layer V_0; Eq. (11) applies in V.

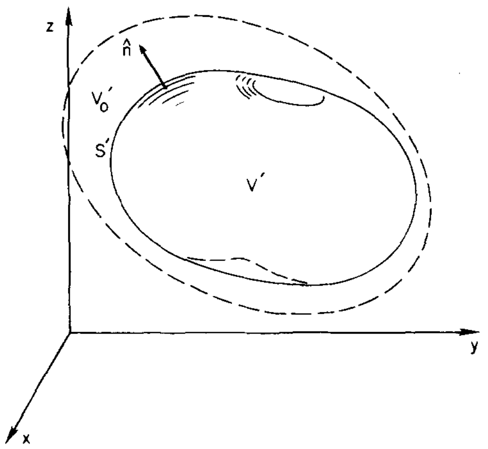

FIG. 5. Volume V' containing a local divv disturbance, with bounding surface S' and outward unit normal $\hat{n}$. Equations (25) apply on S' and in the external layer V_0'.

For complete specification, divg must also be specified in V; we choose

$$\text{div}\mathbf{g} = 0 \quad (\text{in } V). \tag{35d}$$

As in Eq. (29a), we take

$$\mathbf{g} = \text{grad}\psi \quad (\text{in } V), \tag{36a}$$

where ψ now satisfies the equation

$$\text{div grad}\psi = 0 \quad (\text{in } V) \tag{36b}$$

with the boundary condition

$$\hat{n}\times\text{grad}\psi = \hat{n}\times\mathbf{f} \quad (\text{on } S). \tag{36c}$$

This boundary condition can be integrated to determine ψ on S from a knowledge of $\mathbf{f}$; or, if the potential function ξ is known in V_0, the boundary condition (36c) may be written

$$\psi = \xi + \text{const} \quad (\text{on } S). \tag{37}$$

Therefore, $\mathbf{g}$ is determined by a scalar potential that satisfies Laplace's equation in V with Dirichlet boundary conditions on S. Equations (8), (11), (34), and (36) constitute our formulation for the movement of local $\mathbf{c}$-disturbances generated by Eq. (1).

VII. PROPAGATION OF A LOCAL D-DISTURBANCE

In this section the $\mathbf{c}$-field may be arbitrarily disturbed; however, the disturbance in the D-field is taken to be a *local* D-disturbance that is contained within volume V', bounded by closed surface S', as shown in Fig. 5. Equations (25) apply on S' and in an external layer of finite thickness termed V_0'; therefore, the interface S' between V' and V_0' is considered to be included in V_0'.

In order to determine the propagation velocity of this D-disturbance, we may integrate Eq. (19) over V'; however, as stated before, contributions from the arbitrary vector field $\mathbf{A}$ must be made to vanish. To accomplish this, we first restrict $\mathbf{A}$ to that subset $\mathbf{B}$ of arbitrary vector fields that satisfy the condition

$$\text{curl}(\mathbf{B}-\mathbf{M}) = 0 \quad (\text{in } V_0'), \tag{38}$$

and Eq. (19) becomes

$$-D\mathbf{W} + \mathbf{f} = \text{curl}\mathbf{B}. \tag{39}$$

A consequence of this restriction is that

$$D\mathbf{W} = 0 \quad (\text{in } V_0'); \tag{40}$$

therefore, this restriction suppresses the static D-field contributions discussed in Sec. V.

The other step in suppressing the arbitrary contributions is to construct a vector function $\mathbf{h}$ that satisfies the conditions

$$\mathbf{h} = \text{curl}\mathbf{P} \quad (\text{in } V' \text{ and } V_0'), \tag{41a}$$

$$\mathbf{h} = \mathbf{f} \quad (\text{in } V_0'), \tag{41b}$$

and to subtract Eq. (41a) from (39):

$$-D\mathbf{W} + \mathbf{f} - \mathbf{h} = \text{curl}(\mathbf{B}-\mathbf{P}). \tag{42}$$

Since both sides of Eq. (42) vanish in V_0', we obtain upon integration over V'

$$-\iiint D\mathbf{W}dV' + \iiint(\mathbf{f}-\mathbf{h})dV' = 0. \tag{43}$$

Application of Eq. (23) in (43) then gives

$$\dot{\mathbf{R}}_E = E^{-1}\iiint(\mathbf{f}-\mathbf{h})dV', \tag{44a}$$

where by Eq. (22)

$$\mathbf{R}_E = E^{-1}\iiint\mathbf{r}D\,dV' \tag{44b}$$

and

$$E = \oint\!\!\oint \hat{n}\cdot\mathbf{v}dS' = \iiint D\,dV'. \tag{44c}$$

Finally, we must construct the vector field $\mathbf{h}$ in V'. Conditions (41) will be satisfied if

$$\mathbf{h} = \mathbf{f} \quad (\text{in } V_0'), \tag{45a}$$

$$\hat{n}\cdot[\mathbf{h}] = 0 \quad (\text{on } S'), \tag{45b}$$

$$\text{div}\mathbf{h} = 0 \quad (\text{in } V'). \tag{45c}$$

To complete the specification, curlh must also be specified in V'; we choose

$$\text{curl}\mathbf{h} = 0 \quad (\text{in } V'). \tag{45d}$$

Hence, $\mathbf{h}$ may be written

$$\mathbf{h} = \text{grad}\chi \quad (\text{in } V'), \tag{46a}$$

where χ satisfies the equation

$$\text{div grad}\chi = 0 \quad (\text{in } V') \tag{46b}$$

with the boundary condition

$$\frac{\partial\chi}{\partial n} = \hat{n}\cdot\mathbf{f} \quad (\text{on } S'). \tag{46c}$$

Therefore, $\mathbf{h}$ is determined by a scalar potential that satisfies Laplace's equation in V' with Neumann boundary conditions on S'. Equations (19), (44), and (46) constitute a complete formulation for the movement of local D-disturbances generated by Eq. (1).

VIII. EXAMPLES FROM ELECTRODYNAMICS

Movement of electric charge: The propagation of a local disturbance in electric charge density ρ may be determined from the Maxwell equations

$$\frac{\partial \mathscr{D}}{\partial t} + \mathscr{J} - \mathrm{curl}\mathscr{H} = 0,$$
$$\mathrm{div}\mathscr{D} = \rho$$

by identifying [Eq. (1)]

$$\mathbf{v} = \mathscr{D}, \quad \mathbf{f} = \mathscr{J} - \mathrm{curl}\mathscr{H},$$
$$\mathbf{c} = \mathrm{curl}\mathscr{D}, \quad D = \rho.$$

Thus for the locally disturbed volume V', the net charge is given by [Eqs. (44) and (46)]

$$E = \iiint \rho \, dV',$$

its centroid by

$$\mathbf{R}_E = E^{-1} \iiint \mathbf{r}\rho \, dV',$$

and the velocity of the centroid by

$$\dot{\mathbf{R}}_E = E^{-1} \iiint \mathscr{J} dV' - E^{-1} \iint \hat{n}\chi \, dS',$$

where the curl$\mathscr{H}$ term in $\mathbf{f}$ has been suppressed in accordance with Eq. (19), and where χ is obtained by solving

$$\mathrm{div\ grad}\chi = 0 \quad (\text{in } V')$$

with the boundary condition

$$\frac{\partial \chi}{\partial n} = \hat{n}\cdot\mathscr{J} \quad (\text{on } S').$$

The associated formulas for the propagation of a local disturbance in curl$\mathscr{D}$ are given by [Eqs. (34) and (36)]

$$\mathbf{G} = \oiint \hat{n}\times\mathscr{D} \, dS,$$

$$\left.\begin{matrix}\mathbf{R}_{G\perp}\\ \mathbf{R}_{G\parallel}\end{matrix}\right\} = \frac{\mathbf{G}}{G^2}\left\{\begin{matrix}\times\\ \cdot\end{matrix}\right\} \iiint \mathbf{r}\times\mathrm{curl}\mathscr{D} \, dV,$$

$$\left.\begin{matrix}\dot{\mathbf{R}}_{G\perp}\\ \dot{\mathbf{R}}_{G\parallel}\end{matrix}\right\} = -\frac{\mathbf{G}}{G^2}\left\{\begin{matrix}\times\\ \cdot\end{matrix}\right\} \iiint \mathscr{J} \, dV$$
$$+ \frac{\mathbf{G}}{G^2}\left\{\begin{matrix}\times\\ \cdot\end{matrix}\right\} \oiint (\hat{n}\times\mathscr{H} + \hat{n}\psi) \, dS \quad (V\in T), \tag{A1}$$

where

$$\mathrm{div\ grad}\psi = 0 \quad (\text{in } V), \tag{A2}$$

$$\hat{n}\times\mathrm{grad}\psi = \hat{n}\times(\mathscr{J} - \mathrm{curl}\mathscr{H}) \quad (\text{on } S). \tag{A3}$$

Here it is unneccessary to solve Eq. (A2) because the values of ψ required for insertion in Eq. (A1) may be obtained directly by integrating Eq. (A3) on S.

Movement of magnetic flux: The propagation of a local disturbance in the magnetic field strength $\mathscr{B}$ may be determined from the relations for the electrodynamic potentials $\mathscr{A}$ and φ

$$\frac{\partial \mathscr{A}}{\partial t} + \mathscr{E} + \mathrm{grad}\varphi = 0,$$
$$\mathrm{curl}\mathscr{A} = \mathscr{B},$$
$$\mathrm{div}\mathscr{A} = -\frac{1}{\mathscr{c}_c^2}\frac{\partial \varphi}{\partial t},$$

by identifying [Eq. (1)]

$$\mathbf{v} = \mathscr{A}, \quad \mathbf{f} = \mathscr{E} + \mathrm{grad}\varphi$$
$$\mathbf{c} = \mathscr{B}, \quad D = \mathrm{div}\mathscr{A}.$$

Thus for the locally disturbed volume V, we have, by Eqs. (34) and (36)

$$\mathbf{G} = \iiint \mathscr{B} \, dV,$$

$$\left.\begin{matrix}\mathbf{R}_{G\perp}\\ \mathbf{R}_{G\parallel}\end{matrix}\right\} = \frac{\mathbf{G}}{G^2}\left\{\begin{matrix}\times\\ \cdot\end{matrix}\right\} \iiint \mathbf{r}\times\mathscr{B} \, dV,$$

$$\left.\begin{matrix}\dot{\mathbf{R}}_{G\perp}\\ \dot{\mathbf{R}}_{G\parallel}\end{matrix}\right\} = -\frac{\mathbf{G}}{G^2}\left\{\begin{matrix}\times\\ \cdot\end{matrix}\right\} \iiint \mathscr{E} \, dV$$
$$+ \frac{\mathbf{G}}{G^2}\left\{\begin{matrix}\times\\ \cdot\end{matrix}\right\} \oiint \hat{n}\psi \, dS \quad (V\in T),$$

where the gradφ term in $\mathbf{f}$ has been suppressed in accordance with Eq. (8), and where ψ satisfies

$$\mathrm{div\ grad}\psi = 0 \quad (\text{in } V) \tag{A4}$$

and

$$\hat{n}\times\mathrm{grad}\psi = \hat{n}\times\mathscr{E} \quad (\text{on } S). \tag{A5}$$

Again Eq. (A4) need not be solved, since Eq. (A5) is sufficient for determining ψ on S.

The associated formulas for the propagation of a local disturbance in div$\mathscr{A}$ are given by [Eqs. (44) and (46)]

$$E = \oiint \hat{n}\cdot\mathscr{A} \, dS',$$

$$\mathbf{R}_E = E^{-1} \iiint \mathbf{r}(\mathrm{div}\mathscr{A}) \, dV',$$

$$\dot{\mathbf{R}}_E = E^{-1} \iiint \mathscr{E} \, dV' + E^{-1} \oiint \hat{n}(\varphi - \chi) \, dS',$$

where

$$\mathrm{div\ grad}\chi = 0 \quad (\text{in } V')$$

$$\frac{\partial \chi}{\partial n} = \hat{n}\cdot\mathscr{E} + \frac{\partial \varphi}{\partial n} \quad (\text{on } S').$$

Alternatively, we may write

$$\frac{\partial}{\partial t}\mathrm{grad}\varphi + \mathrm{grad}(\mathscr{c}_c^2\,\mathrm{div}\mathscr{A}) = 0$$

and identify [Eq. (1)]

$$\mathbf{v} = \mathrm{grad}\varphi, \quad \mathbf{f} = \mathrm{grad}(\mathscr{c}_c^2\,\mathrm{div}\mathscr{A}),$$
$$\mathbf{c} = 0, \quad D = \mathrm{div\ grad}\varphi.$$

The formulas for propagation of a local disturbance in div gradφ are given by [Eqs. (44) and (46)]

$$E = \oiint \frac{\partial \varphi}{\partial n} \, dS',$$

$$\mathbf{R}_E = E^{-1}\iint\left(\mathbf{r}\frac{\partial\varphi}{\partial n} - \hat{n}\varphi\right)dS',$$

$$\mathbf{R}_E = E^{-1}\iint (\upsilon_c^2\,\mathrm{div}\mathscr{A} - \chi)\hat{n}dS',$$

where

$$\mathrm{div\,grad}\chi = 0 \quad (\text{in } V'),$$

$$\frac{\partial\chi}{\partial n} = \frac{\partial}{\partial n}(\upsilon_c^2\,\mathrm{div}\mathscr{A}) \quad (\text{on } S').$$

In this example, only fields external to the disturbance are required to calculate E, $\mathbf{R}_E$, and $\dot{\mathbf{R}}_E$.

Finally, the propagation of a local disturbance in curl$\mathscr{B}$ may be obtained from the equations

$$\frac{\partial\mathscr{B}}{\partial t} + \mathrm{curl}\mathscr{E} = 0,$$

$$\mathrm{div}\mathscr{B} = 0$$

by identifying [Eq. (1)]

$$\mathbf{v} = \mathscr{B}, \quad \mathbf{f} = \mathrm{curl}\mathscr{E},$$

$$\mathbf{c} = \mathrm{curl}\mathscr{B}, \quad D = 0.$$

Thus for the locally disturbed volume V, we have by Eqs. (34) and (36)

$$\mathbf{G} = \oiint \hat{n}\times\mathscr{B}dS,$$

$$\left.\begin{matrix}\mathbf{R}_{G\perp}\\ \mathbf{R}_{G\parallel}\end{matrix}\right\} = \frac{\mathbf{G}}{G^2}\left\{\begin{matrix}\times\\ \cdot\end{matrix}\right\}\iiint \mathbf{r}\times\mathrm{curl}\mathscr{B}\,dV,$$

$$\left.\begin{matrix}\dot{\mathbf{R}}_{G\perp}\\ \dot{\mathbf{R}}_{G\parallel}\end{matrix}\right\} = -\frac{\mathbf{G}}{G^2}\left\{\begin{matrix}\times\\ \cdot\end{matrix}\right\}\oiint(\hat{n}\times\mathscr{E} - \hat{n}\psi)dS \quad (V\in T),$$

where ψ satisfies

$$\mathrm{div\ grad}\psi = 0 \quad (\text{in } V),$$

$$\hat{n}\times\mathrm{grad}\psi = \hat{n}\times\mathrm{curl}\mathscr{E} \quad (\text{on } S).$$

In this example, only the fields external to the disturbance must be known in order to calculate $\mathbf{G}$, $\dot{\mathbf{R}}_{G\perp}$, and $\dot{\mathbf{R}}_{G\parallel}$.

IX. DISCUSSION

As shown in the previous examples, our results allow one to calculate how local disturbances in a vector field will propagate without integrating forward in time. That is, if the fields in the vicinity of a local **c**-disturbance (D-disturbance) are measured at some epoch, we can calculate the instantaneous velocity of the centroid $\dot{\mathbf{R}}_G$ ($\dot{\mathbf{R}}_E$) provided $\mathbf{G}$ (E) is definite and, additionally for **c**-disturbances, provided the individual **c**-lines undergo pure translations. The velocities $\dot{\mathbf{R}}_G$ and $\dot{\mathbf{R}}_E$ are interpreted as the propagation velocities for these local disturbances. A word of caution is appropriate here, for since $\mathbf{c}$ (D) may take on positive and negative values, the centroid $\mathbf{R}_G$ ($\mathbf{R}_E$) may lie or move outside the disturbed volume V (V').

Some local disturbances, such as hurricanes, consist of both **c**- and D-disturbances superimposed. Since the formulas for $\dot{\mathbf{R}}_G$ and $\dot{\mathbf{R}}_E$ are different, the longevity of such a combined disturbance should depend upon agreement between the calculated values of these velocities.

X. CONCLUSIONS

The purpose of this theoretical investigation is to determine the propagation velocities of local disturbances that may occur in the curl and divergence of a vector field $\mathbf{v}$ that satisfies the equation $\partial\mathbf{v}/\partial t + \mathbf{f} = 0$, where $\mathbf{f}$ is a general vector field. Equations of this form frequently occur in physics and mechanics. A disturbance in the curl**v** (div**v**) field is defined by $\partial\mathrm{curl}\mathbf{v}/\partial t \neq 0$ ($\partial\,\mathrm{div}\mathbf{v}/\partial t \neq 0$). Such a disturbance is termed a local disturbance if it is surrounded by a quiescent layer of finite thickness.

Equations are derived for the velocity $\mathbf{U}(\mathbf{r},t)$ of curl**v** lines and the velocity $\mathbf{W}(\mathbf{r},t)$ of div**v** elements. These velocities conserve, respectively, the circulation of **v** about arbitrary circuits and the efflux of **v** through arbitrary closed surfaces. These conservation conditions are not sufficient to completely specify the two velocity fields, and each contains an arbitrary term. More convenient for three-dimensional applications than the circulation is the gyration **G** (spatial integral of curl**v**); and **U** is found to also conserve **G** in regions T where the curl**v** lines retain their individual shapes and orientations as they move, and our consideration of curl**v** disturbances is confined to such regions.

The centroid of efflux is defined for arbitrary volumes, and the centroid of gyration for volumes in T. The movement of these centroids is linked to the velocities **U** and **W** by the conservation relations. Contributions from the arbitrary term in **U** (**W**) are suppressed by integration over a local curl**v** (div**v**) disturbance and by construction of a scalar potential field that satisfies Laplace's equation in the interior of the disturbance with Dirichlet (Neumann) conditions related to **f** on the closed boundary.

Formulas for the velocities of the centroids are thus obtained for local curl**v** and div**v** disturbances as functions of **f**, and these velocities are interpreted as propagation velocities of the disturbances. Only one type of disturbance need be localized for its formula to apply. These formulas enable us to calculate how such local disturbances will propagate without integrating forward in time and in certain cases only the fields outside a disturbance are required for this calculation. Applications to Maxwell's electrodynamic equations are presented as examples.

[1]Lyman Spitzer, Jr., *Physics of Fully Ionized Gases* (Interscience, New York, 1962), 2nd ed., pp. 40–42.

[2]Arnold Sommerfield, *Mechanics of Deformable Bodies* (Academic, New York, 1950), pp. 285–87.

[3]R.C. Costen, NASA TN D-5964 (1970).

[4]R.C. Costen, J. Aircraft **9**, 406–12 (1972).

[5]James Serrin, Hanb. Phys. **8-1**, 125–263 (1959).

Theory of strings and membranes in an external field. I. General formulation

A. Aurilia

Blackett Laboratory, Imperial College, London SW7 2BZ, England and Istituto Nazionale di Fisica Nucleare, Sezione di Trieste, Italy [a)]

D. Christodoulou[b)]

Max Planck Institut für Physik und Astrophysik, München, Federal Republic of Germany

(Received 8 November 1978)

Motivated by recent developments in particle physics we give a mathematical formulation of the classical theory of extended objects under the influence of an external force field. Our geometric approach unifies the treatment of strings and membranes. We prove the existence and uniqueness of local solutions to the Cauchy problem associated with the object equations. We demonstrate the preservation of the topology of the object during its evolution. This allows us to extract the dynamical content of the theory. Some general properties of maximal timelike submanifolds, which represent worldtracks of extended objects when in free motion, are also deduced.

I. INTRODUCTION AND SUMMARY

In recent years, the search for classical solutions of theories containing extended objects has become a major topic of investigation in particle physics in view of the observed properties of hadronic systems. It is now generally accepted that strongly interacting particles are composite objects whose structure may be explained in terms of more fundamental entities (quarks) confined to the interior of the hadron. Several models were proposed in an attempt to describe the basic features of hadron spectroscopy in terms of objects that are extended in space.[1-3] Thus, within the string picture of hadrons, the states of motion of a relativistic string with open ends are commonly interpreted as mesonic resonances while the strong interaction background (the Pomeron) is usually associated with the states of motion of a closed string.[3,4] Extended objects of higher dimensionality, in particular the membranes, originally studied by Dirac in the context of electrodynamics, qualify equally well as candidates for possible formulation of hardons dynamics.[2,5,6] By analogy with the mechanics of point particles, it is assumed that extended objects, when in free motion, sweep out a maximal area in spacetime.[1,5] However, in spite of the growing literature on the subject, a general, unified, and precise mathematical formulation is, in our opinion, still lacking. It is the aim of this paper to provide such a formulation for the classical dynamics of relativistic extended objects in a given external force field. In an accompanying article we shall treat the case where the external field is static and uniform, we shall classify the solutions according to symmetry groups, and we shall exhibit all solutions possessing a sufficiently high symmetry.[7] In the present paper the theory is developed in complete analogy with classical electrodynamics of point charges. Our approach is mostly geometric, allowing for a unified, up to a certain point, treatment of strings and membranes.

The plane of the paper is as follows. We start our formulation in Sec. IIA with a precise definition of an extended object as a geometric entity in spacetime. The external field is introduced as a differential form satisfying a generalization of Maxwell's equations. We then introduce the action functional, the invariance properties of which are discussed in Sec. IIB. The principle of stationary action leads in Sec. IIIA to the generalized Lorentz force, or object, equations. The identities which result from the reparametrization invariance and the conserved quantities which result from the Poincare invariance are also derived. The introduction of the harmonic condition in Sec. IIIB reduces the object equations to a system of weakly coupled quasilinear hyperbolic partial differential equations and allows us to establish existence and uniqueness of local solutions to the Cauchy problem. Other authors who have studied the string case reduce the object equations to a semilinear hyperbolic system by introducing isothermal coordinates. This is possible for the string, due to the special circumstance that two-dimensional manifolds are locally conformally flat, but cannot be generalized to higher dimensions. Our approach, on the other hand, applies equally well to strings and membranes, indeed to any r-dimensional submanifold of any n-dimensional globally hyperbolic spacetime ($1 \leqslant r \leqslant n-1$). In Sec. IVA we prove a proposition which establishes the preservation of the topology of the extended object during its evolution. This proposition allows us to give a dynamical formulation of the theory in Sec. IVB and to reduce the object equations to two dynamical or evolution equations in the case of the string and to a single such equation in the case of the membrane. The paper ends with two propositions which establish some general properties of the free motion of extended objects.

[a)]Permanent address.
[b)]Alexander von Humboldt fellow.

II. FORMULATION

A. Action

Let us denote the spacetime by (R^4,η), the manifold R^4 endowed with the Minkowski metric η. We define the world track of the extended object as an r-dimensional timelike submanifold $\xi(K)$ of Minkowski space, the image through an embedding ξ of a connected orientable r-manifold K (without boundary), which serves as the topological model. The space $E'(K,R^4)$ of timelike embeddings of K into R^4 is a cone in the linear space $C^\infty(K,R^4)$ of differential maps of K into R^4. The embedding ξ induces on K the pullback metric $g=\xi_*\eta$ and the condition that ξ be timelike is equivalent to g being a Lorentzian metric. We introduce in addition an external force field F which is a differential form on R^4 of degree equal to $r+1$. We shall assume that F is conservative, that is, there exists a potential r-form A such that $F=dA$. If Σ is any compact subdomain in K, we define the action of the object in Σ to be the functional

$$S_\Sigma(\xi)=\rho\int_\Sigma \mu_g + c\int_{\xi(\Sigma)} A, \tag{2.1}$$

where μ_g denotes the canonical measure of the metric g while ρ and c are given real constants, generalized mass and generalized charge, respectively, with $\rho>0$. An important example of an external force field is a generalized Maxwell field which satisfies the generalized Maxwell equations:

$$dF=d\,{*F}=0, \tag{2.2}$$

the first of which, in view of the simple connectedness of R^4, is equivalent to the condition that F be conservative. In the case $r=2$ we shall call the object "string" and in the case $r=3$ we shall name it "membrane." For $r=1$ and F a Maxwell field, our formulation reduces to standard electrodynamics of point charges. If $r=2$, $*F$ is a 1-form and the second of Eqs. (2.2) implies

$$*F=d\varphi,$$

where φ is a scalar function. Then the first of Eq. (2.2) gives $d{*d\varphi}=0$, which is equivalent to

$$\Box\varphi=0.$$

If $r=3$, $*F$ is a scalar function and the first of Eqs. (2.2) is an identity. The second then gives

$$*F=\text{const}.$$

Thus, in the string case the generalized Maxwell tensor $F_{\mu\nu\rho}$ possesses a single radiative degree of freedom,[4] while in the membrane case $F_{\mu\nu\rho\sigma}$ contains no radiation field at all.[5] The idea underlying our geometric formulation is that antisymmetric tensor fields play a central role in particle physics. This idea originates not only from quantum electrodynamics but also from more recent investigations in string models of strongly interacting particles.[1,3,4] As the standard Maxwell tensor $F_{\mu\nu}$ mediates electromagnetic interactions between point charges, the tensor $F_{\mu\nu\rho}$ mediates the string–string interaction via the exchange of massless spinless quanta.[4,7] In consistency with the established string–vortex analogy,[4,9] a massless longitudinal radiation is also associated to the hydrodynamic motion of a relativistic vortex in a superfluid.

The role of the antisymmetric tensor field $F_{\mu\nu\rho\sigma}$ associated with the relativistic membrane is altogether different and was investigated in a previous communication.[5] It acts as a pressure and, unlike the case of point charges and strings, leads to a finite self-energy proportional to the volume enclosed by the membrane. In the so-called bag formulation of hadron dynamics the presence of such a pressure term is responsible for the confinement of the hadronic constituents to the interior of the bag.

B. Invariance properties of the action

The symmetry properties of the action functional defined by Eq. (2.1) play an important role in the theory of extended objects. On $E'(K,R^4)$ acts the group $\mathscr{D}(K)$ of diffeomorphisms of K on the right: The action

$$E'(K,R^4)\times\mathscr{D}(K)\to E'(K,R^4)$$

sends

$$(\xi,f)\to\xi\circ f. \tag{2.3}$$

This action is free of fixed points, since $\xi\circ f=\xi, f\neq$ identity, would contradict the injectivity of ξ. A subgroup of $\mathscr{D}(K)$ is the group $\mathscr{D}^+(K)$ of orientation preserving diffeomorphisms. The space of timelike oriented images of K in (R^4,η) is the quotient: $E'(K,R^4)/\mathscr{D}^+(K)$. The action functional S_Σ is invariant under $\mathscr{D}^+(K)$ but not under the complete group $\mathscr{D}(K)$,

$$\forall f\in\mathscr{D}^+(K): S_\Sigma(\xi\circ f)=S_{f(\Sigma)}(\xi),$$

but

$$\forall f\in\mathscr{D}(K)-\mathscr{D}^+(K): S_\Sigma(\xi\circ f;\rho,c)=S_{f(\Sigma)}(\xi;\rho,-c). \tag{2.4}$$

Thus, an orientation altering diffeomorphism reverses the sign of the generalized charge. The orbit O_ξ of $\mathscr{D}(K)$ through ξ is the image of the map

$$O_\xi:\mathscr{D}(K)\to E'(K,R^4),$$

where

$$O_\xi(f)=\xi\circ f, \tag{2.5}$$

and the tangent to this map at the identity element in $\mathscr{D}(K)$ sends each vector field η on K into:

$$\eta\to L_\eta\xi=\eta(\xi)=\xi{*\eta}, \tag{2.6}$$

where L_η is the Lie derivative of ξ with respect to η and $\xi*\eta$ is the pushout of η on the spacetime R^4. Since ξ is an immersion, $L_\eta\xi=0$ implies $\eta=0$. The Poincare group P of Minkowski space acts on $E'(K,R^4)$ on the left: The action

$$P\times E'(K,R^4)\to E'(K,R^4)$$

sends

$$(p,\xi)\to p\circ\xi. \tag{2.7}$$

If there is a subgroup P' of the Poincare group which leaves invariant the external potential A ($\forall p'\in P': P'_* A=A$), then the action S_Σ is itself invariant under the same subgroup:

$$S_\Sigma(p'\circ\xi)=S_\Sigma(\xi). \tag{2.8}$$

Finally the action is invariant under the group of generalized gauge transformations on the external potential: $A\to A+d\chi$, χ an $r-1$ form such that $\chi/_{\partial\Sigma}=0$.

III. THE PROBLEM

A. Object equations

The requirement that S_Σ be stationary with respect to variations of ξ which vanish on $\partial\Sigma$, leads to the generalized Lorentz force equation:

$$\rho H^\mu = c f^\mu, \tag{3.1}$$

where H^μ is the mean curvature vector of $\xi(K)$ given by:

$$H^\mu = \Box_g \xi^\mu \tag{3.2}$$

where f^μ is the generalized Lorentz force:

$$f^\mu = F^\mu{}_{|\nu_0\cdots\nu_{r-1}|}\dot{\xi}^{\nu_0\cdots\nu_{r-1}}/\|\dot{\xi}\|. \tag{3.3}$$

In the above equation, $\dot{\xi}$ is the r-vector representing the tangent element to the embedded submanifold $\xi(K)$:

$$\dot{\xi} = \frac{\partial\xi}{\partial s^0}\wedge\cdots\wedge\frac{\partial\xi}{\partial s^{r-1}}, \tag{3.4}$$

expressed in local coordinates $(s^0\cdots s^{r-1})$ on K. Further, $\|\dot{\xi}\|$ denotes the norm of $\dot{\xi}$:

$$\|\dot{\xi}\| = |\dot{\xi}_{|\nu_0\cdots\nu_{r-1}|}\dot{\xi}^{\nu_0\cdots\nu_{r-1}}|^{1/2} = |\det g|^{1/2}. \tag{3.5}$$

The invariance property of the action S_Σ under those $f\in\mathscr{D}^+(K)$ which preserve $\Sigma\,[f(\Sigma)=\Sigma]$ is in infinitesimal form expressed by the statement that for any vector field η on K which vanishes on $\partial\Sigma$ [cf. Eq. (2.6)],

$$D_\xi S_\Sigma\cdot L_\eta\xi = \int \eta^a\frac{\partial\xi^\mu}{\partial s^a}(-\rho H_\mu + c f_\mu)\mu_g = 0, \tag{3.6}$$

which in turn implies that Eqs. (3.1) satisfy the identities

$$(\rho H^\mu - c f^\mu)\frac{\partial\xi_\mu}{\partial s^a} = 0, \tag{3.7}$$

where in fact each of $H^\mu\partial\xi^\mu/\partial s_a, f^\mu\partial\xi_\mu/\partial s^a$ individually vanish. If X [vector field in (R^4,η)] is any element of the Lie algebra of a subgroup P' of the Poincare group P leaving invariant the potential A, then the invariance property of S_Σ under P' yields:

$$D_\xi S_\Sigma\cdot X\circ f = \int_\Sigma X^\mu(-\rho H_\mu + c f_\mu)\mu_g + \int_{\xi(\partial\Sigma)} X\neg(\rho p + cA) = 0, \tag{3.8}$$

where p is the normalized r-form corresponding to $\dot{\xi}$:

$$p_{\mu_0\cdots\mu_{r-1}} = \dot{\xi}_{\mu_0\cdots\mu_{r-1}}/\|\dot{\xi}\|. \tag{3.9}$$

As the first of the integrals in Eq. (3.8) vanishes in virtue of Eq. (3.1), the boundary integral itself vanishes, a fact which will yield conservation laws in the next section. As a consequence of the symmetry properties of the action under orientation altering diffeomorphisms [Eq. (2.4)], if $\xi(K)$ is a timelike submanifold satisfying Eq. (3.1), the same submanifold with the reversed orientation will satisfy the equation obtained from Eq. (3.1) by replacing c by $-c$.

B. The Cauchy problem

We wish now to formulate a local Cauchy problem for Eq. (3.1). Let V be a bounded coordinate neighborhood in K and let Δ, a bounded open set in R^{r-1}, be the surface $s^0=0$ on V. Let us give data $\xi^\mu(0,s^1,\ldots,s^{r-1})$, $\partial\xi^\mu/\partial s^0(0,s^1,\ldots,s^{r-1})$ on Δ such that: (1) ξ/Δ is an embedding, (2) the metric g_{AB} $(A,B=1,\ldots,r-1)$ induced on Δ is uniformly positive definite, and (3) the supremum in Δ of the determinant of g_{ab} $(a,b,=0,\ldots,r-1)$ is negative. We wish then to know whether for some neighborhood $U\subset V$ of Δ there exists one and only one timelike submanifold $\xi(U)$ satisfying Eqs. (3.1) and conforming with the data. For the analytical study of this problem Eqs. (3.1) are not suitable. This is due to the fact that they are not hyperbolic but degenerate, satisfying the identitites (3.7) which manifest the freedom involved in choosing coordinates on Σ. Writing the operator $\Box_g$ which acts on ξ^μ in Eqs. (3.1) in the form,

$$\Box_g\xi^\mu = g^{ab}\frac{\partial^2\xi^\mu}{\partial s^a\,\partial s^b} - \Gamma^a\frac{\partial\xi^\mu}{\partial s^a}, \tag{3.10}$$

where

$$\Gamma^a = g^{bc}\Gamma^a_{bc} = \frac{1}{|\det g|^{1/2}}\frac{\partial}{\partial s^b}(|\det g|^{1/2}g^{ab}), \tag{3.11}$$

and recalling that

$$g_{ab} = \frac{\partial\xi^\mu}{\partial s^a}\frac{\partial\xi_\mu}{\partial s^b},$$

we see that the principal (second order) part of expression (3.10) is not totally contained in the hyperbolic first term but appears also in Γ^a. To reduce the Lorentz force equations to a hyperbolic system we impose the condition

$$\Gamma^a = 0, \tag{3.12}$$

which is equivalent to requiring that the coordinates s^a are harmonic:

$$\Box_g s^a = 0. \tag{3.12'}$$

Thus we obtain the reduced Lorentz force equations:

$$\rho\, g^{ab}\frac{\partial^2\xi^\mu}{\partial s^a\,\partial s^b} = cf^\mu, \tag{3.13}$$

constituting a (weakly coupled) system of quasilinear hyperbolic partial differential equations having the same principal part. In the following two lemmas we shall show that Eqs. (3.1) and Eqs. (3.13) are locally equivalent:

Lemma 1: For any timelike embedding ξ: $V\to R^4$ there exists a unique coordinate transformation f: $U\to V$ such that the timelike embedding $\xi' = \xi\circ f: U\to R^4$ satisfies the harmonic condition and has the same Cauchy data as ξ.

Proof: The new coordinates f^a must satisfy the equations:

$$\Box_{g'} f^a = 0, \tag{3.14}$$

where $g' = \xi'_*\eta = f_*\,g$, $g = \xi_*\eta$. Each component of the above equations is a scalar equation. In the original coordinates they reduce to

$$\Box_g f^a = 0, \tag{3.15}$$

which, since g is given, are linear hyperbolic equations for the functions f^a. Thus we have unique solutions for given

Cauchy data on Δ in any open U' which is a development of Δ. Let us pose the following data on Δ:

$$f^0(0,s^1,...,s^{r-1}) = 0, \quad f^A(0,s^1,...,s^{r-1}) = s^A,$$
$$(A = 1,...,r-1),$$
$$\frac{\partial f^0}{\partial s^0}(0,s^1,...,s^{r-1}) = 1, \quad \frac{\partial f^A}{\partial s^0}(0,s^1,...,s^{r-1}) = 0. \tag{3.16}$$

Then f restricted to Δ is the identity (Δ is carried into itself) and the tangent to f at any point on Δ is also the identity. From the implicit function theorem[10] it follows that there is a neighborhood $U \subset U'$ of Δ in which f is a diffeomorphism.

Finally conditions (3.16) are obviously sufficient to ensure that ξ and ξ' have the same Cauchy data. That they are also necessary is implied by the injectivity of ξ and the fact that ξ is an immersion.

Lemma 2: Any timelike embedding ξ which satisfies the reduced Eqs. (3.13) also satisfies Eqs. (3.1).

Proof: We consider identities (3.7) and taking into account Eqs. (3.13) we replace $\rho H^\mu - cf^\mu$ by $-\rho\Gamma^b\, \partial\xi^\mu/\partial s^b$ [cf. Eq. (3.10)]. We obtain:

$$\Gamma^b \frac{\partial \xi^\mu}{\partial s^b}\frac{\partial \xi_\mu}{\partial s^a} = \Gamma^b g_{ab} = 0,$$

which, since g is nondegenerate implies the harmonic condition $\Gamma^a = 0$, and the lemma is established. Leray's theory[11] of strictly hyperbolic systems can directly be applied to Eqs. (3.13) to yield, with the use of Lemma 2, the following theorem:

Theorem 1: For any Cauchy data on Δ satisfying the three conditions stated above, there is a neighborhood U of Δ (development of Δ) for which there exists one and only one timelike embedding ξ: $U \to R^4$ satisfying Eqs. (3.13). This embedding defines both a submanifold $\xi(U)$ satisfying Eqs. (3.1) and also a harmonic coordinate system on this submanifold.

The question of uniqueness of the solutions to Eqs. (3.1) is answered by:

Theorem 2: If ξ_1 and ξ_2 are two solutions of Eqs. (3.1) having the same Cauchy data, there is a diffeomorphism f such that $\xi_2 = \xi_1 \circ f$.

Proof: According to Lemma 1 there exist unique diffeomorphisms h_1 and h_2 such that $\xi_1' = \xi_1 \circ h_1$ and $\xi_2' = \xi_2 \circ h_2$ satisfy both the harmonic condition and have the same Cauchy data as ξ_1 and ξ_2, respectively. Hence ξ_1' and ξ_2' are solutions of the reduced Eqs. (3.13) with identical Cauchy data. By Theorem 1, ξ_1' and ξ_2' must be identical, and defining $f = h_1 \circ h_2^{-1}$ we obtain $\xi_2 = \xi_1 \circ f$.

We wish finally to note that in the case $r = 1$ the concept of harmonic coordinates reduces to the concept of an affine parameter along the world line.

IV. THE DYNAMICS

A. Preservation of topology

The introduction of the harmonic condition allowed us in the previous section to establish in a unified way the existence and uniqueness properties of the solutions of the object equations. However, the full meaning of these equations as describing the evolution of extended objects under the influence of external force fields has not been revealed thus far in our exposition. Of central importance to the understanding of the dynamical content of the theory of extended objects is the following proposition:

Proposition 1: Let k be a connected r-manifold, without boundary, and ξ: $K \to R^4$ an embedding of K as a timelike submanifold in Minkowski space (R^4,η). Let in addition $\xi(K)$ be a closed subset of R^4. Then there is a connected $r-1$ manifold M, without boundary, such that $K \cong R \times M$.

Proof: Consider the differentiable function ξ^0 on K. Since K is connected, $\xi^0(K)$ is a connected subset of R. Thus $\xi^0(K)$ is either R itself or an interval: open, closed, or half-open. We shall now show that the function ξ^0 has no critical points. For, if q is a critical point of ξ^0, then

$$(d\xi^0)_q = 0.$$

But this implies that the metric $g = \xi_*\eta$ induced on K is positive-semidefinite at q:

$$(g)_q = \sum_{i=1}^{3} (d\xi^i \otimes d\xi^i)_q,$$

in contradiction to the assumption that g is Lorentzian everywhere on K. The fact that ξ^0 is a differentiable function without critical points implies through the implicit function theorem that the level surfaces $M_t = (\xi^0)^{-1}(t)$, $t \in \xi^0(K)$ are C^∞ closed submanifolds of K. Further, $\xi^0(K)$ cannot be a closed or half-open interval of R since, in that case, the endpoints of the interval would be extrema, and hence critical points of ξ^0. Consequently, $\xi^0(K)$ is either an open interval in R or R itself, in either case diffeomorphic to R. The lemma would follow if we show that $K = \cup_{t\in\xi^0(k)} M_t$ is a foliation, namely that the level surfaces of ξ^0 are diffeomorphic to each other. Consider the C^∞ vector field u on K defined by:

$$g(u) = d\xi^0.$$

This vector field is nowhere vanishing and is orthogonal to the surfaces M_t. Therefore, the pushout ξ^*u is orthogonal to $\xi(M_t)$, namely to the surfaces formed by the intersection of the $x^0 = t$ hyperplanes in R^4 with the embedded manifold $\xi(K)$. These intersections are spacelike; hence ξ^*u is timelike, and the same is true for u. We now define on K the vector field n by:

$$n = (1/\varphi)u, \tag{4.1}$$

where

$$\varphi = g(u,u).$$

This vector field is timelike and C^∞, since $\varphi < 0$. We shall demonstrate that the vector field n is complete. Let μ_q: $I_a =]-a,a[\to K$ be an integral curve of n through some point

$q \in K$, whose domain cannot be extended. Let $\{t_n\}$ be a sequence in I_a converging to a. Then the sequence $\{\xi^0(\mu_q(t_n))\}$ converges to $\xi^0(q) + a$. This follows from

$$\frac{\partial}{\partial s}(\xi^0 \circ \mu_q) = (n, d\xi^0) = 1,$$

which implies ($\forall s \in I_a$)

$$\xi^0 \circ \mu_q(s) = \xi^0 \circ \mu_q(0) + s = \xi^0(q) + s.$$

Now μ_q is a timelike curve and so is $\xi \circ \mu_q$. Hence, $\mathrm{Im}\xi \circ \mu_q$ is contained in the light cone at $\xi(q)$. Therefore, $\{\xi(\mu_q(t_n))\}$ is a bounded sequence in R^4 and we can extract a subsequence $\{\xi(\mu_q(t_{n_j}))\}$ converging to a point $x' \in R^4$. Clearly, $x' \in \xi(K)$. On the other hand, $x' \notin \xi(K)$, otherwise μ_q could be extended beyond I_a (K having no boundary). Thus $\xi(K)$ is not closed and we have a contradiction. The same argument holds for the lower limit of I_a. We conclude that the domain of μ_q can be extended to R and the vector field n is complete. Also we obtain $\xi^0(K) = \xi^0(\mathrm{Im}\mu_q) = R$. The complete C^∞ vector field n generates the one-parameter group of diffeomorphisms[12]

$\{f_s : K \cong K \mid s \in R\}$, where

$$\forall\ s \in R, \quad q \in K: \quad f_s(q) = \mu_q(s),$$

whose action on the function ξ^0 is given by

$$f_{s*}\xi^0(q) = \xi^0 \circ \mu_q(s) = \xi^0(q) + s,$$

whence

$$f_{s*}\xi^0 = \xi^0 + s.$$

It follows that the restriction $f_s|_{M_t}$ is a diffeomorphism of M_t into M_{t+s} and the surfaces M_t are diffeomorphic to each other. Consequently, $K = \cup_{t \in R} M_t$ is a foliation and the map

$$h : K \to R \times M_0, \quad q \to (\xi^0(q), f_{-\xi^0(q)}(q)), \tag{4.2}$$

is a diffeomorphism.

Consider now the case M_0 compact and take the compact domain Σ in Eq. (3.8) to be $\Sigma = \cup_{a \leqslant t \leqslant b} M_t$. Then $\partial\Sigma = M_b - M_a$ and Eq. (3.8) implies that the quantity

$$\mathscr{E}_X = \int_{\xi(M_t)} X \neg (\rho p + cA) \tag{4.3}$$

is independent of t, that is, conserved.

B. Dynamical formulation

From the proof of Proposition 1 we conclude that given a closed timelike embedding $\xi : K \to R^4$ we can define an equivalent embedding $x : R \times M_0 \to R^4$ by $x = \xi \circ h^{-1}$ where [cf. Eq. (4.2)] $h^{-1} : R \times M_0 \cong K$ is given by: $(t,p) \to f_t(p)$. The time component of x is trivial: $x^0(t,p) = t$, while the space component $\bar{x}_t = \bar{x}(t,\cdot) = \bar{\xi} \circ f_t$ is a curve in the space $E(M_0, R^3)$ of embeddings of M_0 into R^3. The space $E(M_0, R^3)$ is a cone in the linear space $C^\infty(M_0, R^3)$ of differentiable maps of M_0 into R^3. The action, on $E(M_0, R^3)$, of the group $\mathscr{D}(M_0)$ of diffeomorphisms of M_0 is analogous to the action of $\mathscr{D}(K)$ on $E'(K, R^4)$. Similarly, the action of the Euclidean group on $E(M_0, R^3)$ is analogours to the action of the Poincare group on $E'(K, R^4)$. We endow R^3 with the Euclidean metric e, the inner product of which will be denoted by "$\cdot$" and the norm by "$|\cdot|$", as usual. The tangent to the velocity [cf. Eq. (4.1)]

$$\bar{v}_t = \frac{\partial \bar{x}_t}{\partial t} = (\bar{\xi}^* n)|_{M_t}, \tag{4.4}$$

which is at each t a vector field on $\bar{x}_t(M_0)$.

Proposition 2: The velocity vector field $\bar{v}_t$ is orthogonal to the submanifold $\bar{x}_t(M_0)$.

Proof: Let ψ_0 be any vector field on M_0. We define a vector field ψ on K tangent to the foliation $\{M_t \mid t \in R\}$ by posing, for each $t \in R$, $\psi \mid M_t = \psi_t = f_t^* \psi_0$. Then, since the vector field n is orthogonal to the foliation, we have:

$$g(n, \psi) = \eta(\xi^* n, \xi^* \psi) = 0.$$

Further, since ψ leaves the function ξ^0 invariant, we have $\xi^{0*}\psi = 0$. It follows that

$$\bar{\xi}^* n \cdot \bar{\xi}^* \psi = \eta(\xi^* n, \xi^* \psi) = 0.$$

Restricting the above to each M_t we obtain

$$\bar{v}_t \cdot (\bar{x}_t{}^* \psi_0) = 0, \tag{4.5}$$

which shows that $\bar{v}_t$ is orthogonal to any vector field tangent to $\bar{x}_t(M_0)$. In local coordinates $\{\sigma^A \mid A = 1,...,r-1\}$ on M_0, Eq. (4.5) reduces to

$$\bar{v}_t \cdot \frac{\partial \bar{x}_t}{\partial \sigma^A} = 0. \tag{4.6}$$

As a consequence of Proposition 1, the metric $g = x_* \eta$ (we use the same symbol g from now on to denote the metric induced on $R \times M_0$) assumes the form

$$g = -N^2\, dt \otimes dt + \gamma, \tag{4.7}$$

where for each t

$$N_t = (1 - |\bar{v}_t|^2)^{1/2} \tag{4.8}$$

and

$$\gamma_t = \bar{x}_t{}^* e. \tag{4.9}$$

The condition that the embedding x be timelike is seen to be equivalent to the requirement that $|\bar{v}_t| < 1$ everywhere on $\bar{x}_t(M_0)$ and for all t.

The energy $\mathscr{E}$ of the object is the quantity $\mathscr{E}_X$ associated by Eq. (4.3) to the generator of time translations $X = \partial/\partial x^0$. It can be written as the sum $\mathscr{E} = \mathscr{E}_k + \mathscr{E}_p$ of kinetic and potential contributions:

$$\mathscr{E}_k = \rho \int_{x(M_t)} p_\perp = \rho \int_{M_0} \frac{\mu_{\gamma_t}}{N_t}, \tag{4.10}$$

where if ω is an n-form on R^4, $\omega_\perp$ is the $(n-1)$-form $\partial/\partial x^0 \neg \omega$ namely the projection of ω normal to the hyperplane $x^0 = t$, and

$$\mathscr{E}_p = c \int_{x(M_t)} A_\perp. \tag{4.11}$$

In the case that there exists an (r-dimensional) submanifold V_t of the hyperplane $x^0 = t$ such that $\partial V_t = x(M_t)$, we can express the potential energy also in the form:

$$\mathscr{E}_p = c \int_{V_t} F^\perp. \tag{4.12}$$

The equivalence of expressions (4.11) and (4.12) follows from the formula $d(X \neg A) = L_X A + X \neg dA$, the hypothesis that $L_X A = 0$ for $X = \partial/\partial x^0$, and Stokes' theorem.

In the case of the string, we can apply the theory of space curves of differential geometry and define at each point on the curve $\bar{x}_t(M_0)$ the orthonormal triad $(\bar{l},\bar{n},\bar{b})$ where l is the unit tangent:

$$\bar{l} = \frac{\partial \bar{x}}{\partial s} = \frac{1}{|\det\gamma|^{1/2}} \frac{\partial \bar{x}}{\partial \sigma^1}, \tag{4.13}$$

s being the arc length, $\bar{n}$ is the principal normal:

$$\frac{\partial^2 \bar{x}}{\partial s^2} = \Delta_\gamma \bar{x} = k\bar{n}, \tag{4.14}$$

k being the curvature, and $\bar{b}$ is the binormal, namely the unique vector such that

$$\|\bar{l},\bar{n},\bar{b}\| = \bar{l}\cdot(\bar{n}\times\bar{b}) = 1. \tag{4.15}$$

As a consequence of Eq. (4.6), the expansion of $\bar{v}$ in the basis $(\bar{l},\bar{n},\bar{b})$ has the form:

$$\bar{v} = u\bar{n} + w\bar{b}. \tag{4.16}$$

In the case of the membrane, the mean curvature vector of the surface $\bar{x}_t(M_0)$, being normal to the surface, is expressed as

$$\bar{K} = \Delta_\gamma \bar{x} = K\bar{n}, \tag{4.17}$$

where K is the mean curvature (trace of the second fundamental form, sum of the principal curvatures) and $\bar{n}$ is the unit normal:

$$\bar{n} = \left(\frac{\partial \bar{x}}{\partial \sigma^1} \times \frac{\partial \bar{x}}{\partial \sigma^2}\right)\left(\left|\frac{\partial \bar{x}}{\partial \sigma^1} \times \frac{\partial \bar{x}}{\partial \sigma^2}\right|\right)^{-1}. \tag{4.18}$$

Further, in the case of the string $*F$ is a 1-form which we can write as:

$$*F = B dx^0 - \sum_{i=1}^{3} E^i dx^i, \tag{4.19}$$

and in the case of the membrane, $*F$ is a 0-form (function)

$$*F = \alpha.$$

We note that $E^i = \epsilon^{ijk}(F_\perp)_{jk}$ and $\alpha = \epsilon^{ijk}(F_\perp)_{ijk}$. Thus [cf. Eq. (4.12)] the potential energies of the string and membrane are given by:

$$\mathscr{E}_p = c\int_{S_t} \bar{E}\cdot d\bar{S} \tag{4.20}$$

and

$$\mathscr{E}_p = c\int_{V_t} \alpha\, d^3x, \tag{4.21}$$

respectively, where at a given time t, S_t is any surface having the string as boundary while V_t is the volume enclosed by the membrane.

In view of identities (3.7) the time component of Eq. (3.1) for the embedding x is a consequence of the space component. This time component gives the rate of change ζ of the kinetic energy density [cf. Eq. (4.10)]

$$\zeta = \frac{1}{|\det\gamma|^{1/2}} \frac{\partial}{\partial t}\left(\rho \frac{|\det\gamma|^{1/2}}{N}\right). \tag{4.22}$$

In the case of the string we have

$$\zeta = c\|\bar{E},\bar{v},\bar{l}\| = c(E_n w - E_b u), \tag{4.23}$$

where E_n and E_b are the components of $\bar{E}$ along $\bar{n}$ and $\bar{b}$, respectively. In the case of the membrane we have

$$\zeta = c\alpha v, \tag{4.24}$$

where, in view of Eq. (4.6),

$$\bar{v} = v\bar{n}. \tag{4.25}$$

The identities (3.7) imply also that the projection of the space component of Eq. (3.1) on the tangent space to $\bar{x}_t(M_0)$ is an identity. Thus in the case of the string, Eqs. (3.1) are equivalent to the following pair of evolution equations obtained by projecting the space component on the plane spanned by $\bar{n}$ and $\bar{b}$:

$$\rho\left(\frac{\partial u}{\partial t} + \nu w - N^2 k\right) = cF_n N, \tag{4.26a}$$

$$\rho\left(\frac{\partial w}{\partial t} + \nu u\right) = cF_b N, \tag{4.26b}$$

where

$$F_n = -(1-u^2)E_b - uwE_n - Bw, \tag{4.27a}$$

$$F_b = (1-w^2)E_n + uwE_b + Bu. \tag{4.27b}$$

Here, ν is the normal angular velocity of the osculating plane of the string:

$$\nu = \bar{n}\cdot\frac{\partial \bar{b}}{\partial t}. \tag{4.28}$$

In the case of the membrane, Eqs. (3.1) are equivalent to the single evolution equation obtained by projecting the space component along the unit normal $\bar{n}$:

$$\rho\left(\frac{\partial v}{\partial t} - N^2 K\right) = c\alpha N^3. \tag{4.29}$$

In deriving Eqs. (4.26) and (4.29) use is made of the following formula:

$$\frac{1}{|\det\gamma|^{1/2}} \frac{\partial}{\partial t}|\det\gamma|^{1/2} = -(\Delta_\gamma \bar{x})\cdot\bar{v},$$

which follows when Eq. (4.6) is taken into account.

Finally, if external fields are absent, Eqs. (4.23) and (4.24) give $\zeta = 0$, which yields

$$\rho|\det\gamma|^{1/2}/N = \epsilon, \tag{4.30}$$

where ϵ is a fixed density on M_0. From Eq. (4.30) the following two propositions follow:

Proposition 3: The length of any compact $\Delta \subset M_0$ is bounded, that is, there exists a positive real number L such that $\forall\, t \in R$:

$$\int_\Delta \mu_{\gamma_t} \leqslant L.$$

Proposition 4: If for some $t \in R$, $\bar{x}_t$ is singular at p, that is, $\det\gamma_t(p) = 0$, then $|\bar{v}_t(p)| = 1$ and the curve $x(\cdot,p)$ (in Minkowski space) is null at the point $x(t,p)$.

[1] Y. Nambu, Lectures for the Copenhagen Summer Symposium, 1970; T. Goto, Prog. Theor. Phys. **46**, 1560 (1971); P. Goddard, T. Goldstone, C. Rebbi, and C.B. Thorn, Nucl. Phys. B **56**, 109 (1973).

[2] A. Chodos *et al.*, Phys. Rev. D **9**, 3471; **10**, 2599 (1974); W.A. Bardeen *et al.*, Phys. Rev. D **11**, 1094 (1974); P. Hasenfratz and J. Kuti, Phys. Rep. C **40**, 75 (1978).

[3] S. Mandelstam, Phys. Rep. **13**, 260 (1974); J. Scherk, Rev. Mod. Phys. **47**, 123 (1975).

[4] M. Kalb and P. Ramond, Phys. Rev. D **9**, 2273 (1974); Y. Nambu, "Strings, Vortices and Gauge Fields," talk presented at the Rochester Symposium on Quark Confinement, June, 1976; F. Lund and T. Regge, Phys. Rev. D **14**, 1524 (1976); A. Aurilia and D. Christodoulou, Phys. Lett. B **71**, 90 (1977).

[5] A. Aurilia, D. Christodoulou, and F. Legovini, Phys. Lett. B **73**, 429 (1978). A. Aurilia and D. Christodoulou, Phys. Lett. B **78**, 589 (1978).

[6] P.A.M. Dirac, Proc. R. Soc. London, Ser. A **268**, 57 (1962); Y. Nambu, Phys. Rep. C **23**, 250 (1976); P.A. Collins and R.W. Tucker, Nucl. Phys. B **112**, 150 (1976); A. Aurilia and F. Legovini, Phys. Lett. B **67**, 299 (1977).

[7] A. Aurilia and D. Christodoulou, "Theory of strings and membranes in an external field. II. The string," to be published in J. Math. Phys.

[8] E. Cremmer and J. Scherk, Nucl. Phys. B **72**, 117 (1974).

[9] H. Nielsen and P. Olesen, Nucl. Phys. B **61**, 45 (1973); A. Aurilia, Nucl. Phys. B **92**, 241 (1975).

[10] J.T. Schwartz, "Nonlinear Functional Analysis," Courant Institute of Mathematical Sciences (1963–1964); Theorem 1.20.

[11] J. Leray, "Hyperbolic differential equations," Institute for Advanced Study (Lecture Notes, 1953); L. Gårding, *Solution directe du probléme de Cauchy pour les équations aux dérivées partielles*, Colloque International (CNRS, Nancy, France, 1956); P. Dionne, "Sur les problémes de Cauchy bien posés," J. Anal. Math. Jérusalem **10**, 1–90 (1962).

[12] R. Abraham, *Foundations of Mechanics* (Benjamin, New York, 1967), §7.

Asymptotic eigensolutions of fourth and sixth rank octahedral tensor operators

William G. Harter[a)]

Joint Institute for Laboratory Astrophysics, University of Colorado and National Bureau of Standards, Boulder, Colorado 80309

Chris W. Patterson

University of California, Los Alamos Scientific Laboratory, Los Alamos, New Mexico 87545
(Received 23 August 1978)

Qualitative and quantitative features of high quantum rotational spectra are discussed by appealing to geometrical and topographical representations of the tensor operators. Approximate formulas are derived for level-cluster energies. The approximate conditions for the occurrence of "anomalous" fourfold clusters are given.

I. INTRODUCTION

Eigensolutions of fourth-rank octahedral tensor operators (T^4) have become very useful for analyzing high-resolution laser spectra of heavy spherical top molecules[1,2] such as CF_4, SiF_4, and SF_6.[3] The operators model the centrifugal distortion effects of the molecules which show up in the fine structure patterns in the rotational or rovibrational spectra. Because of high rotational inertia of these molecules their most easily observed lines belong to high rotational quanta typically $J = 10$–100 and higher in some cases. The high J lines are arranged into surprising patterns of spectral "clusters," and the understanding of the clusters has led to simpler theories and better understanding of these molecules.

It is interesting to note that the first observations of clusters were made in computer studies of crystal field splitting of rare earth atomic levels. Lea, Leask, and Wolf[4] used a computer to investigate the effect of adding varying amounts of the sixth rank (T^6) tensor operator to the fourth rank tensor (T^4). They noticed some unexpected triple point degeneracies in their energy level diagrams even for low angular momentum states. (They treated $J \leqslant 8$ only.) Ten years later Dorney and Watson[5] diagonalized T^4 for $J \leqslant 20$ and noted many nearly degenerate clusters for the higher J, and gave a classical model to explain some properties of them. Finally, Fox, Galbraith, Krohn, and Louck[6] performed computer diagonalization of T^4 for $J = 2$–100 and observed many extraordinary properties of clusters. This led to the quantum theory of spectral clusters, parts of which will be reviewed briefly below.[7-11]

The purpose of this article is to continue the original investigation by Lea *et al.*[4] of the effects of adding varying amounts of sixth rank tensor T^6 to T^4. The results of cluster theory will be used and asymptotic formulas will be derived for the limiting cases of high J. The analysis of effects due to tensors of rank six or higher is important for the light spherical tops such as CH_4 (methane), SiH_4 (silane), or GeH_4 (germane) which have appreciable amounts of T^6 in their model Hamiltonians. While there are far fewer light (hydride) tops we believe they may make up for their small number with interesting spectroscopic effects. We shall try to emphasize the intuitive physical nature of the cluster eigenstates and indicate what effects one might predict.

II. TENSOR OPERATORS AND THEIR EIGENVALUES

Rank-k irreducible tensorial operators $T^k_q(q = k, k-1, \ldots, -k)$ of the rotational group R_3 are sets of $2k+1$ operators which transform as follows[12]

$$R(\alpha\beta\gamma)T^k_q R^{-1}(\alpha\beta\gamma) = \sum_{q'} T^k_{q'} \mathscr{D}^k_{q'q}(\alpha\beta\gamma), \qquad (1)$$

where $\mathscr{D}^k_{q'q}$ are irreducible representations of rotation operators in $R_3 = \{\cdots R(\alpha\beta\gamma)\cdots\}$. Certain combinations of certain T^k_q are invariant to the subgroup of octahedral rotations. Besides the trivial case T^0_0 we shall consider the two lowest rank octahedral invariants

$$\hat{T}^4 = (7/12)^{1/2} T^4_0 + (5/24)^{1/2}(T^4_4 + T^4_{-4}) \qquad (2)$$

and

$$\hat{T}^6 = (1/8)^{1/2} T^6_0 - (\sqrt{7}/4)(T^6_4 + T^6_{-4}). \qquad (3)$$

Here we have chosen axes of quantization to be the fourfold symmetry axes of the octahedral subgroup. Many texts[13] give procedures for deriving invariant operators. Instead of giving derivations we shall show ways to picture the operator so that their octahedral symmetry is obvious.

The operators T^4 and T^6 can be expressed in terms of polynomials of coordinate operators x, y, and z or else angular momentum operators J_x, J_y, and J_z. This is done by replacing each T^k_q with spherical harmonics $Y^k_q(\theta,\phi) = Y^k_q(x/r,y/r,z/r)$ or else angular momentum harmonics $Y^k_q(J_x,J_y,J_z)$. Using spherical harmonics we obtain

$$\hat{T}^4 = (21/\pi)^{1/2}[(35\cos^4\theta - 30\cos^2\theta + 3) + 5\sin^4\theta\cos^4\phi]/32 \qquad (4a)$$

$$= (7/3\pi)^{1/2}(15)[x^4 + y^4 + z^4 - (3/5)r^4]/8r^4 \qquad (4b)$$

and

[a)]Presently at The School of Physics, Georgia Institute of Technology, Atlanta, Georgia 30332.

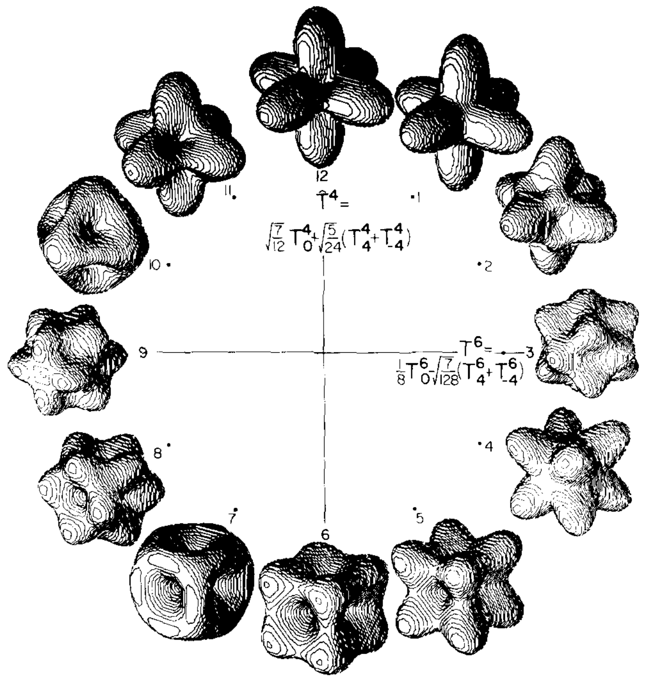

FIG. 1. Geometrical representations of octahedral tensor $T(\mu) = \hat{T}^4 \cos\mu + T^6 \sin\mu$ $[T^6 = (1/\sqrt{8})\,\hat{T}^6]$. Each combination $T(\mu)$ for $\mu = 0, \pi/6, 2\pi/6, \ldots, 11\pi/6$ is drawn in spherical coordinates $[r = 1 + 1.2T(\mu,\theta,\phi)]$ using (2), (3), (4a), and (5a). Drawings were made on a computer using codes and algorithms written by Chela Kunasz and Thomas Wright.

$$\hat{T}^6 = (104/\pi)^{1/2}[(231\cos^6\theta - 315\cos^4\theta + 105\cos^2\theta - 5) + 21\sin^4\theta(11\cos^4\theta - 1)\cos^4\phi]/256 \quad (5a)$$

$$= (13/2\pi)^{1/2}21[x^6 + y^6 + z^6 - 5(x^4y^2 + y^4x^2 + x^4z^2 + y^4z^2 + z^4x^2 + z^4y^2) + 70x^2y^2z^2 - (5/21)r^2]/40r^6. \quad 5b)$$

Figure 1 shows solid spherical plots of the function

$$T(\mu) = \hat{T}^4\cos\mu + T^6\sin\mu, \quad (6a)$$

where angle μ ranges from zero to 2π in steps of $2\pi/12$ and the nonnormalized operator

$$T^6 \equiv (1/8)^{1/2}\,\hat{T}^6 \quad (6b)$$

is used. The plots of Eqs. (4) to (6) were done on computer by Chela Kunasz using solid graphics software developed at the National Center for Atmospheric Research by Thomas Wright. The octahedral symmetry of the figures is evident if one ignores the "wood grain" which merely stands for the limit of the computer storage resolution. (To obtain hidden line drawings the entire function is first stored numerically.) It is useful to think of each figure as a potential surface. One can imagine the hills and valleys correspond to high and low energies in the octahedrally anisotropic Hamiltonian. Note that diametrically opposed numerals on the clock, i.e., 12 and 6, 1 and 7,··· etc., correspond to potentials that differ only by overall sign, that is, hilltops are interchanged with valley bottoms.

If one is interested in moelcular centrifugal distortion operators then it is appropriate to imagine that each of the twelve objects in Fig. 1 are plotted in $\{J_x J_y J_z\}$ space. For example the 12 o'clock object corresponds to molecule whose energy is highest when rotating around the J_x, J_y, or J_z axes which go through the hilltops, and lowest when the J-vector points out of the valley in the (1,1,1) direction. Each surface gives the energy as a function of *J direction* for fixed total angular momentum

$$J_x^2 + J_y^2 + J_z^2 = \text{const.} \quad (7)$$

An octahedral (XY_6) molecule would be distorted the *least* by centrifugal force when rotating around a fourfold x, y, or z axis since then the force is along or perpendicular to the six stronger radial bonds. Rotation around the (1,1,1) or threefold axis affects the weaker bending bonds and causes the greatest distortion. Greater distortion corresponds to more rotational inertia and hence lower energy. This in turn corresponds to a valley on the potential surface. Therefore the surfaces around the 12 o'clock position are apt to describe octahedral XY_6 pure rotational distortion, while their "negatives" around 6 o'clock are more apt to describe cubic (XY_8) or tetrahedral (XY_4) rotational distortion. Note that tetrahedral symmetry alone would allow a third-rank invariant

$$T^3 = J_x J_y J_z, \quad (8)$$

but this is excluded according to time reversal symmetry.

The matrices of tensor operators are made using the Wigner–Echart theorem

$$\langle {}^J_{M'}|T^k_q|{}^J_M\rangle = (-1)^{J-M}\begin{pmatrix} J & k & J \\ -M' & q & M\end{pmatrix}(J\|T^k\|J), \quad (9)$$

where the reduced matrix element $(J\|T^k\|J)$ can be taken outside of the matrix since it does not depend on M or M'. We shall set it equal to unity here, but its value can be computed once the exact form of the tensor operator is established. The other factor in (9) is the Wigner 3-j coefficient. The following values of this coefficient are needed:

$$\begin{pmatrix} J & 4 & J \\ -M & 0 & M\end{pmatrix} = (-1)^{J-M}[6(J+2:-1) - 10M^2(6J^2+6J-5) + 70M^4]/[(2J+5:-3)]^{1/2}, \quad (10a)$$

$$\begin{pmatrix} J & 4 & J \\ -(M-4) & -4 & M\end{pmatrix} = (-1)^{J-M}[70(J+M+0:-3)(J-M+4:1)]^{1/2}/[(2J+5:-3)]^{1/2}, \quad (10b)$$

$$\begin{pmatrix} J & 6 & J \\ -M & 0 & M\end{pmatrix} = (-1)^{J-M}[-20(J+3:-2) + 84M^2(5J^4+10J^3-20J^2-25J+14) - 420M^4(3J^2+3J-7) + 924M^6]/[(2J+7:-5)]^{1/2} \quad (10c)$$

$$\begin{pmatrix} J & 6 & J \\ -(M-4) & -4 & M \end{pmatrix}$$

$$= 12(-1)^{J-M}(50-44M+11M^2-J^2-J)$$

$$\times [(7/2)(J+M+0:-3)(J-M+4:1)]^{1/2}/$$

$$[(2J+7:-5)]^{1/2}. \tag{10d}$$

Here we use the notation

$$(X+a:b) = (X+a)(X+a-1)(\cdots)(X+b+1)(X+b). \tag{11}$$

The angular momentum representation of the (4,6)-octahedral operator

$$T(\nu) = \widehat{T}^4 \cos\nu + \widehat{T}^6 \sin\nu \tag{12}$$

in the $|{}^J_M\rangle$ basis is a $(2J+1)\times(2J+1)$ matrix. Using octahedral symmetry adapted bases[1,2] it is possible to reduce it to a direct sum of smaller block diagonal matrices belonging to each symmetry species A_1, A_2, E, T_1, and T_2. For example for $J=30$ one can reduce the (61×61) matrix to two (3×3) A_1 and A_2 blocks, *a* (5×5) E block, *a* (7×7) T_1 block, and an (8×8) T_2 block. Finally 26 distinct eigenvalues (the E- and T-levels are, doubly and triply degenerate, respectively) are obtained in the $(J=30)$ case by diagonalizing the sub blocks. These blocks can be produced and diagonalized by computer[14] and the results for $J=30$ are shown in Fig. 2 as a function of parameter ν $(0\leqslant\nu\leqslant\pi)$. It is only necessary to carry this expensive computer calculation half way around the ν-clock. The eigenvectors of a matrix are unchanged by an overall (-1) factor and the eigenvalues are merely inverted.

We now see ways to understand the results of the diagonalizations and derive simple approximate formulas for the eigenvalues.

III. APPROXIMATIONS FOR TENSOR OPERATOR SPECTRA

The first thing one notes in Fig. 2 is that many of the 26

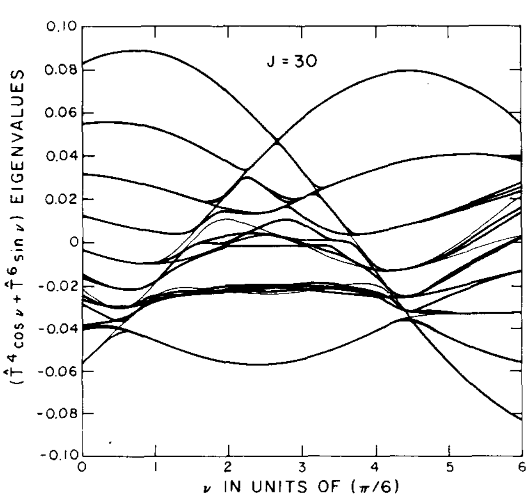

FIG. 2. $(J=30)$ Eigenvalue spectrum of $T(\nu)' = \widehat{T}^4 \cos\nu + \widehat{T}^6 \sin\nu$.

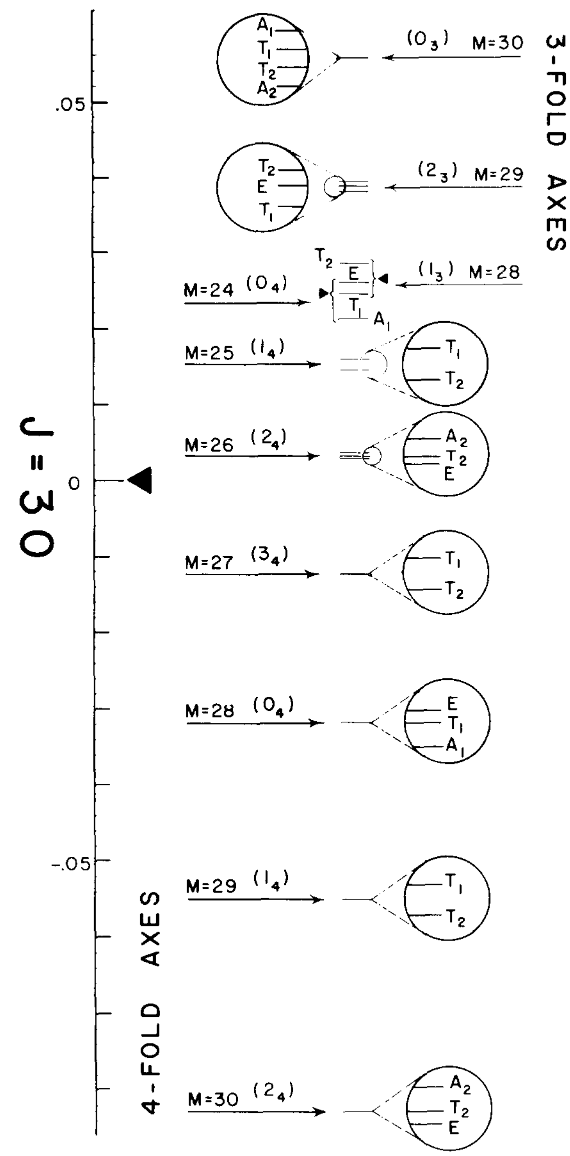

FIG. 3. $(J=30,\ \nu=\pi)$ Spectrum of $T(\nu)$. Magnified view of the level clusters are shown to exhibit the separate octahedral symmetry species.

eigenvalues are "clustering" for most ν values. The spectrum is a good deal simpler than it would have been if all 26 values had been randomly deployed. Figure 3 shows the spectral detail for the extreme right-hand side of Fig. 2, i.e., $\nu=\pi$... [the extreme left-hand side $(\nu=0)$ is just the negative of this]. The lower portion of the spectrum exhibits alternative $(A\oplus T\oplus E)$ and $(T_1\oplus T_2)$ clusters. Each of these clusters contain six levels: either $1+3+2=6$ or $3+3=6$. As explained in Refs. 7–9 each set of six levels corresponds to six more-or-less well-defined localized rotation states in each of the six valleys on fourfold axes at 6 o'clock $(\nu=\pi=\mu)$ in Fig. 1. The upper portion of the spectra exhibits $(A_1\oplus T_1\oplus T_2\oplus A_2)$ and $(T_1\oplus E\oplus T_2)$ clusters each containing eight levels. These belong to states sitting in each of eight hills on three-fold axes at 6 o'clock or in each of eight valleys at 12 o'clock $(\nu=0=\mu)$. Note that the fourfold axial hills or valleys are deeper than the threefold valleys or hills at the 12 o'clock or 6 o'clock positions. Therefore there are more fourfold or six-level clusters than threefold or eight-level clusters at these positions. However, at just before 4 o'clock ($\mu = 120° = 2\pi/3$, $\nu = \tan^{-1}[(1/8)^{1/2}\tan\mu] = 148.5°$ $= 9.9\pi/12$) or else at 10 o'clock in Fig. 1 the threefold hills or valleys predominate. In the corresponding neighborhood between $\nu = 8\pi$ and $\nu = 10\pi$ in Fig. 2 the threefold clusters dominate the spectrum.

The splitting of a level cluster is determined by how easily tunneling occurs between the valleys or hills associated with the cluster. As explained in Refs. 6–9 and 11 the nearest neighbor tunneling amplitude ($-S$) appears in the $(A_1 \oplus T_1 \oplus E)$ cluster eigenvalues

$$E(E) = H + 2S, \tag{13a}$$

$$E(T_1) = H, \tag{13b}$$

$$E(A_1) = H - 4S, \tag{13c}$$

and predicts the observed 2:1 splitting ratio between E-T_1 and T_1-A_1 intervals in the fourfold region of Fig. 3. (Here H is the cluster center-of-gravity which will be discussed shortly.) Similarly, threefold cluster eigenvalues are given by

$$E(A_1) = H + 3S + 3T, \tag{14a}$$

$$E(T_1) = H + S - T, \tag{14b}$$

$$E(T_2) = H - S - T, \tag{14c}$$

$$E(A_2) = H - 3S + 3T, \tag{14d}$$

for the $(A_1 \oplus T_1 \oplus T_2 \oplus A_2)$ cluster and for the $(T_1 \oplus E \oplus T_2)$ cluster by

$$E(T_2) = H + 2S - T, \tag{15a}$$

$$E(E) = H + 3T, \tag{15b}$$

$$E(T_1) = H - 2S - T, \tag{15c}$$

where S is nearest neighbor tunneling amplitude and T is the *next* nearest neighbor tunneling amplitude. For most threefold clusters the T amplitude is negligible, however there are certain cases when S goes through zero with T nonzero as will be discussed at the end of this section.

The fourfold clusters dominate the spectrum around 7 o'clock in Fig. 1 and, of course, also at the antipodal position of 1 o'clock (1 o'clock corresponds to $\mu = 30°$ or $\nu = 11.53° = 0.77\pi/12$). Note that the "pass" between fourfold valleys becomes higher as one goes from 5 o'clock to 6 o'clock. Finally, just before 7 o'clock the pases get cut off and 6 perfect "craters" are formed around the fourfold valleys. This is close to the ν-value for which the spectrum in Fig. 2 is all fourfold clusters. By 8 or 9 o'clock the "passes" have grown up to form twelve mountains. On the opposite side at 2 or 3 o'clock twelve valleys are visible in Fig. 1. These twelve extrema are the source of the twelve-level two-fold clusters $(A_1 \oplus E_1 \oplus T_1 \oplus 2T_2)$ and $(A_2 \oplus E \oplus T_2 \oplus 2T_1)$ discussed in Ref. 8. (They also can arise in a ν_3-type coriolis spectrum.[15] The twofold axes are located at saddle points or "passes" on the T^4 surface (12 and 6 o'clock) and these points correspond to crossover regions in the spectrum of Fig. 3 between threefold clusters on the high side and four-fold clusters on the low side. Saddle points do not give rise to clusters since it is possible to travel globally between them without having to rise or fall in energy. Clusters correspond to localized states, only.

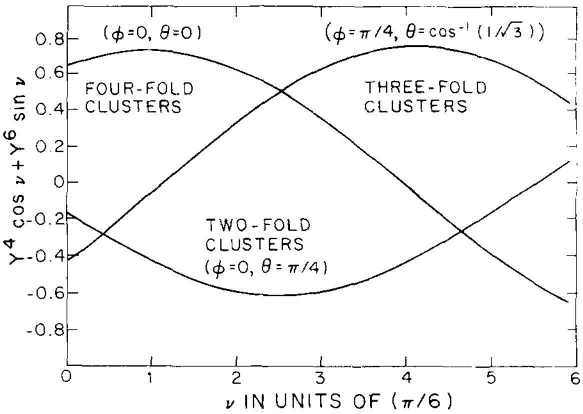

FIG. 4. Local $T(\nu)$ potential values at two-, three-, and fourfold axes. Values of potential $T(\nu) = \hat{T}^4 \cos\nu + \hat{T}^6 \sin\nu$ are plotted versus ν using (4a) and (5a) for select (θ,ϕ) values corresponding to the three kinds of octahedral symmetry axes.

It is instructive to plot the harmonic form of the tensor $T(\nu)$ (12) using (4a) and (5a) for fourfold ($\phi = 0$, $\theta = 0$), threefold [$\phi = \pi/4$, $\theta = \cos^{-1}(1/3)^{1/2}$)], and twofold ($\phi = 0$, $\theta = \pi/4$) axes as in Fig. 4. The qualitative form of the spectrum in Fig. 2 is imitated to some extent. The two-, three-, and fourfold curves mark the regions in which the respective clusters exist. The curves take turns serving as cluster boundaries, i.e., hill tops or valley bottoms, and crossover boundaries, i.e., "passes" at saddle points.

A more accurate approximation of the spectrum is obtained using the matrix elements in (10). If the operators are represented in the appropriate basis then excellent cluster energy approximations are given just by the diagonal components. Even more accurate results are obtained by perturbation. The operators in (2) and (3) are set up in the fourfold basis already. The following zeroth approximation follows:

$$\langle T(\nu)\rangle^0_{4\text{-fold}}$$

$$= (7/12)^{1/2}\begin{pmatrix} J & 4 & J \\ -M & 0 & M \end{pmatrix}(-1)^{J-M}\cos\nu$$

$$+ (1/8)^{1/2}\begin{pmatrix} J & 6 & J \\ -M & 0 & M \end{pmatrix}(-1)^{J-M}\sin\nu, \tag{16}$$

where $M = K_4$ is the fourfold axial momentum associated each cluster. (Recall Fig. 3.) One can do the same for the threefold cluster by first representing (2) and (3) in a threefold axial basis as follows

$$\hat{T}^4 = -2[(7/12)^{1/2}T^4_0 + 2(5/24)^{1/2}(T^4_{-3} - T^4_3)]/3, \tag{17}$$

$$\hat{T}^6 = 2[(8)^{1/2}T^6_0 + (70/24)^{1/2}(T^6_3 - T^6_{-3})$$

$$+ (77/24)^{1/2}(T^6_6 + T^6_{-6})]/9. \tag{18}$$

Then the zeroth threefold cluster approximation is

$$\langle T(\nu)\rangle^0_{3\text{-fold}}$$

$$= -(2/3)(7/12)^{1/2}\begin{pmatrix} J & 4 & J \\ -M & 0 & M \end{pmatrix}(-1)^{J-M}\cos\nu,$$

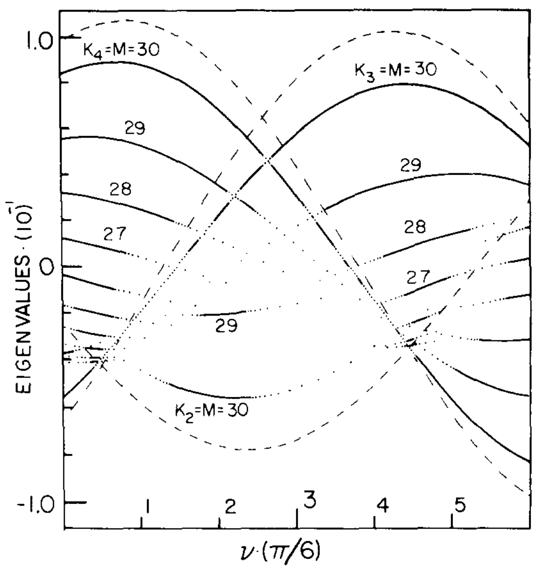

FIG. 5. Approximate ($J=30$) spectrum of $T(\nu)$. Approximations (16), (19), and (20) are plotted versus ν. Solid lines indicate where there is agreement with the exact results in Fig. 2. Dots indicate where clusters split and approximations break down. Dashed curves corresponds to classical potentials obtained by using (21a) with $P_k=1$.

$$+(2/9)(8)^{1/2}\begin{pmatrix} J & 6 & J \\ -M & 0 & M\end{pmatrix}(-1)^{J-M}\sin\nu, \qquad (19)$$

where $M=K_3$ is the threefold cluster momentum. Finally, the zeroth approximation for the twofold clusters is

$$\langle T(\nu)\rangle^0_{2\text{-fold}}$$
$$=-(1/4)(7/12)^{1/2}\begin{pmatrix} J & 4 & J \\ -M & 0 & M\end{pmatrix}(-1)^{J-M}\cos\nu,$$

$$-(13/8))(1/8)^{1/2}\begin{pmatrix} J & 6 & J \\ -M & 0 & M\end{pmatrix}(-1)^{J-M}\sin\nu, \qquad (20)$$

where $M=K_2$ is the twofold cluster momentum. The functions of ν given by (16), (19), and (20) are plotted together in Fig. 5. The solid curves indicate where the agreement is within 2% of the exact results in Fig. 2. This occurs practically everywhere that a cluster exists. In other words, the error is the same order of magnitude as the cluster splitting.

It is interesting to use Edmond's[16] approximate expression

$$\begin{pmatrix} J & k & J \\ -M & 0 & M\end{pmatrix}(-1)^{J-M}=P_k(\cos\theta)/(2J+1)^{1/2}, \qquad (21a)$$

where P_k is a Legendre polynomial and

$$\cos\theta=M/[J(J+1)]^{1/2}\sim M/(J+1/2). \qquad (21b)$$

Its accuracy is about 0.6% at $M=J=30$ and about 0.05% at $M=J=100$. It is very useful for helping to understand the classical limit. Let us imagine that the angular momentum in state $|^J_M\rangle$ is represented by a cone of altitude $\langle J_z\rangle=M$ and slant height $\langle J\cdot J\rangle^{1/2}=[J(J+1)]^{1/2}\sim J+1/2$ about the axis of quantization. According to (21b) the apex half-angle is θ, and according to (21a) the eigenvalue of tensor T^k is proportional to the value of the kth harmonic at angle θ, i.e.,

$$\langle T^k\rangle\propto P_k(\cos\theta)/(2J+1)^{1/2}$$
$$=Y^{k*}_0(\theta)[4\pi/(2k+1)(2J+1)]^{1/2}. \qquad (22)$$

The harmonic valley bottom occurs at $\theta=0$ where $P_k(1)=1$. Therefore, if we replace the $3-j$ coefficient and phase in (16), (19), and (20) by $(2J+1)^{-1/2}$ we obtain a better representation than Fig. 4 for the fourfold, threefold, and twofold valley bottoms, hilltops or passes, whichever they might be. These are plotted as dashed lines in Fig. 5. Note that their form is similar apart from the overall magnitude to the topographical features represented by Fig. 4. Any difference is due to the factor $[4\pi/(2k+1)]^{1/2}$ in Eq. (22), which is close to unity for $k=4$ and 6. Note that even "top clusters" for which $J=M$ have some "zero-point" energy with respect to their valley bottoms or hilltops. This becomes less and less significant as $J\to\infty$.

While studying the asymptotic functional forms of angular coefficients, Ponzano and Regge[17] have deduced "potentials" for Schrödinger-like equations for certain Racah coefficients. Schulten and Gordon[18] have developed a theory of these "potentials" and applied it to approximate derivations of coupling coefficients. The cluster properties seem to point out the physical "reality" of these potentials.

Improved accuracy of the cluster approximation is obtained by using off-diagonal matrix elements in perturbation formulas. This has been done successfully for several cases of fourth-rank tensor operators.[10,15,19] It is interesting to observe that the contribution from the sixth-rank tensor can cause key off-diagonal components to vanish. For example, the key component relating $K_4=M$ and $K_4=M-4$ fourfold cluster states will vanish when

$$(5/24)^{1/2}\begin{pmatrix} J & 4 & J \\ -(M-4) & 4 & M\end{pmatrix}\cos\nu(0)$$

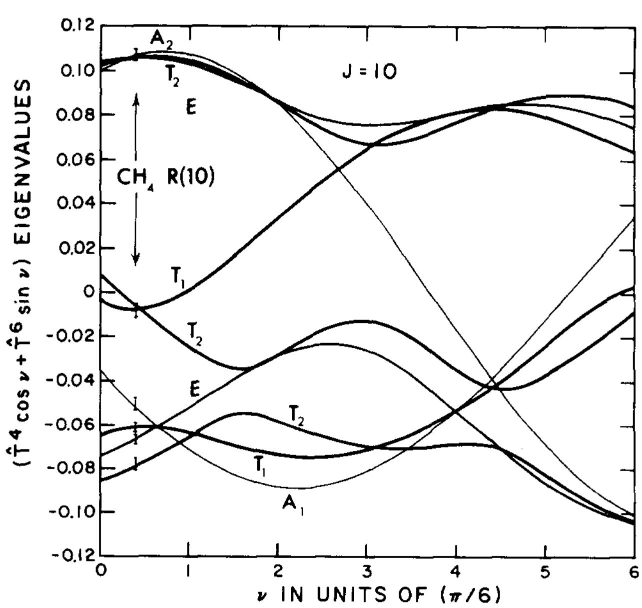

FIG. 6. ($J=10$) Eigenvalue spectrum of $T(\nu)$.

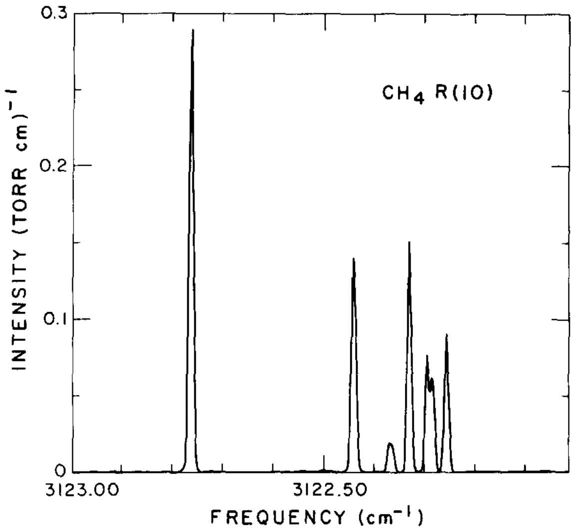

FIG. 7. R (10) Laser spectra of CH_4 (courtesy of Allen S. Pine, MIT Lincoln Laboratory). Symmetry species can be identified by fitting the spectrum with Fig. 6 ($\nu \cong 0.5/6\pi$). This is further verified by the heights of the lines which approximately correspond to the well-known statistical weights: 5 for A_1, 3 for T_1 or T_2, and 2 for E.

$$-(7/16)^{1/2}\begin{pmatrix} J & 6 & J \\ -(M-4) & -4 & M \end{pmatrix} \sin\nu(0) = 0. \tag{23}$$

Solving this using (10) we have

$$\tan\nu_M^J(0) = (5/3)[(2J+7)(2J+6)(2J-4)(2J-5)/42]^{1/2} / [50 + 11M(M-4) - J(J+1)]. \tag{24}$$

For example for $J = 10$ one finds the predictions

$$\nu_{10}^{10}(0) = 9.98^\circ, \quad \upsilon_9^{10}(0) = 13.06^\circ, \quad \nu_8^{10}(0) = 17.74^\circ,$$

which agree rather well with the cluster formations on the left-hand side of Fig. 6. Indeed, it is remarkable that the all fourfold cluster tunneling amplitudes vanish at certain points for each $(A \oplus T \oplus E)$ cluster to give the triple point cluster degeneracies first noticed by Lea *et al.*[4] [Note that the tunneling amplitude corresponding to "reversal"

$$\left(\left| \begin{matrix} J \\ M \end{matrix} \right\rangle \to \left| \begin{matrix} J \\ -M \end{matrix} \right\rangle \right)$$

is not included in (13)–(15).] The "triple points" for $J = 30$ clusters in Fig. 2 are not resolved on the scale of the graph. However, (24) predicts correctly the values $\nu_{20}^{30}(0) = 16.8^\circ$, $\nu_{21}^{30}(0) = 15.1^\circ$, $\nu_{22}^{30}(0) = 13.7^\circ$, $\nu_{23}^{30}(0) = 12.4^\circ, \ldots, \nu_{30}^{30}(0) = 7.6^\circ$ for which the low-M clusters exist near the fourfold "crater" region around 1 or 7 o'clock (see also Fig. 5).

For threefold clusters there appear points at which the nearest-neighbor tunneling amplitude (S) vanishes while the *next*-nearest-neighbor tunneling amplitude (T) is small but nonzero. The residual (T) exists because of the extra $T^6_{\pm 6}$ terms in (18). This causes the $(T_1 \oplus E \oplus T_2)$ cluster to have a crossing of T_1 and T_2 just below E in the upper right-hand corner of Fig. 6. This was predicted by (15) for $S = 0$ and T small. Similarly, (14) predicts a coincident crossing of A_1 with A_2 and T_1 with T_2 in the $(A_1 \oplus T_1 \oplus T_2 \oplus A_2)$ cluster when $S = 0$. This is seen at ($\nu = 4.4\pi/6$, $E = -0.04$) in Fig. 6. Note just below that crossing there are two levels E and T_2 getting together. This corresponds to part of a cluster falling together in the threefold dominated 4 o'clock (or 10 o'clock) region. The same thing happens to incomplete or "leftover" fourfold clusters in the 1 o'clock (or 7 o'clock) region.

One concludes that sixth rank centrifugal tensors (T^6) can make fourfold clusters anomalously "tight" or degenerate even for lower J, and reorder the structure of threefold clusters. This is important for high resolution spectroscopy since it makes anomalous or case (2) hyperfine structure possible in more accessible regions of the spectrum.[20] The mixing of hyperfine states and symmetry species that inevitably occurs at high J may also happen in select portions of the low J spectrum.

As an example, consider the $J = 10$ spectrum of CH_4 by Pine[21] which is shown in Fig. 7. This fits rather well with the level diagram at about $\nu = 0.9/12\pi$ in Fig. 6. Note that the first ($K_4 = 10$) cluster is very nearly degenerate at this point. The first two fourfold cluster splittings are not resolved in Pine's spectra. Note that the heights of the spectral lines correspond approximately to the well-known nuclear spin statistical weights: 5 for A_1 or A_2, 3 for T_1 or T_2, and 2 for E. The weights "pile up" for clusters so $A_2 + T_2 + E$ is about ten units high in Fig. 7.

More detailed treatments of the molecular Hamiltonian including Coriolis and off-diagonal tensor are needed to fit CH_4 spectra properly, [the P(10) pattern does not fit Fig. 6 well for any value of ν.] One may need also a small amount of eighth rank tensor. However, it appears to be possible to simplify any tetrahedral Hamiltonian calculation by casting it into an appropriate cluster basis for high angular momentum,[11,15,19] and high momentum lines provide the most accurate determination of molecular Hamiltonian constants. Furthermore, the ease with which one can understand and discuss the tensor spectra should be evident. This alone should motivate further development and applications.

ACKNOWLEDGMENTS

We gratefully acknowledge the skilled assistance of Chela Kunasz which was needed to produce the first figure of this article. We also thank the JILA Chairman Dr. David Hummer for making special funds available for this work.

[1]J.D. Louck, thesis, Ohio State University, 1958 (unpublished).
[2]J.M. Bailly, Can. J. Phys. **15**, 237 (1961).
[3]R. McDowell, in *Laser Spectroscopy III*, edited by J.L. Hall and J.L. Carlsten (Springer-Verlag, New York, 1977), p. 102.
[4]K.R. Lea, M.J.M. Leask, and W.P. Wolf, J. Phys. Chem. Solids **23**, 1381 (1962).
[5]A.J. Dorney and J.K.G. Watson, J. Mol. Spec. **42**, 1 (1972).
[6]K. Fox, H.W. Galbraith, B.J. Krohn, and J.D. Louck, Phys. Rev. A **15**, 1363 (1977).
[7]W.G. Harter and C.W. Patterson, Phys. Rev. Lett. **38**, 224 (1977).
[8]W.G. Harter and C.W. Patterson, J. Chem. Phys. **66**, 4872 (1977).
[9]W.G. Harter and C.W. Patterson, Int. J. Quant. Chem. Symp. **11**, 479 (1977).
[10]C.W. Patterson and W.G. Harter, J. Chem. Phys. **11**, 4886 (1977).

[11]W.G. Harter, C.W. Patterson, and F.J. daPaixao, Rev. Mod. Phys. **50**, 37 (1978).
[12]U. Fano and G. Racah, *Irreducible Tensorial Sets* (Academic, New York, 1959).
[13]M. Hammermesh, *Group Theory and Its Applications to Physical Problems* (Addison Wesley, Reading, Massachusetts, 1960).
[14]B.J. Krohn, Los Alamos Report LA-6554-MS, Los Alamos, New Mexico (1976).
[15]W.G. Harter, C.W. Patterson, and H.W. Galbraith, J. Chem. Phys. **69**, 4888–907 (1978).
[16]A.R. Edmonds, *Angular Momentum in Quantum Mechanics* (Princeton U.P., Princeton, New Jersey, 1957), p. 122.
[17]G. Ponzano and T. Regge, in *Spectroscopic and Group Theoretical Methods in Physics: Racah Memorial Volume* (North-Holland, Amsterdam, 1968), p. 1.
[18]K. Schulten and R.G. Gordon, J. Math. Phys. **16**, 196 (1975).
[19]H.W. Galbraith, C.W. Patterson, B.J. Krohn, and W.G. Harter, J. Mol. Spectrosc. **73**, 475–93 (1978).
[20]W.G. Harter and C.W. Patterson, in *Advances in Laser Chemistry*, edited by A.H. Zewail (Springer-Verlag, New York, 1978); "Theory of hyperfine and superfine levels in symmetric polyatomic molecules," Phys. Rev. A (submitted).
[21]A.S. Pine, J. Opt. Soc. Am. **66**, 97 (1976).

Note on the linear representations of any dimensional Lorentz group and their matrix elements

Takayoshi Maekawa

Department of Physics, Kumamoto University, Kumamoto, Japan

(Received 20 November 1978)

The linear representations of the group SO(n,1) are studied in the space of functions given on the maximal compact subgroup SO(n) without making use of a special parametrization for elements of the group. The scalar products, which are invariant under the transformation by any element of SO(n,1), are introduced into the space together with the action of any representation operator of SO(n,1) on the function. The computation formula for the representation matrix elements is obtained.

In previous papers[1,2] the linear representations of the group SO(n,1) are discussed in the space of functions given on the maximal compact subgroup SO(n) with the Euler angle parametrization and the computation formulas for the matrix elements of all representations are given. Though the representations are discussed for a special element of SO(n,1), i.e., the boost to the nth direction, it is easy to generalize the discussion to any element of SO(n,1). The purpose of this note is to generalize the action of the representation operator on the functions given on SO(n) in such a way that the action of the representation operator for any element of SO(n,1) on the function is given independent of the parametrization for elements of the group, and to examine the properties of the matrix elements.

Let us first introduce the Iwasawa decomposition[3] for an element of SO(n,1) as follows

$$g^{(n,1)} = n(\xi)^b\, t^{(n+1)}_{n+1n}(\eta) g^{(n)}$$
$$= \begin{bmatrix} 1 & \xi & \xi \\ -\xi^t & 1-\Xi & -\Xi \\ \xi^t & \Xi & 1+\Xi \end{bmatrix} \times \begin{bmatrix} 1 & 0 & 0 \\ 0 & h^0_{nn} & h^0_{nn+1} \\ 0 & h^0_{n+1n} & h^0_{n+1n+1} \end{bmatrix} \times \begin{bmatrix} g^{(n)} & 0 \\ 0 & 1 \end{bmatrix}, \tag{1}$$

where $g^{(n)} \in \mathrm{SO}(n)$, ${}^b t^{(n+1)}_{n+1n}(\eta)$ denotes a boost to the nth direction with the parametrization, e.g., $h^0_{nn} = h^0_{n+1n+1} = \cosh\eta$, $h^0_{nn+1} = h^0_{n+1n} = \sinh\eta$, and $n(\xi) \in N$, a nilpotent subgroup of SO(n,1) with $\Xi = (\xi^2_i)/2$. According to Refs. 1 and 2, we can construct a computation formula for the elements of the irreducible representation of SO(n,1) in the representation space of SO(n) in a way independent of the special choice of parameters. In what follows, we show this briefly. Unless stated otherwise, the same notations as in Ref. 2. will be used.

Considering a transformation with $h \in \mathrm{SO}(n,1)$,

$$g^{(n,1)} \xrightarrow{h} g^{(n,1)'} = g^{(n,1)}h = n(\xi')^b t^{(n+1)}_{n+1n}(\eta') g^{(n)'}, \tag{2}$$

we obtain

$$g^{(n)'}_{nj} = \frac{h_{n+1j} + \Sigma g^{(n)}_{ni} h_{ij}}{h_{n+1n+1} + \Sigma g^{(n)}_{ni} h_{in+1}} \quad (j = 1,2,\ldots,n), \tag{3}$$

$$g^{(n)}_{ij} = \frac{1}{h_{n+1n+1} + \Sigma g^{(n)}_{ni} h_{in+1}} \times \Big[\Big(h_{n+1n+1} + \sum g^{(n)}_{nl} h_{1n+1} \Big) \Big] \times \sum g^{(n)}_{ik} h_{kj} - \sum g^{(n)}_{il} h_{ln+1} \Big(h_{n+1j} + \sum g^{(n)}_{nk} h_{kj} \Big) \Big]$$
$$(i = 1,2,\ldots,n-1; j = 1,2,\ldots,n), \tag{4}$$

$$h^{0'}_{n+1n+1} + h^{0'}_{nn+1} = (h^0_{n+1n+1} + h^0_{nn+1}) \times \Big(h^0_{n+1n+1} + \sum g^{(n)}_{nj} h_{jn+1} \Big),$$

$$\xi'_j = \xi_j + \frac{\Sigma g^{(n)}_{ni} h_{in+1}}{(h^0_{n+1n+1} + h^0_{nn+1})(h_{n+1n+1} + \Sigma g^{(n)}_{ni} h_{in+1})}$$
$$(j = 1,2,\ldots,n-1). \tag{5}$$

It follows from (2) that in order to express $g^{(n)}$ in terms of $g^{(n)'}$ the replacements $g^{(n)} \rightleftarrows g^{(n)}$ and $h \to h^{-1}$ in (3) and (4) lead to the desired results.

The transformation property of the Haar measure (dV_n) of SO(n) under (2) is easily determined by taking into account that the measure is invariant under the rotation (shift) by the element of SO(n) and is transformed under h^0 [$= t^{(n+1)}_{n+ln}(\xi)$] into the form

$$dV^0_n = \frac{dV_n}{(h^0_{n+1n+1} + g^{(n)}_{nn} h^0_{nn+1})^{n-1}}. \tag{6}$$

In order to obtain the explicit expression for this, it is noted that h can be parametrized in a form

 0022-2488/79/071460-04$01.00

$$h = g_1 h^{\circ} g_2, \tag{7}$$

where g_1 and $g_2 \in SO(n)$. On the other hand, it follows from (2) and (7) that the following relations hold with obvious notations:

$$\begin{aligned} & n(\xi)^b t^{(n+1)}_{n+1n}(\eta) g^{(n)} h \\ & = n(\xi')^b t^{(n+1)}_{n+1n}(\eta') g^{(n)'} = n(\xi)^b t^{(n+1)}_{n+1n}(\eta) \bar{g}^{(n)} h^{\circ} g_2 \\ & = n(\xi')^b t^{(n+1)}_{n+1n}(\eta') \bar{g}^{(n)'} g_2 = n(\xi')^b t^{(n+1)}_{n+1n}(\eta') g^{(n)'}. \end{aligned} \tag{8}$$

The following relations are obtained by considering (6), (8), and invariance of dV_n under the rotation by the element of SO(n):

$$dV_n = d\bar{V}_n, \quad dV'_n = d\bar{V}'_n = \bar{\bar{V}}'_n,$$

$$d\bar{V}'_n = \frac{d\bar{V}_n}{(h^0_{n+1n+1} + \bar{g}^{(n)}_{nn} h^0_{nn+1})^{n-1}}.$$

Taking into account the above results and the relations which are easily seen from (2) and (7),

$$h_{n+1n+1} = h^0_{n+1n+1}, \quad \bar{g}^{(n)}_{nn} = (g^{(n)} g_1)_{nn} = \sum g^{(n)}_{nj} (g_1)_{jn},$$

$$h_{jn+1} = (g_1 h^{\circ} g_2)_{jn+1} = (g_1)_{jn} h^0_{nn+1},$$

$$\bar{g}^{(n)}_{nn} h^0_{nn+1} = \sum g^{(n)}_{nj} (g_1)_{jn} h^0_{nn+1} = \sum g^{(n)}_{nj} h_{jn+1},$$

we obtain the transformation of dV_n under (2) as follows,

$$dV'_n = \frac{dV_n}{(h_{n+1n+1} + \Sigma g^{(n)}_{nj} h_{jn+1})^{n-1}}. \tag{9}$$

It is noted that (9) holds for the replacements $dV_n \rightleftarrows dV'_n$ and $h \to h^{-1}$. It is obvious that a similar discussion to the above can be made with the left shift by changing the order of the factors on the right-hand side of the decomposition (1).

Let E_n be a linear space of functions given on SO(n) and taking values in the linear space $\mathscr{H}$. It is easily seen that the action of the representation operator $R(h)$ corresponding to h on the function $\phi \in E_n$ can be given as follows[1,2]

$$\begin{aligned} & R(h) \Phi(g^{(n)}) \\ & = \Big(h_{n+1n+1} + \sum g^{(n)}_{nj} h_{jn+1}\Big)^{\rho_{n+1}} \Phi(g^{(n,1)'}), \end{aligned} \tag{10}$$

where ρ_{n+1} is a complex number and $g^{(n)'}$ on the right-hand side is given by (3) and (4).

Applying a representation operator $R(h')$ to (10), we obtain

$$\begin{aligned} R(h')R(h)\Phi(g^{(n)}) = & \Big(h'_{n+1n+1} + \sum g^{(n)}_{nj} h'_{jn+1}\Big)^{\rho_{n+1}} \\ & \times \Big(h_{n+1n+1} + \sum \bar{g}^{(n)'}_{nj} h_{jn+1}\Big)^{\rho_{n+1}} \\ & \times \Phi(g^{(n,1)''}), \end{aligned} \tag{11}$$

where

$$\bar{g}^{(n)'}_{nj} = \frac{h'_{n+1j} + \Sigma g^{(n)}_{ni} h'_{ij}}{(h'_{n+1n+1} + \Sigma g^{(n)}_{ni} h'_{in+1})},$$

and $g^{(n)''}$ means the transformation under h' after that of h. It is straightward to show the relations

$$\begin{aligned} & \Big(h'_{n+1n+1} + \sum g^{(n)}_{nj} h'_{jn+1}\Big)\Big(h_{n+1n+1} + \sum \bar{g}^{(n)'}_{nj} h_{jn+1}\Big) \\ & = (h'h)_{n+1n+1} + \sum g^{(n)}_{nj} (h'h)_{jn+1}, \end{aligned}$$

$$g^{(n)''}_{nj} = \frac{(h'h)_{n+1j} + \Sigma g^{(n)}_{ni} (h'h)_{ij}}{(h'h)_{n+1n+1} + \Sigma g^{(n)}_{ni} (h'h)_{in+1}} \quad (j = 1,2,\ldots,n),$$

$$\begin{aligned} g^{(n)''}_{ij} = & \frac{1}{(h'h)_{n+1n+1} + \Sigma g^{(n)}_{nk} (h'h)_{kn+1}} \\ & \times \Big\{\Big[(h'h)_{n+1n+1} + \sum g^{(n)}_{nl} (h'h)_{ln+1}\Big] \sum g^{(n)}_{ip} (h'h)_{pj} \\ & - \sum g^{(n)}_{il} (h'h)_{ln+1} \Big[(h'h)_{n+1j} + \sum g^{(n)}_{np} (h'h)_{pj}\Big]\Big\} \\ & (i = 1,2,\ldots,n-1; \; j = 1,2,\ldots,n). \end{aligned}$$

Thus, (11) gives the relation

$$R(h')R(h)\Phi(g^{(n)}) = R(h'h)\Phi(g^{(n)}), \tag{12}$$

which means that the representation condition is satisfied for any complex number ρ_{n+1}.

It is noted that (10) means the right-regular representation[4] for $h \in SO(n)$. Similarly, we can construct the left-regular representation for the elements of SO(n). The bases of the regular representation can be classified by the group chain $SO(n) \supset SO(n-1) \supset \cdots \supset SO(2)$ and given as follows,[2]

$$\phi^{(\lambda_n)}_{\{\lambda'_{n-1}\}\{\lambda_{n-1}\}}(g^{(n)}) = \sqrt{\frac{N(\lambda_n)}{V_n}}\, D^{(\lambda_n)}_{\{\lambda'_{n-1}\}\{\lambda_{n-1}\}}(g^{(n)}), \tag{13}$$

where V_n is a volume of SO(n), $N(\lambda_n)$ the dimension of the representation (λ_n), and $D^{(\lambda_n)}$ the D matrix elements of SO(n). The functions (13) form a complete orthonormal system on SO(n) for the invariant measure dV_n on this group.[4] The orthonormality means

$$\begin{aligned} & \int_{SO(n)} \overline{\phi^{(\lambda_n)}_{\{\lambda'_{n-1}\}\{\lambda_{n-1}\}}(g^{(n)})}\, \phi^{(\lambda'_n)}_{\{\lambda'''_{n-1}\}\{\lambda''_{n-1}\}}(g^{(n)})\, dV_n \\ & = \delta_{\lambda_n \lambda'_n} \delta_{\{\lambda'_{n-1}\}\{\lambda'''_{n-1}\}} \delta_{\{\lambda_{n-1}\}\{\lambda''_{n-1}\}}. \end{aligned} \tag{14}$$

The numbers $\{\lambda'_{n-1}\}$ and $\{\lambda_{n-1}\}$ are, of course, characterized by invariants of SO(j) ($j = 2,3,\ldots,n-1$) for the first and second parameter groups, respectively.[2] The explicit expressions for the generators of the representation with respect to the Euler angles together with their matrix elements are known.[1,2] However, the representation matrix elements of SO(n) together with those for the generators may be regarded as known independently of the special parameters of the group because (13) and (14) hold regardless of the choice of parameters, and the generators between the standard and parameter representations are intimately connected with each other.[2]

In Refs. 1 and 2, in which the Euler angles are adopted as the parameters of SO(n), the bases of E_n have the form

$$\Phi^{(\rho_{n+1},\Lambda_{n-1})}_{\{\lambda_n\}}(g^{(n)}) = N(\rho_{n+1};\lambda_n)\phi^{(\lambda_n)}_{\{\Lambda_{n-1}\}\{\lambda_{n-1}\}}(g^{(n)}), \tag{15}$$

with

$$\frac{N(\rho_{n+1};\lambda_n)}{N(\rho_{n+1};\lambda'_n)} = \left(\prod_{j=1}^{[n/2]}\Gamma(m_{nj}-\rho_{n+1}-j+1)\right.$$
$$\times\Gamma(m'_{nj}+\rho_{n+1}-j+n)/$$
$$\Gamma(m_{nj}+\rho_{n+1}-j+n)$$
$$\left.\times\Gamma(m'_{nj}-\rho_{n+1}-j+1)\right)^{1/2}.$$

The bases are classified by the group chain SO(n,1)$\supset$ SO(n)$\supset\cdots\supset$SO(2). That is, the numbers $\{\lambda_j\}$ ($j=2,3,\ldots,n$) are characterized by the invariants of the second parameter group of SO(j) and (ρ_{n+1},Λ_{n-1}) by the invairants of SO(n,1) which can be expressed in terms of the number ρ_{n+1} and the invariants of the first parameter group of SO($n-1$). Thus, whatever the parameter of SO(n) are used, it is expected that the bases (15) give the correct result. Actually, it follows from (10) that the generator (J_{n+1n}) of the representation of the boost for the nth direction has the form

$$J_{n+1n} = -i(-\rho_{n+1}g^{(n)}_{nn} + \text{differential operator}),$$

where the second term on the right-hand side is independent of the number ρ_{n+1} and can explicitly be given in the form $\Sigma f_j(\alpha)\partial/\partial\alpha_j$ when the parameters α_j are chosen. Then it follows as in Ref. 2 that the bases (15) give the correct matrix elements for the generators of SO(n,1).

The matrix elements for the representation corresponding to h are found as follows. We may define the matrix elements for the representation operator $R(h)$ in a form

$$R(h)\Phi^{(\rho_{n+1},\Lambda_{n-1})}_{\{\lambda_n\}}(g^{(n)})$$
$$= \sum D^{(\rho_{n+1},\Lambda_{n-1})}_{\{\lambda'_n\}\{\lambda_n\}}(h)\Phi^{(\rho_{n+1},\Lambda_{n-1})}_{\{\lambda'_n\}}(g^{(n)}). \tag{16}$$

Due to the orthogonality (14) of the D matrix elements, (16) together with (10) gives

$$D^{(\rho_{n+1},\Lambda_{n-1})}_{\{\lambda'_n\}\{\lambda_n\}}(h)$$
$$= \frac{N(\rho_{n+1};\lambda_n)}{N(\rho_{n+1};\lambda'_n)}\int_{\mathrm{SO}(n)} dV_n\overline{\phi^{(\lambda'_n)}_{\{\lambda_{n-1}\}\{\lambda'_{n-1}\}}(g^{(n)})}$$
$$\times\left(h_{n+1n+1} + \sum g^{(n)}_{nj}h_{jn+1}\right)^{\rho_{n+1}}\phi^{(\lambda_n)}_{\{\Lambda_{n-1}\}\{\lambda_{n-1}\}}(g^{(n)\prime}), \tag{17}$$

where the elements of $g^{(n)\prime}$ are given by (3) and (4). (17) is the computation formula for the D matrix elements of SO(n,1).

It follows from (14) and (17) that the following relation holds for $h=g\in$SO(n),

$$D^{(\rho_{n+1},\Lambda_{n-1})}_{\{\lambda'_n\}\{\lambda_n\}}(g) = D^{(\lambda_n)}_{\{\lambda'_{n-1}\}\{\lambda_{n-1}\}}(g)\delta_{\lambda'_n\lambda_n}. \tag{18}$$

The D elements (17) may be written as the product of the two D matrix elements of SO(n) and the d matrix (boost) elements of SO(n,1). It is noted that (12) can explicitly be written in the form

$$D^{(\rho_{n+1},\Lambda_{n-1})}_{\{\lambda'_n\}\{\lambda_n\}}(h'h) = \sum D^{(\rho_{n+1},\lambda_{n-1})}_{\{\lambda'_n\}\{\lambda''_n\}}(h')D^{(\rho_{n+1},\lambda_{n-1})}_{\{\lambda''_n\}\{\lambda_n\}}(h). \tag{19}$$

For the parametrization of h as in (7), thus (17) can be rewritten as follows

$$D^{(\rho_{n+1},\Lambda_{n-1})}_{\{\lambda'_n\}\{\lambda_n\}}(h) = \sum D^{(\lambda'_n)}_{\{\lambda'_{n-1}\}\{\lambda''_{n-1}\}}(g_1)\ {}^{b}d^{(\rho_{n+1},\Lambda_{n-1})}_{\lambda'_n(\lambda''_{n-1})\lambda_n}(h^0).$$
$$\times D^{(\lambda_n)}_{\{\lambda''_{n-1}\}\{\lambda_{n-1}\}}(g_2), \tag{20}$$

where

$${}^{b}d^{(\rho_{n+1},\lambda_{n-1})}_{\lambda'_n(\lambda_{n-1})\lambda_n}(h^0).$$
$$= \frac{N(\rho_{n+1},\lambda_n)}{N(\rho_{n+1},\lambda'_n)}\int_{\mathrm{SO}(n)} dV_n\overline{\phi^{(\Lambda'_n)}_{\{\Lambda_{n-1}\}\{\lambda_{n-1}\}}(g^{(n)})}$$
$$\times(h^0_{n+1n+1} + g^{(n)}_{nn}h^0_{nn+1})^{\rho_{n+1}}\phi^{(\lambda_n)}_{\{\Lambda_{n-1}\}\{\lambda_{n-1}\}}(g^{(n)\prime}).$$

The elements of $g^{(n)\prime}$ are given by (3) and (4) with $h=h^0$. It is noted that the above matrix elements are defined regardless of the existence of the invariant scalar product in the space E_n and thus they must hold for any representation of SO(n,1), i.e., unitary as well as nonunitary. It is obvious that making choice of the Euler angles as the parameters of SO(n) gives the same results as in Refs. 1 and 2.

Let us introduce invariant scalar products under which the representations become unitary.

(i) *Principal series*: The scalar product is defined for $\Phi_1,\Phi_2\in H_1\subset E_n$ as follows

$$\langle\Phi_1,\Phi_2\rangle = \int_{\mathrm{SO}(n)} dV_n\,\overline{\Phi_1(g^{(n)})}\,\Phi_2(g^{(n)}). \tag{21}$$

Then the representation of SO(n,1) becomes unitary for $\rho_{n+1}=(1-n)/2+i\nu_{n+1},\nu_{n+1}$ real, i.e.,

$$\langle R(h)\Phi_1,R(h)\Phi_2\rangle = \langle\Phi_1,\Phi_2\rangle. \tag{22}$$

This can be seen as follows. We have from (10) and (21)

$$\langle R(h)\Phi_1,R(h)\Phi_2\rangle$$
$$= \int_{\mathrm{SO}(n)} dV_n\left(h_{n+1n+1} + \sum g^{(n)}_{nj}h_{jn+1}\right)^{\bar\rho_{n+1}+\rho_{n+1}}$$
$$\times\overline{\Phi_1(g^{(n)\prime})}\,\Phi_2(g^{(n)\prime}). \tag{23}$$

We use the following relations in order to change the integration variables from $g^{(n)}$ to $g^{(n)\prime}$

$$\left(h_{n+1n+1}+\sum g_{nj}^{(n)}h_{jn+1}\right)$$
$$=\left\{(h^{-1})_{n+1n+1}+\sum g_{nj}^{(n)'}(h^{-1})_{jn+1}\right\}^{-1},$$
$$dV_n=\left\{(h^{-1})_{n+1n+1}+\sum g_{nj}^{(n)'}(h^{-1})_{jn+1}\right\}^{1-n}dV'_n.$$

Substitution of these into (23) gives

$$\langle R(h)\Phi_1,R(h)\Phi_2\rangle$$
$$=\int_{\mathrm{SO}(n)}dV'_n\left\{(h^{-1})_{n+1n+1}+\sum g_{nj}^{(n)'}\right.$$
$$\left.\times(h^{-1})_{jn+1}\right\}^{1-n-\bar{\rho}_{n+1}-\rho_{n+1}}\overline{\Phi_1(g^{(n)'})}\Phi_2(g^{(n)'}), \quad (24)$$

which becomes (22) for $\rho_{n+1}=(1-n)/2+i\nu_{n+1}$. It is seen that the unitary irreducible representation corresponds to the principal series.

The bases of the space H_1 are defined by (15) with the above ρ_{n+1}, and the matrix elements for the generators of the representation together with their matrix elements[1,2] are easily obtained. The relations (17)–(20), of course, hold for this series.

(ii) *Complementary series*: The other invariant scalar product can also be defined for $\Phi_1,\Phi_2\in H_2\subset E_n$ as follows[1,2]:

$$\langle\Phi_1\Phi_2\rangle_c$$
$$=\int_{\mathrm{SO}(n)}\int_{\mathrm{SO}(n)}dV'_n\,dV_n\{1-(g^{(n)'}g^{(n)-1})_{nn}\}^{(1-n)/2-\sigma_{n+1}}$$
$$\times K(g^{(n)'}g^{(n)-1})\overline{\Phi_1(g^{(n)'}}\Phi_2(g^{(n)}), \quad (25)$$

where σ_{n+1} is some real constant. It is assumed as in Refs. 1 and 2 that the function K is invariant under the transformation of SO$(n,1)$. At first sight, it seems sufficient for us to fix the function K to some constant, but in this case only special classes of the unitary representations of the complementary series are realized, as seen in Refs. 1 and 2.

Then the representation becomes unitary for $\rho_{n+1}=(1-n)/2+\sigma_{n+1}$, σ_{n+1} real, i.e.,

$$\langle R(h)\Phi_1,R(h)\Phi_2\rangle_c=\langle\Phi_1,\Phi_2\rangle_c. \quad (26)$$

This can be seen as follows. We have from (10) and (25)

$$\langle R(h)\Phi_1,R(h)\Phi_2\rangle_c$$
$$=\int_{\mathrm{SO}(n)}\int_{\mathrm{SO}(n)}dV'_n\,dV_n\{1-(g^{(n)'}g^{(n)-1})_{nn}\}^{(1-n)/2-\sigma_{n+1}}$$
$$\times K(g^{(n)'}g^{(n)-1})\left(h_{n+1n+1}+\sum g_{nj}^{(n)'}h_{jn+1}\right)^{\bar{\rho}_{n+1}}$$
$$\times\left(h_{n+1n+1}+\sum g_{nj}^{(n)}h_{jn+1}\right)^{\rho_{n+1}}\overline{\Phi_1(\bar{g}^{(n)'})}\Phi_2(\bar{g}^{(n)}), \quad (27)$$

where $\bar{g}^{(n)'}$ and $\bar{g}^{(n)}$ denote the transformed quantities of $g^{(n)'}$ and $g^{(n)}$, respectively. Making use of the relations below (23) and

$$1-(g^{(n)'}g^{(n)-1})_{nn}$$
$$=\{1-(\bar{g}^{(n)'}\bar{g}^{(n)-1})_{nn}\}$$
$$\times\left\{(h^{-1})_{n+1n+1}+\sum\bar{g}_{nj}^{(n)'}(h^{-1})_{jn+1}\right\}^{-1}$$
$$\times\left\{(h^{-1})_{n+1n+1}+\sum\bar{g}_{nj}^{(n)}(h^{-1})_{jn+1}\right\}^{-1},$$

we rewrite (27) into the form

$$\langle R(h)\Phi_1,R(h)\Phi_2\rangle_c$$
$$=\int_{\mathrm{SO}(n)}\int_{\mathrm{SO}(n)}d\bar{V}'_n\,d\bar{V}_n$$
$$\times\{1-(\bar{g}^{(n)'}\bar{g}^{(n)-1})_{nn}\}^{(1-n)/2-\sigma_{n+1}}$$
$$\times\left\{(h^{-1})_{n+1n+1}+\sum\bar{g}_{nj}^{(n)'}(h^{-1})_{jn+1}\right\}^{(1-n)/2+\sigma_{n+1}-\bar{\rho}_{n+1}}$$
$$\times\left\{(h^{-1})_{n+1n+1}+\sum\bar{g}_{nj}^{(n)}(h^{-1})_{jn+1}\right\}^{(1-n)/2+\sigma_{n+1}-\rho_{n+1}}$$
$$\times K(\bar{g}^{(n)'}\bar{g}^{(n)-1})\overline{\Phi_1(\bar{g}^{(n)'})}\Phi_2(\bar{g}^{(n)}), \quad (28)$$

where use has been made of the invariance of K. This gives (26) for $\rho_{n+1}=(1-n)/2+\sigma_{n+1}$, and it is seen that the unitary irreducible representations correspond to the complementary series.

The bases of the space H_2 are also defined by (15) with $\rho_{n+1}=(1-n)/2+\sigma_{n+1}$, and the representation matrix elements together with those for the generators are easily obtained with respect to the bases.[1,2] The relations (17)–(20), of course, hold even in this complementary series.

[1] T. Maekawa, J. Math. Phys. **19**, 2028 (1978).
[2] T. Maekawa, J. Math. Phys. (to be published).
[3] K. Iwasawa, Ann. Math. **50**, 507 (1949).
[4] N.J. Vilenkin, *Special Functions and the Theory of Group Representations*, Math. Monographs Vol. 22, translated by V.N. Singh (Am. Math. Soc., Providence, Rhode Island, 1968).

Noncontinuous gauge potentials without magnetic monopoles[a)]

Francisco Antonio Doria

Departamento de Física Matemática, Instituto de Física, Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brazil

(Received 3 August 1978)

We give examples for Abelian and non-Abelian gauge fields without magnetic Dirac-like monopoles and with essentially discontinuous potentials.

1. INTRODUCTION

Wu and Yang[1,2] have recently discussed the relationship between noncontinuous gauge potentials and Dirac monopoles for such fields. Table I in the original Wu and Yang paper actually suggests an equivalence between noncontinuous gauge potentials and the existence of Dirac monopoles for gauge fields.[2] In the present note we show that there are noncontinuous potentials for an electromagnetic field without monopoles; the discontinuity observed is an essential fact that cannot be eliminated with the help of a global gauge transformation in the Wu and Yang sense.[1] Our construction is valid for spacetime manifolds (originally taken without monopoles) where the second De Rham cohomology group $D^2(M)$ is nontrivial.[3]

We give below two examples of such fields. Section 2 deals with the Abelian case while Sec. 3 deals with the non-Abelian one. Our example consists of a monopoleless field defined on a spacetime manifold with De Rham group $D^2(M) = \mathbb{R} \neq 0$.

2. THE ABELIAN CASE

Consider the four-dimensional manifold $S^2 \times \mathbb{R}^2$ embedded in 5-space $\mathbb{R}^5$ with coordinates as follows: $S^2 = (z,\varphi)$ where $z = \rho \exp i\theta$, ρ and θ being polar coordinates for this Cartesian plane $(x^1,x^2,0,0,0)$, and φ the usual azimuthal angle of spherical coordinates. The remaining coordinates are the Cartesian ones. Note that: (a) $M = S^2 \times \mathbb{R}^2$ admits a Lorentz metric[3,4]; (b) the origin $(0,0,0,0,0) \notin M$, and (c) $D^2(M) = \mathbb{R}$. Now the 2-form

$$\omega = \frac{1}{z} dz \wedge d\varphi \tag{2.1}$$

is closed, that is,

$$d\omega = 0 \tag{2.2}$$

all over M. If one takes ω to describe an electromagnetic field on M, Eq. (2.2) ensures the nonexistence of monopoles for such a field. ω, however, has no continuous potentials over the whole of $S^2 \times \mathbb{R}^2$, since

$$\omega = d\alpha \quad \text{and} \quad \alpha = (\mathrm{Re} \log z + i(\mathrm{Im} \log z)\, d\varphi, \tag{2.3}$$

and Im log$z$ is a "multivalued function," that is, no function in the strict sense. An equivalent example may be obtained out of

$$\omega' = (xdy - ydx) \wedge \frac{d\varphi}{(x^2 + y^2)}, \tag{2.4}$$

restricted to $S^2 \times \mathbb{R}^2$, where (x,y) are Cartesian coordinates related to (ρ,θ) in the standard way. Again

$$\omega' = d\alpha' \quad \text{and} \quad \alpha' = \mathrm{Im} \log z\, d\varphi, \tag{2.5}$$

that is, α' is not a potential in the standard mathematical sense. And there are no monopoles, for

$$d\omega' = 0 \tag{2.6}$$

is valid all over M. Thus, the existence of monopoles is not necessarily associated to noncontinuous potentials on a given spacetime M.

3. THE NON-ABELIAN CASE

Now for the non-Abelian case: Take as potentials for an SO(3) gauge field

$$\alpha^1 = \alpha^2 = 0, \tag{3.1}$$

$$\alpha^3 = \alpha', \tag{3.2}$$

α' given above in (2.5). Calculate the field intensities for α^a from the standard definition,

$$\omega^a = d\alpha^a + \lambda^{\ a}_{bc} \alpha^b \wedge \alpha^c, \tag{3.3}$$

where the $\lambda^{\ a}_{bc}$ are SO(3) Lie algebra constants. Field components are given by

$$\omega^1 = \omega^2 = 0 \tag{3.4}$$

and

$$\omega^3 = \omega', \tag{3.5}$$

where ω' is given by (2.4). SO(3) gauge fields admit the existence of Dirac-like monopoles.[1,5] But here there will be none, since on M we have

$$\hat{d}\omega^a = d\omega^a = 0, \quad a = 1,2,3, \tag{3.6}$$

where operator $\hat{d}$ includes the gauge connection contribution.

4. DISCUSSION

Dirac magnetic monopoles are characterized by a local breakdown of the differential Bianchi identities on the spacetime manifold.[6] If one excludes from the manifold the region where such a breakdown occurs, the Bianchi condition will be valid all over the remaining manifold, but the field's potential will be discontinuous as Wu and Yang have shown.[1] The examples given in Secs. 2 and 3 of the present paper arise

[a)]Partially supported by FINEP.

 0022-2488/79/071464-02$01.00

out of the following construction: We form a function which is a cocycle all over $\mathbb{R}^5$ but at the origin; in particular such a function is *always* a cocycle on the manifold $S^2 \times \mathbb{R}^2 \subset \mathbb{R}^5$. That is, we have placed the singularity out of the spacetime manifold. And so there are no monopoles on physical spacetime. But on the sphere $S^2 = (e^{i\theta}, \varphi, 0,0,0) \subset \mathbb{R}^5$ one has

$$\int_{S^2} \omega = 4\pi^2, \tag{4.1}$$

both for the Abelian and the non-Abelian cases, as a consequence of the nontrivial nature of the field ω.

[1]T.T. Wu and C.N. Yang, Phys. Rev. D **12**, 3845 (1975).
[2]C.N. Yang, preprint ITP/SB 77–14.
[3]F.A. Doria and S.M. Abrahão, "Mesonic Test Fields and Spacetime Cohomology," J. Math. Phys. **19**, 1650 (1978).
[4]L. Markus, Ann. Math. **62**, 3, 411 (1955).
[5]Z.F. Ezawa and H.C. Tze, J. Math. Phys. **17**, 2228 (1976).
[6]P.A.M. Dirac, Phys. Rev. **74**, 817 (1948).

Discontinuities of the Green's functions across the characteristic surfaces of field equations

Teymour Darkhosh

Physics Department, William Paterson College of New Jersey, 300 Pompton Road, Wayne, New Jersey 07470

(Received 28 August 1978)

A general method of calculating the retarded and advanced Green's functions for the second order hyperbolic partial differential equations in the neighborhood of the characteristic surfaces is developed. This method is then applied to two examples: the Proca wave equation in an external electromagnetic field and the Proca wave equation in an external symmetrical tensor field. The second example is of special interest since it exhibits acausal behavior. For this case, it is shown that the field commutators do not vanish for all spacelike separations and that the commutator of the energy tensor with itself is nonzero in the region bounded by the light cone and the spacelike characteristic surface.

INTRODUCTION

In this paper, we will develope a general method for calculating the retarded and advanced Green's functions for a system of second order hyperbolic differential equations in the neighborhood of the characteristic surfaces. We will do this by expanding the solution in terms of different derivatives of the Dirac delta function. In general, there will also be a remainder which in turn can be expanded in terms of the different powers of the characteristic variable. We will then substitute the expansion into the original equation and separate the coefficients of different singular functions as well as different powers of the characteristic variable. This procedure will yield a system of first order differential equations which can be solved within an unknown constant that is determined by considering the right-hand side of the original equation. However, because in some cases the first order differential equations are singular, we have to employ a slightly different method. This special problem will be demonstrated in the Proca wave equation in a constant external electromagnetic field where we find the first term of the expansion exactly. We will demonstrate the general method as it is applied to the Proca wave equation in a symmetrical tensor field. Since this problem exhibits the Velo–Zwanziger acausality phenomenon, it is of special interest to calculate the Green's functions and compare them with the canonical commutation relations. We will show that the field commutators do not vanish for all spacelike separations and that the commutator of the energy tensor with itself is nonzero in the spacelike region bounded by the light cone and the extraordinary characteristic surface.

In the subsequent calculation, our notation is that of Bjorken and Drell.[1] The metric tensor, $g^{\mu\nu}$, is used with the components $g^{00}=1, g^{11}=g^{22}=g^{33}=-1$. The Greek letters are used as the Lorentz indices and range from 0 to 3, while the Latin letters range from 1 to 3. Unless otherwise stated, the Einstein summation rule is used throughout the calculation, and $x\cdot y$ is defined as $x^{\alpha}y_{\alpha}$.

I. GENERAL FORMULATION

Consider a system of second order linear hyperbolic differential equations in space-time $x^{\mu}=(x^0,x^1,x^2,x^3)$:

$$L^{\mu\alpha}(x)\,D^{R}_{\alpha\nu}(x)=[A^{\mu\alpha}(x)_{\beta\gamma}\,\partial^{\beta}\partial^{\gamma}+B^{\mu\alpha}(x)_{\beta}\,\partial^{\beta}+C^{\mu\alpha}(x)]\times D^{R}_{\alpha\nu}(x)=g^{\mu}{}_{\nu}\,\delta^4(x), \qquad (1.1)$$

where $\partial^{\alpha}=\partial/\partial x_{\alpha}$ and $A^{\mu\alpha}(x)_{\beta\gamma}$, $B^{\mu\alpha}(x)_{\beta}$, and $C^{\mu\alpha}(x)$ are smooth functions of space–time. The function, $D^{R}_{\alpha\nu}(x)$, is the retarded Green's function,

$$D^{R}_{\alpha\nu}(x)=0,\quad x^0<0, \qquad (1.1')$$

and in general has the form

$$D^{R}_{\alpha\nu}(x)=\sum_{j=1}^{n}\left(\sum_{k=N}^{0}\delta^{k}(u_j)\,E^{kj}_{\alpha\nu}(x)+\Theta(u_j)\,G^{j}_{\alpha\nu}(x)\right), \qquad (1.2)$$

where $u_j=0$ is the jth characteristic surface defining the future cone, $\delta^{k}(u_j)$ is the kth derivative of the Dirac delta function, and $\Theta(u_j)$ is the step function:

$$\Theta(u_j)=0,\quad u_j<0,$$
$$\Theta(u_j)=1,\quad u_j>0.$$

Substituting (1.2) into (1.1) and separating the coefficients of the different singularities yield[2–4]:

$$\begin{aligned}&\mathscr{A}^{\mu\alpha}_{j}E^{N_j}_{\alpha\nu}=0,\quad j=1,2,\dots,n\\ &\mathscr{A}^{\mu\alpha}_{j}E^{N_j-1}_{\alpha\nu}+O^{\mu\alpha}_{j}E^{N_j}_{\alpha\nu}=0 \qquad (1.3)\\ &\mathscr{A}^{\mu\alpha}_{j}E^{N_j-2}_{\alpha\nu}+O^{\mu\alpha}_{j}E^{N_j-1}_{\alpha\nu}+L^{\mu\alpha}E^{N_j}_{\alpha\nu}=0\quad\text{or}\quad g^{\mu}{}_{\nu}\,\delta^3(r)\\ &\vdots\qquad\vdots\qquad\vdots\qquad\qquad\vdots\qquad\vdots\qquad=0,\end{aligned}$$

where N_j is the highest degree of singularity on the jth surface and the equations above are not summed over j. In the following calculation, we will consider a particular j and omit it from our equations.

To analyze Eqs. (1.3), we start with the first equation,

$$\mathscr{A}^{\mu\alpha}E^{N}_{\alpha\nu}=0\quad\text{or}\quad A^{\mu\alpha}(x)_{\beta\gamma}\,n^{\beta}n^{\gamma}E^{N}_{\alpha\nu}=0 \qquad (1.4)$$

 0022-2488/79/071466-07$01.00

with $\partial^\mu u = n^\mu$. The characteristic matrix $\mathscr{A}^{\mu\alpha}$ is singular, i.e.,

$$\det|\mathscr{A}| = 0 \tag{1.5}$$

and, in general, has the following form[5]:

$$\det|\mathscr{A}| = Q_1^s \cdot Q_2^t \cdot \cdots = 0, \tag{1.6}$$

where s and t are integers. Each factor in (1.6) determines a surface, and its exponent indicates the degeneracy of that surface. For example, the first factor, $Q_1^s = 0$, determines $u_1 = 0$ with s-fold degeneracy. In other words, there are s linearly independent vectors that satisfy

$$\mathscr{A}^{\mu\alpha} r^i_\alpha = 0,$$
$$l^i_\mu \mathscr{A}^{\mu\alpha} = 0, \quad i = 1,2,\ldots,s. \tag{1.7}$$

Then one can write $E^N_{\alpha\nu}(x)$ as

$$E^N_{\alpha\nu}(x) = \sum_{i=1}^{s} r^i_\alpha \sigma_\nu^{iN}. \tag{1.8}$$

The functions σ_ν^{iN} are unknown and, to calculate them, we consider the second equation in (1.3),

$$\mathscr{A}^{\mu\alpha} E^{N-1}_{\alpha\nu} + O^{\mu\alpha} E^N_{\alpha\nu} = 0. \tag{1.9}$$

Multiplying (1.9) by l^i_μ and summing over μ, we get

$$l^i_\mu O^{\mu\alpha} E^N_{\alpha\nu} = 0, \tag{1.10}$$

which is a first order differential equation and has the following form

$$l^j_\mu \frac{\mathscr{A}^{\mu\alpha}}{\partial n_\lambda} r^i_\alpha \partial_\lambda \sigma_\nu^{iN} + l^j_\mu \frac{\partial \mathscr{A}^{\mu\alpha}}{\partial n_\lambda} (\partial_\lambda r^i_\alpha) + l^j_\mu$$
$$\times (L^{\mu\alpha} u - C^{\mu\alpha} u)\, r^i_\alpha \sigma_\nu^{iN} = 0. \tag{1.10a}$$

The first term above can be shown to be equal to $M^{ij}\partial/\partial v$, i.e.,

$$l^i_\mu \frac{\partial \mathscr{A}^{\mu\alpha}}{\partial n_\lambda} r^j_\alpha \partial_\lambda = M^{ij} \frac{\partial}{\partial v}, \tag{1.11}$$

where v is a parameter along the bicharacteristics. Therefore, Eq. (1.10) becomes a first order differential equation in one variable, v, i.e.,

$$M^{ij} \frac{\partial \sigma^{jN}}{\partial v} + f^{ij} \sigma^{jN}_\nu = 0, \tag{1.12}$$

which can be integrated. However, if M^{ij} is singular, there is no solution unless f^{ij} is also singular and can be written as $f^{ij} = M^{ik} g^{kj}$. This will be shown in the first example below.

From Eq. (1.11), we see that $\mathscr{A}^{\mu\alpha}$ and $O^{\mu\alpha} E^N_{\alpha\nu}$ are orthogonal to l^i_μ, linearly independent vectors. Therefore, for any j

$$O^{\mu\alpha} r^i_\alpha \sigma_\nu^{iN} = \mathscr{A}^{\mu\alpha} F^N_{\alpha\nu}, \tag{1.13}$$

where $F^N_{\alpha\nu}$ are known functions.

Using (1.13) and (1.9), we can write

$$\mathscr{A}^{\mu\alpha}(E^{N-1}_{\alpha\nu} + F^N_{\alpha\nu}) = 0, \tag{1.14}$$

and solve for $E^{N-1}_{\alpha\nu}$:

$$E^{N-1}_{\alpha\nu} = \sum_{i=1}^{s} r^i_\alpha \sigma_\nu^{iN-1} - F^N_{\alpha\nu}. \tag{1.15}$$

Substituting (1.15) into the third equation in (1.3) and multiplying it by l^i_μ, we get

$$M^{ij} \frac{\partial \sigma_\nu^{jiN-1}}{\partial v} + K^{ij} \sigma_\nu^{jiN-1} = g^i_\nu(x). \tag{1.16}$$

The functions, k^{ij} and g^i_ν are known. Therefore, all the unknown functions can be calculated from first order partial differential equations of the form, (1.16), within unknown constants which are determined from the initial conditions, namely from the coefficient of the delta function on the right-hand side of Eq. (1.1).

This procedure can also be used to determine the coefficient of the step function, $G_{\alpha\nu}(x)$, by expanding it in a Taylor series about the characteristic surfaces, i.e.,

$$G(x) = G(0) + u_j \frac{\partial G(0)}{\partial u_j} + \frac{1}{2} u_j^2 \frac{\partial^2 G(0)}{\partial u_j^2} + \cdots. \tag{1.17}$$

By $G(0)$, we mean $G(u_j = 0)$.

The advanced Green's function is calculated in exactly the same manner, except that the characteristic surfaces define the past cone,

$$D^A_{\alpha\nu}(x) = 0, \quad x^0 > 0. \tag{1.18}$$

Expressions and equations similar to (1.4), (1.12), and (1.16) are obtained by considering a second order differential equation, $L\phi = 0$, where L is the operator of (1.1), and by calculating the propagation of the discontinuities in the second and higher derivatives of ϕ.[6]

II. EXAMPLE I: THE PROCA WAVE EQUATION IN AN EXTERNAL ELECTROMAGNETIC FIELD

In this section, we will apply the above formalism to a particular external field problem, the Proca wave equation in an external electromagnetic field. The equation which we are concerned with is

$$[(\Pi^2 + m^2) g^{\mu\alpha} - \Pi^\mu \Pi^\alpha + ie(\lambda - 1) F^{\mu\alpha}]\, D^R_{\alpha\nu}(x,y)$$
$$= 4\pi g^\mu{}_\nu \delta^4(x-y), \tag{2.1}$$

where $\Pi^\mu = \partial^\mu - ieA^\mu$, $F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu$, and λ is the magnetic moment strength. $A^\mu(x)$ is the electromagnetic vector potential. For simplicity, we will take a constant $F^{\mu\nu}$.

The characteristic matrix of (2.1) is

$$\mathscr{A}^{\mu\alpha} = g^{\mu\alpha} n^2 - n^\mu n^\alpha, \tag{2.2}$$

where n^α is $\partial^\alpha u$, and $u = 0$ is the characteristic surface. The determinant of (2.2) is identically zero, and to find the characteristic surfaces, we will apply the Velo–Zwanziger method,[7] i.e., we will multiply (2.1) by Π_ν and substitute the result back into (2.1). The result is

$$\{g^{\mu\alpha}(\Pi^2 + m^2) + [ie(\lambda+1)/m^2]\Pi^\mu \Pi \cdot F^\alpha$$
$$+ ie(\lambda - 1) F^{\mu\alpha}\}\, D^R_{\alpha\nu}(x,y)$$
$$= 4\pi\Big(g^\mu{}_\nu + \frac{1}{m^2} \Pi^\mu \Pi_\nu\Big)\, \delta^4(x-y). \tag{2.3}$$

Instead of solving (2.3), we will solve the closely related equation

$$\{g^{\mu\alpha}(\Pi^2+m^2)+[ie(\lambda+1)/m^2]\,\Pi^\mu\Pi\cdot F^\alpha + ie(\lambda+1)\,F^{\mu\alpha}\}\,\Delta^R_{\alpha\nu}(x,y) = 4\pi g^\mu{}_\nu\delta^4(x-y). \quad (2.4)$$

Once we find $\Delta_{\alpha\nu}$, we can find $D_{\alpha\nu}$ by means of the relation

$$D_{\alpha\nu}(x,y)=(g^\mu{}_\nu+(1/m^2)\,\Pi^\lambda{}_y\Pi_{y\nu})\Delta_{\alpha\lambda}(x,y), \quad (2.5)$$

where $\Pi^\lambda{}_y=\partial_y-ieA(y)$, and ∂y is the derivative with respect to the y variable. We will show that a relation (2.5) holds in general.

Let us assume that $\Delta_{\alpha\nu}$ satisfies

$$L^{\mu\alpha}(x)\Delta_{\alpha\nu}(x,y)=g^\mu{}_\nu\,\delta^4(x-y) \quad (2.6)$$

and that $D_{\alpha\nu}(x,y)$ satisfies

$$L^{\mu\alpha}(x)\,D_{\alpha\nu}(x,y)=A^{\mu\alpha}(x)I_{\alpha\nu}(\partial_x)\delta^4(x-y), \quad (2.7)$$

where $A^{\mu\alpha}(x)$ is a smooth function of x, and $I_{\alpha\nu}$ is some polynomial of the partial derivatives. The right-hand side of (2.7) can be written as

$$A^{\mu\alpha}(x)I_{\alpha\nu}(-\partial_y)\delta^4(x-y) \quad (2.8)$$

or

$$A^{\mu\alpha}(x)\delta^4(x-y)\,I_{\alpha\nu}(-\overleftarrow{\partial}y). \quad (2.9)$$

Using the property of the delta function, we get

$$A^{\mu\alpha}(y)\delta^4(x-y)\,I_{\alpha\nu}(-\overleftarrow{\partial}_y). \quad (2.10)$$

Multiplying (2.6) from the right by $A^{\nu\sigma}(y)\,I_{\sigma\lambda}(-\overleftarrow{\partial}_y)$, we have

$$L^{\mu\alpha}(x)\Delta_{\alpha\nu}(x,y)A^{\nu\sigma}(y)\,I_{\sigma\lambda}(-\overleftarrow{\partial}_y) = g^\mu{}_\nu\delta^4(x-y)\,A^{\nu\sigma}(y)\,I_{\sigma\lambda}(-\overleftarrow{\partial}_y) \quad (2.11)$$

or

$$L^{\mu\alpha}(x)\Delta_{\alpha\nu}(x,y)A^{\nu\sigma}I_{\sigma\lambda}(-\overleftarrow{\partial}_y) = A^{\mu\sigma}(x)\,I_{\sigma\lambda}(\partial_x)\delta^4(x-y). \quad (2.12)$$

The comparison of (2.12) and (2.7) gives

$$D_{\alpha\nu}(x,y)=\Delta_{\alpha\lambda}(x,y)\,A^{\mu\sigma}(y)I_{\sigma\nu}(-\overleftarrow{\partial}_y). \quad (2.13)$$

In terms of the notations used in the previous section, we have the following expressions:

$$\mathscr{A}^{\mu\alpha}=g^{\mu\alpha}n^2+[ie(\lambda+1)/m^2]n^\mu n\cdot F^\alpha,$$

$$O^{\mu\alpha}=2g^{\mu\alpha}(n\cdot\Pi+1/r) + [ie(\lambda+1)/m^2][n^\mu\Pi\cdot F^\alpha+\Pi^\mu n\cdot F^\alpha], \quad (2.14)$$

and

$$L^{\mu\alpha}=g^{\mu\alpha}(\Pi^2+m^2)+[ie(\lambda+1)/m^2]\Pi^\mu\Pi\cdot F^\alpha + ie(\lambda-1)\,F^{\mu\alpha},$$

where

$$r=[(x_1-y_1)^2+(x_2-y_2)^2+(x_3-y_3)^2]^{1/2}.$$

The determinant of the characeristic matrix $\mathscr{A}^{\mu\alpha}$ is

$$\det|\mathscr{A}^{\mu\alpha}|=-(n^2)^3\{n^2[ie(\lambda+1)/m^2]\;n\cdot F\cdot n\}. \quad (2.15)$$

Setting it equal to zero, we get

$$-(n^2)^3(n^2)=0. \quad (2.16)$$

Since $F^{\mu\nu}$ is antisymmetric, $n\cdot F\cdot n$ vanishes. The conoid solution to (2.16) is $u=t-r=0$, the light cone, which indicates that this wave equation is causal. t is defined as x^0-y^0. Therefore, the characteristic matrix with $u=0$ is given by

$$\mathscr{A}^{\mu\alpha}=[ie(\lambda+1)/m^2]\;n^\mu n\cdot F^\alpha, \quad (2.17)$$

which has three linearly independent right-hand solutions and three left-hand solutions as follows:

$$r^1=\begin{pmatrix} n\cdot F^1\\ n\cdot F^0\\ 0\\ 0\end{pmatrix},\quad r^2=\begin{pmatrix} n\cdot F^2\\ 0\\ n\cdot F^0\\ 0\end{pmatrix},\quad r^3=\begin{pmatrix} n\cdot F^3\\ 0\\ 0\\ n\cdot F^0\end{pmatrix} \quad (2.18)$$

and

$$l^1=(n^1,n^0,0,0),\quad l^2=(n^2,0,n^0,0),\quad l^3=(n^3,0,0,n^0). \quad (2.19)$$

From Eqs. (1.8), (1.10), and the second equation of (2.14), we have

$$2l^{j\alpha}(n\cdot\Pi+1/r)\,r^i_\alpha\sigma^{iN}_\nu=0,\quad j=1,2,3. \quad (2.20)$$

The differential operator of (2.20) is given by

$$n\cdot\partial=2\frac{\partial}{\partial v}, \quad (2.21)$$

where $v=t+r$ is the parameter along the bicharacteristics. The differential equation (2.20) is singular, i.e., $l^{j\alpha}r^i_\alpha$ is a singular matrix, and the general solution is

$$r^i_\alpha\sigma^{iN}_\nu=n_\alpha(e^{i\varphi}/r)\,\sigma^N_\nu+(e^{i\varphi}/r)C^N_{\alpha\nu}(0), \quad (2.22)$$

where $\varphi=(e/2)\int_0^v A\cdot n\,dv'$ and $C^N_{\alpha\nu}(0)$ is a constant in v. There are two unknown functions to be determined, $\sigma^N{}_\nu$ and $C^N_{\alpha\nu}$.

Using Eqs. (1.14) and (1.15) of Sec. 1, we can determine $E^{N-1}_{\alpha\nu}$ as

$$E^{N-1}_{\alpha\nu}=r^i_\alpha\sigma^{iN-1}_\nu+\Pi^\alpha\frac{e^{i\varphi}}{r}\sigma^N{}_\nu-\frac{n\cdot F_\alpha}{i\gamma|n\cdot F|^2} \times\left(4\frac{e^{i\varphi}}{r}\frac{\partial}{\partial v}\sigma^N{}_\nu+i\gamma\Pi\cdot F^\lambda C^N_{\lambda\nu}(0)\frac{e^{i\varphi}}{r}\right), \quad (2.23)$$

where $|n\cdot F|^2$ is defined as $n\cdot F^\alpha n\cdot F_\alpha$ and $\gamma=e(\lambda+1)/m^2$.

Next, we consider

$$l^j{}_\mu O^{\mu\alpha}E^{N-1}_{\alpha\nu}+l^j{}_\mu L^{\mu\alpha}E^N_{\alpha\nu}. \quad (2.24)$$

Because of the $1/r$ dependence of $E^N_{\alpha\nu}$, we get the following singular term:

$$4\pi l^j{}_\mu C^{\mu N}_\nu(0)\delta^3(r). \quad (2.25)$$

This either has to be zero or equal to $4\pi l^j{}_\mu g^\mu{}_\nu\delta^3(r)$. The latter is not possible since $C^{\mu\nu}(0)$ is orthogonal to $n\cdot F_\mu$ while $g^{\mu\nu}$ is not. Therefore, we set $C^{\mu\nu}(0)=0$.

We substitute (2.22) and (2.23) into (2.24), and we get

$$2l^{j\alpha}\left(n\cdot\Pi+\frac{1}{r}\right)r^i_\alpha\sigma^{iN}_\nu-\frac{16}{i\gamma}l^{j\alpha}\frac{n\cdot F_\alpha}{|n\cdot F|^2}\frac{e^{i\varphi}}{r}\frac{\partial^2\sigma^N{}_\nu}{\partial v^2} - ie(\lambda+1)l^{j\alpha}n\cdot F_\alpha\frac{e^{i\varphi}}{r}\sigma^N{}_\nu=0. \quad (2.26)$$

Multiplying the above equation by $n\cdot F^j$ and summing over j yields

$$\frac{\partial^2\sigma^N{}_\nu}{\partial v^2}-\beta^2\sigma^N{}_\nu=0, \quad (2.27)$$

where $\beta = m\gamma|n\cdot F|/4$. The solution to (2.27) is

$$\sigma^N{}_\nu(v) = A^N_\nu(0)\sinh\beta v + B^N_\nu(0)\cosh\beta v \qquad (2.28)$$

with $B^N_\nu(0) = 0$ since $\sigma^N_\nu(0) = 0$. Solving for $r^i_\alpha\sigma^{iN-1}_\nu$ with $\sigma^N{}_\nu$ from (2.28), we find

$$r^i_\alpha\sigma^{iN}_\nu = n_\alpha(e^{i\varphi}/r)\sigma^{iN}_\nu + C^{N-1}_{\alpha\nu}(0)\,e^{i\varphi}/r. \qquad (2.29)$$

Now we consider

$$\mathscr{A}^{\mu\alpha}E^{N-2}_{\alpha\nu} + O^{\mu\alpha}E^{N-1}_{\alpha\nu} + L^{\mu\alpha}E^N_{\alpha\nu} = 0, \qquad (2.30)$$

which gives

$$E^{N-2}_{\alpha\nu} = r^i_\alpha\sigma^{iN-2}_\nu + \Pi_\alpha\frac{e^{i\varphi}}{r}\sigma^{N-1}_\nu + \frac{4}{i\gamma}\Pi\cdot F_\alpha\frac{1}{|n\cdot F|^2}\frac{e^{i\varphi}}{r}\frac{\partial}{\partial v}\sigma^N{}_\nu$$
$$-\frac{n\cdot F_\alpha}{i\gamma|n\cdot F|^2}\left[4\frac{e^{i\varphi}}{r}\frac{\partial}{\partial v}\sigma^{N-1}_\nu + k\frac{e^{i\varphi}}{r}\sigma^N{}_\nu\right.$$
$$\left.+4\frac{n\cdot F^\lambda\cdot(\partial n)\cdot F_\lambda}{|n\cdot F|^2}\frac{e^{i\varphi}}{r}\frac{\partial\sigma^N{}_\nu}{\partial v} + i\gamma\Pi\cdot F\cdot C^{N-1}_\nu\frac{e^{i\varphi}}{r}\right], \qquad (2.31)$$

where $k = [\Pi^2 + m^2 + (e\gamma/2)|F|^2]$ with

$$|F|^2 = F^{\mu\alpha}F_{\mu\alpha}.$$

Our next step is to consider

$$l^j{}_\mu O^{\mu\alpha}E^{N-2}_{\alpha\nu} + l^j{}_\mu L^{\mu\alpha}E^{N-1}_{\alpha\nu}. \qquad (2.32)$$

When $L^{\mu\alpha}$ is operated on $E^{N-1}_{\alpha\nu}$, we get the following singular terms:

$$4\pi l^{j\alpha}C^{N-1}_{\alpha\nu}(0)\delta^3(r) - 16\pi l^{j\alpha}\frac{n\cdot F_\alpha}{|n\cdot F|^2}\frac{\partial\sigma^N{}_\nu(0)}{\partial v}\delta^3(r). \qquad (2.33)$$

Multiply by $n\cdot F^j$, we get

$$-16\pi\frac{\partial\sigma^N{}_\nu}{\partial v}(0)\,\delta^3(r), \qquad (2.34)$$

which cannot be zero. Therefore, it has to equal the singularity on the right-hand side, i.e.,

$$4\pi n\cdot F^j\,l^j{}_\mu\,g^\mu{}_\nu\,\delta^3(r) = 4\pi n\cdot F_\nu\,\delta^3(r). \qquad (2.35)$$

Equating (2.34) and (2.35), we completely determine $E^N_{\alpha\nu}$:

$$E^N_{\alpha\nu} = \frac{-in_\alpha n\cdot F_\alpha}{|n\cdot F|}\frac{e^{i\varphi}}{r}\sinh\beta v. \qquad (2.36)$$

We also determine N since $E^{N-1}_{\alpha\nu}$ is the coefficient of $\delta(u)$ implying that $N-1=0$. Therefore, the series starts with the first derivative of the delta function. The next term in the series is

$$E^0_{\alpha\nu} = n_\alpha\frac{e^{i\varphi}}{r}\sigma^{N-1}{}_\nu + g_{\alpha\nu}\frac{e^{i\varphi}}{r}$$
$$+\Pi_\alpha\frac{e^{i\varphi}}{r}\left(\frac{-in\cdot F_\nu}{m|n\cdot F|}\right)\sinh\beta v$$
$$-\frac{n\cdot F_\alpha n\cdot F_\nu}{|n\cdot F|^2}(1-\cosh\beta v)\frac{e^{i\varphi}}{r}. \qquad (2.37)$$

where σ^{N-1}_ν satisfies the following second order differential equation and with the initial condition $\sigma^{N-1}{}_\nu(0) = 0$:

$$-\frac{16}{i\gamma}\frac{\partial^2\sigma^{N-1}{}_\nu}{\partial v^2} - ie(\lambda+1)|n\cdot F|^2\frac{e^{i\varphi}}{r}\sigma^{N-1}{}_\nu$$
$$= -h\cdot F^\alpha\left[g_{\alpha\nu}(\Pi^2+m^2) + ie(\lambda-1)F_{\alpha\nu}\right]\frac{e^{i\varphi}}{r}$$
$$-2n\cdot F^\alpha\left(n\cdot\Pi + \frac{1}{r}\right)\Pi\cdot F_\alpha\;\frac{n\cdot F_\nu}{|n\cdot F|^2}$$
$$\times\frac{e^{i\varphi}}{r}(1-\cosh\beta v)$$
$$+ie(\lambda+1)n\cdot F^\alpha\Pi\cdot F_\alpha\frac{e^{i\varphi}}{r}\sigma^N{}_\nu$$
$$+\frac{2}{i\gamma}\left(n\cdot\Pi + \frac{1}{r}\right)k\frac{e^{i\varphi}}{r}\sigma^N{}_\nu$$
$$-2\left(n\cdot\Pi+\frac{1}{r}\right)\frac{F^\lambda\cdot(\partial n)\cdot F_\lambda}{|n\cdot F|^2}\frac{e^{i\varphi}}{r}n\cdot F_\nu$$
$$+2\left(n\cdot\Pi+\frac{1}{r}\right)\Pi\cdot F_\nu\frac{e^{i\varphi}}{r}$$
$$+n\cdot F^\alpha(\Pi^2+m^2)\frac{n\cdot F_\alpha n\cdot F_\nu}{|n\cdot F|^2}(1-\cosh\beta v)\frac{e^{i\varphi}}{r}. \qquad (2.38)$$

All the singular terms are subtracted from the right-hand side of (2.38).

Therefore, with this procedure, we reduce second order partial differential equations to a system of second order ordinary differential equations on the characteristic surfaces that have the following form:

$$\begin{aligned}
&\frac{d^2\sigma^N_\nu}{dv^2} - \beta^2\sigma^N_\nu = 0,\\
&\frac{d^2\sigma^{N-1}_\nu}{dv^2} - \beta^2\sigma^{N-1}_\nu = f^{N-1}_\nu,\\
&\frac{d^2\sigma^{N-1}_\nu}{dv^2} - \beta^2\sigma^{N-2}_\nu = f^{N-2}_\nu,\\
&\vdots\qquad\vdots\qquad\vdots
\end{aligned} \qquad (2.39)$$

III. EXAMPLE II: THE PROCA WAVE EQUATION IN A[8,9] SYMMETRICAL TENSOR FIELD

This problem structurally resembles the massive neutral spin-1 particle with an external quadrupole moment interaction.[10] The equation for the retarded Green's function is

$$[(\partial^2+m^2)g^{\mu\alpha} - \partial^\alpha + \lambda T^{\mu\alpha}]\,D^R_{\alpha\nu}(x,y)$$
$$= 4\pi g^\mu{}_\nu\,\delta^4(x-y), \qquad (3.1)$$

where $T^{\mu\nu} = T^{\nu\mu}$ and λ is a constant. To solve (3.1), we convert it to

$$[(\partial^2+m^2)\,g^{\mu\alpha} + (\lambda/m^2)\partial^\mu(\cdot T^\alpha + \lambda T^{\mu\alpha}]\,D^R_{\alpha\nu}(x,y)$$
$$= 4\pi(g^\mu{}_\nu + (1/m^2)\partial^\mu\partial_\nu)d^4(x-y) \qquad (3.2)$$

by the method mentioned in the second section, Eqs. (2.2) and (2.3).

As before, instead of Eq. (3.2), we will consider

$$[(\partial^2 + m^2) g^{\mu\alpha} + (\lambda/m^2)\partial^\mu\partial\cdot T^\alpha + \lambda T^{\mu\alpha}]\, \Delta^R_{\alpha\nu}(x,y) = 4\pi g^\mu{}_\nu \delta^4(x-y) \tag{3.3}$$

and use the relation

$$D^R_{\alpha\nu}(x,y) = [g^\lambda{}_\nu + (1/m^2)\partial^\lambda_y \partial_{y\nu}]\, \Delta^R_{\alpha\nu}(x,y) \tag{3.4}$$

to determine $D^R_{\alpha\nu}$.

The characteristic matrix for (3.2) is

$$\mathscr{A}^{\mu\alpha} = g^{\mu\alpha} n^2 + (\lambda/m^2) n^\mu n\cdot T^\alpha \tag{3.5}$$

with the determinant

$$\det|\mathscr{A}| = -(n^2)^3 [n^2 + (\lambda/m^2) n\cdot T\cdot n]. \tag{3.6}$$

Setting $\det|\mathscr{A}| = 0$, we see that there are two distinct surfaces: $u_1 = t - r = 0$, which results from $(n^2)^3 = 0$, is threefold degenerate; $u_2 = 0$, which corresponds to $[n^2 + (\lambda/m^2) n\cdot T\cdot n] = 0$, is not degenerate. To simplify the calculation and to obtain a well-defined cone for u_2, we will consider a special case where $T^{\mu\nu}$ has only one component, T^{00}. Then, $u_2 = Vt - r$, where V, the speed of propagation, is given by

$$V = \frac{1}{(1 + (\lambda/m^2)T)^{1/2}}. \tag{3.7}$$

Now, if we take T^{00} to satisfy

$$-1 < (\lambda/m^2) T^{00} < 0, \tag{3.8}$$

V becomes larger than the speed of light, and causes the surface, u_2, to lie outside the light cone. The lower limit in (3.8) is a necessary condition for the hyperbolic differential equation.

Assuming the general solution to (3.3) to be in the form of (1.2), we can write

$$\mathscr{A}_1^{\mu\alpha} = (\lambda/m^2) T^{00} n_1^\mu \delta^{0\alpha},$$

$$O_1^{\mu\alpha} = 2g^{\mu\alpha}(n_1\cdot\partial + 1/r) + (\lambda/m^2) T^{00}(\partial^\mu n_1^0 \delta^{0\alpha} + n_1^\mu \delta^{0\alpha}\partial^0),$$

$$L^{\mu\alpha} = g^{\mu\alpha}(\partial^2 + m^2) + (\lambda/m^2) T^{00} \partial^\mu \partial^0 \delta^{0\alpha} + \lambda T^{00}\delta^{0\alpha}\delta^{0\alpha} \tag{3.9}$$

on the $u_1 = 0$ surface, where $n_1^\mu = \partial^\mu u_1 = (1,\hat{r})$ is a constant vector along the bicharacteristics, and $\delta^{\mu\alpha}$ is 1 for $\mu = \alpha$ and zero otherwise.

Similarly, for the surface $u_2 = 0$, we can write

$$\mathscr{A}_2^{\mu\alpha} = (\lambda/m^2) T^{00}(-g^{\mu\alpha} + n_2^\mu \delta^{0\alpha}/V),$$

$$O_2^{\mu\alpha} = 2g^{\mu\alpha}(n_2\cdot\partial + 1/r) + (\lambda/m^2) T^{00}(\partial^\mu n_2^0 \delta^{0\alpha} + n_2^\mu \delta^{0\alpha}\partial^0). \tag{3.10}$$

$L^{\mu\alpha}$ is the same as (3.9), and $n_2^\mu = \partial^\mu u_2 = (V,\hat{r})$.

Since the $u_1 = 0$ is a threefold degenerate surface, there are three linearly independent vectors that are right-hand side solutions to the characteristic matrix:

$$r^1_\alpha = \delta^{1\alpha}, \quad r^2_\alpha = \delta^{2\alpha}, \quad r^3_\alpha = \delta^{3\alpha}. \tag{3.11}$$

Similarly, three left-hand solutions are

$$l^1 = \begin{pmatrix} n^1 \\ -1 \\ 0 \\ 0 \end{pmatrix}, \quad l^2 = \begin{pmatrix} n^2 \\ 0 \\ -1 \\ 0 \end{pmatrix}, \quad l^3 = \begin{pmatrix} n^3 \\ 0 \\ 0 \\ -1 \end{pmatrix}. \tag{3.12}$$

For the $u_2 = 0$ surface, we have one right-hand solution vector and one left-hand vector as follows:

$$r^4 = \begin{pmatrix} V \\ n_1 \\ n_2 \\ n_3 \end{pmatrix} = n_{2\alpha}, \quad l^4 = \frac{\delta^{0\alpha}}{V}. \tag{3.13}$$

Therefore, on $u_1 = 0$, we can write

$$E^{N_1}_{\alpha\nu} = \sum_{i=1}^{3} r^i_\alpha \sigma^{iN_1}{}_\nu \tag{3.14}$$

and on $u_2 = 0$

$$E^{N_2}_{\alpha\nu} = r^4_\alpha \sigma^{N_2}{}_\nu. \tag{3.15}$$

Following the method outlined in the first section, and using Eqs. (1.9) and (1.10) for each surface, we get

$$E^{N_1}_{\alpha\nu} = r^i_\alpha C^{iN_1}_\nu / r \tag{3.16}$$

and

$$E^{N_2}_{\alpha\nu} = n_{2\alpha} C^{iN_2}_\nu / r \tag{3.17}$$

where $C^{iN_1}_\nu$ is a constant independent of $v_1 = t + r$ and $C^{iN_2}_\nu$ is a constant independent of $v_2 = Vt + r$.

Since $E^{N_1}_{\alpha\nu}$ and $E^{N_2}_{\alpha\nu}$ have $1/r$ dependence, they both contribute to the singularity in the following way:

$$r^i_\alpha C^{iN_1}_\nu + n_{2\alpha} C^{N_2}_\nu = g_{\alpha\nu} \tag{3.18}$$

or

$$r^i_\alpha C^{iN_1}_\nu = g_{\alpha\nu} - n_{1\alpha} C^{N_2}_\nu. \tag{3.19}$$

Notice that we substituted $n_{1\alpha}$ for $n_{2\alpha}$ in (3.19) since

$$\delta^3(r)\partial_\alpha \Theta(u_1) = \delta^3(r)\partial_\alpha \Theta(u_2). \tag{3.20}$$

From Eq. (3.19), it follows that $N_1 = N_2 = 0$, *and* $C^0_\nu = \delta^0_\nu$.

The expansion on both surfaces starts with the delta function and results in terms such as

$$E^0_{\alpha\nu} = (g_{\alpha\nu} - n_{1\alpha}\delta_{0\nu})/r,$$
$$G^0_{\alpha\nu} = \mp \tfrac{1}{2} m^2 (g_{\alpha\nu} - n_{1\alpha}\delta_{0\nu}) - \partial_\alpha \delta_{0\nu}/r, \tag{3.21}$$
$$G^1_{\alpha\nu} = \tfrac{1}{8} m^4 (g_{\alpha\nu} - n_{1\alpha}\delta_{0\nu}) r - m^2(\delta_{0\alpha} - n_{1\alpha})\delta_{0\nu}/r,$$
$$\cdots\cdots$$

on the $u_1 = 0$ surface, and

$$E^0_{\alpha\nu} = \frac{n_{2\alpha}\delta_{0\nu}}{r},$$
$$G^0_{\alpha\nu} = \mp \frac{m^2}{2V^2} n_{2\alpha}\delta_{0\nu} + \partial_\alpha \frac{1}{r}\delta_{0\nu}, \tag{3.22}$$
$$G^1_{\alpha\nu} = \frac{m^4}{8V^4} n_{2\alpha}\delta_{0\nu} r + \frac{m^2}{V}\frac{\delta_{0\alpha}\delta_{0\nu}}{r} - \frac{m^2}{V^2} n_{2\alpha}\frac{\delta_{0\alpha}}{r}$$

on the $u_2 = 0$ surface.

In the above equations, the upper sign and $n^\mu = \partial^\mu u$ are for the retarded while the lower sign and $n^\mu = \partial^\mu v$ are for the advanced Green's functions.

Using Eq. (3.4), we get the following result for $D_{\alpha\nu}$:

$$D^R_{\alpha\nu} = \frac{V}{m^2}\partial_\alpha\partial_\nu\frac{\delta(u_2)}{r} + \partial_\alpha\frac{\Theta(u_2)}{r}\delta_{0\nu} + \delta_{0\alpha}\partial_\nu\frac{\Theta(u_2)}{r} - \partial_\alpha\frac{\Theta(u_1)}{r}\delta_{0\nu} - \delta_{0\alpha}\partial_\nu\frac{\Theta(u_1)}{r} - \frac{1}{2V}\partial_\alpha\partial_\nu\Theta(u_2) + \cdots \qquad (3.23)$$

and

$$D^A_{\alpha\nu} = \frac{V}{m^2}\partial_\alpha\partial_\nu\frac{\delta(v_2)}{r} + \partial_\alpha\frac{\Theta(v_2)}{r}\delta_{0\nu} + \delta_{0\alpha}\partial_\nu\frac{\Theta(v_2)}{r} - \partial_\alpha\frac{\Theta(v_2)}{r}\delta_{0\nu} - \delta_{0\alpha}\partial_\nu\frac{\Theta(v_2)}{r} + \frac{1}{2V}\partial_\alpha\partial_\nu\Theta(v_2) + \cdots. \qquad (3.24)$$

Note that the higher singular terms are due to the extraordinary surface, $u_2 = Vt - r$.

From Eqs. (3.23) and (3.24), we see that $D^R_{\alpha\nu}$ or $D^A_{\alpha N}$ does not vanish outside the light cone. This phenomenon contradicts the assumption made in the canonical quantization.

To show this, we write the Lagrangian density

$$\mathscr{L} = -\tfrac{1}{4}G^{\mu\nu}G_{\mu\nu} + \tfrac{1}{2}m^2\varphi\cdot\varphi + \tfrac{1}{2}\lambda\varphi\cdot T\cdot\varphi, \qquad (3.25)$$

where $G^{\mu\nu} = \partial^\mu\varphi^\nu - \partial^\nu\varphi^\mu$.

The canonical quantization imposes the following commutation relations on an arbitrary spacelike surface σ:

$$n_\mu[\Pi^{\mu i}(x),\varphi^j(y)]_\sigma = -\delta^{ij}\delta^3_\sigma(x-y)$$

and

$$n_\mu n_\nu[\Pi^{\mu i}(x),\Pi^{\nu j}(y)]_\sigma = [\varphi^i(x),\varphi^j(y)]_\sigma = 0, \qquad (3.26)$$

where n^μ is the normal vector to the surface σ and $\Pi^{\mu i} = G^{i\mu}$ which for $\mu = 0$ is the momentum conjugate to the field $\varphi^i(x)$. The delta function in (3.26) is defined by

$$\int g(x)\delta^3_\sigma(x-y)d\sigma = g(y). \qquad (3.27)$$

If we take the spacelike surface to be the $u_2 = Vt - r = 0$ surface, then from Peierl's quantization[11,12]

$$[\varphi_i(x),\varphi_j(y)]_{u_2=0} = D^R_{ij}(x,y)_{u_2=0}. \qquad (3.28)$$

Now according to (3.26), the right-hand side is zero, but the left hand side is

$$D^R_{ij}|_{u_2=0} = \frac{V}{m^2}\partial_i\partial_j\frac{\delta(u_2)}{r} - \frac{1}{2V}\partial_i\partial_j\Theta(u_2) - \frac{1}{V}\frac{1}{r}\partial_i\partial_j u_2\Theta(u_2), \qquad (3.29)$$

which does not vanish.

Similarly, we can write

$$n_\mu[\Pi^{\mu i}(x),\varphi^j(y)]_{u_2=0} = n_\mu\partial^\mu D^{ij}(x,y) - n_\mu\partial^i D^{\mu j}(x,y). \qquad (3.30)$$

The right-hand side can be written as

$$-V\partial^i\partial^j\frac{\Theta(u_2)}{r} - \frac{m^2}{2V}\partial^i\partial^j u_2\Theta(u_2), \qquad (3.31)$$

which is not the same as (3.26).

Next we will examine the Dirac–Schwinger[13,14] causality condition by calculating the symmetric energy–momentum tensor which is

$$\theta^{\mu\nu} = -\tfrac{1}{2}(G^{\mu\alpha}G^\nu{}_\alpha + G^{\nu\alpha}G^\mu{}_\alpha) + \tfrac{1}{2}m^2(\varphi^\mu\varphi^\nu + \varphi^\nu\varphi^\mu) + \tfrac{1}{2}\lambda(T^\mu\cdot\varphi\varphi^\nu + T^\nu\cdot\varphi\varphi^\mu) - \mathscr{L}, \qquad (3.32)$$

where $\mathscr{L}$ is the Lagrangian density.

From (3.32), θ^{00} is given

$$\theta\theta^{00} = -\tfrac{1}{2}G^{0i}G_{0i} + \tfrac{1}{4}G^{ij}G_{ij} + (m^2/2V^2)\varphi^{02} - \tfrac{1}{2}m^2\varphi^i\varphi_i \qquad (3.33)$$

$$\begin{aligned}[\theta^{00}(x),\theta^{00}(y)] &= \{\varphi^0(x),\varphi^0(y)\}\Big[\frac{m^2}{2V}\frac{\delta''(u_2)}{r} - \frac{m^4}{4V^3}\delta'(u_2) + \frac{m^4}{4V^3}\delta'(u_2) + \frac{m^4}{2V^3}\frac{\delta(u_2)}{r} - \frac{m^6}{2V^5}\Theta(u_2) + \frac{m^6}{4V^4}\Theta(u_1) + \frac{m^6}{2V^3}u_2\frac{\Theta(u_2)}{r} + \cdots\Big] \\ &\quad + \{\varphi^i(x),\varphi_i(y)\}\Big[\frac{m^4}{2}\frac{\delta(u_1)}{r} + \frac{m^6}{4}\Theta(u_1) + \cdots\Big] + \{\varphi^i_{(x)},\varphi^j_{(y)}\} \\ &\quad + \frac{m^4}{2}\partial_i\partial_j\frac{u_1\Theta(u_1)}{r} - \frac{m^4}{2V}\partial_i\partial_j\frac{u_2\Theta(u_2)}{r} + \cdots\Big],\end{aligned} \qquad (3.34)$$

where $\{\varphi^i,\varphi^j\} = \varphi^i\varphi^j + \varphi^j\varphi^i$ and $\delta'(u)$ is the derivative of the delta function with respect to its argument. The energy density is a measurable quantity, and its measurements at two different points have to be independent if the separation of the points is spacelike. We see, however, from (3.34) that the commutator does not vanish in the spacelike region bounded by the surface u_2 and the light cone.

ACKNOWLEDGMENT

I would like to thank Professor Daniel Zwanziger of New York University for his guidance, his encouragement, and his part in many stimulating discussions.

[1]J.D. Bjorken and S.D. Drell, *Relativistic Quantum Fields* (McGraw-Hill, New York, 1965).
[2]R. Courant and D. Hilbert, *Methods of Mathematical Physics, II* (Wiley, New York, 1962), p. 618.
[3]For a more rigorous derivation of Eqs. (1.3) using generalized functions, see T. Darkhosh, *Acausal Behavior of the Proca Equation in an External Field*, Ph.D. thesis, New York University (1976).
[4]D. Zwanziger, "Method of Characteristics in the External Field Problem," *Invariant Wave Equations, Proc. Erice* 1977, *Lecture Notes in Physics*, Vol. 73 (Springer, New York, 1978), pp. 143–64.
[5]Courant and Hilbert, Ref. 2, p. 596.
[6]J. Madore and W. Tait, "Propagation of Shock Waves in Interacting Higher Spin Wave Equations," Commun. Math. Phys. **30**, 201–09 (1973).
[7]G. Velo and D. Zwanziger, "Propagation and Quantization of Rarita-Schwinger Waves in an External Electromagnetic Potential," Phys. Rev. **186**, 1337 (1969).
[8]Darkhosh, Ref. 3.
[9]G. Velo, "An Existence Theorem for Massive Spin One Particles in an External Tensor Field," Ann. Inst. H. Poincaré **XXII**, 249 (1975).
[10]G. Velo and D. Zwanziger, "Noncausality and Other Defects of Interaction Lagrangian for Particles with Spin One and Higher," Phys. Rev. **188**, 2218 (1969).
[11]R.E. Peierls, "The Commutation Laws of Relativistic Field Theory," Proc. Roy. Soc. (London) Ser. A **124**, 143 (1952).
[12]A.S. Wightman, "Relativistic Wave Equations as Singular Hyperbolic Systems," Proc. Symp. in Pure Math. (Berkeley **XXXIII**, 1971).
[13]P.A.M. Dirac, "The Conditions for a Quantum Field Theory to be Relativistic," Rev. Mod. Phys. **34**, 592 (1962).
[14]J. Schwinger, "Commutation Relations and Conservation Laws," Phys. Rev. **130**, 406 (1963).

Coupled inertial and gravitational effects in the proper reference frame of an accelerated, rotating observer[a)]

Wann-Quan Li and Wei-Tou Ni

Department of Physics, National Tsing Hua University, Hsinchu, Taiwan, Republic of China

(Received 10 July 1978; revised manuscript received 11 January 1979)

To the second order in metric and the first order in equations of motion in the local coordinates of an accelerated rotating observer, the inertial effects and gravitational effects are simply additive. To look into the coupled inertial and gravitational effects, we derive the third-order expansion of the metric and the second-order expansion of the equations of motion in local coordinates. Besides purely gravitational (purely curvature) effects, the equations of motion contain, in this order, the following coupled inertial and gravitational effects: redshift corrections to electric, magnetic, and double-magnetic type curvature forces; velocity-induced special relativistic corrections; and electric, magnetic, and double-magnetic type coupled inertial and gravitational forces. An example is provided with a static observer in the Schwarzschild spacetime.

I. INTRODUCTION

In 1973, Misner, Thorne, and Wheeler (MTW)[1] defined a natural coordinate system for an accelerated, rotating observer (laboratory). This is the local coordinate system, suitable for analyzing experiments performed in the proper reference frame of an accelerated, rotating observer. Using these coordinates, they calculated the first-order expansion of the metric, and obtained the inertial accelerations on the world line of the observer. To this order, there is no curvature (gravitational) effect, as required by Einstein's principle of equivalence. Higher order inertial effects are treated in special relativity by Ni,[2] Li and Ni,[3] and DeFacio, Dennis, and Retzloff.[4]

Extending MTW's work, Ni and Zimmermann[5] derived the second-order expansion of the metric and the first-order expansion of the coordinate acceleration of a freely falling particle. To this order, the inertial effects and the gravitational effects are clearly separable and there are no coupled terms. Using a geodesic observer, Mashhoon[6] has obtained the second-order gravitational effects in coordinate acceleration.

To study the coupled inertial and gravitational effects, we will derive the third-order expansion of the metric and the second-order expansion of the equations of motion in local coordinates. In this order, there are two types of terms: (i) purely gravitational, and (ii) coupled inertial and gravitational effects. The coupled effects are listed in Table I. As an example for illustration, we consider a static observer in the Schwarzschild geometry. This case is important because Schwarzschild geometry is the most general spherical-symmetric exterior solution of general relativity. It would also be helpful in understanding the local difference between Hawking effect and the effect of constant acceleration in Minkowski spacetime if one carries the radiation calculation further to include coupled inertial and gravitational effects.

[a)]Supported in part by the National Science Council of the Republic of China.

II. QUADRATIC TERMS IN THE AFFINITY AND CUBIC TERMS IN THE METRIC

Consider an observer moving along the world line $P_0(\tau)$ with 4-velocity $\vec{u}(\tau)$, and 4-rotation $\vec{\omega}(\tau)$ in a gravitational field with Riemann tensor $R^{\mu}{}_{\nu\alpha\beta}(\tau)$ along the world line. The orthonormal tetrad $\{\vec{e}_{\hat{\alpha}}\}$ which the observer carries transports according to[7]

$$\frac{d\vec{e}_{\hat{\alpha}}}{d\tau} = -\vec{\vec{\Omega}}\cdot\vec{e}_{\hat{\alpha}}, \tag{1}$$

where[8]

$$\Omega^{\mu\nu} \equiv a^{\mu}u^{\nu} - a^{\nu}u^{\mu} + u_{\alpha}\,\omega_{\beta}\,\epsilon^{\alpha\beta\mu\nu}, \tag{2}$$

$$\vec{a}(\tau) \equiv \nabla_{\vec{u}}\vec{u}, \tag{3}$$

and τ is the proper time along the world line.

TABLE I. Various coupled inertial and gravitational effects in coordinate acceleration.

Effects	Terms in coordinate acceleration $d^2x^{\hat{i}}/dx^{\hat{0}^2}$
1. "Electric-type" redshift correction	$-2\mathbf{a}\cdot\mathbf{x}R_{\hat{0}\hat{i}\hat{0}\hat{l}}x^{\hat{l}}$
2. High-velocity correction to term 1	$\frac{8}{3}\mathbf{a}\cdot\mathbf{x}R_{\hat{0}\hat{j}\hat{0}\hat{l}}x^{\hat{l}}v^{i}v^{j}$
3. "Magnetic-type" redshift correction	$\frac{8}{3}\mathbf{a}\cdot\mathbf{x}R_{\hat{i}\hat{j}\hat{0}\hat{l}}x^{\hat{l}}v^{j}$
	$-\frac{2}{3}\mathbf{a}\cdot\mathbf{x}R_{\hat{i}\hat{0}\hat{l}\hat{j}}x^{\hat{l}}v^{j}$
4. High-velocity correction to term 3	$\frac{4}{3}\mathbf{a}\cdot\mathbf{x}R_{\hat{0}\hat{j}\hat{k}\hat{l}}x^{\hat{l}}v^{i}v^{j}v^{k}$
5. "Double-magnetic" redshift correction	$\frac{4}{3}\mathbf{a}\cdot\mathbf{x}R_{\hat{i}\hat{j}\hat{k}\hat{l}}x^{\hat{l}}v^{j}v^{k}$
6. "Electric-type" coupled effect	$-\frac{2}{3}v^{i}\mathbf{a}\cdot\mathbf{v}R_{\hat{0}\hat{l}\hat{0}\hat{m}}x^{\hat{l}}x^{\hat{m}}$
7. "Magnetic-type" coupled effect	$-\frac{2}{3}\mathbf{a}\cdot\mathbf{v}R_{\hat{0}\hat{l}\hat{i}\hat{m}}x^{\hat{l}}x^{\hat{m}}$
	$+\frac{1}{3}v^{i}a^{\hat{k}}R_{\hat{0}\hat{l}\hat{k}\hat{m}}x^{\hat{l}}x^{\hat{m}}$
8. "Double-magnetic" coupled effect	$\frac{1}{3}a^{\hat{k}}R_{\hat{i}\hat{m}\hat{k}\hat{l}}x^{\hat{l}}x^{\hat{m}}$

 0022-2488/79/071473-08$01.00

Following §13.6 of MTW, at any event $P_0(\tau)$ we send out geodesics $P(\tau;\vec{n};s)$ orthogonal to $\vec{u}(\tau)$, where $\vec{n}$ is the unit vector tangent to a particular geodesic at $P_0(\tau)$, and $\vec{n}\cdot\vec{u}(\tau)=0$. An event a distance s out along any geodesic $\vec{n}$ is then assigned the coordinates $x^{\hat{0}}\equiv\tau$, $x^{\hat{j}}=s\vec{n}\cdot e_{\hat{j}}$. These coordinates are well defined in the neighborhood of the world-line $P_0(\tau)$, and are called local coordinates. For a discussion of when these coordinates are well defined, see Ref. 5.

In the local coordinate system we decompose a 4-vector $\vec{V}$ as $\vec{V}=(V^{\hat{0}};V^{\hat{i}})\equiv(V^{\hat{0}};\mathbf{V})$. Now defining $\vec{b}\equiv\nabla_{\vec{u}}\,a$, $\vec{c}\equiv\nabla_{\vec{u}}\,\vec{b}$, $\vec{\eta}\equiv\nabla_{\vec{u}}\,\vec{\omega}$, $\vec{\zeta}\equiv\nabla_{\vec{u}}\vec{\eta}$ and using (1) and (2) we have

$$b^{\hat{0}}=\mathbf{a}\cdot\mathbf{a},\quad \mathbf{b}=\frac{d\mathbf{a}}{d\tau}+\boldsymbol{\omega}\times\mathbf{a},\quad c^{\hat{0}}=3\mathbf{a}\cdot\mathbf{b},\quad \mathbf{c}=\frac{d\mathbf{b}}{d\tau}+\boldsymbol{\omega}\times\mathbf{b}+\mathbf{a}(\mathbf{a}\cdot\mathbf{a}),$$
$$\eta^{\hat{0}}=\boldsymbol{\omega}\cdot\mathbf{a},\quad \boldsymbol{\eta}=\frac{d\boldsymbol{\omega}}{d\tau},\quad \zeta^{\hat{0}}=2\mathbf{a}\cdot\boldsymbol{\eta}+\boldsymbol{\omega}\cdot\mathbf{b},\quad \boldsymbol{\zeta}=\frac{d\boldsymbol{\eta}}{d\tau}+\boldsymbol{\omega}\times\boldsymbol{\eta}+\mathbf{a}(\mathbf{a}\cdot\boldsymbol{\omega}). \tag{4}$$

For Minkowski spacetime, these definitions reduce to those of Refs. 2 and 3.

Along $P_0(\tau)$, MTW derived the connection coefficients and the first-order partial derivatives of the metric to be

$$\Gamma^{\hat{0}}{}_{\hat{0}\hat{0}}=\Gamma^{\hat{\alpha}}{}_{\hat{j}\hat{k}}=0,\quad \Gamma^{\hat{0}}{}_{\hat{j}\hat{0}}=a^{\hat{j}},\quad \Gamma^{\hat{j}}{}_{\hat{k}\hat{0}}=-\omega^{\hat{i}}\epsilon^{\hat{i}\hat{j}\hat{k}},\quad \text{all along } P_0(\tau), \tag{5}$$
$$g_{\hat{\alpha}\hat{\beta},\hat{0}}=g_{\hat{i}\hat{j},\hat{l}}=0,\quad g_{\hat{0}\hat{0},\hat{j}}=-2a_{\hat{j}},\quad g_{\hat{0}\hat{j},\hat{k}}=-\epsilon^{\hat{j}\hat{k}\hat{l}}\omega^{\hat{l}},\quad \text{all along } P_0(\tau). \tag{6}$$

For the first-order partial derivatives of the connection coefficients and the second-order partial derivatives of the metric, Ni and Zimmermann[5] obtained the following formulas:

$$\begin{aligned}&\Gamma^{\hat{0}}{}_{\hat{0}\hat{0},\hat{0}}=\Gamma^{\hat{\alpha}}{}_{\hat{j}\hat{k},\hat{0}}=0,\quad \Gamma^{\hat{0}}{}_{\hat{j}\hat{0},\hat{0}}=b^{\hat{j}}(\tau)+\epsilon^{\hat{j}\hat{k}\hat{l}}a^{\hat{k}}(\tau)\omega^{\hat{l}}(\tau),\quad \Gamma^{\hat{i}}{}_{\hat{j}\hat{0},\hat{0}}=-\eta^{\hat{k}}(\tau)\epsilon^{\hat{i}\hat{j}\hat{k}},\\ &\Gamma^{\hat{0}}{}_{\hat{0}\hat{0},\hat{i}}=b^{\hat{i}}(\tau)+2a^{\hat{j}}(\tau)\omega^{\hat{k}}(\tau)\epsilon^{\hat{i}\hat{j}\hat{k}},\quad \Gamma^{\hat{j}}{}_{\hat{0}\hat{0},\hat{i}}=R_{\hat{0}\hat{j}\hat{0}\hat{i}}-\eta^{\hat{k}}\epsilon^{\hat{j}\hat{i}\hat{k}}+a^{\hat{i}}a^{\hat{j}}+\omega^{\hat{i}}\omega^{\hat{j}}-\delta_{\hat{i}}{}^{\hat{j}}(\omega^{\hat{l}})^2,\\ &\Gamma^{\hat{0}}{}_{\hat{j}\hat{0},\hat{i}}=R_{\hat{0}\hat{j}\hat{0}\hat{i}}-a^{\hat{i}}a^{\hat{j}},\quad \Gamma^{\hat{j}}{}_{\hat{k}\hat{0},\hat{i}}=R_{\hat{j}\hat{k}\hat{i}\hat{0}}+a^{\hat{k}}\omega^{\hat{l}}\epsilon_{\hat{l}\hat{j}\hat{i}},\quad \Gamma^{\hat{\mu}}{}_{\hat{i}\hat{j},\hat{k}}=-\tfrac{1}{3}(R^{\hat{\mu}}{}_{\hat{i}\hat{j}\hat{k}}+R^{\hat{\mu}}{}_{\hat{j}\hat{i}\hat{k}}),\end{aligned} \tag{7}$$

all along $P_0(\tau)$;

$$\begin{aligned}&g_{\hat{\alpha}\hat{\beta},\hat{0}\hat{0}}=0,\quad g_{\hat{j}\hat{k},\hat{l}\hat{0}}=0,\quad g_{\hat{0}\hat{0},\hat{j}\hat{0}}=-2(b^{\hat{j}}+\epsilon^{\hat{j}\hat{k}\hat{l}}a^{\hat{k}}\omega^{\hat{l}}),\quad g_{\hat{0}\hat{j},\hat{k}\hat{0}}=-\epsilon_{\hat{j}\hat{k}\hat{l}}\eta^{\hat{l}},\\ &g_{\hat{0}\hat{0},\hat{j}\hat{k}}=-R^{\hat{0}}{}_{\hat{j}\hat{k}\hat{0}}-2a^{\hat{j}}a^{\hat{k}}+2\delta^{\hat{j}}{}_{\hat{k}}(\omega^{\hat{l}})^2-2\omega^{\hat{j}}\omega^{\hat{k}},\quad g_{\hat{0}\hat{i},\hat{j}\hat{k}}=-\tfrac{2}{3}(R_{\hat{0}\hat{k}\hat{i}\hat{j}}+R_{\hat{0}\hat{j}\hat{i}\hat{k}}),\quad g_{\hat{l}\hat{m},\hat{i}\hat{j}}=-\tfrac{1}{3}(R_{\hat{i}\hat{l}\hat{j}\hat{m}}+R_{\hat{i}\hat{m}\hat{j}\hat{l}}),\end{aligned} \tag{8}$$

all along $P_0(\tau)$.

Differentiating (7) along the trajectory with respect to τ and using (4), we have

$$\begin{aligned}&\Gamma^{\hat{0}}{}_{\hat{0}\hat{0},\hat{0}\hat{0}}=\Gamma^{\hat{\alpha}}{}_{\hat{j}\hat{k},\hat{0}\hat{0}}=0,\quad \Gamma^{\hat{0}}{}_{\hat{j}\hat{0},\hat{0}\hat{0}}=c^{\hat{j}}-2\epsilon^{\hat{j}\hat{m}\hat{n}}\omega^{\hat{m}}b^{\hat{n}}-a^{\hat{j}}\mathbf{a}^2+\omega^{\hat{j}}\mathbf{a}\cdot\boldsymbol{\omega}-a^{\hat{j}}\omega^2+\epsilon^{\hat{j}\hat{k}\hat{l}}a^{\hat{k}}\eta^{\hat{l}},\\ &\Gamma^{\hat{i}}{}_{\hat{j}\hat{0},\hat{0}\hat{0}}=-\epsilon^{\hat{i}\hat{j}\hat{k}}\zeta^{\hat{k}}+\epsilon^{\hat{i}\hat{j}\hat{k}}a^{\hat{k}}\mathbf{a}\cdot\boldsymbol{\omega}+\omega^{\hat{i}}\eta^{\hat{j}}-\omega^{\hat{j}}\eta^{\hat{i}},\quad \Gamma^{\hat{0}}{}_{\hat{0}\hat{0},\hat{i}\hat{0}}=c^{\hat{i}}-3\epsilon^{\hat{i}\hat{m}\hat{n}}\omega^{\hat{m}}b^{\hat{n}}-a^{\hat{i}}\mathbf{a}^2+2\omega^{\hat{i}}\mathbf{a}\cdot\boldsymbol{\omega}-2a^{\hat{i}}\boldsymbol{\omega}^2+2\epsilon^{\hat{i}\hat{j}\hat{k}}a^{\hat{j}}\eta^{\hat{k}},\\ &\Gamma^{\hat{j}}{}_{\hat{0}\hat{0},\hat{i}\hat{0}}=R_{\hat{0}\hat{j}\hat{0}\hat{i}}+a^{\hat{l}}(R_{\hat{0}\hat{i}\hat{l}\hat{j}}+R_{\hat{0}\hat{j}\hat{l}\hat{i}})-(\epsilon^{\hat{l}\hat{j}\hat{n}}R_{\hat{0}\hat{l}\hat{0}\hat{i}}+\epsilon^{\hat{l}\hat{i}\hat{n}}R_{\hat{0}\hat{j}\hat{0}\hat{l}})\omega^{\hat{n}}-\epsilon^{\hat{j}\hat{i}\hat{k}}\zeta^{\hat{k}}+\epsilon^{\hat{j}\hat{i}\hat{k}}a^{\hat{k}}\boldsymbol{\omega}\cdot\mathbf{a}+a^{\hat{i}}b^{\hat{j}}+a^{\hat{j}}b^{\hat{i}}-\epsilon^{\hat{j}\hat{m}\hat{n}}\omega^{\hat{m}}a^{\hat{n}}a^{\hat{i}}\\ &\qquad-\epsilon^{\hat{i}\hat{m}\hat{n}}\omega^{\hat{m}}a^{\hat{n}}a^{\hat{j}}+2\eta^{\hat{i}}\omega^{\hat{j}}-2\delta_{\hat{i}\hat{j}}\,\omega^{l}\eta^{l},\\ &\Gamma^{\hat{0}}{}_{\hat{j}\hat{0},\hat{i}\hat{0}}=R_{\hat{0}\hat{j}\hat{0}\hat{i};\hat{0}}+a^{\hat{l}}(R_{\hat{0}\hat{i}\hat{l}\hat{j}}+R_{\hat{0}\hat{j}\hat{l}\hat{i}})-(\epsilon^{\hat{l}\hat{j}\hat{n}}R_{\hat{0}\hat{l}\hat{0}\hat{i}}+\epsilon^{\hat{l}\hat{i}\hat{n}}R_{\hat{0}\hat{j}\hat{0}\hat{l}})\omega^{\hat{n}}-a^{\hat{j}}b^{\hat{i}}-a^{\hat{i}}b^{\hat{j}}+\epsilon^{\hat{i}\hat{m}\hat{n}}\omega^{\hat{m}}a^{\hat{n}}a^{\hat{j}}+\epsilon^{\hat{j}\hat{m}\hat{n}}\omega^{\hat{m}}a^{\hat{n}}a^{\hat{i}},\\ &\Gamma^{\hat{j}}{}_{\hat{k}\hat{0},\hat{i}\hat{0}}=-R_{\hat{0}\hat{i}\hat{j}\hat{k};\hat{0}}-a^{\hat{l}}R_{\hat{l}\hat{i}\hat{j}\hat{k}}-a^{\hat{j}}R_{\hat{0}\hat{i}\hat{0}\hat{k}}+a^{\hat{k}}R_{\hat{0}\hat{i}\hat{0}\hat{j}}+(\epsilon^{\hat{l}\hat{i}\hat{n}}R_{\hat{0}\hat{l}\hat{j}\hat{k}}+\epsilon^{\hat{l}\hat{j}\hat{n}}R_{\hat{0}\hat{i}\hat{l}\hat{k}})\omega^{\hat{n}}+\epsilon^{\hat{l}\hat{k}\hat{n}}\omega^{\hat{n}}R_{\hat{0}\hat{i}\hat{j}\hat{l}}+\epsilon^{\hat{l}\hat{j}\hat{i}}b^{\hat{k}}\omega^{\hat{l}}+\epsilon^{\hat{l}\hat{j}\hat{i}}a^{\hat{k}}\eta^{\hat{l}}-\epsilon^{\hat{l}\hat{j}\hat{i}}(\boldsymbol{\omega}\times\mathbf{a})^{\hat{k}}\omega\\ &\Gamma^{\hat{j}}{}_{\hat{l}\hat{m},\hat{n}\hat{0}}=P_{lm}[\tfrac{1}{3}R_{\hat{j}\hat{m}\hat{n}\hat{l};\hat{0}}+\tfrac{1}{3}a^{\hat{j}}R_{\hat{0}\hat{l}\hat{n}\hat{m}}+\tfrac{1}{3}a^{\hat{l}}(R_{\hat{0}\hat{j}\hat{m}\hat{n}}+R_{\hat{0}\hat{n}\hat{m}\hat{j}})+\tfrac{1}{3}a^{\hat{n}}R_{\hat{0}\hat{m}\hat{j}\hat{l}}-\tfrac{1}{3}\epsilon^{\hat{k}\hat{j}\hat{i}}R_{\hat{k}\hat{l}\hat{n}\hat{m}}\omega^{\hat{i}}\\ &\qquad-\tfrac{1}{3}\epsilon^{\hat{k}\hat{n}\hat{i}}\omega^{\hat{i}}R_{\hat{k}\hat{m}\hat{j}\hat{l}}-\tfrac{1}{3}\epsilon^{\hat{k}\hat{l}\hat{i}}(R_{\hat{k}\hat{j}\hat{m}\hat{n}}+R_{\hat{k}\hat{n}\hat{m}\hat{j}})\omega^{\hat{i}}],\\ &\Gamma^{\hat{0}}{}_{\hat{l}\hat{m},\hat{n}\hat{0}}=P_{lm}[-\tfrac{1}{3}R_{\hat{0}\hat{m}\hat{n}\hat{l};\hat{0}}-\tfrac{1}{3}a^{\hat{k}}R_{\hat{k}\hat{l}\hat{n}\hat{m}}+\tfrac{1}{3}a^{\hat{l}}R_{\hat{0}\hat{m}\hat{0}\hat{n}}-\tfrac{1}{3}a^{\hat{n}}R_{\hat{0}\hat{l}\hat{0}\hat{m}}-\tfrac{1}{3}\epsilon^{\hat{k}\hat{l}\hat{j}}\omega^{\hat{j}}(R_{\hat{0}\hat{k}\hat{m}\hat{n}}+R_{\hat{0}\hat{m}\hat{k}\hat{n}})+\tfrac{1}{3}\epsilon^{\hat{k}\hat{n}\hat{j}}\omega^{\hat{j}}R_{\hat{0}\hat{l}\hat{k}\hat{m}}],\end{aligned} \tag{9}$$

all along $P_0(\tau)$, where the symbol $P_{i_1\cdots i_r}$ indicates that the expression following it is to be summed over all $r!$ permutations of $i_1,\ldots,i_r$.

From the definition of the Riemann tensor,

$$\Gamma^{\hat{\alpha}}{}_{\hat{\mu}\hat{0},\hat{\nu}}=R^{\hat{\alpha}}{}_{\hat{\mu}\hat{\nu}\hat{0}}+\Gamma^{\hat{\alpha}}{}_{\hat{\mu}\hat{\nu},\hat{0}}+(\Gamma^{\hat{\sigma}}{}_{\hat{\mu}\hat{\nu}}\Gamma^{\hat{\alpha}}{}_{\hat{\sigma}\hat{0}}-\Gamma^{\hat{\sigma}}{}_{\hat{\mu}\hat{0}}\Gamma^{\hat{\alpha}}{}_{\hat{\sigma}\hat{\nu}}). \tag{10}$$

Differentiating this expression, we get

$$\Gamma^{\hat{\alpha}}{}_{\hat{\mu}\hat{0},\hat{\nu}\hat{\lambda}}=R^{\alpha}{}_{\hat{\mu}\hat{\nu}\hat{0},\hat{\lambda}}+\Gamma^{\hat{\alpha}}{}_{\hat{\mu}\hat{\nu},\hat{0}\hat{\lambda}}+\Gamma^{\hat{\sigma}}{}_{\hat{\mu}\hat{\nu},\hat{\lambda}}\Gamma^{\hat{\alpha}}{}_{\hat{\sigma}\hat{0}}+\Gamma^{\hat{\sigma}}{}_{\hat{\mu}\hat{\nu}}\Gamma^{\hat{\alpha}}{}_{\hat{\sigma}\hat{0},\hat{\lambda}}-\Gamma^{\hat{\sigma}}{}_{\hat{\mu}\hat{0},\hat{\lambda}}\Gamma^{\hat{\alpha}}{}_{\hat{\sigma}\hat{\nu}}-\Gamma^{\hat{\sigma}}{}_{\hat{\mu}\hat{0}}\Gamma^{\hat{\alpha}}{}_{\hat{\sigma}\hat{\nu},\hat{\lambda}}. \tag{11}$$

To express the partial derivatives of $R^{\alpha}{}_{\beta\gamma\delta}$ in terms of covariant derivatives, we use the formula

$$R^{\alpha}{}_{\beta\gamma\delta,\lambda}=R^{\alpha}{}_{\beta\gamma\delta;\lambda}+R^{\epsilon}{}_{\beta\gamma\delta}\Gamma^{\alpha}{}_{\epsilon\lambda}-R^{\alpha}{}_{\epsilon\gamma\delta}\Gamma^{\epsilon}{}_{\beta\lambda}-R^{\alpha}{}_{\beta\epsilon\delta}\Gamma^{\epsilon}{}_{\gamma\lambda}-R^{\alpha}{}_{\beta\gamma\epsilon}\Gamma^{\epsilon}{}_{\delta\lambda}. \tag{12}$$

Now combining (11) with (5), (7), (9), and (12), we find

$$\Gamma^{\hat{0}}{}_{\hat{0}\hat{0},\hat{m}\hat{n}}=P_{mn}(\tfrac{1}{2}R_{\hat{0}\hat{m}\hat{0}\hat{n};\hat{0}}-2\epsilon^{\hat{i}\hat{j}\hat{m}}a^{\hat{i}}a^{\hat{n}}\omega^{\hat{j}}+\tfrac{1}{3}a^{\hat{k}}R_{\hat{0}\hat{n}\hat{k}\hat{m}}-a^{\hat{m}}b^{\hat{n}}),$$

$$\Gamma^{\hat 0}{}_{\hat 0\hat l,\hat m\hat n} = P_{lmn}\big(-\tfrac{1}{6}R_{\hat 0\hat l\hat 0\hat n;\hat m} - \tfrac{1}{3}a^{\hat l}R_{\hat 0\hat m\hat 0\hat n}\big) + P_{mn}\big(a^{\hat l}a^{\hat m}a^{\hat n} + \tfrac{1}{3}\epsilon^{\hat k\hat n\hat i}\omega^{\hat i}(R_{\hat 0\hat l\hat m\hat k} + R_{\hat 0\hat k\hat m\hat l})\big),$$

$$\Gamma^{\hat j}{}_{\hat 0\hat 0,\hat m\hat n} = P_{mn}\big(-\tfrac{1}{2}R_{\hat 0\hat n\hat 0\hat m;\hat j} + R_{\hat 0\hat j\hat 0\hat m;\hat n} + 2a^{\hat m}R_{\hat 0\hat j\hat 0\hat n} - \tfrac{1}{3}a^{\hat k}R_{\hat j\hat m\hat k\hat n} + \epsilon^{\hat j\hat m\hat i}b^{\hat n}\omega^{\hat i} + 2\epsilon^{\hat k\hat m\hat i}R_{\hat 0\hat n\hat j\hat k}\omega^{\hat i} - 2\epsilon^{\hat j\hat m\hat i}(\boldsymbol{\omega}\times\mathbf{a})^{\hat n}\omega^{\hat i}\big), \tag{13}$$

$$\Gamma^{\hat j}{}_{\hat 0\hat l,\hat m\hat n} = P_{mn}\big(-\tfrac{1}{6}R_{\hat 0\hat l\hat j\hat m;\hat n} - \tfrac{1}{2}R_{\hat 0\hat m\hat j\hat l;\hat n} + \tfrac{1}{6}R_{\hat 0\hat m\hat j\hat n;\hat l} - \tfrac{1}{3}a^{\hat m}(R_{\hat 0\hat n\hat j\hat l} + R_{\hat 0\hat j\hat n\hat l}) + \tfrac{1}{3}a^{\hat l}R_{\hat 0\hat m\hat j\hat n} + \epsilon^{\hat j\hat m\hat i}R_{\hat 0\hat l\hat 0\hat n}\omega^{\hat i}$$
$$+ \tfrac{1}{3}\epsilon^{\hat k\hat m\hat i}(R_{\hat k\hat n\hat j\hat l} + R_{\hat k\hat j\hat n\hat l})\omega^{\hat i} - \epsilon^{\hat l\hat j\hat m}a^{\hat n}a^{\hat i}\omega^{\hat i}\big),$$

all along $P_0(\tau)$.

To calculate $\Gamma^{\hat\mu}{}_{\hat i\hat j;\hat k\hat l}$, we use the geodesic deviation equation

$$\frac{d^2N^\mu}{ds^2} + 2\frac{dN^\sigma}{ds}\Gamma^\mu{}_{\sigma\alpha}U^\alpha + N^\sigma U^\alpha U^\beta R^{\ \mu}{}_{\alpha\sigma\beta} + N^\sigma U^\alpha U^\beta(\Gamma^\mu{}_{\sigma\alpha,\beta} + \Gamma^\tau{}_{\sigma\alpha}\Gamma^\mu{}_{\tau\beta} - \Gamma^\mu{}_{\sigma\tau}\Gamma^\tau{}_{\alpha\beta}) = 0, \tag{14}$$

where $\vec N = \partial/\partial N$ and $\vec U = \partial/\partial s$ of a one-parameter family of geodesics $\mathscr{R}(N,s)$, and where s is an affine parameter along the geodesic $\mathscr{R}(N,s)$ for N fixed. The family of geodesics we want to consider is $P(\tau;\alpha^{\hat i};s)\equiv P(\tau;\vec n;s)$ where $\vec n = \alpha^{\hat i}\vec e_{\hat i}$. The case $\vec N = \partial/\partial\alpha^{\hat i}$ leads to the desired results. In this case $\vec N = \partial/\partial\alpha^{\hat i} = s(\partial/\partial x^{\hat i})$, hence $N^{\hat\mu} = s\delta_{\hat i}{}^{\hat\mu}$. Expanding the geodesic deviation equation in powers of s, we have

$$s(3\Gamma^{\hat\mu}{}_{\hat i\hat j,\hat k} + R^{\ \hat\mu}{}_{\hat j\hat i\hat k})\,|_{P_0(\tau)}\alpha^{\hat j}\alpha^{\hat k} + s^2(2\Gamma^{\hat\mu}{}_{\hat i\hat j,\hat k\hat l} + R^{\ \hat\mu}{}_{\hat j\hat i\hat k,\hat l} + \Gamma^{\hat 0}{}_{\hat i\hat j,\hat l}\Gamma^{\hat\mu}{}_{\hat 0\hat k} - \Gamma^{\hat\mu}{}_{\hat 0\hat i}\Gamma^{\hat 0}{}_{\hat j\hat k,\hat l})\,|_{P_0(\tau)}\alpha^{\hat j}\alpha^{\hat k}\alpha^{\hat l} + O[(s^3)] = 0, \tag{14'}$$

where $O(x)$ means of the order of x.

For (14′) to hold, the coefficient of every power of s must vanish. Hence

$$(3\Gamma^{\hat\mu}{}_{\hat i\hat j,\hat k} + R^{\ \hat\mu}{}_{\hat j\hat i\hat k})\,|_{P_0(\tau)}\alpha^{\hat j}\alpha^{\hat k} = 0, \tag{15}$$

$$(2\Gamma^{\hat\mu}{}_{\hat i\hat j,\hat k\hat l} + R^{\ \hat\mu}{}_{\hat j\hat i\hat k,\hat l} + \Gamma^{\hat 0}{}_{\hat i\hat j,\hat l}\Gamma^{\hat\mu}{}_{\hat 0\hat k} - \Gamma^{\hat\mu}{}_{\hat 0\hat i}\Gamma^{\hat 0}{}_{\hat j\hat k,\hat l})\,|_{P_0(\tau)}\alpha^{\hat j}\alpha^{\hat k}\alpha^{\hat l} = 0. \tag{16}$$

(15) leads to the last one of (7).[9] Since $\alpha^{\hat i}$ can be arbitrary, (16) leads to

$$(4\Gamma^{\hat\mu}{}_{\hat i\hat j,\hat k\hat l} + 4\Gamma^{\hat\mu}{}_{\hat i\hat k,\hat l\hat j} + 4\Gamma^{\hat\mu}{}_{\hat i\hat l,\hat j\hat k} + P_{jkl}R^{\ \hat\mu}{}_{\hat j\hat i\hat k,\hat l} + P_{jkl}\Gamma^{\hat 0}{}_{\hat i\hat j,\hat l}\Gamma^{\hat\mu}{}_{\hat 0\hat k})\,|_{P_0(\tau)} = 0. \tag{17}$$

From the definition of the Riemann tensor and (5), we get

$$R^{\ \hat\mu}{}_{\hat i\hat j\hat k,\hat l\,|P_0(\tau)} = \Gamma^{\hat\mu}{}_{\hat i\hat k,\hat j\hat l} - \Gamma^{\hat\mu}{}_{\hat i\hat j,\hat k\hat l} + \Gamma^{\hat\mu}{}_{\hat 0\hat j}\Gamma^{\hat 0}{}_{\hat i\hat k,\hat l} - \Gamma^{\hat\mu}{}_{\hat 0\hat k}\Gamma^{\hat 0}{}_{\hat i\hat j,\hat l},$$
$$R^{\ \hat\mu}{}_{\hat i\hat j\hat l,\hat k\,|P_0(\tau)} = \Gamma^{\hat\mu}{}_{\hat i\hat l,\hat j\hat k} - \Gamma^{\hat\mu}{}_{\hat i\hat j,\hat l\hat k} + \Gamma^{\hat\mu}{}_{\hat 0\hat j}\Gamma^{\hat 0}{}_{\hat i\hat l,\hat k} - \Gamma^{\hat\mu}{}_{\hat 0\hat l}\Gamma^{\hat 0}{}_{\hat i\hat j,\hat k}. \tag{18}$$

Combining (17), (18), and (12), we obtain

$$\Gamma^{\hat\mu}{}_{\hat i\hat j,\hat k\hat l\,|P_0(\tau)} = \tfrac{1}{24}P_{ij}P_{kl}(5R^{\ \hat\mu}{}_{\hat i\hat k\hat j;\hat l} - R^{\ \hat\mu}{}_{\hat k\hat i\hat l;\hat j}) - \tfrac{1}{3}P_{ij}P_{kl}\Gamma^{\hat\mu}{}_{\hat 0\hat k}R^{\hat 0}{}_{\hat i\hat l\hat j}. \tag{19}$$

From the compatibility condition

$$g_{\mu\nu;\alpha} = g_{\mu\nu,\alpha} - g_{\mu\sigma}\Gamma^\sigma{}_{\nu\alpha} - g_{\sigma\nu}\Gamma^\sigma{}_{\mu\alpha} = 0, \tag{20}$$

we find by differentiations that

$$g_{\hat\mu\hat\nu,\hat\alpha\hat\beta\hat\lambda\,|P_0(\tau)} = g_{\hat\mu\hat\sigma}\Gamma^{\hat\sigma}{}_{\hat\nu\hat\alpha,\hat\beta\hat\lambda} + g_{\hat\sigma\hat\nu}\Gamma^{\hat\sigma}{}_{\hat\mu\hat\alpha,\hat\beta\hat\lambda} + g_{\hat\mu\hat\sigma,\hat\beta}\Gamma^{\hat\sigma}{}_{\hat\nu\hat\alpha,\hat\lambda} + g_{\hat\sigma\hat\nu,\hat\beta}\Gamma^{\hat\sigma}{}_{\hat\mu\hat\alpha,\hat\lambda} + g_{\hat\mu\hat\sigma,\hat\lambda}\Gamma^{\hat\sigma}{}_{\hat\nu\hat\alpha,\hat\beta} + g_{\hat\sigma\hat\nu,\hat\lambda}\Gamma^{\hat\sigma}{}_{\hat\mu\hat\alpha,\hat\beta} + g_{\hat\mu\hat\sigma,\hat\beta\hat\lambda}\Gamma^{\hat\sigma}{}_{\hat\nu\hat\alpha}$$
$$+ g_{\hat\sigma\hat\nu,\hat\beta\hat\lambda}\Gamma^{\hat\sigma}{}_{\hat\mu\hat\alpha}. \tag{21}$$

Substituting (4)–(9), (13), and (19) into Eq. (21), we obtain the third-order derivatives of the metric along $P_0(\tau)$ as follows:

$$\begin{aligned}
&g_{\hat\alpha\hat\beta,\hat 0\hat 0\hat 0} = 0,\quad g_{\hat j\hat k,\hat l\hat 0\hat 0} = 0,\quad g_{\hat 0\hat 0,\hat j\hat 0\hat 0} = -2c^{\hat j} + 4\epsilon^{\hat j\hat i\hat k}\omega^{\hat i}b^{\hat k} + 2a^{\hat j}\mathbf{a}^2 - 2\omega^{\hat j}\mathbf{a}\cdot\boldsymbol{\omega} + 2a^{\hat j}\boldsymbol{\omega}^2 - 2\epsilon^{\hat j\hat k\hat l}a^{\hat k}\eta^{\hat l},\\
&g_{\hat 0\hat j,\hat k\hat 0\hat 0} = -\epsilon^{\hat j\hat k\hat l}\zeta^{\hat l} + \epsilon^{\hat j\hat k\hat l}a^{\hat l}\mathbf{a}\cdot\boldsymbol{\omega} + \omega^{\hat j}\eta^{\hat k} - \omega^{\hat k}\eta^{\hat j},\\
&g_{\hat 0\hat 0,\hat j\hat k\hat 0} = -2P_{jk}\big(\tfrac{1}{2}R_{\hat 0\hat j\hat 0\hat k;\hat 0} + a^{\hat l}R_{\hat 0\hat k\hat l\hat j} - \epsilon^{\hat l\hat j\hat i}R_{\hat 0\hat l\hat 0\hat k}\omega^{\hat i} + a^{\hat j}b^{\hat k} - \epsilon^{\hat j\hat m\hat n}\omega^{\hat m}a^{\hat n}a^{\hat k} - \delta^{\hat j}{}_{\hat k}\omega^{\hat l}\eta^{\hat l} + \eta^{\hat j}\omega^{\hat k}\big),\\
&g_{\hat 0\hat i,\hat j\hat k\hat 0} = -\tfrac{2}{3}P_{jk}\big(R_{\hat 0\hat j\hat i\hat k;\hat 0} + a^{\hat l}R_{\hat l\hat j\hat i\hat k} + a^{\hat i}R_{\hat 0\hat k\hat 0\hat j} - a^{\hat j}R_{\hat 0\hat k\hat 0\hat i} + \epsilon^{\hat l\hat j\hat m}R_{\hat 0\hat l\hat k\hat i}\omega^{\hat m} - \epsilon^{\hat l\hat i\hat m}R_{\hat 0\hat k\hat l\hat j}\omega^{\hat m} - \epsilon^{\hat i\hat k\hat n}R_{\hat 0\hat j\hat i\hat l}\omega^{\hat n}\big),\\
&g_{\hat l\hat m,\hat i\hat j\hat 0} = -\tfrac{1}{6}P_{lm}P_{ij}\big(R_{\hat i\hat l\hat j\hat m;\hat 0} + 2a^{\hat i}R_{\hat 0\hat l\hat j\hat m} + 2a^{\hat l}R_{\hat 0\hat i\hat m\hat j} - 2\epsilon^{\hat k\hat l\hat n}R_{\hat k\hat i\hat m\hat j}\omega^{\hat n}\big),\\
&g_{\hat 0\hat 0,\hat l\hat m\hat n} = -\tfrac{1}{3}P_{lmn}\big(R_{\hat 0\hat l\hat 0\hat n;\hat m} + 4a^{\hat m}R_{\hat 0\hat l\hat 0\hat n} - 4\epsilon^{\hat k\hat l\hat i}R_{\hat 0\hat m\hat k\hat n}\omega^{\hat i}\big),\\
&g_{\hat 0\hat j,\hat l\hat m\hat n} = -P_{lmn}\big(\tfrac{1}{4}R_{\hat 0\hat l\hat j\hat m;\hat n} + \tfrac{1}{3}a^{\hat l}R_{\hat 0\hat m\hat j\hat n} - \tfrac{1}{3}\epsilon^{\hat k\hat l\hat i}R_{\hat k\hat m\hat j\hat n}\omega^{\hat i}\big),\quad g_{\hat i\hat j,\hat l\hat m\hat n} = -\tfrac{1}{6}P_{lmn}R_{\hat i\hat l\hat j\hat m;\hat n},
\end{aligned} \tag{22}$$

all along $P_0(\tau)$.

From (22), we find the third-order expansion of the metric at the point $P(x^{\hat 0},x^{\hat j})$ as

$$ds^2{}_{|P(x^{\hat 0},x^{\hat j})} = -(dx^{\hat 0})^2\big[1 + 2a^{\hat j}x^{\hat j} + (a^{\hat l}x^{\hat l})^2 + (\omega^{\hat l}x^{\hat l})^2 - (\boldsymbol{\omega})^2x^{\hat l}x^{\hat l} + R_{\hat 0\hat l\hat 0\hat m}x^{\hat l}x^{\hat m} + \tfrac{1}{3}R_{\hat 0\hat l\hat 0\hat m;\hat n}x^{\hat l}x^{\hat m}x^{\hat n} + \tfrac{4}{3}\mathbf{a}\cdot\mathbf{x}R_{\hat 0\hat l\hat 0\hat m}x^{\hat l}x^{\hat m}$$
$$+ \tfrac{4}{3}(\boldsymbol{\omega}\times\mathbf{x})^{\hat k}R_{\hat 0\hat l\hat k\hat m}x^{\hat l}x^{\hat m}\big] + 2dx^{\hat 0}dx^{\hat i}\big[\epsilon^{\hat i\hat j\hat k}\omega^{\hat j}x^{\hat k} - \tfrac{2}{3}R_{\hat 0\hat l\hat i\hat m}x^{\hat l}x^{\hat m} - \tfrac{1}{4}R_{\hat 0\hat l\hat i\hat m;\hat n}x^{\hat l}x^{\hat m}x^{\hat n} - \tfrac{1}{3}\mathbf{a}\cdot\mathbf{x}R_{\hat 0\hat l\hat i\hat m}x^{\hat l}x^{\hat m}$$

$$-\tfrac{1}{3}(\omega\times\mathbf{x})^{\hat{k}}R_{\hat{i}\hat{l}\hat{k}\hat{m}}x^{\hat{l}}x^{\hat{m}}] + dx^{\hat{i}}dx^{\hat{j}}(\delta_{\hat{i}\hat{j}} - \tfrac{1}{3}R_{\hat{i}\hat{l}\hat{j}\hat{m}}x^{\hat{l}}x^{\hat{m}} - \tfrac{1}{6}R_{\hat{i}\hat{l}\hat{j}\hat{m};\hat{n}}x^{\hat{l}}x^{\hat{m}}x^{\hat{n}}) + O\,[dx^{\hat{\alpha}}dx^{\hat{\beta}}x^{\hat{i}}x^{\hat{j}}x^{\hat{k}}x^{\hat{l}}], \tag{23}$$

where $a^{\hat{j}}$, $\omega^{\hat{l}}$, $R_{\hat{\alpha}\hat{\beta}\hat{\mu}\hat{\nu}}$, and $R_{\hat{\alpha}\hat{\beta}\hat{\mu}\hat{\nu};\hat{\lambda}}$ are all evaluated on the world line at time $x^{\hat{0}}$.

There are three types of third-order terms appearing in the metric expansion (23): (i) those involving the covariant derivatives of the curvature tensor; (ii) those involving both $a_{\hat{j}}$ and the curvature tensor; (iii) those involving both $\omega^{\hat{l}}$ and the curvature tensor. The first type of terms are purely gravitational in origin and are due to the inhomogeneities in the curvature properties of spacetime; they are absent for locally symmetric spacetimes. The second and third type of terms are coupled inertial and gravitational terms and are our main interest in this paper.

Note that since the second-order equal-time expansion of the metric is exact in flat spacetimes,[3] there are no third or higher order purely inertial terms (i.e., those not involving the curvature tensors) in the metric. Note also that the space metric depends only upon the curvature properties, as it must because of the definitions of the local coordinates.

III. SECOND-ORDER EXPANSION OF THE GEODESIC EQUATIONS AND COUPLED INERTIAL-GRAVITATIONAL EFFECTS

To calculate the coordinate acceleration of a freely falling body, we use the geodesic equation in the following form:

$$\frac{d^2x^{\hat{i}}}{dx^{\hat{0}2}} + \left(\Gamma^{\hat{i}}{}_{\hat{\mu}\hat{\nu}} - \Gamma^{\hat{0}}{}_{\hat{\mu}\hat{\nu}}\frac{dx^{\hat{i}}}{dx^{\hat{0}}}\right)\frac{dx^{\hat{\mu}}}{dx^{\hat{0}}}\frac{dx^{\hat{\nu}}}{dx^{\hat{0}}} = 0 \tag{24}$$

and substitute into it the second-order expansion of the Γ's. Defining $w^{\hat{i}}\equiv dx^{\hat{i}}/dx^{\hat{0}}$, the velocity measured by the accelerated rotating observer, the resulting coordinate acceleration is

$$\begin{aligned}
\frac{d^2x^{\hat{i}}}{dx^{\hat{0}2}} = {} & -(\Gamma^{\hat{i}}{}_{\hat{0}\hat{0}|P_0(x^{\hat{0}})} + \Gamma^{\hat{i}}{}_{\hat{0}\hat{0},\hat{l}|P(x^{\hat{0}})}x^{\hat{l}} + \tfrac{1}{2}\Gamma^{\hat{i}}{}_{\hat{0}\hat{0},\hat{l}\hat{m}|P_0(x^{\hat{0}})}x^{\hat{l}}x^{\hat{m}}) - 2(\Gamma^{\hat{i}}{}_{\hat{j}\hat{0}|P_0(x^{\hat{0}})} + \Gamma^{\hat{i}}{}_{\hat{j}\hat{0},\hat{l}|P_0(x^{\hat{0}})}x^{\hat{l}} + \tfrac{1}{2}\Gamma^{\hat{i}}{}_{\hat{j}\hat{0},\hat{l}\hat{m}|P_0(x^{\hat{0}})}x^{\hat{l}}x^{\hat{m}})w^{j} \\
& -(\Gamma^{\hat{i}}{}_{\hat{j}\hat{k}|P_0(x^{\hat{0}})} + \Gamma^{\hat{i}}{}_{\hat{j}\hat{k},\hat{l}|P_0(x^{\hat{0}})}x^{\hat{l}} + \tfrac{1}{2}\Gamma^{\hat{i}}{}_{\hat{j}\hat{k},\hat{l}\hat{m}|P_0(x^{\hat{0}})}x^{\hat{l}}x^{\hat{m}})w^{j}w^{k} + (\Gamma^{\hat{0}}{}_{\hat{0}\hat{0}|P_0(x^{\hat{0}})} + \Gamma^{\hat{0}}{}_{\hat{0}\hat{0},\hat{l}|P_0(x^{\hat{0}})}x^{\hat{l}} + \tfrac{1}{2}\Gamma^{\hat{0}}{}_{\hat{0}\hat{0},\hat{l}\hat{m}|P_0(x^{\hat{0}})}x^{\hat{l}}x^{\hat{m}})w^{i} \\
& + 2(\Gamma^{\hat{0}}{}_{\hat{j}\hat{0}|P_0(x^{\hat{0}})} + \Gamma^{\hat{0}}{}_{\hat{j}\hat{0},\hat{l}|P_0(x^{\hat{0}})})x^{\hat{l}} + \tfrac{1}{2}\Gamma^{\hat{0}}{}_{\hat{j}\hat{0},\hat{l}\hat{m}|P_0(x^{\hat{0}})}x^{\hat{l}}x^{\hat{m}})w^{i}w^{j} + (\Gamma^{\hat{0}}{}_{\hat{j}\hat{k}|P_0(x^{\hat{0}})} + \Gamma^{\hat{0}}{}_{\hat{j}\hat{k},\hat{l}|P_0(x^{\hat{0}})}x^{\hat{l}} \\
& + \tfrac{1}{2}\Gamma^{\hat{0}}{}_{\hat{j}\hat{k},\hat{l}\hat{m}|P_0(x^{\hat{0}})}x^{l}x^{m})w^{i}w^{j}w^{k} \\
= {} & -(1+\mathbf{a}\cdot\mathbf{x})a^{\hat{i}} - [\boldsymbol{\omega}\times(\boldsymbol{\omega}\times\mathbf{x})]^{\hat{i}} - (\boldsymbol{\eta}\times\mathbf{x})^{\hat{i}} - 2(\boldsymbol{\omega}\times\mathbf{w})^{i} + 2(\mathbf{a}\cdot\mathbf{w})(\boldsymbol{\omega}\times\mathbf{x})^{\hat{i}} + w^{i}[\mathbf{b}\cdot\mathbf{x} + 2\mathbf{a}\cdot(\boldsymbol{\omega}\times\mathbf{x}) + 2\mathbf{a}\cdot\mathbf{w}(1-\mathbf{a}\cdot\mathbf{x})] \\
& + (\mathbf{b}\cdot\mathbf{x})(\boldsymbol{\omega}\times\mathbf{x})^{\hat{i}} + 2\mathbf{a}\cdot(\boldsymbol{\omega}\times\mathbf{x})(\boldsymbol{\omega}\times\mathbf{x})^{\hat{i}} - 2\mathbf{a}\cdot\mathbf{w}\mathbf{a}\cdot\mathbf{x}(\boldsymbol{\omega}\times\mathbf{x})^{\hat{i}} - w^{i}\mathbf{a}\cdot\mathbf{x}\mathbf{b}\cdot\mathbf{x} + 2w^{i}\mathbf{a}\cdot\mathbf{w}(\mathbf{a}\cdot\mathbf{x})^2 - 2w^{i}\mathbf{a}\cdot\mathbf{x}\mathbf{a}\cdot(\boldsymbol{\omega}\times\mathbf{x}) \\
& - x^{\hat{l}}R_{\hat{0}\hat{i}\hat{0}\hat{l}} - 2x^{\hat{l}}w^{j}R_{\hat{i}\hat{j}\hat{l}\hat{0}} + \tfrac{2}{3}x^{\hat{l}}w^{j}w^{k}R_{\hat{i}\hat{j}\hat{k}\hat{l}} + 2x^{\hat{l}}w^{i}w^{j}R_{\hat{0}\hat{j}\hat{0}\hat{l}} + \tfrac{2}{3}x^{\hat{l}}w^{i}w^{j}w^{k}R_{\hat{0}\hat{j}\hat{k}\hat{l}} \\
& + \tfrac{1}{3}a^{\hat{k}}R_{\hat{i}\hat{l}\hat{k}\hat{m}}x^{\hat{l}}x^{\hat{m}} + 2x^{\hat{l}}(\boldsymbol{\omega}\times\mathbf{x})^{\hat{k}}R_{\hat{0}\hat{l}\hat{i}\hat{k}} - 2x^{\hat{l}}\mathbf{a}\cdot\mathbf{x}R_{\hat{0}\hat{i}\hat{0}\hat{l}} + \tfrac{2}{3}\mathbf{a}\cdot\mathbf{x}x^{\hat{l}}w^{j}(R_{\hat{0}\hat{l}\hat{i}\hat{j}} + R_{\hat{0}\hat{i}\hat{l}\hat{j}}) \\
& - \tfrac{2}{3}\mathbf{a}\cdot\mathbf{w}R_{\hat{0}\hat{l}\hat{i}\hat{m}}x^{\hat{l}}x^{\hat{m}} + 2x^{\hat{l}}w^{j}(\boldsymbol{\omega}\times\mathbf{x})^{\hat{i}}R_{\hat{0}\hat{j}\hat{0}\hat{l}} + \tfrac{2}{3}x^{\hat{l}}w^{j}(\boldsymbol{\omega}\times\mathbf{x})^{\hat{k}}(R_{\hat{k}\hat{l}\hat{i}\hat{j}} + R_{\hat{k}\hat{i}\hat{l}\hat{j}}) + \tfrac{2}{3}x^{\hat{m}}w^{j}w^{k}(\omega\times x)^{\hat{i}}R_{\hat{0}\hat{k}\hat{j}\hat{m}} + \tfrac{1}{3}a^{\hat{k}}R_{\hat{0}\hat{m}\hat{k}\hat{l}}w^{i}x^{\hat{l}}x^{\hat{m}} \\
& - \tfrac{4}{3}x^{\hat{m}}w^{i}w^{j}\mathbf{a}\cdot\mathbf{x}R_{\hat{0}\hat{j}\hat{0}\hat{m}} - \tfrac{2}{3}x^{\hat{l}}x^{\hat{m}}w^{i}\mathbf{a}\cdot\mathbf{w}R_{\hat{0}\hat{l}\hat{0}\hat{m}} + \tfrac{2}{3}x^{\hat{l}}w^{i}w^{j}(\boldsymbol{\omega}\times\mathbf{x})^{\hat{k}}(R_{\hat{0}\hat{j}\hat{k}\hat{l}} + R_{\hat{0}\hat{k}\hat{j}\hat{l}}) - \tfrac{2}{3}w^{i}w^{j}w^{k}\mathbf{a}\cdot\mathbf{x}R_{\hat{0}\hat{k}\hat{j}\hat{m}}x^{\hat{m}} \\
& - \tfrac{1}{2}(R_{\hat{i}\hat{l}\hat{m}\hat{0};\hat{0}} + R_{\hat{i}\hat{0}\hat{l}\hat{0};\hat{m}})x^{\hat{l}}x^{\hat{m}} + \tfrac{1}{3}(R_{\hat{i}\hat{l}\hat{j}\hat{m},\hat{0}})x^{\hat{l}}x^{\hat{m}}w^{j} - R_{\hat{i}\hat{j}\hat{l}\hat{0};\hat{m}}x^{\hat{l}}x^{\hat{m}}w^{j} + \tfrac{1}{12}(5R_{\hat{i}\hat{k}\hat{j}\hat{l};\hat{m}} + R_{\hat{i}\hat{l}\hat{j}\hat{m};\hat{k}})x^{\hat{l}}x^{\hat{m}}w^{j}w^{k} + \tfrac{1}{2}R_{\hat{0}\hat{l}\hat{0}\hat{m};\hat{0}}x^{\hat{l}}x^{\hat{m}}w^{i} \\
& + R_{\hat{0}\hat{j}\hat{0}\hat{l};\hat{m}}x^{\hat{l}}x^{\hat{m}}w^{i}w^{j} + \tfrac{1}{3}R_{\hat{0}\hat{l}\hat{j}\hat{m};\hat{0}}x^{\hat{l}}x^{\hat{m}}w^{i}w^{j} + \tfrac{1}{12}(5R_{\hat{0}\hat{k}\hat{j}\hat{l};\hat{m}} + R_{\hat{0}\hat{l}\hat{j}\hat{m};\hat{k}})x^{\hat{l}}x^{\hat{m}}w^{i}w^{j}w^{k} + O\,[(x^{\hat{i}})^3].
\end{aligned} \tag{25}$$

To express $d^2x^{\hat{i}}/dx^{\hat{0}2}$ in terms of the velocity $\mathbf{v}$ observed in the local coordinates $(x^{0'},x^{i'})$ of an unaccelerated nonrotating observer, we will find the relation between $\mathbf{w}$ and $\mathbf{v}$. Note that both $\mathbf{w}$ and $\mathbf{v}$ are functions of $x^{\hat{0}}$. In the proper frame of an unaccelerated nonrotating observer, the components of $\mathbf{v}$, $v^{i'}$, are constant to first order in $x^{\mu'}$, i.e.,

$$v^{i'} = \text{const} + O\,[(x^{\mu'})^2]. \tag{26}$$

To calculate the components of $\mathbf{v}$ in the accelerated rotating frame, we use the Lorentz transformation to get

$$\frac{dv^{\hat{i}}}{dx^{\hat{0}}} = -a^{\hat{i}} + v^{\hat{i}}\mathbf{a}\cdot\mathbf{v} - (\boldsymbol{\omega}\times\mathbf{v})^{\hat{i}} + O\,[(x^{\hat{k}})], \tag{27}$$

where $O\,[(x^{\hat{k}})]$ does not contain purely inertial terms. Using (27), Eq. (25) to zeroth-order can be written as

$$\frac{d^2x^{\hat{i}}}{dx^{\hat{0}2}} = -a^{\hat{i}} + 2v^{\hat{i}}\mathbf{a}\cdot\mathbf{v} - 2(\boldsymbol{\omega}\times\mathbf{v})^{\hat{i}} + O\,[(x^{\hat{k}})] = \frac{d}{dx^{\hat{0}}}\{v^{\hat{i}} + v^{\hat{i}}\mathbf{a}\cdot\mathbf{x} - (\boldsymbol{\omega}\times\mathbf{x})^{\hat{i}} + O\,[(x^{\hat{k}})^2]\}. \tag{28}$$

Integrating (28), we obtain

$$\mathbf{w} = \mathbf{v}(1+\mathbf{a}\cdot\mathbf{x}) - \boldsymbol{\omega}\times\mathbf{x} + O\,[(x^{\hat{i}})^2] \tag{29}$$

which is (21) derived in Ref. 5. (29) without $O\,[(x^{\hat{i}})^2]$ is exact in special relativity. Since to second-order in (29) the inertial terms and curvature terms are additive and both $\mathbf{v}$ and $\mathbf{w}$ involve the same curvature terms, (29) is valid to second-order, i.e.,

$$\mathbf{w} = \mathbf{v}(1+\mathbf{a}\cdot\mathbf{x}) - \boldsymbol{\omega}\times\mathbf{x} + O\,[(x^{\hat{i}})^3]. \tag{30}$$

Explicitly, from Ref. 5,

$$\frac{d^2x^{\hat i}}{dx^{\hat 0^2}} = -(1+\mathbf{a}\cdot\mathbf{x})a^{\hat i} + 2v^i\mathbf{a}\cdot\mathbf{v}(1+\mathbf{a}\cdot\mathbf{x}) + v^i\mathbf{b}\cdot\mathbf{x} - 2(\boldsymbol{\omega}\times\mathbf{v})^{\hat i}(1+\mathbf{a}\cdot\mathbf{x}) + [\boldsymbol{\omega}\times(\boldsymbol{\omega}\times\mathbf{x})]^{\hat i} - (\boldsymbol{\eta}\times\mathbf{x})^{\hat i}$$
$$- R_{\hat 0\tilde{\hat i}\hat 0\hat l}x^{\hat l} - 2R_{\tilde{\hat i}\hat j\hat l\tilde{\hat 0}}x^{\hat l}v^j + \tfrac{2}{3}R_{\tilde{\hat i}\hat j\hat k\hat l}x^{\hat l}v^jv^k + 2R_{\hat 0\hat j\tilde{\hat 0}\hat l}x^{\hat l}v^iv^j + \tfrac{2}{3}R_{\hat 0\hat j\hat k\hat l}x^{\hat l}v^iv^jv^k + O[(x^{\hat k})^2], \tag{31}$$

where $v^i \equiv v^{i'} = v^{\hat i}$ at $P(x^{\hat 0}, x^{\hat i})$ by the coincidence of the two coordinate systems we chose. Hence

$$\frac{dv^{i'}}{dx^{0'}} = \frac{d^2x^{i'}}{dx^{0'^2}} = -R_{0'i'0'l'}x^{l'} - 2R_{i'j'l'0'}x^{l'}v^j + \tfrac{2}{3}R_{0'j'k'l'}x^{l'}v^iv^jv^k + 2R_{0'j'0'l'}x^{l'}v^iv^j + \tfrac{2}{3}R_{i'j'k'l'}x^{l'}v^jv^k + O[(x^{k'})^2]. \tag{32}$$

Combining (27), (32), and

$$\frac{dx^{\hat 0}}{dx^{0'}} = 1 - \mathbf{a}\cdot\mathbf{x} + O[(x^{\hat i})^2], \tag{33}$$

we obtain

$$\frac{dv^{\hat i}}{dx^{\hat 0}} = -a^{\hat i} + v^i\mathbf{a}\cdot\mathbf{v} - (\boldsymbol{\omega}\times\mathbf{v})^{\hat i} - R_{\hat 0\tilde{\hat i}\hat 0\hat l}x^{\hat l} - 2x^{\hat l}R_{\tilde{\hat i}\hat j\hat l\hat 0}v^j + \tfrac{2}{3}R_{\hat 0\hat j\hat k\hat l}x^{\hat l}v^iv^jv^k + 2R_{\hat 0\hat j\tilde{\hat 0}\hat l}x^{\hat l}v^iv^j + \tfrac{2}{3}R_{\tilde{\hat i}\hat j\hat k\hat l}x^{\hat l}v^jv^k + O[(x^{\hat k})^2]. \tag{34}$$

Compare (34) with (31), we obtain (30) readily.

Substituting (30) into (25), we find

$$\frac{d^2x^{\hat i}}{dx^{\hat 0^2}} = -(1+\mathbf{a}\cdot\mathbf{x})a^{\hat i} + 2v^i\mathbf{a}\cdot\mathbf{v}(1+\mathbf{a}\cdot\mathbf{x}) + v^i\mathbf{b}\cdot\mathbf{x} - 2(\boldsymbol{\omega}\times\mathbf{v})^{\hat i}(1+\mathbf{a}\cdot\mathbf{x}) + [\boldsymbol{\omega}\times(\boldsymbol{\omega}\times\mathbf{x})]^{\hat i} - (\boldsymbol{\eta}\times\mathbf{x})^{\hat i} - x^{\hat l}R_{\hat 0\tilde{\hat i}\hat 0\hat l} - 2x^{\hat l}\mathbf{a}\cdot\mathbf{x}R_{\hat 0\tilde{\hat i}\hat 0\hat l}$$
$$+ 2x^{\hat l}R_{\hat 0\hat j\tilde{\hat 0}\hat l}v^iv^j + \tfrac{8}{3}\mathbf{a}\cdot\mathbf{x}x^{\hat l}R_{\hat 0\hat j\tilde{\hat 0}\hat l}v^iv^j - \tfrac{2}{3}v^i\mathbf{a}\cdot\mathbf{v}R_{\hat 0\tilde{\hat l}\hat 0\hat m}x^{\hat l}x^{\hat m} + 2x^{\hat l}R_{\tilde{\hat i}\hat j\hat 0\hat l}v^j + \tfrac{8}{3}x^{\hat l}\mathbf{a}\cdot\mathbf{x}R_{\tilde{\hat i}\hat j\hat 0\hat l}v^j - \tfrac{2}{3}x^{\hat l}\mathbf{a}\cdot\mathbf{x}R_{\tilde{\hat i}\hat 0\hat l\hat j}v^j - \tfrac{2}{3}\mathbf{a}\cdot\mathbf{v}R_{\hat 0\hat l\tilde{\hat i}\hat m}x^{\hat l}x^{\hat m}$$
$$+ \tfrac{2}{3}x^{\hat l}R_{\hat 0\hat j\hat k\hat l}v^iv^jv^k + \tfrac{4}{3}x^{\hat l}\mathbf{a}\cdot\mathbf{x}R_{\hat 0\hat j\hat k\hat l}v^iv^jv^k + \tfrac{1}{3}a^{\hat k}R_{\hat 0\hat m\hat k\hat l}x^{\hat l}x^{\hat m}v^{\hat i} + \tfrac{2}{3}x^{\hat l}R_{\tilde{\hat i}\hat j\hat k\hat l}v^jv^k + \tfrac{4}{3}\mathbf{a}\cdot\mathbf{x}x^{\hat l}R_{\tilde{\hat i}\hat j\hat k\hat l}v^jv^k + \tfrac{1}{3}a^{\hat k}R_{\hat i\hat l\hat k\hat m}x^{\hat l}x^{\hat m}$$
$$- \tfrac{1}{2}(R_{\tilde{\hat i}\hat 0\hat l\hat 0;\hat m} + R_{\hat i\hat l\hat m\hat 0;\hat 0})x^{\hat l}x^{\hat m} + \tfrac{1}{3}R_{\hat i\hat l\hat j\hat m;\hat 0}x^{\hat l}x^{\hat m}v^j - R_{\tilde{\hat i}\hat j\hat l\hat 0;\hat m}x^{\hat l}x^{\hat m}v^j + \tfrac{1}{12}(5R_{\hat i\hat k\hat j\hat l;\hat m} + R_{\hat i\hat l\hat j\hat m;\hat k})x^{\hat l}x^{\hat m}v^jv^k$$
$$+ \tfrac{1}{2}R_{\hat 0\hat l\hat 0\hat m;\hat 0}x^{\hat l}x^{\hat m}v^i + R_{\hat 0\hat j\tilde{\hat 0}\hat l;\hat m}x^{\hat l}x^{\hat m}v^iv^j + \tfrac{1}{3}R_{\hat 0\hat l\hat j\hat m;\hat 0}x^{\hat l}x^{\hat m}v^iv^j + \tfrac{1}{12}(5R_{\hat 0\hat k\hat j\hat l;\hat m} + R_{\hat 0\hat l\hat j\hat m;\hat k})x^{\hat l}x^{\hat m}v^iv^jv^k + O((x^{\hat k})^3). \tag{35}$$

Following a similar procedure as from (26) to (34) and integrating Eq. (35), we obtain

$$\frac{dv^{\hat i}}{dx^{\hat 0}} = -a^{\hat i} + v^i\mathbf{a}\cdot\mathbf{v} - (\boldsymbol{\omega}\times\mathbf{v})^{\hat i} + \tfrac{1}{3}v^ia^{\hat k}R_{\hat 0\hat m\hat k\hat l}x^{\hat l}x^{\hat m} + \tfrac{1}{3}a^{\hat k}R_{\hat i\hat l\hat k\hat m}x^{\hat l}x^{\hat m} + O[(x^{\hat k})^3], \tag{36}$$

$$w^{\hat i} = v^i(1+\mathbf{a}\cdot\mathbf{x}) - (\boldsymbol{\omega}\times\mathbf{x})^{\hat i} - \tfrac{2}{3}v^i\mathbf{a}\cdot\mathbf{x}R_{\hat 0\hat l\hat 0\hat m}x^{\hat l}x^{\hat m} - \tfrac{2}{3}\mathbf{a}\cdot\mathbf{x}R_{\hat 0\hat l\tilde{\hat i}\hat m}x^{\hat l}x^{\hat m} + O[(x^{\hat i})^4]. \tag{37}$$

Note that in (35) there are no coupled angular velocity and curvature terms. This fact can be seen clearly by the following arguments. Consider first a free falling observer. In the local coordinates $x^{0'}$, $x^{i'}$ of this observer, (35) becomes

$$\frac{d^2x^{i'}}{dx^{0'^2}} = -x^{l'}R_{0'i'0'l'} + 2x^{l'}R_{0'j'0'l'}v^iv^j + 2x^{l'}R_{i'j'0'l'}v^j + \tfrac{2}{3}x^{l'}R_{0'j'k'l'}v^iv^jv^k$$
$$+ \tfrac{2}{3}x^{l'}R_{i'j'k'l'}v^jv^k - \tfrac{1}{2}(R_{i'0'l'0';m'} + R_{i'l'm'0';0'})x^{l'}x^{m'} + \tfrac{1}{3}R_{i'l'j',m';0'}x^{l'}x^{m'}v^j$$
$$- R_{i'j'l'0';m'}x^{l'}x^{m'}v^j + \tfrac{1}{12}(5R_{i'k'j'l';m'} + R_{i'l'j'm';k'})x^{l'}x^{m'}v^jv^k + \tfrac{1}{2}R_{0'l'0'm';0'}x^{l'}x^{m'}v^i$$
$$+ R_{0'j'0'l';m'}v^iv^jx^{l'}x^{m'} + \tfrac{1}{3}R_{0'l'j'm';0'}x^{l'}x^{m'}v^iv^j + \tfrac{1}{12}(5R_{0'k'j'l';m'} + R_{0'l'j'm';k'})x^{l'}x^{m'}v^iv^jv^k + O(x^{\hat k^3}). \tag{38}$$

Note that in (35) there are no coupled angular velocity and curvature terms. This fact can be seen clearly by the following arguments. Consider first a free falling observer. In the local coordinates $x^{0'}$, $x^{i'}$ of this observer, (35) becomes

$$x^{0'} = x^{0''} = \tau \tag{39}$$

and that the spatial coordinates are related by Euclidean rotations

$$x^{i''} = e^{i''}{}_{j'}(x^{0'})x^{j'}, \tag{40}$$

where $x^{0''}, x^{i''}$ are local coordinates of the freely falling rotating observer and $e^{i''}{}_{j'}e^{k''}{}_{j'} = \delta^{i''k''}$. Differentiating (40) twice,

$$\frac{d^2x^{i''}}{dx^{0''^2}} = \frac{d^2e^{j''}{}_{j'}}{dx^{0''^2}}x^{j'} + 2\frac{de^{i''}{}_{j'}}{dx^{0''}}\frac{dx^{j'}}{dx^{0''}} + e^{i''}{}_{j'}\frac{d^2x^{j'}}{dx^{0''^2}}. \tag{41}$$

Suppose at the time we are considering, the rotating tetrad and the nonrotating tetrad coincides. Then, at this time,

$$x^{i''} = x^{i'}, \quad e^{i''}{}_{j'} = \delta_{ij}, \quad \frac{de^{i''}{}_{j'}}{dx^{0''}} = -\epsilon^{ikj}\omega^{k''}, \quad \frac{d^2e^{i''}{}_{j'}}{dx^{0''^2}} = \epsilon^{ijk}\eta^{k''} + \omega^{i''}\omega^{j''} - \delta^i{}_j\omega^2. \tag{42}$$

Substituting (42) and (38) into (41), we have

$$
\begin{aligned}
\frac{d^2x^{i''}}{dx^{0''2}} &= -2(\omega\times\mathbf{v})^{i''} + [\omega\times(\omega\times\mathbf{x})]^{i''} - (\eta\times\mathbf{x})^{i''} + \frac{d^2x^{i'}}{dx^{0'2}} \\
&= -2(\omega\times v)^{i''} + [\omega\times(\omega\times\mathbf{x})]^{i} - (\eta\times\mathbf{x})^{i''} - x^{l''}R_{0''i''0''l''} + 2x^{l''}R_{0''j''0''l''}v^iv^j + 2x^{l''}R_{i''j''0''l''}v^j \\
&\quad + \tfrac{2}{3}x^{l''}R_{0''j''k''l''}v^iv^jv^k + \tfrac{2}{3}x^{l''}R_{i''j''k''l''}v^jv^k - \tfrac{1}{2}(R_{i''0''l''0'';m''} + R_{i''l''m''0'';0''})x^{l''}x^{m''} \\
&\quad + \tfrac{1}{3}R_{i''l''j''m'';0''}x^{l''}x^{m''}v^j - R_{i''j''l''0'';m''}x^{l''}x^{m''}v^j + \tfrac{1}{12}(5R_{i''k''j''l'';m''} \\
&\quad + R_{i''l''j''m'';k''})x^{l''}x^{m''}v^jv^k + \tfrac{1}{2}R_{0''l''0''m'';0''}x^{l''}x^{m''}v^i + R_{0''j''0''l'';m''}v^iv^jx^{l''}x^{m''} \\
&\quad + \tfrac{1}{3}R_{0''l''j''m'';0''}x^{l''}x^{m''}v^iv^j + \tfrac{1}{12}(5R_{0''k''j''l'';m''} + R_{0''l''j''m'';k''})x^{l''}x^{m''}v^iv^jv^k + O(x^{k''3}).
\end{aligned}
\tag{43}
$$

In (43), there are no coupled angular velocity and curvature terms. (35) must reduce to (43) if we set $a = 0$, hence there must be no coupled angular velocity and curvature terms in Eq. (35). Q.E.D.

From Eq. (43), we can also see that there must be no purely angular velocity and curvature terms to every order in the expansion similar to (35) (i.e., if we use v^i instead of w^i).

Equation (35) is exact in purely inertial terms.[3] If we use the relation

$$w^i = v^i + v^i\mathbf{a}\cdot\mathbf{x} - (\omega\times\mathbf{x})^{\hat{i}} \tag{44}$$

to replace v^i, the inertial terms in (35) can be replaced by [3,4]

$$-a^{\hat{i}}(1+\mathbf{a}\cdot\mathbf{x}) + \left(\frac{\mathbf{w}+\omega\times\mathbf{x}}{1+\mathbf{a}\cdot\mathbf{x}}\right)^{\hat{i}}[\mathbf{b}\cdot\mathbf{x} + 2\mathbf{a}\cdot\mathbf{w} + 2\mathbf{a}\cdot(\omega\times\mathbf{x})] - 2(\omega\times\mathbf{w})^{\hat{i}} - [(\omega\times(\omega\times\mathbf{x})]^{\hat{i}} - (\eta\times\mathbf{x})^{\hat{i}}. \tag{45}$$

(45) are the only purely inertial terms in the w representation.

There are two types of second-order terms in (35): (i) purely gravitational terms, (ii) coupled inertial and gravitational terms. The first type of terms are due to the inhomogeneities in the curvature properties. They agree with those derived by Mashhoon[6] for a geodesic observer. The second type of terms are interpreted in Table I.

IV. STATIC OBSERVERS IN SCHWARZSCHILD SPACETIME

The line element of the Schwarzschild space time geometry in Schwarzschild coordinates is

$$ds^2 = -\left(1 - \frac{2M}{r}\right)dt^2 + \frac{dr^2}{1-2M/r} + r^2(d\theta^2 + \sin^2\theta\, d\phi^2). \tag{46}$$

A static observer in Schwarzschild coordinates has constant r, θ, and ϕ: $r = R$, $\theta = \Theta$, $\phi = \Phi$. The tetrad $e_{\hat{\mu}}$ in the proper reference frame of such an observer (i.e., static orthonormal frame) is related to the Schwarzschild tetrad e_α by

$$e_{\hat{0}} = \left(1 - \frac{2M}{R}\right)^{-1/2}e^0, \quad e_{\hat{1}} = \left(1 - \frac{2M}{R}\right)^{1/2}e^1, \quad e_{\hat{2}} = R^{-1}e_2, \quad e^{\hat{3}} = (R\sin\theta)^{-1}e_3, \tag{47}$$

all along the observer's world line $P_0(\tau)$.

Using (46) and (47), we calculate $R_{\hat{\alpha}\hat{\beta}\hat{\gamma}\hat{\delta}}$ and $R_{\hat{\alpha}\hat{\beta}\hat{\gamma}\hat{\delta};\hat{\epsilon}}$ along the world line $P_0(\tau)$ to be

$$
\begin{aligned}
&R_{\hat{0}\hat{1}\hat{0}\hat{1}} = -2M/R^3, \quad R_{\hat{0}\hat{2}\hat{0}\hat{2}} = R_{\hat{0}\hat{3}\hat{0}\hat{3}} = M/R^3, \quad R_{\hat{2}\hat{3}\hat{2}\hat{3}} = 2M/R^3, \quad R_{\hat{1}\hat{2}\hat{1}\hat{2}} = R_{\hat{1}\hat{3}\hat{1}\hat{3}} = -M/R^3, \\
&R_{\hat{0}\hat{1}\hat{0}\hat{1};\hat{1}} = (6M/R^4)(1-2M/R)^{1/2}, \quad R_{\hat{0}\hat{2}\hat{0}\hat{2};\hat{1}} = R_{\hat{0}\hat{3}\hat{0}\hat{3};\hat{1}} = R_{\hat{0}\hat{2}\hat{0}\hat{1};\hat{2}} = R_{\hat{0}\hat{3}\hat{0}\hat{1};\hat{3}} = -(3M/R^4)(1-2M/R)^{1/2}, \\
&R_{\hat{0}\hat{l}\hat{0}\hat{m};\hat{0}} = R_{\hat{0}\hat{l}\hat{l}\hat{m};\hat{\mu}} = 0, \quad R_{\hat{1}\hat{3}\hat{1}\hat{3};\hat{1}} = R_{\hat{1}\hat{2}\hat{1}\hat{2};\hat{1}} = (3M/R^4)(1-2M/R)^{1/2} \\
&R_{\hat{2}\hat{1}\hat{2}\hat{3};\hat{3}} = R_{\hat{3}\hat{1}\hat{3}\hat{2};\hat{2}} = -(3M/R^4)(1-2M/R)^{1/2}, \quad R_{\hat{2}\hat{3}\hat{2}\hat{3};\hat{1}} = -(6M/R^4)(1-2M/R)^{1/2}, \quad R_{\hat{i}\hat{l}\hat{j}\hat{m};\hat{0}} = 0;
\end{aligned}
\tag{48}
$$

all other components vanish except those obtainable from the above by symmetries of the Riemann tensor or its covariant derivatives.

For the static observer, his 4-velocity is

$$u^\mu = \frac{dx^\mu}{d\tau} = ((1-2M/R)^{-1/2};0,0,0). \tag{49}$$

From (49), we obtain

$$a^\mu = \nabla_{\vec{u}}\vec{u} = (0,M/R^2,0,0), \quad b^\mu = ((M/R^2)^2/(1-2M/R)^{3/2},0,0,0), \quad c^\mu = (0,(M/R^2)^3/(1-2M/R),0,0). \tag{50}$$

Therefore, by the transformation law (47),

$$u^{\hat\mu}=(1,0,0,0),\quad a^{\hat\mu}=(0,(M/R^2)/(1-2M/R)^{1/2},0,0),$$

$$b^{\hat\mu}=((M/R^2)^2/(1-2M/R),0,0,0)=(\mathbf{a}\cdot\mathbf{a},0,0,0),\quad c^{\hat\mu}=(0,(M/R^2)^3/(1-2M/R)^{3/2},0,0)=(0,a^{\hat 1}\mathbf{a}\cdot\mathbf{a},0,0). \tag{51}$$

Also, from (46), we see

$$\omega^{\hat\mu}=\eta^{\hat\mu}=\zeta^{\hat\mu}=0. \tag{52}$$

Substituting (48), (51) into (23), (35) we obtain the line element and the freely falling coordinate acceleration to be

$$\begin{aligned}
ds^2 = &-(dx^{\hat 0})^2\Big[\Big(1+\frac{M}{R^2(1-2M/R)^{1/2}}x^{\hat 1}\Big)^2-\frac{2M}{R^3}x^{\hat 1 2}+\frac{M}{R^3}(x^{\hat 2 2}+x^{\hat 3 2})+\frac{2M}{R^4}\Big(1-\frac{2M}{R}\Big)^{1/2}(x^{\hat 1 3}-x^{\hat 1}x^{\hat 2 2}-x^{\hat 1}x^{\hat 3 2})\\
&+\frac{4}{3}\frac{M^2}{R^5(1-2M/R)^{1/2}}(-2x^{\hat 1 3}+x^{\hat 1}(x^{\hat 2 2}+x^{\hat 3 2}))\Big]+dx^{\hat i}dx^{\hat i}+dx^{\hat 1 2}\Big[\frac{1}{3}\frac{M}{R^3}(x^{\hat 2 2}+x^{\hat 3 2})\\
&-\frac{1}{2}\frac{M}{R^3}\Big(1-\frac{2M}{R}\Big)^{1/2}x^{\hat 1}(x^{\hat 2 2}+x^{\hat 3 2})\Big]+dx^{\hat 2 2}\Big[\frac{1}{3}\frac{M}{R^3}(x^{\hat 1 2}-2x^{\hat 3 2})-\frac{1}{2}\frac{M}{R^4}\Big(1-\frac{2M}{R}\Big)^{1/2}x^{\hat 1}(x^{\hat 1 2}-3x^{\hat 3 2})\Big]\\
&+dx^{\hat 3 2}\Big[\frac{1}{3}\frac{M}{R^3}(x^{\hat 1 2}-2x^{\hat 2 2})-\frac{1}{2}\frac{M}{R^4}\Big(1-\frac{2M}{R}\Big)^{1/2}x^{\hat 1}(x^{\hat 1 2}-3x^{\hat 2 2})\Big]\\
&+dx^{\hat 1}dx^{\hat 2}\Big[-\frac{1}{3}\frac{M}{R^3}x^{\hat 1}x^{\hat 2}+\frac{1}{2}\frac{M}{R^4}\Big(1-\frac{2M}{R}\Big)^{1/2}x^{\hat 1 2}x^{\hat 2}\Big]+dx^{\hat 1}dx^{\hat 3}\Big[-\frac{1}{3}\frac{M}{R^3}x^{\hat 1}x^{\hat 3}\\
&+\frac{1}{2}\frac{M}{R^4}\Big(1-\frac{2M}{R}\Big)^{1/2}x^{\hat 1 2}x^{\hat 3}\Big]+dx^{\hat 2}dx^{\hat 3}\Big[\frac{2}{3}\frac{M}{R^3}x^{\hat 2}x^{\hat 3}-\frac{2M}{R^4}\Big(1-\frac{2M}{R}\Big)^{1/2}x^{\hat 1}x^{\hat 2}x^{\hat 3}\Big]+O\,[dx^{\hat\alpha}dx^{\hat\beta}x^{\hat i}x^{\hat j}x^{\hat k}x^{\hat l}\,],
\end{aligned} \tag{53}$$

$$\begin{aligned}
&\frac{d^2x^{\hat 1}}{dx^{\hat 0 2}}\\
&=-\Big(1+\frac{1}{(1-2M/R)^{1/2}}\frac{M}{R^2}x^{\hat 1}\Big)\frac{1}{(1-2M/R)^{1/2}}\frac{M}{R^2}+2\Big(\frac{M}{R^2}\frac{1}{(1-2M/R)^{1/2}}v^1\Big)\Big(1+\frac{Mx^{\hat 1}}{R^2(1-2M/R)^{1/2}}\Big)v^1\\
&+\frac{2M}{R^3}x^{\hat 1}-\frac{4M}{R^3}x^{\hat 1}v^{12}+\frac{2M}{R^3}v^1(v^2x^{\hat 2}+v^3x^{\hat 3})+\frac{2}{3}\frac{M}{R^3}(x^{\hat 1}v^2v^2+x^{\hat 1}v^3v^3-x^{\hat 2}v^1v^2-x^{\hat 3}v^1v^3)\\
&+4\frac{M^2}{R^5(1-2M/R)^{1/2}}x^{\hat 1 2}+\frac{8}{3}\frac{M^2}{R^5(1-2M/R)^{1/2}}(-2x^{\hat 1 2}v^{12}+x^{\hat 1}x^{\hat 2}v^1v^2+x^{\hat 1}x^{\hat 3}v^1v^3)\\
&-\frac{2}{3}\frac{M^2}{R^5(1-2M/R)^{1/2}}v^{12}[-2x^{\hat 1 2}+(x^{\hat 2 2}+x^{\hat 3 2})]+\frac{4}{3}\frac{M^2}{R^5(1-2M/R)^{1/2}}(x^{\hat 1 2}v^{22}+x^{\hat 1 2}v^{32}-x^{\hat 1}x^{\hat 2}v^1v^2-x^{\hat 1}x^{\hat 3}v^1v^3)\\
&+\frac{1}{3}\frac{M^2}{R^5(1-2M/R)^{1/2}}(x^{\hat 2 2}+x^{\hat 3 2})-\frac{3}{2}\frac{M}{R^4}\Big(1-\frac{2M}{R}\Big)^{1/2}(2x^{\hat 1 2}-x^{\hat 2 2}-x^{\hat 3 2})+\frac{1}{4}\frac{M(1-2M/R)^{1/2}}{R^4}\\
&\times[(x^{\hat 2 2}+x^{\hat 3 2})v^{12}+(-5x^{\hat 1 2}+4x^{\hat 3 2})v^{22}+(5x^{\hat 1 2}+4x^{\hat 2 2})v^{32}+4x^{\hat 1}x^{\hat 2}v^1v^2+4x^{\hat 1}x^{\hat 3}x^1v^3-8x^{\hat 2}x^{\hat 3}v^2v^3]\\
&+\frac{3M(1-2M/R)^{1/2}}{R^4}[(2x^{\hat 1 2}-x^{\hat 2 2}-x^{\hat 3 2})v^{12}-2x^{\hat 1}x^{\hat 2}v^1v^2-2x^{\hat 1}x^{\hat 3}v^1v^3]+O\,[(x^{\hat k})^3],
\end{aligned} \tag{54}$$

$$\begin{aligned}
\frac{d^2x^{\hat 2}}{dx^{\hat 0 2}}=\;&2\Big(\frac{M}{R^2(1-2M/R)^{1/2}}\Big)v^1\Big(1+\frac{M}{R^2(1-2M/R)^{1/2}}x^{\hat 1}\Big)v^2-\frac{M}{R^3}x^{\hat 2}-\frac{4M}{R^3}x^{\hat 1}v^1v^2+\frac{2M}{R^3}(x^{\hat 2}v^{22}+x^{\hat 3}v^2v^3)\\
&+\frac{2}{3}\frac{M}{R^3}(x^{\hat 2}v^{12}-2x^{\hat 2}v^{32}-x^{\hat 1}v^1v^2+2x^{\hat 3}v^2v^3)-\frac{2M^2}{R^5(1-2M/R)^{1/2}}x^{\hat 1}x^{\hat 2}+\frac{8}{3}\frac{M^2}{R^5(1-2M/R)^{1/2}}\\
&\times(x^{\hat 1}x^{\hat 2}v^{22}-2x^{\hat 1 2}v^1v^2+x^{\hat 1}x^{\hat 3}v^2v^3)-\frac{2}{3}\frac{M^2}{R^5(1-2M/R)^{1/2}}v^2v^1(-2x^{\hat 1 2}+x^{\hat 2 2}+x^{\hat 3 2})\\
&+\frac{4}{3}\frac{M^2}{R^5(1-2M/R)^{1/2}}(x^{\hat 1}x^{\hat 2}v^{12}-2x^{\hat 1}x^{\hat 2}v^{32}-x^{\hat 1 2}v^1v^2+2x^{\hat 1}x^{\hat 3}v^2v^3)+\frac{1}{3}\frac{M^2}{R^5(1-2M/R)^{1/2}}x^{\hat 1}x^{\hat 2}\\
&-\frac{3}{2}\frac{M}{R^4}\Big(1-\frac{2M}{R}\Big)^{1/2}(2x^{\hat 2 2}-x^{\hat 1}x^{\hat 2})+\frac{1}{4}\frac{M}{R^4}\Big(1-\frac{2M}{R}\Big)^{1/2}[-6x^{\hat 1}x^{\hat 2}v^{12}+16x^{\hat 1}x^{\hat 2}v^{32}\\
&+(6x^{\hat 1 2}+2x^{\hat 2}x^{\hat 3})v^1v^2-10x^{\hat 1}x^{\hat 3}v^2v^3-8x^{\hat 3 2}v^1v^2]+\frac{3M}{R^4}\Big(1-\frac{2M}{R}\Big)^{1/2}\\
&\times[-2x^{\hat 1}x^{\hat 2}v^{22}-2x^{\hat 1}x^{\hat 3}v^2v^3+(2x^{\hat 1 2}-x^{\hat 2 2}-x^{\hat 3 2})v^1v^2]+O\,[(x^{\hat k})^3],
\end{aligned} \tag{55}$$

$$\frac{d^2x^{\hat 3}}{dx^{\hat 0 2}}=(2\leftrightarrow 3)\quad\text{in}\quad\frac{d^2x^{\hat 2}}{dx^{\hat 0 2}}. \tag{56}$$

V. DISCUSSIONS

The results presented in this paper give explicit understanding of the coupled inertial-gravitational effect and may be useful in analysis of tidal deformation and tidal radiation of objects in close encounters. A Newtonian physicist can think about the terms in Eq. (35) or Table I as simple Newtonian forces, as described in Box 37.1 of MTW. Moreover, a Newtonian physicist can use the equation of motion (25) or (35) to analyze mechanical apparatus in an experimental laboratory. All he needs to do is multiply this equation by the mass of a mass element in his apparatus, and add it linearly onto the forces that would be present if the apparatus were at rest in an inertial reference frame.

[1]C.W. Misner, K.S. Thorne, and J.A. Wheeler, *Gravitation* (Freeman, San Francisco, 1973), cited henceforth as MTW.

[2]W.-T. Ni, Chinese J. Phys. **15**, 51 (1977).

[3]W.-O. Li and W.-T. Ni, "On an Accelerated Observer with Rotating Tetrad in Special Relativity," Chinese J. Phys. **16**, No. 4 (1978).

[4]B. DeFacio, P.W. Dennis, and D. G. Retzloff, "Presymmetry of Classical Relativistic Particles," Phys. Rev. D **18**, 2813 (1978).

[5]W.-T. Ni and M. Zimmermann, Phys. Rev. D **17**, 1473 (1978).

[6]B. Mashhoon, Astrophys. J. **216**, 591 (1977).

[7]Throughout this paper we use MTW notations and conventions.

[8]$\epsilon^{\alpha\beta\mu\nu}$ is the completely antisymmetric tensor with $\epsilon^{0123} = -1$ in an inertial frame.

[9]E.K. Manasse and C.W. Misner, J. Math. Phys. **4**, 735 (1963).

Toroidal black holes?

P. C. Peters

Department of Physics, University of Washington, Seattle, Washington 98195
(Received 25 September 1978)

Recently Thorne presented a class of toroidal solutions which are cylindrically symmetric near the ring singularity. Although the class is a three-parameter family, there is a one-parameter constraint in order to eliminate an unwanted singularity. We show that for parameters satisfying a different one-parameter constraint, the ring singularity is replaced by a nonsingular toroidal horizon. We also show that such a constraint is inconsistent with Thorne's constraint, by finding the latter constraint explicitly in the limit opposite to that considered by Thorne. We exhibit the two-parameter class of solutions with nonsingular toroidal horizons which do not violate the black hole uniqueness theorems, since there is a singularity outside the horizon.

I. INTRODUCTION

In a recent paper Thorne[1] presented a new toroidal solution of the vacuum field equations of general relativity, which, unlike previous toroidal solutions,[2-4] is locally cylindrically symmetric near the ring singularity. The solution was obtained by the Weyl technique[5] for generating static, axially symmetric vacuum metrics from axially symmetric solutions (potentials) of the flat space Laplace equation. For Thorne's metric the generating potential satisfies the boundary conditions that the normal derivative of the potential vanishes on each of two disks of radius b, where the disks are separated by a distance $2a$, and where a uniform line singularity (source) exists on the common axis between the two disks. As the cylindrical radial coordinate ρ approaches 0, the potential approaches $(M/a)\ln\rho$, and, from Gauss' law, this implies that the potential approaches $-M/r$ for large values of the spherical radial coordinate r. The parameter M in the solution is then the Newtonian mass as determined from the asymptotic form of the metric. The toroidal nature of the solution is then induced by identifying the outer surfaces of the two disks and also identifying the inner surfaces of the two disks, so that, for example, the coordinate along a line between the disks which is parallel to the axis is actually a cyclical coordinate in the metric. Thus the line singularity in the potential becomes a ring singularity in the metric.

The potential is determined by the three parameters M, a, and b, and therefore, so is the metric. However, for an arbitrary choice of parameters, there is in general an additional unwanted singularity in the metric,[1] generated through the Weyl procedure, arising from the behavior of the potential near the edge of the disks ($\rho \to b$). In the metric this singularity appears as a ring singularity (not locally cylindrically symmetric) which lies inside the original ring singularity in the same plane. However, if the potential ψ near the edge of the disk has the behavior

$$\psi \to \pm [2(1-\rho/b)]^{1/2} + \text{const}$$

as $\rho \to b$ along the disk, then the unwanted singularity is no longer present and one is left with a toroidal metric with only the original ring singularity.[1] For a fixed a and b it is always possible to find a value of M such that ψ has the above behavior since ψ is proportional to M. Therefore, there are in fact only two independent parameters which specify a given Thorne solution.[1]

Given the flexibility of a two-parameter solution, it would seem reasonable that one could choose a one-parameter constraint, e.g., $M/a = 1$. However, with this choice, as we will show, the geometry in the vicinity of $\rho = 0$ becomes nonsingular, and in fact the singularity is replaced by a toroidal horizon, analogous to the spherical horizon of the Schwarzchild metric. The generation of a nonsingular horizon for a special choice of parameters also occurs in the generation of the Schwarzchild metric through the Weyl formalism, where the generating potential is that of a rod of uniform mass per unit length.[6] A metric with a toroidal horizon and no external singularity would violate the black hole uniqueness theorems of Israel, Carter, Hawking, and others.[7] Therefore, it is desirable to study in further detail the constraint on the parameters M, a, and b to produce a solution free of the unwanted singularity in order to see if the choice $M/a = 1$ is consistent with that constraint. In Sec. II we show that there is a lower bound of 2 for the ratio M/a so that the choice $M/a = 1$ necessarily introduces an unwanted singularity outside the horizon. In Sec. III we give a procedure by which the constraint equation on M, a and b can be expressed in a power series in (b/a), useful for small (b/a), which complements the form of the constraint equation given by Thorne, which is valid for large (b/a). In Sec. IV we illustrate some of the properties of the metric with $M/a = 1$, the only static, axially symmetric metric with a nonsingular toroidal horizon (albeit with an external singularity) that has been exhibited to date.

II. LOWER BOUND ON M/a IN THORNE'S SOLUTION

The generating potential ψ for Thorne's solution is defined as a unique solution to a particular boundary value problem, but it does not appear possible to find a closed form expression for ψ, or for the constraint equations. By solving approximate boundary value problems in three regions and matching solutions in adjacent regions, Thorne investigated the case of a thin-ring torus,[1] in which $b \gg a$. The constraint

 0022-2488/79/071481-05$01.00

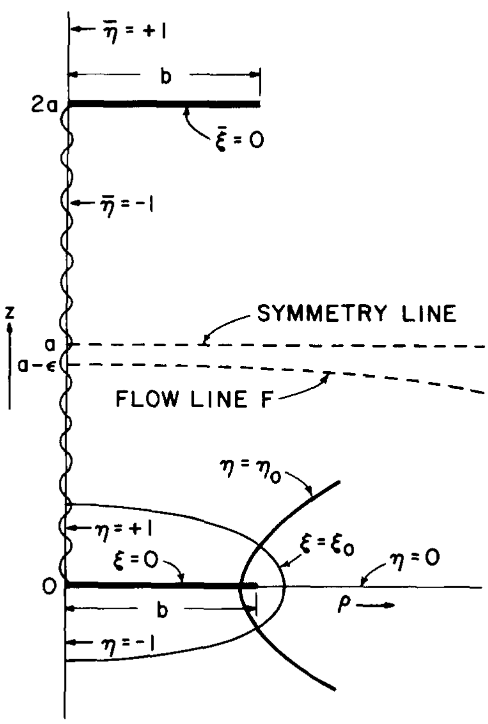

FIG. 1. Orientation of the two disks and line source for the generating potential ψ. Oblate spheroidal coordinates are defined centered on each disk, where the azimuthal coordinate ϕ has been suppressed for simplicity. One flow line F is shown to illustrate the point that $\partial\psi/\partial z \leqslant 0$, for $z < a$.

equation that the parameters must satisfy in order to generate a metric free of the unwanted singularity is[1]

$$\frac{M}{a} = \left(\frac{\pi b}{a}\right)^{1/2}, \quad b \gg a. \tag{1}$$

The relation (1) is valid only for large values of M/a, since $b/a \gg 1$. In particular nothing can be said from (1) about the possibility of choosing $M/a = 1$.

Consider the opposite limit, $b/a \ll 1$, the extreme case of $b/a = 0$ occurring when the disks are separated by an infinite distance, in which case the behavior of the potential on one disk is determined by the solution of a boundary value problem involving only one disk. The solution of Laplace's equation for a single disk is most easily carried out in oblate spheroidal coordinates.[8] Choosing the disk to be located at $z = 0$, $\rho \leqslant b$, and defining coordinates ξ and η (see Fig. 1) by

$$z = b\xi\eta, \quad \rho = b[(1-\eta^2)(1+\xi^2)]^{1/2}, \tag{2}$$

the disk becomes the coordinate surface $\xi = 0$, the positive (negative) z-axis becomes $\eta = +1$ ($\eta = -1$), and the condition of vanishing normal derivative of ψ on the disk implies $\partial\psi/\partial\xi = 0$, for $\xi = 0$. Laplace's equation is separable in oblate spheroidal coordinates[8] with the axially symmetric solutions being linear combinations of $P_n(\eta)$ [or $Q_n(\eta)$] times $P_n(i\xi)$ [or $Q_n(i\xi)$].

The potential ψ must satisfy the boundary condition that there be a line source at $\rho = 0$ for $z > 0$ but no line source at $\rho = 0$, $z < 0$. On the other hand, the particular separation of variables is most useful if there are equal (but opposite) line sources for $z > 0$ and $z < 0$. We can achieve this by writing $\psi = (M/2a)\ln\rho + \psi'$, where the normal derivative condition on ψ' is unchanged and where ψ' has line singularities on the $+z$ and $-z$ axis such that as the axes are approached,

$$\psi' \to \pm \frac{M}{2a}\ln\rho \to \pm \frac{M}{4a}\ln(1 \mp \eta), \tag{3}$$

where $\pm$ refers to the $\pm z$ axes, also defined by $\eta = \pm 1$. The potential ψ' then involves only one term of the otherwise infinite series, which gives the exact potential ψ, denoted now by $\psi^{(0)}$ for a single disk,

$$\psi^{(0)} = \frac{M}{2a}(\ln\rho - Q_0(\eta)) = \frac{M}{2a}\left(\ln\rho - \tfrac{1}{2}\ln\left(\frac{1+\eta}{1-\eta}\right)\right). \tag{4}$$

For a nonsingular behavior of the metric at $\rho = b$, $z = 0$, is is necessary that as $\rho \to b$ on the disk (or $\eta \to 0$)

$$\psi \to \pm [2(1-\rho/b)]^{1/2} + \text{const} = \pm \eta + \text{const}. \tag{5}$$

Therefore, in the expression for $\psi(\xi,\eta)$, if we set $\xi = 0$ and let $\eta \to 0$, the magnitude of the coefficient of η in the resulting expression must be 1. Applied to the potential (4), this implies that the metric will be free of the unwanted singularity in the limit $b/a \ll 1$ if the ratio M/a has the value

$$\frac{M}{a} = 2 \quad (b/a = 0). \tag{6}$$

We now show that the value $M/a = 2$ for $b/a = 0$ is in fact a lower bound to M/a, so that for $b/a \geqslant 0$, $M/a \geqslant 2$. Let us consider the potential to describe the velocity potential of incompressible fluid flow. Choose coordinates so that one constraining disk is at $z = +2a$ and the other is at $z = 0$, as shown in Fig. 1. By symmetry $z = a$ is a flow line with zero slope, i.e., $\partial\psi/\partial z = 0$. In the region $z > a$ all flow lines have positive slope $\partial\psi/\partial z > 0$ and in the region $z < a$ all flow lines have negative slope $d\psi/dz < 0$. To see this we note that $\phi \equiv \partial\psi/\partial z$ is a solution of Laplace's equation with boundary values $\phi = 0$ for $z = a$; no singularity at $\rho = 0$, $a < z < 2a$; $\phi = 0$ for $z = 2a$, $\rho < b$, and $\phi \to M(z-a)/R^3$ as $R \to \infty$, where $R = [\rho^2 + (z-a)^2]^{1/2}$. There is, of course, a singularity (source) at $\rho = b$, $z = 2a$. However, near the singularity ϕ is positive, so that $\phi \geqslant 0$ on the boundary. Then, since ϕ cannot achieve its minimum at any point on the interior of the boundary (by the mean value theorem), this means that $\phi \geqslant 0$ on the interior, or $\partial\psi/\partial z \geqslant 0$, for $z \geqslant a$. For the same reasons, $\partial\psi/\partial z \leqslant 0$, for $z < a$.

Consider now a flow line F which starts at $\rho = 0$, $z = a - \epsilon$ where $\epsilon \ll a$ (see Fig. 1). This flow line, from the above, monotonically moves away from the horizontal symmetry line $z = a$, becoming, for large r, radial as $\rho \gg b$. Define a surface S by rotating F about the z-axis. Then the potential ψ satisfies the boundary conditions in the region below S that $\psi \to (M/a)\ln\rho$ as $\rho \to 0$, $0 < z < a - \epsilon$; $\partial\psi/\partial n = 0$ on S; and $\partial\psi/\partial z = 0$, for $z = 0$, $\rho < b$. Now compare that potential with the potential $\tilde{\psi}$ which satisfies the same boundary conditions as above except that $\partial\tilde{\psi}/\partial z = 0$ for $z = a - \epsilon$ rather than $\partial\psi/\partial n = 0$ on S. Physically this means that the flow $\tilde{\psi}$ is less constricted than the flow ψ, which reduces the flow everywhere, including that near the edge of the disk at $z = 0$. Alternatively, $\delta\psi = \tilde{\psi} - \psi$ must satisfy $\partial\delta\psi/\partial z \geqslant 0$ on the surface $z = a - \epsilon$ since $\partial\psi/\partial z \leqslant 0$ there. Since $\delta\psi$ has no source except at ∞, flow lines of $\delta\psi$ must come in from infinity and cross the surface $z = a - \epsilon$. These lines are in the opposite direction to the flow lines of ψ near the disk at $z = 0$, again implying that the flow is reduced near the edge of the disk.

The flow $\tilde{\psi}$ is, when reflected by the plane $z = a - \epsilon$, the solution of a boundary value problem with two disks, one at $z = 0$, the other at $z = 2(a - \epsilon)$, (for the same b), where $z = a - \epsilon$ is now a symmetry plane. However, the flow potential still approaches the same value $(M/a)\ln\rho$ near $\rho = 0$ as did the solution $\tilde{\psi}$, although the mass $\tilde{M}$ associated with $\tilde{\psi}$ is given by $\tilde{M} = M(1 - \epsilon/a)$. Suppose that the original ψ had the property that as we approached the edge of one disk along the disk, the coefficient of $\sqrt{1 - \rho/b}$ was $\pm\sqrt{2}$, generating a metric free of the unwanted singularity. Then for the solution $\tilde{\psi}$, the coefficient would have a magnitude less than $\sqrt{2}$, since the flow is reduced. Therefore, to generate a potential for which the coefficient had a magnitude $\sqrt{2}$, we would have to increase the mass M to $M + \Delta M$, which means that $\tilde{\psi} \to \tilde{\psi}(1 + \Delta M/M)$. Near the source the flow potential is increased from $(M/a)\ln\rho$ to $(M/a + \delta)\ln\rho$, where $\delta = \Delta M/a$. Thus keeping b fixed, if we decrease a, the strength of the line source (in general M/a) must be increased to give a metric free of the $\rho = b$ singularity. On dimensional grounds M/a must be a function of (b/a), so that from the above dependence of M/a on a, keeping b fixed, we conclude that M/a is a monotonically increasing function of b/a which achieves its minimum value $M/a = 2$ at $b/a = 0$. Thus M/a must be greater than or equal to 2 for a metric free of the $\rho = b$ singularity. This means that the $M/a = 1$ metrics must have a singularity outside the toroidal horizon, so that there is no violation of the black hole uniqueness theorems.

III. CONSTRAINT EQUATION FOR SMALL b/a

The deviation of M/a from 2 as a function of b/a can be obtained by an iterative procedure, giving M/a in a power series in (b/a) to as high an order as desired. Let oblate spheroidal coordinates $\bar{\xi}$ and $\bar{\eta}$, centered on the disk at $z = 2a$ (see Fig. 1), be defined by

$$z - 2a = b\bar{\xi}\bar{\eta}, \quad \rho = b[(1 - \bar{\eta}^2)(1 + \bar{\xi}^2)]^{1/2}. \tag{7}$$

Then a solution of Laplace's equation which has no singularity on the z axis for $z < 0$ or $z > 2a$, and which approaches $(M/a)\ln\rho$ as $\rho \to 0$ for $0 < z < 2a$, is

$$\psi_T^{(0)} = \psi^{(0)} - \bar{\psi}^{(0)} = \frac{M}{4a}\left[\ln\left(\frac{(1 + \bar{\eta})(1 - \eta)}{(1 - \bar{\eta})(1 + \eta)}\right)\right], \tag{8}$$

where $\psi^{(0)} \equiv \psi^{(0)}(\xi,\eta)$, $\bar{\psi}^{(0)} \equiv \psi^{(0)}(\bar{\xi},\bar{\eta})$, and the function $\psi^{(0)}(\xi,\eta)$ is given by Eq. (4). The expression (8) is, however, not a solution of our boundary value problem, in spite of the correct behavior on the z axis, because $\partial\psi^{(0)}/\partial z$ does not vanish on the disk $z = 2a, \rho \leqslant b$ and $\partial\bar{\psi}^{(0)}/\partial z$ does not vanish on the disk $z = 0, \rho \leqslant b$.

Consider for the moment a single disk at $z = 0$. Define the Neumann Green's function $G_N(\xi,\eta,\xi',\eta')$ to be the potential at ξ,η of a unit ring source at ξ',η' (so that $G_N \to 1/r$ for large r), subject to the vanishing of the normal derivative of G_N on the disk, i.e., $\partial G_N/\partial\xi = 0$, for $\xi = 0$. Then we find explicitly, by standard techniques, that

$$G_N(\xi,\eta,\xi',\eta') = \frac{i}{b}\sum_{n=0}^{\infty}(2n + 1)P_n(\eta)P_n(\eta')\Big(P_n(i\xi_<)\times Q_n(i\xi_>) - \frac{P_n'(0)}{Q_n'(i0)}Q_n(i\xi)Q_n(i\xi')\Big). \tag{9}$$

Using Green's theorem we can find the potential with no sources given the normal derivative of the potential on the surface of the disk (assuming the potential vanishes at infinity):

$$\psi(\xi,\eta) = -\frac{b}{2}\int_{-1}^{1}\left(\frac{\partial\psi}{\partial\xi'}(\xi',\eta')\,G_N(\xi,\eta,\xi',\eta')\right)_{\xi'=0}d\eta'. \tag{10}$$

The analogous Green's function associated with the disk at $z = 2a$ will be $\bar{G}_N \equiv G_N(\bar{\xi},\bar{\eta},\bar{\xi}',\bar{\eta}')$, and the solution of a boundary value problem on the disk $z = 2a$ will be given by the barred form of (10).

Consider now the problem with the two disks at $z = 0$ and $z = +2a$. We can use the Green's function (9) and solution to the boundary value problem (10), together with their barred expressions, to add a homogeneous solution of Laplace's equation to (8) so that the normal derivative of ψ vanishes on both disks to as high an order as desired. Specifically, define the sequence of functions

$$\psi_U^{(n+1)} = -\psi_L^{(n)} - \frac{b}{2}\int_{-1}^{1}\left(\frac{\partial\psi_L^{(n)}}{\partial\bar{\xi}'}\bar{G}_N\right)_{\bar{\xi}'=0}d\bar{\eta}',$$

$$\psi_L^{(n+1)} = -\psi_U^{(n)} - \frac{b}{2}\int_{-1}^{1}\left(\frac{\partial\psi_U^{(n)}}{\partial\xi'}G_N\right)_{\xi'=0}d\eta', \tag{11}$$

where $\partial\psi_U^{(n)}/\partial\bar{\xi} = 0$ on the upper disk ($\bar{\xi} = 0$) and $\partial\psi_L/\partial\xi = 0$ on the lower disk ($\xi = 0$), and where $\psi_U^{(0)} = \bar{\psi}^{(0)}$, $\psi_L^{(0)} = \psi^{(0)}$. The potential ψ, which satisfies the same boundary conditions as does (8) on the z axis, and which also has a vanishing normal derivative on each of the two disks, is then given by

$$\psi(\xi,\eta) = \lim_{n\to\infty}\psi_T^{(n)}(\xi,\eta), \tag{12}$$

where $\psi_T^{(n)} = \psi_L^{(n)} - \psi_U^{(n)}$. For small b/a the sequence of operations in (11) converges quickly.

We will illustrate how the iteration scheme works by considering the lowest order correction to the potential (8) in powers of (b/a). In (11) there are derivatives of barred variables with respect to unbarred variables and vice versa. These are best treated by returning to the defining relations (2) and (7), expressing quantities in terms of ρ and z. Starting from (4), the derivative $\partial\psi^{(0)}/\partial z$ on the disk $z = +2a, \rho \leqslant b$, is approximately

$$\frac{\partial\psi^{(0)}}{\partial z} = -\frac{M}{2a}\frac{1}{1 - \eta^2}\frac{\partial\eta}{\partial z} = -\frac{M}{2ab}\frac{\xi}{\xi^2 + \eta^2} \approx \frac{M}{4a^2}, \tag{13}$$

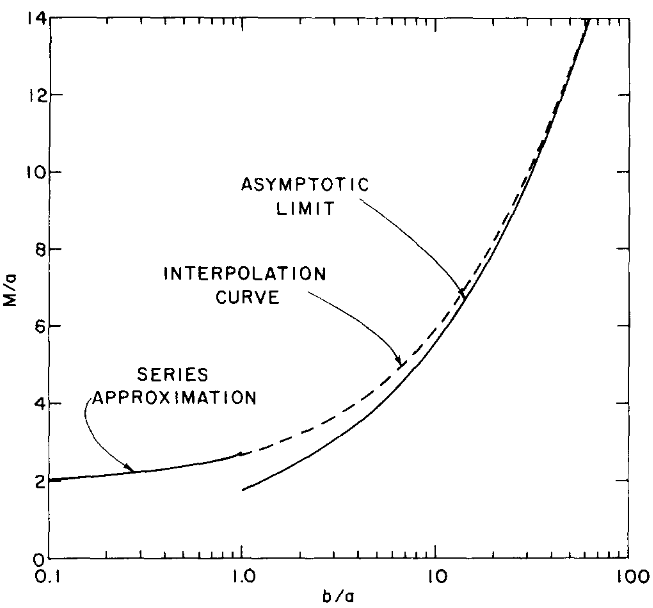

FIG. 2. The constraint relation M/a as a function of b/a to produce a solution free of the $\rho = b$ singularity. The solid line for small b/a gives the results of the calculation described in Sec. III; the solid line for large b/a gives the leading term in a large b/a expansion, as computed by Thorne; a possible interpolation curve is shown as a dashed line.

independent of ρ, where we have used in the last approximation $\eta \approx 1$, $\xi \approx 2a/b \gg 1$. Then, from (11), $\psi_U^{(1)}$ is given by

$$\psi_U^{(1)} = -\psi^{(0)} + \frac{Mb^2}{8a^2} \int_{-1}^{1} [\bar{\eta}' \bar{G}_N]_{\bar{\xi}'=0} \, d\bar{\eta}' . \qquad (14)$$

In the expression for $\bar{G}_N$, obtained from (9) by taking the bar of all variables, only the $n = 1$ term in the infinite sum survives the integration in (14) because of the orthogonality of $P_n(\bar{\eta}')$. This simplifies (14) to

$$\psi_U^{(1)} = -\psi^{(0)} - \frac{Mb\bar{\eta}}{2\pi a^2} [\bar{\xi} \cot^{-1}\bar{\xi} - 1], \qquad (15)$$

and similarly

$$\psi_L^{(1)} = -\bar{\psi}^{(0)} - \frac{Mb\eta}{2\pi a^2} [\xi \cot^{-1}\xi - 1], \qquad (16)$$

where $\psi_T^{(1)} = \psi_L^{(1)} - \psi_U^{(1)}$.

Consider now the behavior of $\psi_T^{(1)}$ along the disk $\xi = 0$ as $\eta \to 0$. There are two terms which contribute to a term in $\psi_T^{(1)}$ proportional to η—the first term on the right side of (15) and the second term on the right side of (16). Setting $\xi = 0$, and taking the small η limit, gives

$$\psi_T^{(1)}(0,\eta) \approx \text{const} - \frac{M}{2a}\eta\left(1 - \frac{b}{\pi a}\right) + O(\eta^2). \qquad (17)$$

Since the coefficient of η must be ± 1 for a metric free of the edge singularity, this means that $M/2a(1 - b/\pi a) = 1$, or to this order

$$\frac{M}{a} = 2\left(1 + \frac{b}{\pi a}\right), \quad \frac{b}{a} \ll 1, \qquad (18)$$

which increases linearly away from the minimum value.

We have computed the potential to one more iteration $\psi^{(2)}$, which also necessitates calculating $\psi^{(1)}$ to higher order in b/a. Considering the behavior of $\psi^{(2)}$ on the disk $\xi = 0$ as $\eta \to 0$ then gives the constraint relation on the parameters (valid to 6th order in b/a)

$$\frac{M}{a} = 2 \Big/ \left[1 - \frac{1}{\pi}\left(\frac{b}{a}\right) g\left(\frac{b}{a}\right)\right], \qquad (19)$$

where

$$g(x) = 1 - \frac{1}{3}x^2 + \frac{1}{6\pi}x^3 + \frac{1}{5}x^4 - \frac{11}{60\pi}x^5 + O(x^6). \qquad (20)$$

In Fig. 2 we display the behavior of M/a as a function of (b/a), the asymptotic form (1) which Thorne computed[1] (only the leading term was calculated in the asymptotic series), and a possible interpolation curve between the two which is monotonic in (b/a).

IV. METRICS WITH A TOROIDAL HORIZON

The axially symmetric static metric generated through the Weyl formalism can be cast in the form[2,5]

$$ds^2 = e^{2\psi}dt^2 - e^{2(\gamma - \psi)}(d\rho^2 + dz^2) - \rho^2 e^{-2\psi} d\phi^2 , \qquad (21)$$

where ψ is the generating potential and γ can be found, once ψ is known, by integration. For the boundary conditions on ψ appropriate to a line source between two disks as stated in Sec. I, the behavior of ψ and γ near the $\rho = 0$ singularity between the disks is[1]

$$\psi \to (M/a) \ln\rho + \psi_0, \quad \gamma \to (M/a)^2 \ln\rho + \gamma_0, \qquad (22)$$

where ψ_0 and γ_0 are constants. There is a singularity at $\rho = 0$ in general, which is a ring singularity because z is a cyclical coordinate between the disks. However, if we take the special case $(M/a) = 1$, the metric (21), near $\rho = 0$, becomes

$$ds^2 = e^{2\psi_0}\rho^2 dt^2 - e^{-2\psi_0}[e^{2\gamma_0}(d\rho^2 + dz^2) + d\phi^2]. \qquad (23)$$

The nature of $\rho = 0$ in this case is best exhibited by changing to coordinates u and v, defined by

$$u = \alpha\rho \cosh\beta t, \quad v = \alpha\rho \sinh\beta t, \qquad (24)$$

with $\alpha = e^{\gamma_0 - \psi_0}$, $\beta = e^{2\psi_0 - \gamma_0}$, in which case the metric (23) becomes

$$ds^2 = dv^2 - du^2 - e^{-2\psi_0}[e^{2\gamma_0}dz^2 + d\phi^2]. \qquad (25)$$

Since $u^2 - v^2 = \alpha^2\rho^2$, the locus of points $\rho = 0$ is actually the null surface $u = \pm v$ of the flat space–time metric (25). Since this surface is generated by radial null geodesics and since g_{00} in (23) vanishes at $\rho = 0$, this surface is both a horizon and surface of infinite redshift, the same roles played by the Schwarzschild radius in the Schwarzschild metric.[9] Unlike the surface of a torus in three dimensions, which has nonvanishing two-dimensional curvature, the toroidal horizon is flat, and in fact has the proper surface area defined by a rectangular coordinate range $0 \leqslant z < 2a$, $0 \leqslant \phi < 2\pi$,

$$A = 4\pi a e^{\gamma_0 - 2\psi_0}, \qquad (26)$$

since both z and ϕ are cyclical coordinates.

In the limit that $b/a \to \infty$ the metric functions (22) become exact solutions, and (25) is then an exact flat space–time metric with vanishing curvature tensor and vanishing curvature invariants. Suppose b/a is large, but finite. Then, using standard techniques for finding γ, (22), with $M/a = 1$, becomes

$$\begin{aligned}\psi &= \ln\rho + \psi_0 + \delta\psi, \\ \gamma &= \ln\rho + \gamma_0 + 2\delta\psi + O\,[(\delta\psi)^2],\end{aligned} \tag{27}$$

where $\delta\psi \to 0$ as $a/b \to 0$. Substituting in the metric (21), making the coordinate change (24) and ignoring terms of order $(\delta\psi)^2$, we obtain the generalization of (25)

$$ds^2 = e^{2\delta\psi}(dv^2 - du^2) - e^{-2(\psi_0 + \delta\psi)} \times \{e^{2(\gamma_0 + 2\delta\psi)}\, dz^2 + d\phi^2\}. \tag{28}$$

The general source-free solution of Laplace's equation which is regular as $\rho \to 0$, and cyclic in z and which has a symmetry plane at $z = a$ is

$$\delta\psi = \sum_{n=1}^{\infty} C_n I_0\left(\frac{n\pi}{a}\rho\right)\cos\frac{n\pi(z-a)}{a}, \tag{29}$$

so that for small ρ, $\delta\psi$ is of the form

$$\delta\psi = f_1(z) + \rho^2 f_2(z) + O(\rho^4), \tag{30}$$

where f_1 and f_2 are cyclic and bounded. Since ρ^2 can be expressed as $(u^2 - v^2)/\alpha^2$, that means, as far as the linear terms in (28) are concerned, second derivatives of the metric coefficients are finite at $u = \pm v$. Since the metric (28) has no coordinate singularity at $u = \pm v$, this means that the curvature invariants are finite on the horizon, including linear perturbations arising from the edge of the disk being a finite distance from $\rho = 0$. If we estimate the quadratic terms in γ in (27), using the form (30) for $\delta\psi$, we find that those terms are proportional to ρ^2, so that the location of the horizon is unchanged and the finiteness of the curvature invariants is preserved. Therefore, for finite b/a there is a nonsingular horizon at $\rho = 0$.

The ρ,z,ϕ,t coordinates are not suitable for displaying the metric inside the horizon. The u,v,z,t coordinate system carries one smoothly past the horizon, playing the same role as Kruskal coordinates.[9] In the limit $b \to \infty$, (25) is in fact a valid metric everywhere. However, if b is finite, the exact form of the metric will not be known and the nature of the internal singularities cannot be discussed. However, from the form (28) we can see that the metric for $v > u$ will be time dependent since ρ will become a time coordinate and the metric depends on ρ.

The external singularity at $\rho = b$, $z = 0$ (or $2a$) cannot be removed by a coordinate transformation, since it is characterized by infinite curvature invariants. Because we have imposed only one condition $M/a = 1$ on the three parameters M, a, and b, the metric with a toroidal horizon and external ring singularity is a two-parameter family, since no further constraint removes the singularity. This solution probably has no reasonable physical or astrophysical application, however, since if the ring singularity were replaced by a reasonable physical matter distribution, no toroidal horizon could exist by the black hole uniqueness theorems.[7]

[1] K.S. Thorne, J. Math. Phys. **16**, 1860 (1975).
[2] H. Weyl, Ann. Physik **54**, 117 (1917).
[3] R. Bach and H. Weyl, Math. Z. **13**, 134 (1922).
[4] M. Misra, Proc. Nat. Inst. Sci. India A **27**, 373 (1961).
[5] See, for example, the discussion in J.L. Synge, *Relativity: The General Theory* (North-Holland, Amsterdam, 1960), Chapter VIII.
[6] G. Erez and N. Rosen, Bull. Res. Council Israel F **8**, 47 (1959).
[7] See, for example, S.W. Hawking and G.F.R. Ellis, *The Large Scale Structure of Space-Time* (Cambridge U.P., Cambridge, 1973).
[8] P.M. Morse and H. Feshbach, *Methods of Theoretical Physics* (McGraw-Hill, New York, 1953), Vols. I, II.
[9] See, for example, discussions in C.W. Misner, K.S. Thorne, and J.A. Wheeler, *Gravitation* (Freeman, San Francisco, 1973).

Generation of the Lewis solutions from the Weyl solutions

Yukio Tanabe

Department of Physics, Nagoya University, Nagoya 464, Japan
(Received 12 October 1978)

A method is presented which can be used to generate an infinite series of new solutions of stationary axially symmetric Einstein–Maxwell equations. It is shown that if we start with the Weyl solutions the Lewis solutions are generated as a member of this series.

I. INTRODUCTION

One of the excellent ideas for finding new solutions of the Einstein–Maxwell equations is to seek methods which allow us to generate new solutions from old ones. Such methods are useful not only in increasing the number of exact solutions but also, when the generated solutions are already known, in simplifying the derivation of the solutions. For example, applying the hitherto discovered generating methods[1–11] to the Kerr solution,[12] we can obtain the charged Kerr solutions[3], the NUT-like Kerr solution,[13] the charged NUT-like Kerr solution,[6] the electric/magnetic dipole solution,[7,8] the five-parameter solution,[6,14,15] and others.[16,17] Among the above solutions, the charged and the NUT-like Kerr solutions are first discovered by other methods which require more sophisticated techniques.[18,19]

The purpose of this paper is to formulate and apply a generating method which may be regarded, in a sense, as a refinement of the method of Catenacci and Alonso.[20] Our method applies to a stationary axially symmetric system which can be described by a single complex function satisfying a equation formally equivalent to the original Ernst equation[21] (i.e., the equation satisfied by the $\mathscr{E}$ function). Using the method, we can generate an infinite series of exact solutions. We show that if we start with the Weyl solution,[22] one of the members of the series becomes the Lewis solution,[23] and therefore our method provides a simple derivation of the Lewis solution.

II. GENERATION OF NEW SOLUTIONS

Consider a stationary axially symmetric system described by the metric

$$ds^2 = f(dt - \omega d\phi)^2 - f^{-1}[\rho^2 d\phi^2 + e^{2\gamma}(d\rho^2 + dz^2)], \tag{2.1}$$

where f, ω, and γ are functions of ρ and z only. We write the Ernst equation in the form[21]

$$(\mathrm{Re}\,\mathscr{K})\nabla^2\mathscr{K} = \nabla\mathscr{K}\cdot\nabla\mathscr{K}, \tag{2.2}$$

$$\mathscr{K} = k(\rho,z) + ih(\rho,z), \tag{2.3}$$

where ∇ denotes the gradient operator with respect to the three-dimensional metric $d\sigma^2 = \rho^2 d\phi^2 + d\rho^2 + dz^2$. We leave the functions k and h unspecified because there are several ways to identify fields $\mathscr{K} = k + ih$ which satisfy Ernst's equation.[20] Our method can be applied to all such identifications.

Now from a known solution $\mathscr{K} = k + ih$ we generate a new solution $\mathscr{K}' = k' + ijh'$ in the following way. First we note that Eq. (2.2) can be rewritten in the form

$$(\mathrm{Re}\,\mathscr{K})\rho^{-1}\widehat{D}\cdot(\rho\widehat{D}\mathscr{K}) = \widehat{D}\mathscr{K}\cdot\widehat{D}\mathscr{K}, \tag{2.4}$$

where we use the notations

$$\widehat{D} = \left(\frac{\partial}{\partial\rho},\frac{\partial}{\partial z}\right),\quad \widehat{D}f\cdot\widehat{D}g = \frac{\partial f}{\partial\rho}\frac{\partial g}{\partial\rho} + \frac{\partial f}{\partial z}\frac{\partial g}{\partial z},$$

$$\widehat{D}\cdot\hat{A} = \frac{\partial}{\partial\rho}A_\rho + \frac{\partial}{\partial z}A_z \quad \text{for } \hat{A} = (A_\rho, A_z). \tag{2.5}$$

Since $\mathscr{K} = k + ih$, Eq. (2.4) is equivalent to the simultaneous equations

$$\widehat{D}\cdot\left(\frac{\rho}{k}\widehat{D}k\right) + \frac{\rho}{k^2}\widehat{D}h\cdot\widehat{D}h = 0, \tag{2.6}$$

$$\widehat{D}\cdot\left(\frac{\rho}{k^2}\widehat{D}h\right) = 0. \tag{2.7}$$

Equation (2.7) is the integrability conditions for the existence of the function h' defined by

$$\widetilde{D}h' = (\rho/k^2)\widehat{D}h, \tag{2.8}$$

where

$$\widetilde{D} = \left(\frac{\partial}{\partial z}, -\frac{\partial}{\partial\rho}\right). \tag{2.9}$$

Introducing a new function k' defined by

$$k' = \rho/k, \tag{2.10}$$

we obtain

$$\widehat{D}h = (\rho/k'^2)\widetilde{D}h'. \tag{2.11}$$

This equation implies

$$\widehat{D}\cdot\left(\frac{\rho}{k'^2}\widehat{D}h'\right) = \widetilde{D}\cdot\left(\frac{\rho}{k'^2}\widetilde{D}h'\right) = \widetilde{D}\cdot\widehat{D}h = 0. \tag{2.12}$$

On the other hand, the substitution of Eqs. (2.10) and (2.11) into Eq. (2.6) yields

$$\widehat{D}\cdot\left(\frac{\rho}{k'}\widehat{D}k'\right) - \frac{\rho}{k'^2}\widehat{D}h'\cdot\widehat{D}h' = 0. \tag{2.13}$$

Comparing Eqs. (2.12) and (2.13) with Eqs. (2.6) and (2.7), we see that the complex function

$$\mathscr{K}' = k' + ijh' \tag{2.14}$$

satisfies Ernst's Eq. (2.2), where j is the symbol introduced in a previous paper[9] which has the following properties:

$$j^2 = -1, \quad j^* = j. \tag{2.15}$$

As is fully illustrated in the Appendix of Ref. 9, the symbol j can be interpreted as the imaginary unit contained in the functions k and h. [That is, the function k and h can be complex without contradiction to the complex potential formalism, if we replace the imaginary unit in the functions k and h by the symbol j.] The new solution k_{new} and h_{new} determined by the i- complex[24] function $\mathscr{K}_{\text{new}} = \mathscr{K}'$ is clearly j-complex and the j-complexity of k_{new} and h_{new} means that the metric and the electromagnetic field become complex in the usual sense. However, in the later applications, the imaginary unit in the new solution can be absorbed into the parameters in the solution. Therefore, we shall not confront problems of physical interpretation.

It is easy to check that the iteration of the above operation twice yields the original solution:

$$(\mathscr{K}')' = \mathscr{K}. \tag{2.16}$$

This means that we can obtain only one new solution from one old solution by the transformation $\mathscr{K} \to \mathscr{K}'$. However, it is well known[4,10,20] that if $\mathscr{K}$ is a solution of Eq. (2.2), the function $\overline{\mathscr{K}}$ defined by

$$\overline{\mathscr{K}} = \frac{a\mathscr{K} + ib}{ic\mathscr{K} + d}, \tag{2.17}$$

$$a,b,c,d = i\text{-real},^{24} \quad ad + bc = 1 \tag{2.18}$$

is also a solution of Eq. (2.2). Applying the two transformations alternately, we obtain the following infinite series of new solutions:

$$\cdots\mathscr{K}_{-2}—\mathscr{K}_{-1}\cdots\mathscr{K}_0—\mathscr{K}_1\cdots\mathscr{K}_2—\mathscr{K}_3\cdots, \tag{2.19}$$

where $\mathscr{K}_0$ is a known solution and a dotted (solid) line denotes generation by means of the transformation $\mathscr{K} \to \mathscr{K}'$ ($\mathscr{K} \to \overline{\mathscr{K}}$).

III. STATIONARY AXISYMMETRIC GRAVITATIONAL FIELDS

In this section we apply the method formulated in the preceding section to the case in which the complex function $\mathscr{K}$ is the usual Ernst potential;[21] $\mathscr{E} = f + i\varphi$. This case corresponds to the stationary axisymmetric source-free gravitations fields and the metric (2.1) is determined by integrating the equations

$$\widetilde{D}\omega = \frac{\rho}{f^2}\widehat{D}\varphi, \tag{3.1}$$

$$\frac{\partial\gamma}{\partial\rho} = \frac{\rho}{4f^2}\left[\left(\frac{\partial f}{\partial\rho}\right)^2 - \left(\frac{\partial f}{\partial z}\right)^2 + \left(\frac{\partial\varphi}{\partial\rho}\right)^2 - \left(\frac{\partial\varphi}{\partial z}\right)^2\right],$$
$$\frac{\partial\gamma}{\partial z} = \frac{\rho}{2f^2}\left(\frac{\partial f}{\partial\rho}\frac{\partial f}{\partial z} + \frac{\partial\varphi}{\partial\rho}\frac{\partial\varphi}{\partial z}\right). \tag{3.2}$$

Using the general formulas given in Sec. II, we proceed as follows. If we set $k = f$ and $h = \varphi$ in Eqs. (2.8) and (2.10), we obtain

$$k' = \frac{\rho}{f}, \quad \widetilde{D}h' = \frac{\rho}{f^2}\widehat{D}\varphi. \tag{3.3}$$

The second equation together with Eq. (3.1) shows that we may identify

$$h' = \omega, \tag{3.4}$$

without serious loss in generality. Hence the new function $\mathscr{E}'$ becomes

$$\mathscr{E}' = f' + i\varphi' = (\rho/f) + ij\omega. \tag{3.5}$$

For later convenience we note that Eq. (2.16) implies

$$\mathscr{E} = f + i\varphi = (\rho/f') + ij\omega', \tag{3.6}$$

where

$$\widetilde{D}\omega' = (\rho/f'^2)\widehat{D}\varphi'. \tag{3.7}$$

Combining the transformation $\mathscr{E} \to \mathscr{E}'$ with the transformation of the type (2.17), we finally obtain an infinite series of exact solutions

$$\cdots\mathscr{E}_{-2}—\mathscr{E}_{-1}\cdots\mathscr{E}_0—\mathscr{E}_1\cdots\mathscr{E}_2—\mathscr{E}_3\cdots. \tag{3.8}$$

We show that the series (3.8) has an interesting property. Let f_n, ω_n and γ_n be the metric functions determined by $\mathscr{E}_n$ ($n = 0, \pm 1, \pm 2 \ldots$). Then Eqs. (3.5) and (3.6) imply

$$\mathscr{E}_{2m} = (\rho/f_{2m-1}) + ij\omega_{2m-1},$$
$$\mathscr{E}_{2m+1} = (\rho/f_{2m+2}) + ij\omega_{2m+2}. \tag{3.9}$$

Since the functions $\mathscr{E}_{2m}$ and $\mathscr{E}_{2m+1}$ are related by the transformation of the type (2.17), we can conclude from the result of the Appendix that two metrics determined by $\mathscr{E}_{2m-1}$ and $\mathscr{E}_{2m+2}$ are connected by a coordinate transformation of the type

$$\begin{pmatrix} t \\ \phi \end{pmatrix} = \begin{pmatrix} d & jb \\ jc & a \end{pmatrix}\begin{pmatrix} \bar{t} \\ \bar{\phi} \end{pmatrix}, \tag{3.10}$$
$$a,b,c,d = i\text{-real},^{24} \quad ad + bc = 1.$$

IV. THE LEWIS SOLUTION

In the series (3.8), the solutions associated with $\mathscr{E}_1$ is well known. For example, if we start with the Schwarzschild solution, the function $\mathscr{E}_1$ determines the NUT solution.[19] In this section, we investigate the solutions associated with $\mathscr{E}_2$ and show that the Lewis solutions[23] are contained in this class.

Taking account of the property stated in the last paragraph of Sec. III, we first consider the metric functions f_{-1}, ω_{-1}, and γ_{-1} associated with $\mathscr{E}_{-1}$ and then determine the expressions of f_2, ω_2, and γ_2 in terms of the known functions f_0, ω_0, and γ_0. Since

$$\mathscr{E}_0 = f_0 + i\varphi_0 = (\rho/f_{-1}) + ij\omega_{-1},$$

we obtain

$$f_{-1} = \rho/f_0, \quad \omega_{-1} = -j\varphi_0. \tag{4.1}$$

The integration of Eq. (3.2) yields

$$\gamma_{-1} = \gamma_0 + \tfrac{1}{4}\ln[\rho] - \tfrac{1}{2}\ln[f_0]. \tag{4.2}$$

Substituting Eqs. (4.1) and (4.2) into the formulas (A5), (A6), and (A7), we have

$$f_2 = \rho f_0^{-1}[c^2 f_0^2 + (c\varphi_0 - d)^2],$$
$$f_2\omega_2 = j\rho f_0^{-1}[acf_0^2 + (a\varphi_0 + b)(c\varphi_0 - d)], \tag{4.3}$$
$$f_2^{-1}\exp(2\gamma_2) = \rho^{-1/2}\exp(2\gamma_0),$$

where

$$ad + bc = 1. \tag{4.4}$$

Now let us consider the three special cases in which the functions $\mathscr{E}_0$ are given by

$$1.\ \mathscr{E}_0 = e^{2\psi}, \quad 2.\ \mathscr{E}_0 = e^{2j\psi}, \quad 3.\ \mathscr{E}_0 = (1 + ij)\psi^{-1}, \tag{4.5}$$

where an i,j-real[24] function ψ satisfies

$$\nabla^2\psi = 0. \tag{4.6}$$

The first and second cases correspond to the Weyl solution and the third function determines a new solution. Since $\mathscr{E}_0 = f_0 + i\varphi_0$, we obtain the following results:

Case 1. $f_2 = \rho(d_1{}^2 e^{-2\psi} - c_1{}^2 e^{2\psi})$,

$$f_2\omega_2 = \rho(-b_1 d_1 e^{-2\psi} + a_1 c_1 e^{2\psi}), \tag{4.7}$$

$$f_2^{-1}\exp(2\gamma_2) = \rho^{-1/2} e^{2\lambda},$$

Case 2. $f_2 = \rho[(d_2{}^2 - c_2^2)\cos 2\psi + 2d_2 c_2 \sin 2\psi)]$,

$$f_2\omega_2 = [\rho - (b_2 d_2 - a_2 c_2)\cos 2\psi$$

$$- (a_2 d_2 + b_2 c_2)\sin 2\psi], \tag{4.8}$$

$$f_2^{-1}\exp(2\gamma_2) = \rho^{-1/2} e^{-2\lambda},$$

Case 3. $f_2 = \rho(d_1{}^2\psi - 2c_1 d_1)$,

$$f_2\omega_2 = \rho(-b_1 d_1\psi + a_1 d_1 + b_1 c_1), \tag{4.9}$$

$$f_2 \exp(2\gamma_2) = \rho^{-1/2},$$

where

$$a_1 = a, \quad b_1 = jb, \quad c_1 = jc, \quad d_1 = d,$$

$$\sqrt{2}a_2 = a + b, \quad \sqrt{2}b_2 = j(b - a), \quad \sqrt{2}\,c_2 = c + d,$$

$$\sqrt{2}d_2 = j(c - d), \quad a_i d_i - b_i c_i = 1 (i = 1,2),$$

$$\frac{\partial\lambda}{\partial\rho} = \rho\left[\left(\frac{\partial\psi}{\partial\rho}\right)^2 - \left(\frac{\partial\psi}{\partial z}\right)^2\right], \quad \frac{\partial\lambda}{\partial z} = 2\rho\frac{\partial\psi}{\partial\rho}\frac{\partial\psi}{\partial z}. \tag{4.10}$$

The solutions (4.7) and (4.8) agree with the Lewis solutions, and the solution (4.9) is the generalized result of Levy.[25] When $\psi = (\text{const})\times\ln[\rho]$ the above three solutions together describe the exterior gravitational field of a rotating infinite cylinder.[26]

APPENDIX

In this Appendix we consider the relation between the metric functions (f,ω,γ) and $(\bar f,\bar\omega,\bar\gamma)$ when the two complex functions

$$\mathscr{F} = (\rho/f) + ij\omega, \quad \bar{\mathscr{F}} = (\rho/\bar f) + ij\bar\omega, \tag{A1}$$

are related by

$$\bar{\mathscr{F}} = \frac{a\mathscr{F} + ib}{ic\mathscr{F} + d}, \tag{A2}$$

$$a,b,c,d = i\text{-real},^{24} \quad ad + bc = 1.$$

Since Eqs. (A1) and (A2) readily lead to

$$\begin{aligned}\bar f &= f[c^2\rho^2 f^{-2} + (jc\omega - d)^2],\\ \bar f\bar\omega &= jf[ac\rho^2 f^{-2} + (ja\omega + b)(jc\omega - d)],\end{aligned} \tag{A3}$$

the remaining problem is to express $\bar\gamma$ in terms of f, ω, and γ. The function $\bar\gamma$ may be obtained by the integration of Eq. (3.2). However, there exists a more convenient method for this case.[27] Note that by a coordinate transformation

$$\begin{pmatrix} t\\ \phi\end{pmatrix} = \begin{pmatrix} d & jb\\ jc & a\end{pmatrix}\begin{pmatrix}\bar t\\ \bar\phi\end{pmatrix}, \quad ad + bc = 1, \tag{A4}$$

the canonical metric

$$ds^2 = f\,dt^2 - 2f\omega dt d\phi + (f\omega^2 - f^{-1}\rho^2)\,d\phi^2$$

$$- f^{-1}e^{2\gamma}(d\rho^2 + dz^2)$$

changes into

$$ds^2 = f[c^2\rho^2 f^{-2} + (jc\omega - d)^2]\,d\bar t^2 - 2jf[ac\rho^2 f^{-2}$$

$$+ (ja\omega + b)(jc\omega - d)]\,d\bar t\,d\bar\phi + f[-a^2\rho^2 f^{-2}$$

$$- (ja\omega + b)^2]\,d\bar\phi^2 - f^{-1}e^{2\gamma}(d\rho^2 + dz^2).$$

Assuming that the second metric is also of the canonical form

$$\begin{aligned}ds^2 = \bar f d\bar t^2 &- 2\bar f\bar\omega\, d\bar t\, d\bar\phi\\ &+ (\bar f\bar\omega^2 - \bar f^{-1}\rho^2)\,d\bar\phi^2 - \bar f^{-1}\exp(2\bar\gamma)(d\rho^2 + dz^2),\end{aligned}$$

we obtain

$$\bar f = f[c^2\rho^2 f^{-2} + (jc\omega - d)^2], \tag{A5}$$

$$\bar f\bar\omega = jf[ac\rho^2 f^{-2} + (ja\omega + b)(jc\omega - d)], \tag{A6}$$

$$\bar f^{-1}\exp(2\bar\gamma) = f^{-1}\exp(2\gamma), \tag{A7}$$

$$\bar f\bar\omega^2 - \bar f^{-1}\rho^2 = f[-a^2\rho^2 f^{-2} - (ja\omega + b)^2]. \tag{A8}$$

Comparing Eqs. (A5) and (A6) with Eq. (A3), we can see that the transformation (A2) is equivalent to the coordinate transformation (A4). Hence Eq. (A7) is the general relation between the metric functions γ and $\bar\gamma$. Equation (A8) is merely a result of Eqs. (A5) and (A6) and gives us no new knowledge.

ACKNOWLEDGMENTS

I would like to thank Professor T. Takabayasi for his encouragement. I am also indebted to H. Okumura for fruitful discussions.

[1] J. Ehlers, in *Les théories relativistes de la gravitation* (CNRS, Paris, 1959).
[2] B.K. Harrison, J. Math. Phys. **9**, 1744 (1968).
[3] F.J. Ernst, Phys. Rev. **168**, 1415 (1968).
[4] R. Geroch, J. Math. Phys. **12**, 918 (1968).

[5]W. Kinnersley, J. Math. Phys **14**, 651 (1973).
[6]Y. Tanabe, Prog. Theor. Phys. **57**, 840 (1977); **60**, 136 (1978).
[7]W.B. Bonner, Z. Physik **161**, 439 (1961); **190**, 444 (1966).
[8]R.M. Misra *et al.*, Phys. Rev. D 7, 1587 (1973); D8, 1942 (1973).
[9]Y. Tanabe, J. Math. Phys. **19**, 1808 (1978).
[10]D. Kramer and G. Neugebauer, Commun. Math. Phys. **10**, 132 (1968).
[11]G. Neugebauer and D. Kramer, Ann. Phys. (Leipz.) **24**, 62 (1969).
[12]R.P. Kerr, Phys. Rev. Lett **11**, 237 (1963).
[13]C. Reina and A. Trevas, J. Math. Phys. **16**, 834 (1975).
[14]F.P. Esposito and L. Witten, Phys. Rev. D **8**, 3302 (1973); Gen. Rel. Grav. **6**, 387 (1975).
[15]D. Kramer and G. Neugebauer, Ann. Phys. (Leipz.) **24**, 59 (1969).
[16]F.J. Ernst, J. Math. Phys. **17**, 54 (1976).
[17]Y. Tanabe, "New Exact Einstein–Maxwell Fields for a Massless Charge," Gen. Rel. Phys., to be published.
[18]E.T. Newman *et al.*, J. Math. Phys. **6**, 918 (1965).
[19]E.T. Newman, L. Tamburino, and T. Unti, J. Math. Phys. **4**, 915 (1963); M. Demanski and E.T. Newman, Bull. Acad. Pol. Sci. **24**, 653 (1966).
[20]R. Catenacci and D. Alonso, J. Math. Phys. **17**, 2232 (1976).
[21]F.J. Ernst, Phys. Rev. **167**, 1175 (1968).
[22]H. Weyl, Ann. Phys. (Leipz.) **54**, 117 (1917); J.L. Synge, *Relativity, The General Theory* (North-Holland, Amsterdam, 1960), p. 309.
[23]T. Lewis, Proc. Roy. Soc. (London), A **136**, 176 (1932).
[24]We use the term "*i*-complex" or "*i*-real" ("*j*-complex" or "*j*-real") according as the quantity in question contains or does not contain the imaginary unit *i* (the symbol *j*).
[25]H. Levy, Nuovo Cimento B **56**, 253 (1968).
[26]H. Davies and T. A. Caplan, Proc. Camb. Philos. Soc. **69**, 325 (1971); E. Frehland, Commun. Math. Phys. **23**, 127 (1971); F.J. Tipler, Phys. Rev. D **9**, 2203 (1974).
[27]The method is partly suggested by H. Okumura's Masters thesis (Nagoya University, 1978).

A remark about the duality for non-Abelian lattice fields

J. Bellissard[a)]

Université de Provence et C.N.R.S., Marseille, France
(Received 1 December 1978)

We give an example of a statistical mechanical model of spins with value in the non-Abelian group $\mathfrak{S}_3$ for which the Kramers–Wannier duality holds and gives rise to self-dual interaction.

I. INTRODUCTION

The idea of duality in statistical mechanics, goes back to Kramers and Wannier[1] who remark that the pressure of the ISING model in two dimensions obeys the following relation:

$$p(\beta) = p(\beta^*) + \mathrm{Log}(\tfrac{1}{4}\sinh 2\beta),$$

where β in the inverse temperature and β^* is defined by

$$(e^{2\beta^*} - 1)(e^{2\beta} - 1) = 2.$$

This transformation connects the pressure at low and high temperatures. In particular, if p has a unique singularity in β, it appears at the point β_c such that

$$\beta_c = \tfrac{1}{2}\mathrm{Log}(1 + \sqrt{2}).$$

This value is precisely the critical temperature given by Onsager[2] for this model.

Very recently the idea became of great physical interest in the context of QCD, in view of the picture proposed by Mandelstam.[3,4] The idea is to investigate a kind of symmetry between electric and magnetic components of the field, which can be classically realized through the interchange of g and $1/g$, where g is the coupling constant. At high g the system is described by electric flux tubes that bind quarks, whereas at small g, the quarks are no longer confined, but there are magnetic flux tubes that bind monopoles. In an $\mathrm{SU}(N)$ gauge field theory the topological quantum numbers which classify these tubes are elements of $\mathbb{Z}_N$. It follows that a good qualitatively relevant model is given by a lattice gauge field with $\mathbb{Z}_N$ as a gauge group.

The duality transformation for such a model is very simple to exhibit.[5] For the simple case of $\mathbb{Z}_2$, it has first been exhibited by Wegner,[6] who proved the existence of a self-dual point as in the Kramers–Wannier work. More recently such a study has been extended for the most general interaction between lattice sites with spins in an Abelian group.[7] This transformation has been used many times[8,9] to get information on the Gibbs states in statistical mechanics.

In the case of the discrete or compact Abelian group the duality transformation is nothing but the Fourier transform on the group.[10,11] In particular, finding a self-dual interaction is equivalent to finding a positive function invariant under the Fourier transform. It is proved that, given a finite Abelian group isomorphic to its dual, the family of self-dual interactions is indexed by $(p - 1)$ parameters, where p is the dimension of the eigenspace of the Fourier operator, corresponding to the eigenvalue 1.

However, a question arises as to what is the duality transformation for a non-Abelian group. If one believes that it is also given in term of the Fourier theory, the dual model would describe interaction between spins taking values in the set of irreducible representations of the initial group (at least if it is compact). An idea consists in looking at the smallest non-Abelian finite group which is the permutation group $\mathfrak{S}_3$ of $\{1,2,3\}$. In the two-dimensional case, for nearest neighbors interaction, we show that it reduces to an Abelian model on the group $\mathbb{Z}_2\times\mathbb{Z}_3$. This remark made the model completely trivial; however it has dual models on $\mathfrak{S}_3$, and there is a one-parameter family of self-dual interaction.

It is helpful to look at the corresponding lattice group theory in four dimensions. Indeed in that case the map between $\mathfrak{S}_3$ and $\mathbb{Z}_2\times\mathbb{Z}_3$ no longer transforms the gauge field interaction into an Abelian gauge field. We get a complicated four-body interaction when expressed in term of $\mathbb{Z}_2\times\mathbb{Z}_3$. This expresses the lack of Abelianess.

II. DUALITY FOR ABELIAN GROUPS[7]

A model of statistical mechanics is conventionally described by giving a lattice, say $\mathbb{Z}^\nu$, and for each site x a random variable called the spin.[12] The model is defined by a family a Hamiltonian (H_Λ), where Λ runs over the finite subsets of $\mathbb{Z}^\nu$ which describes the energy of the configuration of spins inside Λ, given some configuration outside. The thermodynamical properties of such a model can be extracted from the partition function,

$$\mathbb{Z}_\Lambda(\beta) = \sum_{g_\Lambda} \exp(-\beta H_\Lambda) \tag{II.1}$$

(where the sum runs over the set of configurations inside Λ) be considering the pressure

$$p(\beta) = \lim_{\Lambda\uparrow\mathbb{Z}^\nu} \frac{1}{|\Lambda|}\mathrm{Log}\mathbb{Z}_\Lambda(\beta). \tag{II.2}$$

We want to examine the case for which the spin at each site takes its values in an Abelian group G. For simplicity we will restrict ourself to the case of finite groups. However,

[a)]Postal address: C.N.R.S.–Luminy–Case 907, Centre de Physique Théorique, F-13288 Marseille Cedex 2, France.

 0022-2488/79/071490-04$01.00

most of the results below can be extended to the case of locally compact Abelian group. On the other hand, we will consider first the two-dimensional case.

Then the set of configurations is $(G)^{\mathbb{Z}^2}$, and the Hamiltonian is described by:

$$-\beta H_\Lambda(g) = \sum_{\{xy\}\in\Lambda} h(g_x g_y^{-1}), \tag{II.3}$$

where the sum runs over the set of pairs of nearest neighbors included in Λ, and h is a function on G. The corresponding pressure will depend on G and h; it will be denoted $p(G;h)$. It is a convex function of h.

Now let G^* be the dual group of G. If f is a function on G, its Fourier transform is given by the formula[10]:

$$\hat f(\chi) = |G|^{-1/2} \sum_{g\in G} \chi(g) f(g), \tag{II.4a}$$

$$f(g) = |G^*|^{-1/2} \sum_{\chi\in G^*} \overline{\chi(g)}\, \hat f(g). \tag{II.4b}$$

(Here $|G| = |G^*|$ denotes the number of elements of G.)

The following theorem summarizes the main result of this section:

Theorem 1:[6,7] Let h be a real function on G, such that the function h^* on G^* defined by

$$\exp[h^*(\chi)] = (\exp h)\widehat{\ }(\chi) \tag{II.5}$$

is real. Then,

$$p(G^*;h^*) = p(G,h). \tag{II.6}$$

The proof of this theorem is almost standard and can be found in Refs. 6, 7, 13 and 14, for instance. The crucial step is the solution of Kirchoff's law: let (χ_{xy}) be a family of characters of G, indexed by pairs of nearest neighbors in $\mathbb{Z}^2$ and such that $\chi_{xy} = 1$ if $\{xy\}$ is not included in Λ, and

$$\prod_y \chi_{xy} = 1, \quad x\in\mathbb{Z}^2. \tag{II.7}$$

Then, on this dual lattice (the lattice whose sites are the unit squares of $\mathbb{Z}^2$) one can find a family of character χ_{x^*} such that

$$\chi_{xy} = \chi_{x^*}\chi_{y^*}^{-1}. \tag{II.8}$$

Here $\{xy\}$ is the common side of the squares x^* and y^* such that (xy) is directly oriented in x^* and (y,x) in y^*.

On the other hand, some care has to be taken with the boundary conditions. But it is known that the pressure is independent of the choice of boundary condition.[12]

In this theorem, we assume that both h and h^* are real functions. This is equivalent to saying that both e^h and $(e^h)^*$ are positive functions. The following lemma characterizes this class of functions. We need only to recall that an Hamiltonian is called ferromagnetic if it is a real function with positive fourier coefficients except possibly the first one.

Lemma 2: Let Γ be a compact group, and H be a continuous function on Γ; then the following are equivalent:

(1) $\exp\beta H$ is a positive function on Γ with positive type for any $\beta > 0$,

(2) H is a ferromagnetic Hamiltonian on Γ.

The proof of (2)$\Rightarrow$(1) is immediate. The converse can be seen by computing the Fourier coefficients of $\exp\beta H$ and letting β go to zero.

We remark also that (1) implies that the measure $Z^{-1}e^{\beta H}d\gamma$ with $Z = \int e^{\beta H}\, d\gamma$, is a probability measure for which the first Griffiths' inequalities hold.[15] Therefore, the restrictions imposed on h in Theorem 1 are of great physical significance.

We are now ready to investigate the occurence of the self-dual model. We need first to have an isomorphism between G and G^*, if G is finite, it has the same cardinality as G^*, and therefore we can identify each by the other; if G is not finite this is not true in general but in many cases this can be done. For example, if $G = \mathrm{U}(1)\times\mathbb{Z}$, G^* is $\mathbb{Z}\times\mathrm{U}(1)$ and is isomorphic to G. This is the reason why the Berezinski–Villain model works from the point of view of duality.[16,17]

In this case, a self-dual model is given by a function h on $G\sim G^*$ which differs from h^* by a constant since the energy is defined up to a real constant. This means that $\exp h$ is an eigenfunction of the Fourier operator. Since $\exp h$ is positive, the corresponding eigenvalue is necessarily equal to 1, because the Fourier operator is unitary (Parseval's formula[10]). If G is finite the following function is always self-dual:

$$h(g) = |G|^{1/2}\beta_c\,\delta(g) \tag{II.9}$$

with

$$\beta_c = |G|^{-1}\,\mathrm{Log}(1 + |G|^{1/2}), \tag{II.10}$$

which describes the Pott's model with $|G|$-levels.[14] This example shows that the set S of positive Fourier invariant functions on a finite group is never empty: It is the intersection of the cone of positive functions on G with the eigenspace of the Fourier operator corresponding to the eigenvalue 1. Denoting by p the dimension of this eigenspace, we remark that $1\leqslant p\leqslant |G|$. S is then a closed convex cone in a space of dimension p. Taking into account that $\exp h$ and $\exp(h + c)$ define the same physical system, the set of self-dual solutions can be seen as the set of functions h on S, such that $\exp h\in S_1$, where S_1 is a basis of the cone S. It is therefore homeomorphic to a closed convex set of dimension $p - 1$.

To finish with this section let us give some examples: For $G = \mathbb{Z}_2$ or $\mathbb{Z}_3$ one has $p = 1$. In this case there is only one self-dual model,[5] For $G = \mathbb{Z}_4$ one finds $p = 2$.

There are two relevant eigenfunctions of the Fourier operator: $\exp h_0 = 1 + z^2$ and $\exp h_1 = 2 + z + \bar z$; more generally if $N\geqslant 4$ and $G = \mathbb{Z}_N$, then $p\geqslant 2$. This result explains the difference between the $\mathbb{Z}_N$-gauge field on a lattice for $N = 2,3$ and $N\geqslant 4$[5] (cf. Sec. III).

III. EXAMPLE OF A SELF-DUAL MODEL WITH A NON-ABELIAN GROUP

Let us now come to the main example of this paper. We consider the same model in which we choose G to be the permutation group $\mathfrak{S}_3$ of three elements. Since $\mathfrak{S}_3$ is non-Abelian the previous theory cannot be applied directly. However, we will see that it reduced in fact to an Abelian group as far as statistical mechanics is concerned.

Let us first recall that $\mathfrak{S}_3$ has six elements[17]:

$g =$	1	J	$\bar{J}$	R_1	R_2	P_3
$g\{1,2,3\} =$	(1,2,3)	(2,3,1)	(3,1,2)	(1,3,2)	(3,2,1)	(2,1,3)

(III.1)

It has three classes of irreducible representations: the trivial one ϵ, the signature σ, and the two-dimensional one ρ in which $\rho(J)$ is the rotation of angle $2\pi/3$, and $\rho(R_k)$ is the symmetry about the line of argument $2\pi k/3$. It will be called the *real representation* of ρ.

We claim that there is a bijection between $\mathfrak{S}_3$ and $\mathbb{Z}_2\times\mathbb{Z}_3$ defined by the map $g\rightarrow(\sigma(g),z(g))$ where $z(g)$ is defined as follows:

$g =$	1	J	$\bar{J}$	R_1	R_2	R_3
$Z(g) =$	1	$e^{2i\pi/3}$	$e^{-2i\pi/3}$	1	$e^{2i\pi/3}$	$e^{-2i\pi/3}$

(III.2)

The crucial fact comes from the following property of z:

$$z(g_1 g_2^{-1}) + \overline{z(g_1 g_2^{-1})} = z(g_1)\overline{z(g_2)} + \overline{z(g_1)}z(g_2). \tag{III.3}$$

Due to this bijection, every function h on $\mathfrak{S}_3$ can be associated to a function $\tilde{h}$ on $\mathbb{Z}_2\times\mathbb{Z}_3$, by

$$h(g) = \tilde{h}(\sigma(g),z(g)). \tag{III.4}$$

Using the Fourier decomposition both for $\mathfrak{S}_3$ and $\mathbb{Z}_2\times\mathbb{Z}_3$ we get the following correspondance:

$$h(g) = h_\epsilon + h_\sigma\,\sigma(g) + 2\,\mathrm{tr}\left(\begin{bmatrix} h_{11} & h_{12} \\ h_{21} & h_{22}\end{bmatrix}\rho(g)\right) = \tilde{h}_{11} + \tilde{h}_{-1,1}\,\sigma(g) + \tilde{h}_{1j}\,z(g) + \tilde{h}_{1,\bar{j}}\,\bar{z}(g) + \tilde{h}_{1j}\,\sigma(g)z(g) + \tilde{h}_{1\bar{j}}\,\sigma(g)\bar{z}(g), \tag{III.5}$$

with

$$\tilde{h}_{11} = h_\epsilon,\quad \tilde{h}_{1,j} = h_{11} - \mathrm{i}h_{21},\quad \tilde{h}_{-1,j} = h_{22} + 1h_{21},\quad \tilde{h}_{11} = h_\sigma,\quad \tilde{h}_{1,\bar{j}} = h_{11} + \mathrm{i}h_{21},\quad \tilde{h}_{-1,\bar{j}} = h_{22} - 1h_{21}. \tag{III.6}$$

In the following we shall restrict ourselves to the case of ferromagnetic interaction for which

$$h_\sigma \geqslant 0 \text{ and } h_\rho = \begin{bmatrix} h_{11} & h_{12} \\ h_{21} & h_{22}\end{bmatrix} \text{ is a positive matrix.}$$

Without loss of generality we can always choose h_ρ to be diagonal; it is then simple to verify

$$h(g_1 g_2^{-1}) = \tilde{h}(\sigma(g_1)\sigma(g_2), z(g_1)\overline{z(g_2)}) \tag{III.7}$$

[using Eq. (III.3)]. This property is also equivalent to

$$\tilde{h}(\sigma,z) = \tilde{h}(\sigma\bar{z}) = \overline{\tilde{h}(\sigma,z)} \tag{III.8a}$$

or

$$h_\epsilon, h_\sigma, h_{11}, h_{22} \text{ real and } h_{21} = h_{12} = 0. \tag{III.8b}$$

If h is a function on $\mathfrak{S}_3$ for which (III.8b) holds, it will be called real and symmetric.

The main result of this section is summarized below, and is a simple application of (III.7).

Theorem 3: Let h be a real symmetric function on $\mathfrak{S}_3$. Then

$$p(\mathfrak{S}_3;h) = p(\mathbb{Z}_2\times\mathbb{Z}_3;\tilde{h}). \tag{III.9}$$

Equipped with this result, we can use the duality theory on $\mathbb{Z}_2\times\mathbb{Z}_3$. If $\tilde{h}^*$ is the dual function of $\tilde{h}$, it is not hard to see that it satisfies (III.8a) because (III.8a) is invariant by taking the exponential, the Fourier transform, and the logarithm. It follows that one can find a function h^* on $\mathfrak{S}_3$ such that $\widetilde{(h^*)} = (\tilde{h})^*$, which allows us to get the duality relation:

$$p(\mathfrak{S}_3;h) = p(\mathfrak{S}_3;h^*). \tag{III.10}$$

In order to get self-dual models on $\mathfrak{S}_3$, it is then sufficient to classify the self-dual interactions on $\mathbb{Z}_2\times\mathbb{Z}_3$. Coming back to the conclusions of the last section we need to find the eigenspace of the Fourier transform on $\mathbb{Z}_2\times\mathbb{Z}_3$, corresponding to the eigenvalue 1. It turns out that this space is generated by the two following functions (where $j = \exp 2i\pi/3$):

$(\sigma,z) =$	(1,1)	$(-1,1)$	$(1,j)$	$(1,\bar{j})$	$(-1,j)$	$(-1,\bar{j})$
$F_0(\sigma,z) =$	$(\sqrt{2}+1)(\sqrt{3}+1)$	$\sqrt{3}+1$	$\sqrt{2}+1$	$\sqrt{2}+1$	1	1
$F_1(\sigma,z) =$	$(\sqrt{2}-1)(\sqrt{3}-1)$	$1-\sqrt{3}$	$1-\sqrt{2}$	$1-\sqrt{2}$	1	1

(III.11)

The set of functions invariant under the Fourier transform is described by the function

$$\exp h_\alpha(g) = F_2(\sigma(g),z(g)) = \alpha F_1(\sigma(g),z(g)) + (1-\alpha)F_0(\sigma(g),z(g)), \tag{III.12}$$

with

$$2\alpha - 1 \leqslant 1/\sqrt{3}.$$

The particular value $\alpha = \frac{1}{2}$ gives rise to the Pott's model[18] with six levels:

$$h_{1/2}(g) = h_\epsilon + \beta_c[\sigma(g) + 2\,\mathrm{tr}\rho(g)] = h_\epsilon + \sigma\beta_c\delta(g), \quad \beta_c = \tfrac{1}{6}\,\mathrm{Log}(1+\sqrt{6}). \tag{III.13}$$

This is the only self-dual solution which does not break the global left $\mathfrak{S}_3$-symmetry $[(g_n)_{n\in Z^2}$ transformed into $(gg_n)_{n\in Z^2}$ with $g\in\mathfrak{S}_3]$.

IV. CONCLUDING REMARKS

The theory developed in Sec. II can be extended for any kind of interaction in any dimension. An exhaustive study can be found in Ref. 7.

We want to concentrate our attention on the special example given by an Abelian gauge field on a lattice.[8,19] The sites are then replaced by the oriented links of $\mathbb{Z}^\nu$. A plaquette is a set of four links bordering a unit two-dimensional square of the lattice. Given a plaquette $P'(l_1,l_2,l_3,l_4)$ we associate the energy

$$H_P(g) = h\,(g_{l_1}\,g_{l_2}\,g_{l_3}\,g_{l_4}), \tag{IV.1}$$

where we assume $g_l = g_{\bar{l}}^{-1}$ if $\bar{l}$ is the link h with the opposite orientation, and h is a function on the group G. The full Hamiltonian is given by

$$H_\Lambda(g) = \sum_{P\subset\Lambda} H_P(g). \tag{IV.2}$$

As in Sec. II, we can define the pressure which is a function of G, and a convex function of h.

In the special case of dimension *four*, the duality theory applies as well as in the case studied previously. We get

$$p(G;h) = p(G^*;h^*) \quad \text{if } \nu = 4. \tag{IV.3}$$

We can generalize to the case of a non-Abelian compact group by assuming only that h is a central function on G. Unfortunately if $G = \mathfrak{S}_3$, the model obtained is no longer equivalent to a lattice gauge field with respect to the Abelian group $\mathbb{Z}_2\times\mathbb{Z}_3$. This is essentially due to the fact that $\mathrm{Re}\,\mathrm{tr}\rho(g_1g_2g_3g_4)$ cannot be expressed as a function of $\sigma(g_1)\sigma(g_2)\sigma(g_3)\sigma(g_4)$ and $\mathrm{Re}(z(g_1)z(g_2)z(g_3)z(g_4))$ only. For this reason, the corresponding Abelian model obtained after the transformation $g\rightarrow(\sigma(g),z(g))$has a complicated four-body interaction which does not give rise to a lattice field: This is a characterization of the lack of Abelianess in this theory.

A dual model can be constructed, however, and allows us to expect the existence of a small temperature expansion.

ACKNOWLEDGMENTS

I am thankful to C.P. Korthal Altes, C. Itzykson, and A. Messager for many discussions and remarks, and to R. Stora for reading the manuscript.

Note added in the proof: Since the paper was written, J.M. Drouffe, C. Itzykson, and J.B. Zuber have proven that such a work can be extended to any solvable group.[20]

I would also like to thank Professor Gruber for information concerning the proof of duality in the Abelian case.[7,14]

1H.A. Kramers and G.H. Wannier, Phys. Rev. **60**, 252 (1941).
2L. Onsager, Phys. Rev. **65**, 117 (1944).
3S. Mandelstam, "Extended systems in field theory," Phys. Rev. C **23**, 245–9 (1976).
4G. t' Hooft, in *High Energy Physics Conference*, Proceedings of the international EPS Conference in Palermo, June 1975, edited by A. Zichichi (Editrice Compositori, Bologna, 1976).
5C.P. Korthals Altes, "Duality for $Z(N)$ gauge theory," to be published in Nucl. Phys.
6F. Wegner, J. Math. Phys. **12**, 2259 (1971).
7C. Gruber, A. Hintermann, and D. Merlini, "Group analysis of classical lattice systems," in *Lecture Notes in Physics*, No. 60 (Springer, New York, 1977).
8R. Balian, J.M. Drouffe, and C. Itzykson, Phys. Rev. D **11**, 2098 (1975).
9A. Messager and S. Miracle-Sole, Commun. Math. Phys. **40**, 187 (1975).
10A. Weil, *L'intégration dans les groupes topologiques et ses applications* (Hermann, Paris, 1965).
11J.P. Serre, *Représentations Linéaires des groupes finis* (Hermann, Paris, 1971), 2nd printing.
12D. Ruelle, *Statistical mechanic: Rigorous results* (Benjamin, New York, 1969).
13F.Y. Wu and Y.K. Wang, J. Math. Phys. **17**, 439 (1976).
14C. Gruber, "Cours de 3*e* cycle," lecture given at Louvain-la-Neuve, Belgium (1977).
15J. Ginibre, Commun. Math. Phys. **16**, 310 (1970).
16V.L. Berenzinski, Sov. Phys. JETP **32**, 493 (1971).
17J. Villain, J. Phys. (Paris) **36**, 581 (1975).
18R.B. Potts, Proc. Cambridge Philos. Soc. **48**, 106 (1952).
19K. Osterwalder and E. Seiler, Proceedings of Cargése Summer School (1976).
20J.M. Drouffe, C. Itzykson, and J.B. Zuber, "Lattice Model with a solvable Symmetry Group," to appear in Nucl. Phys. B.

Dust distribution (cylindrically symmetric) in nonrigid rotation in Brans–Dicke theory

N. Bandyopadhyay

Physics Department, Presidency College, Calcutta, India
(Received 22 May 1978)

Solutions of the field equations of Brans–Dicke theory for the case of stationary cylindrically symmetric dust distribution in nonrigid rotation are studied. It is seen that singularities invariably occur in contrast to the singularity-free solution obtained by Maitra (1966) in the corresponding case in general relativity.

1. INTRODUCTION

In the general theory of relativity, two distinct types of stationary, cylindrically symmetric solutions are known. The first type is originally due to Van Stockum (1937)[1] and the matter in it is rigidly rotating. If the matter distribution is cut off at a certain radial distance from the axis of symmetry, the corresponding solution can be matched to an exterior Lewis vacuum solution. For a certain range of parameters the solution then becomes globally nonsingular with no closed timelike lines and asymptotically flat in the cylindrical sense. If, however, one considers an unbounded distribution of matter, the corresponding solution contains closed timelike lines and also a singularity at a finite proper radial distance. In an earlier communication, the present author examined the analogous solution in Brans–Dicke (B–D) theory and found that the undesirable features, i.e., closed timelike lines and singularity persist. The second type is due to Maitra (1966).[2] Here the rotation is nonrigid and the solution, even though the matter distribution extends over all space, is free of singularities as well as closed timelike lines. If the presence of universal rotation is considered to be against the Mach principle, then Maitra's solution may be taken as indicating a lack of consistency of the general theory of relativity with Mach's principle. In the present paper, the corresponding solution, again the matter distribution extending over all space, is examined in B–D theory. We find that there are singularities and as such there is no acceptable solution. This leads us to the conclusion that, at least in some respects, the B–D theory is more consistent with Mach's principle than the general theory of relativity.

2. THE FIELD EQUATIONS AND THEIR INTEGRATION

With the stationary, cylindrically symmetric line element

$$ds^2 = dt^2 - e^{2\psi}(dr^2 + dz^2) - l d\Phi^2 - 2m d\Phi dt \qquad (2.1)$$

the geodesic equations of motion and the condition $v^\mu v_\mu = 1$ give the following nonvanishing components of the contravariant velocity 4-vector of matter:

$$v^0 = l_1(l_1^2 - 4lm_1^2 + 4mm_1l_1)^{-1/2},$$

$$v^3 = -2m_1(l_1^2 - 4lm_1^2 + 4mm_1l_1)^{-1/2}. \qquad (2.2)$$

The field equations are

$$\left(\frac{mm_1}{2D}\right)_1 = \frac{4\pi\sqrt{-g}}{\phi}\rho(v^0v_0 - v^3v_3) - \frac{mm_1}{2D}\frac{\phi_1}{\phi} + \frac{1}{2}\frac{\Box\phi}{\phi}\sqrt{-g}, \qquad (2.3)$$

$$\left(\frac{l_1 + mm_1}{2D}\right)_1 = \frac{4\pi\sqrt{-g}}{\phi}\rho(v^3v_3 - v^0v_0) - \frac{l_1 + mm_1}{2D} \times \frac{\phi_1}{\phi} + \frac{1}{2}\frac{\Box\phi}{\phi}\sqrt{-g}, \qquad (2.4)$$

$$\left(\frac{m_1}{2D}\right)_1 = \frac{8\pi\sqrt{-g}}{\phi}\rho v^3 v_0 - \frac{m_1}{2D}\frac{\phi_1}{\phi}, \qquad (2.5)$$

$$\left(\frac{ml_1 - lm_1}{2D}\right)_1 = \frac{8\pi\sqrt{-g}}{\phi}\rho v^0 v_3 - \frac{ml_1 - lm_1}{2D}\frac{\phi_1}{\phi}, \qquad (2.6)$$

$$-D\psi_{11} - D_1\psi_1 = \frac{4\pi\sqrt{-g}}{\phi}\rho + D\psi_1\frac{\phi_1}{\phi} - \frac{1}{2}\frac{\Box\phi}{\phi}\sqrt{-g}, \qquad (2.7)$$

$$-D\psi_{11} + \frac{m_1^2}{2D} + D_1\psi_1 - D_{11} = \frac{4\pi\sqrt{-g}}{\phi}\rho + \omega D\frac{\phi_1^2}{\phi^2} - D\psi_1\frac{\phi_1}{\phi} + D\frac{\phi_{11}}{\phi} - \frac{1}{2}\frac{\Box\phi}{\phi}\sqrt{-g}, \qquad (2.8)$$

and

$$\Box\phi = \frac{8\pi\rho}{3 + 2\omega}, \qquad (2.9)$$

with $D^2 = l + m^2$, and the subscript 1 indicating differentiation with respect to r.

From Eq. (2.3) and (2.4) it follows that

$$[D_1\phi]_1 = \Box\phi\sqrt{-g} = -[D\phi_1]_1,$$

the latter equality following from the definition of $\Box\phi$.

Thus,

$$D\phi = ar \quad (a = \text{const}). \qquad (2.10)$$

Also from Eqs. (2.7), (2.9), and (2.10), it follows

 0022-2488/79/071494-02$01.00

$$\psi = (1+\omega)\ln\phi \tag{2.11}$$

Again Eqs. (2.3), (2.5), and (2.10) give

$$\frac{m_1^2\phi^2/2ar - \frac{1}{2}\Box\phi\sqrt{-g}}{\frac{1}{2}(3+2\omega)\Box\phi\sqrt{-g}} = \frac{[(a^2r^2/\phi^2)_1]^2 + (4a^2r^2/\phi^2)m_1^2}{[(a^2r^2/\phi^2)_1]^2 - (4a^2r^2/\phi^2)m_1^2}. \tag{2.12}$$

But from Eqs. (2.7), (2.8), and (2.11),

$$\frac{r^2\phi_1^2}{\phi^2} - 2r\frac{\phi_1}{\phi} = \frac{m_1^2\phi^2}{2a^2(2+\omega)}. \tag{2.13}$$

Thus, eliminating m_1 between (2.12) and (2.13), one has

$$\frac{d\Theta}{dx}[(3+2\omega)(\Theta^2-2\Theta)+1] = (\Theta^2-2\Theta)[(3+2\omega)(\Theta^2-2\Theta)-1], \tag{2.14}$$

where

$$r\frac{d}{dr} \equiv \frac{d}{dx} \quad \text{and} \quad \frac{d}{dx}[\ln\phi] = \Theta.$$

Equation (2.4) has the integral

$$\left(\frac{\Theta-1-\lambda}{\Theta-1+\lambda}\right)^{2/\lambda}\frac{\Theta}{\Theta-2} = Ar^2, \tag{2.15}$$

where A is the constant of integration and $\lambda = [(4+2\omega)/(3+2\omega)]^{1/2}$. An approximate solution corresponding to Eq. (2.15) is as follows:

$$\phi \approx Br^2/[1 - Ar^{2(1+2\epsilon)}]^{1-2\epsilon},$$

$$m \approx C(1+2\epsilon)\ln r,$$

where B and C are constants and ϵ is a small positive quantity given by

$$\lambda = 1 + \epsilon.$$

Other unknowns are now easily obtainable.

3. DISCUSSION

The general behavior of the solutions corresponding to Eq. (2.15) can be studied without going into the approximate solution derived above. One can study the three distinct possibilities

(i) $A = 0$,

(ii) $A = \infty$,

(iii) $A \neq 0$ and finite.

Case (i): $A = 0$: From Eq. (2.15) it follows that either $\Theta = 0$ or $\Theta = 1 + \lambda$. In the case $\Theta = 0$ and hence $\phi = \text{const}$, Eq. (2.9) makes ρ vanish. With $\Theta = 1 + \lambda$,

$$\phi = c_1 r^{1+\lambda}$$

and

$$m = c_2 r^{-\lambda}.$$

Evidently, $\phi \to 0$ and $m \to \infty$ as $r \to 0$; and $\phi \to \infty$ and $m \to 0$ as $r \to \infty$.

Case (ii): $A = \infty$: From Eq. (2.15), it follows that either $\Theta = 1 - \lambda$ or $\Theta = 2$. Both of these cases admit of singular solutions similar to case (i).

Case (iii): $A \neq 0$ *and finite*: In this case, it follows that at $r = 0$ either $\Theta = 0$ or $\Theta = 1 + \lambda$. So that as $r \to 0$,

either $\Theta \to O(r^2)$

or $\Theta \to 1 + \lambda + O(r^\lambda)$.

In the first case ϕ tends to a constant value while in the second case ϕ tends to vanish as r^λ. The latter case would, therefore, mean a singularity of the ϕ-field. Again as $r \to \infty$,

either $\Theta \to 2 + O(1/r^2)$

or $\Theta \to 1 - \lambda + O(1/r^\lambda)$.

In the first case $\phi \to r^2\exp(K_3/r^2)$ (K_3 being a positive constant), which explodes as $r \to \infty$. In the second case $\phi \to r^{1-\lambda}\exp(K_4/r^\lambda)$ (K_4 being a positive constant) which tends to zero as $r \to \infty$ (since $\lambda > 1$).

The above considerations show that the field equations of B–D theory admit of no singularity free interior solutions for stationary, cylindrically symmetric dust distribution in nonrigid rotation. This situation thus permits us to say that at least in this case, the B–D theory seems more consistent with Mach's principle than the general theory of relativity.

ACKNOWLEDGMENT

The author is thankful to Professor A.K. Raychaudhuri of the Physics Department, Presidency College, Calcutta, for suggesting the problem and extending helpful comments.

[1]S.C. Maitra, J. Math. Phys. 7, 1025 (1966).
[2]W.J. Van Stockum, Proc. Roy. Soc. Edinburgh 57, 135 (1937).

Two-dimensional lumps in nonlinear dispersive systems

J. Satsuma

Department of Applied Mathematics and Physics, Faculty of Engineering, Kyoto University, Kyoto 606, Japan

M. J. Ablowitz

Department of Mathematics, Clarkson College of Technology, Potsdam, New York 13676

(Received 10 November 1978)

Two-dimensional lump solutions which decay to a uniform state in all directions are obtained for the Kadomtsev–Petviashvili and a two-dimensional nonlinear Schrödinger type equation. The amplitude of these solutions is rational in its independent variables. These solutions are constructed by taking a "long wave" limit of the corresponding N-soliton solutions obtained by direct methods. The solutions describing multiple collisions of lumps are also presented.

1. INTRODUCTION

Recent studies of nonlinear waves in dispersive systems have shown that the following two nonlinear evolution equations are generic in describing one-dimensional wave propagation in far fields: the Korteweg–deVries (KdV) equation,

$$u_t + 6uu_x + u_{xxx} = 0, \tag{1.1}$$

for systems of long waves with weak nonlinearity, and the nonlinear Schrödinger equation,

$$iA_t + A_{xx} \pm |A|^2 A = 0, \tag{1.2}$$

for wave systems which have, in a linear approximation, plane waves with high frequency oscillations.[1] Both equations have remarkable properties in that they can be solved exactly as initial value problems by applying the inverse scattering transform (IST). The analytical method shows that stable solitary waves, called solitons, play an important role in the solution of these equations.[2] In particular, it is well known that they have special solutions describing multiple collisions of solitons (N-soliton solutions), which are characteristic of nonlinear evolution equations solvable by IST.

Multidimensional effects relating to both of these equations have been studied by many researchers.[3] A two-dimensional KdV equation was first proposed by Kadomtsev and Petviashvili. They also considered the question of the stability of one-dimensional solitons with respect to transverse perturbations. The equation, which we shall call the Kadomtsev–Petviashvili (K–P) equation, may be written as

$$(u_t + 6uu_x + u_{xxx})_x + \alpha u_{yy} = 0, \tag{1.3}$$

where $\alpha(=\pm 1)$ is a parameter depending on the dispersive property of the system (the upper sign is usually refered to as negative dispersion and vice-versa). A two-dimensional generalization of (1.2) describes, for example, modulated two-dimensional wave packets on the surface of a liquid. In a long wave length limit a reduced form of the equation may be expressed as

$$iA_t - \sigma_1 A_{xx} + A_{yy} = \sigma_2 A|A|^2 + 2\sigma_1\sigma_2 QA, \tag{1.4a}$$

$$\sigma_1 Q_{xx} + Q_{yy} = -(|A|^2)_{xx}, \tag{1.4b}$$

where σ_1 and σ_2 $(=\pm 1)$ are also parameters depending on the dispersive property of the wave system. Hereafter we shall refer to (1.4) as the two-dimensional nonlinear Schrödinger (2DNLS) equation, although we note that this is not the usual 2D nonlinear Schrödinger equation. In the latter case we have the existence of severe nonlinear instabilities.

Recently it has been shown that (1.3) and (1.4) belong to the class of nonlinear evolution equations where IST is applicable.[4,5] Though the inverse scattering has not yet been completely done, the N-solitons may be obtained from the integral equations of the IST scheme or directly by applying Hirota's method to the equations.[4,6,7] These solutions are quasi-one-dimensional; they describe the multiple collisions of N solitons, each of which may propagate in different directions.

It is of interest to find essentially two-dimensional solutions of these equations. In the preceding paper,[8] we have developed a method to obtain rational solutions to the nonlinear evolution equations of IST type by taking a "long wave" limit of the N-soliton solutions. As one of the examples, we discussed some of the rational solutions of (1.3). We have shown that if the parameters of 2-soliton solution are chosen adequately, the long wave limit of the solution gives a two-dimensional permanent nonsingular lump, decaying in all directions.

In this paper we shall investigate rational solutions of (1.3) and (1.4) in detail and show that (1.4) also has a lump solution of envelope hole type. Moreover, we shall construct the solutions describing multiple collisions of lumps from the N-soliton solutions of (1.3) and (1.4). Section 2 is devoted to the case of the K-P equation. This result agrees with that recently announced by Manakov *et al.*[9] In Sec. 3, we shall consider the 2DNLS equation with nonzero asymptotic state. In this case the solution is a two-humped envelope lump on a plane wave field. In the Appendix a proof of the N-soliton solution is given for the 2DNLS equation.

2. KADOMTSEV-PETVIASHVILI EQUATION

Following the preceding paper,[8] we start with the bilin-

 0022-2488/79/071496-08$01.00

ear form of the K–P equation. Through a dependent variable transformation,

$$u = 2(\log f)_{xx'}, \tag{2.1}$$

(1.3) is transformed into the bilinear equation,

$$(D_t D_x + D_x^4 + \alpha D_y^2) f\cdot f = 0, \tag{2.2}$$

where we assume $u \to 0$ as $|x| \to \infty$ and define

$$D_x^l D_y^m D_t^n a\cdot b = (\partial_x - \partial_{x'})^l (\partial_y - \partial_{y'})^m (\partial_t - \partial_{t'})^n a(x,y,t) \times b(x',y',t')|_{x=x', y=y', t=t'}. \tag{2.3}$$

The N-soliton solution of (2.2) can be ascertained by applying Hirota's method.[6] It may be written in the following form:

$$f \equiv f_N = \sum_{\mu=0,1} \exp\Big(\sum_{i<j}^{(N)} \mu_i \mu_j A_{ij} + \sum_{i=1}^{N} \mu_i \eta_i \Big), \tag{2.4}$$

where

$$\eta_i = k_i [x + p_i y - (k_i^2 + \alpha p_i^2) t\,] + \eta_i^{(0)}, \tag{2.5}$$

$$\exp A_{ij} = \frac{3(k_i - k_j)^2 - \alpha(p_i - p_j)^2}{3(k_i + k_j)^2 - \alpha(p_i - p_j)^2}. \tag{2.6}$$

The notation $\Sigma_{\mu=0,1}$ indicates summation over all possible combinations of $\mu_1 = 0,1, \mu_2 = 0,1, \ldots, \mu_N = 0,1$; the $\Sigma_{i<j}^{(N)}$ summation is over all possible combinations of the N elements with the specific condition $i<j$. For example, the first two solutions of (2.4) have the form,

$$f_1 = 1 + \exp\eta_1, \tag{2.7}$$

$$f_2 = 1 + \exp\eta_1 + \exp\eta_2 + \exp(\eta_1 + \eta_2 + A_{12}). \tag{2.8}$$

The method to obtain rational solutions from these soliton solutions relies on the arbitrariness of choosing the phase constant $\eta_i^{(0)}$. For example, taking $\exp(\eta_1^{(0)}) = -1$, we may write (2.7) as

$$f_1 = 1 - \exp\xi_1, \tag{2.9}$$

which corresponds to a singular soliton solution,

$$u = \frac{k_1^2}{2} \operatorname{cosech}^2 \tfrac{1}{2} k_1 \xi_1, \tag{2.10}$$

where

$$\xi_i = k_i [x + p_i y - (k_i^2 + \alpha p_i^2) t\,]. \tag{2.11}$$

Taking the "long wave" limit, $k_1 \to 0$ on (2.9), yields

$$f_1 = -k_1 \theta_1 + O(k_1^2), \tag{2.12}$$

where

$$\theta_i = x + p_i y - \alpha p_i^2 t. \tag{2.13}$$

We may add an arbitrary phase factor to θ_i. Since u is given by (2.1), we have deduced the following rational solution,

$$f_1 = \theta_1, \tag{2.14a}$$

or

$$u = -2/\theta_1^2, \tag{2.14b}$$

which gives a singular solution.

For the 2-soliton solution (2.8), we take $\exp(\eta_i^{(0)}) = -1$ and $k_i \to 0$ for $i = 1,2$ with $k_1/k_2 = O(1)$ and $p_i = O(1)$. Then noticing

$$\exp A_{ij} = 1 + \frac{12 k_i k_j}{\alpha(p_i - p_j)^2} + O(\mathbf{k}^3), \tag{2.15}$$

we find

$$f_2 = k_1 k_2 \Big[\theta_1 \theta_1 + \frac{12}{\alpha(p_i - p_j)^2} + O(k) \Big]. \tag{2.16}$$

Thus we have

$$f_2 \simeq \theta_1 \theta_2 + B_{12}, \tag{2.17}$$

[i.e., with regard to a solution of the K–P equation, f_2 is equivalent to $\theta_1\theta_2 + B_{12}$ by virtue of (2.1)], where

$$B_{ij} = \frac{12}{\alpha(p_i - p_j)^2}. \tag{2.18}$$

Though (2.17) generally gives a singular u at some position, a nonsingular solution is obtained by taking $\alpha = -1$ (in which case the wave system is positive dispersive) and $p_2 = p_1^*$. In this case, we have

$$f_2 = \theta_1 \theta_1^* - 12/(p_1 - p_1^*)^2 > 0. \tag{2.19}$$

Inserting (2.19) into (2.1) and putting $p_1 = p_R + i p_I$, we obtain

$$u = 2 \frac{\partial^2}{\partial x^2} \log[(x' + p_R y')^2 + p_I^2 y'^2 + 3/p_I^2]$$

$$= \frac{4[-(x' + p_R y')^2 + p_I^2 y'^2 + 3/p_I^2]}{[(x' + p_R y')^2 + p_I^2 y'^2 + 3/p_I^2]^2}, \tag{2.20}$$

where

$$x' = x - (p_R^2 + p_I^2) t, \tag{2.21}$$

$$y' = y + 2 p_R t. \tag{2.22}$$

The rational solution (2.20) is a permanent lump solution decaying as $O(1/x^2, 1/y^2)$ for $|x|,|y| \to \infty$ and moving with the velocity $v_x = p_R^2 + p_I^2$ and $v_y = -2p_R$. In Fig. 1, the solution (2.20) is drawn for a particular choice of the constants.

We can readily extend this technique of obtaining rational solutions from soliton solutions, to a general N-soliton state (2.4). At first we take every $\exp(\eta_i^{(0)}) = -1$ in (2.4). Then f_N may be written as

$$f_N = \sum_{\mu=0,1} \prod_{i=1}^{N} (-1)^{\mu_i} \exp(\mu_i \eta_i) \prod_{i<j}^{(N)} \exp(\mu_i \mu_j A_{ij}). \tag{2.23}$$

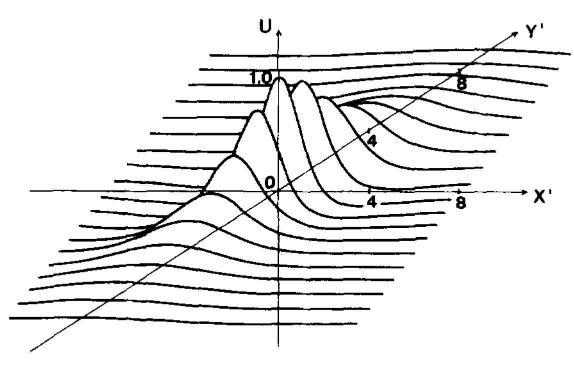

FIG. 1. Lump solution of (2.20) as seen in two dimensions at a fixed time; $p_R = 0$, $p_I^2 = 3/4$, with $\alpha = -1$.

It is easily see that

$$f_N|_{k_1=0} = \sum_{\mu=0,1} (-1)^{\mu_1} f_{N-1} = 0, \tag{2.24}$$

where f_{N-1} is an $(N-1)$-soliton solution with parameters, $k_2, k_3,...,k_N$. Considering the symmetric property of f_N with respect to k_i, we find that f_N is factorized by $\Pi_{i=1}^N k_i$. Thus, if we expand f_N in terms of k_i, the leading terms of (2.23) must be at least of order $\Pi_{i=1}^N k_i$. Taking a limit of $k_i \to 0$ and considering all the k_i to be of the same asymptotic order, we have

$$f_N = \sum_{\mu=0,1} \prod_{i=1}^{N} (-1)^{\mu_i}(1+\mu_i k_i \theta_i)$$

$$\times \prod_{i<j}^{(N)} (1+\mu_i\mu_j k_i k_j B_{ij}) + O(\mathbf{k}^{N+1}). \tag{2.25}$$

The leading terms of (2.25) are given by those of order $\Pi_{i=1}^N k_i$ in $\Pi_{i=1}^N(1+k_i\theta_i)\Pi_{i<j}^{(N)}(1+k_ik_jB_{ij})$. Therefore, a rational solution obtained as a long wave limit of N-soliton solution may be expressed as

$$f_N = \prod_{i=1}^{N} \theta_i + \tfrac{1}{2}\sum_{i,j}^{(N)} B_{ij} \prod_{l\neq i,j}^{N} \theta_l + \cdots + \frac{1}{M!2^M}$$

$$\times \sum_{i,j,...,m,n}^{(N)} \overbrace{B_{ij}B_{kl}\cdots B_{mn}}^{M} \prod_{p\neq i,j,...,m,n}^{N} \theta_p + \cdots, \tag{2.26}$$

where $\Sigma_{i,j,...,m,n}^{(N)}$ means the summation over all possible combinations of $i,j,...,m,n$ which are taken from $1,2,...,N$ and all different. The first four of (2.26) are written as

$$f_1 = \theta_1, \tag{2.27a}$$

$$f_2 = \theta_1\theta_2 + B_{12}, \tag{2.27b}$$

$$f_3 = \theta_1\theta_2\theta_3 + B_{12}\theta_3 + B_{23}\theta_1 + B_{31}\theta_2, \tag{2.27c}$$

$$\begin{aligned} f_4 = {} & \theta_1\theta_2\theta_3\theta_4 + B_{12}\theta_3\theta_4 + B_{13}\theta_2\theta_4 + B_{14}\theta_2\theta_3 + B_{23}\theta_1\theta_4 \\ & + B_{24}\theta_1\theta_3 + B_{34}\theta_1\theta_2 + B_{12}B_{34} + B_{13}B_{24} + B_{14}B_{23}. \end{aligned} \tag{2.27d}$$

We may express (2.26) in the following determinant form:

$$f_N = \begin{vmatrix} \theta_1 & \sqrt{B_{12}} & \cdots & \sqrt{B_{1N}} \\ -\sqrt{B_{12}} & \theta_2 & & \sqrt{B_{2N}} \\ \vdots & & \ddots & \vdots \\ -\sqrt{B_{1N}} & \cdots & \cdots & \theta_N \end{vmatrix}. \tag{2.28}$$

(2.26) or (2.28) generally gives a singular solution. However if we choose $\alpha = -1$ and $p_{M+i} = p_i^*$ $(i = 1,2,...,M)$ for $N = 2M$, we can get a class of nonsingular rational solutions. This may be seen from the determinant form of the solution (2.28). For this choice of parameters, f_{2M} is written as

$$f_{2M} = \begin{vmatrix} C & A \\ -A^* & {}^tC^* \end{vmatrix}, \tag{2.29}$$

where C and A are $M\times M$ matrices defined by

$$C = \begin{vmatrix} \theta_1 & \dfrac{2\sqrt{3}i}{p_1-p_2} & \cdots & \dfrac{2\sqrt{3}i}{p_1-p_M} \\ -\dfrac{2\sqrt{3}i}{p_1-p_2} & \theta_2 & & \dfrac{2\sqrt{3}i}{p_2-p_M} \\ \vdots & & \ddots & \vdots \\ -\dfrac{2\sqrt{3}i}{p_1-p_M} & \cdots & \cdot & \theta_M \end{vmatrix} \tag{2.30}$$

and

$$A = \begin{vmatrix} \dfrac{2\sqrt{3}i}{p_1-p_1^*} & \dfrac{2\sqrt{3}i}{p_1-p_2^*} & \cdots & \dfrac{2\sqrt{3}i}{p_1-p_M^*} \\ \dfrac{2\sqrt{3}i}{p_2-p_1^*} & \dfrac{2\sqrt{3}i}{p_2-p_2^*} & & \dfrac{2\sqrt{3}i}{p_2-p_M^*} \\ \vdots & & \ddots & \vdots \\ \dfrac{2\sqrt{3}i}{p_M-p_1^*} & \cdots & \cdots & \dfrac{2\sqrt{3}i}{p_M-p_M^*} \end{vmatrix} \tag{2.31}$$

Since the determinant

$$f_{2M} = \begin{vmatrix} C & A \\ -A^* & {}^tC^* \end{vmatrix}$$

with A, C given by (2.30) and (2.31) is always positive,[9] $u[=2(\log f_{2M})_{xx}]$ does not become infinite at any position.

Finally, we show that the nonsingular rational solution so obtained describes a multiple collision of M lumps. First we fix ourselves on the phase $|\theta_L|^2 = \text{const.}$, where θ_L is one of $\theta_1, \theta_2,...,\theta_M$. Then in the limit of $t\to\pm\infty$, $|\theta_1|, |\theta_2|,..., |\theta_{L-1}|, |\theta_{L+1}|,..., |\theta_M| = O(t)$ and f_{2M} has the following asymptotic state,

$$f_{2M} \sim |\theta_1|^2|\theta_2|^2\cdots|\theta_M|^2 + B_{L,M+L}|\theta_1|^2\cdots|\theta_{L-1}|^2$$

$$\times |\theta_{L+1}|^2\cdots|\theta_M|^2. \tag{2.32}$$

Noticing (2.1), f_{2M} in (2.32) is equivalent to

$$f_{2M} \to |\theta_L|^2 + B_{L,M+L}, \tag{2.33}$$

which is a lump solution with the phase θ_L. The same argument is possible for all of M lumps, as they have different velocities. Thus, it is concluded that this nonsingular rational solution describes a multiple collision of M lumps. Since the asymptotic forms of a lump are the same at $t = \pm\infty$, there is no phase shift when the lumps collide with each other.

3. TWO-DIMENSIONAL NONLINEAR SCHRÖDINGER EQUATION

The technique to deduce lump type solutions from soliton solutions is also applicable to the 2DNLS equation (1.4). For simplicity, we shall restrict ourselves to the boundary condition of nonzero asymptotic state, i.e., $|A|^2 \to \rho_0^2$ as $|x| \to \infty$. At first we construct the soliton solutions of (1.4) by Hirota's method.

By transforming the dependent variables in (1.4) via

$$A = g/f, \tag{3.1}$$

$$Q = \sigma_2(\log f)_{xx'} \tag{3.2}$$

(f real), we find that A and Q are solutions of (1.4) if f and g satisfy the following bilinear equations,

$$(iD_t - \sigma_1 D_x^2 + D_y^2 - \sigma_2\rho_0^2) g\cdot f = 0, \tag{3.3a}$$

$$(\sigma_1 D_x^2 + D_y^2 - \sigma_2\rho_0^2) f\cdot f = -\sigma_2 g g^*. \tag{3.3b}$$

We can obtain one- and two-soliton solutions of (1.4) by applying a perturbationlike technique on (3.3).[10] Expanding f and g as power series in a parameter ϵ,

$$f = 1 + \epsilon f^{(1)} + \epsilon^2 f^{(2)} + \cdots, \tag{3.4a}$$

$$g = g_0(1 + \epsilon h^{(1)} + \epsilon^2 h^{(2)} + \cdots), \tag{3.4b}$$

and substituting (3.4) into (3.3), we find in the lowest order,

$$(i\partial_t - \sigma_1\partial_x^2 + \partial_y^2 - \sigma_2\rho_0^2) g_0 = 0, \tag{3.5}$$

$$g_0 g_0^* = \rho_0^2. \tag{3.6}$$

g_0 has the solution

$$g_0 = \rho_0 \exp i\xi, \tag{3.7}$$

where

$$\xi = kx + ly - \omega t + \xi^{(0)}, \tag{3.8}$$

$$\omega = -\sigma_1 k^2 + l^2 + \sigma_2\rho_0^2. \tag{3.9}$$

Using (3.7), (3.3) is rewritten as

$$[i(D_t - 2\sigma_1 k D_x + 2lD_y) - \sigma_1 D_x^2 + D_y^2](1 + \epsilon h^{(1)} + \epsilon^2 h^{(2)} + \cdots)\cdot(1 + \epsilon f^{(1)} + \epsilon^2 f^{(2)} + \cdots) = 0, \tag{3.10a}$$

$$(\sigma_1 D_x^2 + D_y^2 - \sigma_2\rho_0^2)(1 + \epsilon f^{(1)} + \epsilon^2 f^{(2)} + \cdots)\cdot(1 + \epsilon f^{(1)} + \epsilon^2 f^{(2)} + \cdots) = -\sigma_2\rho_0^2(1 + \epsilon h^{(1)} + \epsilon^2 h^{(2)} + \cdots)(1 + \epsilon h^{(1)*} + \epsilon^2 h^{(2)*} + \cdots). \tag{3.10b}$$

Collecting the terms with the same power of ϵ, we have

$$L_1^- f^{(1)} + L_1^+ h^{(1)} = 0, \tag{3.11a}$$

$$L_2 f^{(1)} = -\sigma_2\rho_0^2(h^{(1)} + h^{(1)*}), \tag{3.11b}$$

$$L_1^- f^{(2)} + L_1^+ h^{(2)} + [i(D_t - 2\sigma_1 k D_x + 2lD_y) - \sigma_1 D_x^2 + D_y^2] h^{(1)}\cdot f^{(1)} = 0, \tag{3.12a}$$

$$L_2 f^{(2)} + (\sigma_1 D_x^2 + D_y^2 - \sigma_2\rho_0^2) f^{(1)}\cdot f^{(1)} = -\sigma_2\rho_0^2(h^{(2)} + h^{(1)}h^{(1)*} + h^{(2)*}) \tag{3.12b}$$

$$L_1^- f^{(3)} + L_1^+ h^{(3)} + [i(D_t - 2\sigma_1 k D_x + 2lD_y) - \sigma_1 D_x^2 + D_y^2](h^{(1)}\cdot f^{(2)} + h^{(2)}\cdot f^{(1)}) = 0, \tag{3.13a}$$

$$L_2 f^{(3)} + 2(\sigma_1 D_x^2 + D_y^2 - \sigma_2\rho_0^2) f^{(1)}\cdot f^{(2)} = -\sigma_2\rho_0^2(h^{(3)} + h^{(1)}h^{(2)*} + h^{(1)*}h^{(2)} + h^{(3)*}), \tag{3.13b}$$

$$L_1^- f^{(4)} + L_1^+ h^{(4)} + [i(D_t - 2\sigma_1 k D_x + 2lD_y) - \sigma_1 D_x^2 + D_y^2](h^{(1)}\cdot f^{(3)} + h^{(2)}\cdot f^{(2)} + h^{(3)}\cdot f^{(1)}) = 0, \tag{3.14a}$$

$$L_2 f^{(4)} + (\sigma_1 D_x^2 + D_y^2 - \sigma_2\rho_0^2)(2f^{(1)}\cdot f^{(3)} + f^{(2)}\cdot f^{(2)}) = -\sigma_2\rho_0^2(h^{(4)} + h^{(1)}h^{(3)*} + h^{(2)}h^{(2)*} + h^{(1)*}h^{(3)} + h^{(4)*}), \tag{3.14b}$$

.
.
.

where $L_1^{\pm}$ and L_2 are differential operators defined by

$$L_1^+ = i(\partial_t - 2\sigma_1 k\partial_x + 2l\partial_y) - \sigma_1\partial_x^2 + \partial_y^2, \tag{3.15a}$$

$$L_1^- = -i(\partial_t - 2\sigma_1 k\partial_x + 2l\partial_y) - \sigma_1\partial_x^2 + \partial_y^2, \tag{3.15b}$$

$$L_2 = 2(\sigma_1\partial_x^2 + \partial_y^2 - \sigma_2\rho_0^2). \tag{3.15c}$$

The one-soliton solution is obtained by choosing the following pair of starting solutions of (3.11):

$$f^{(1)} = \exp\eta_1, \tag{3.16a}$$

$$h^{(1)} = \exp(\eta_1 + i\phi_1), \tag{3.16b}$$

where

$$\eta_i = P_i x + Q_i y - \Omega_i t + \eta_i^{(0)}. \tag{3.17}$$

Substitution of (3.16) into (3.11) yields

$$\hat{\Omega}_1 \equiv \Omega_1 + 2\sigma_1 k P_1 - 2lQ_1 = \pm(\sigma_1 P_1^2 - Q_1^2) \times \sqrt{2\sigma_2\rho_0^2/(\sigma_1 P_1^2 + Q_1^2) - 1} \tag{3.18}$$

and

$$\sin^2(\phi_1/2) = (\sigma_1 P_1^2 + Q_1^2)/2\sigma_2\rho_0^2. \tag{3.19}$$

For this choice of the starting solution we find that all higher order terms can be taken to be zero. Thus we have an exact solution of (3.3),

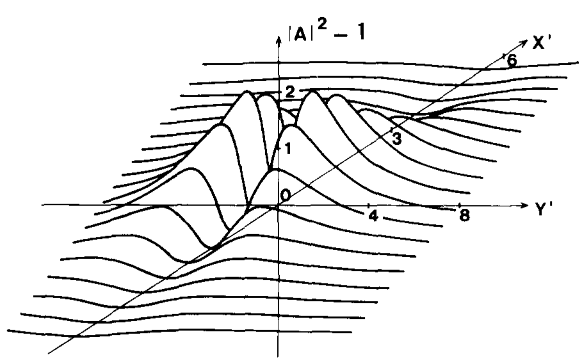

FIG. 2. Envelope lump solution of (3.42) as seen in two dimensions at a fixed time; $R_R = B_R = 0$, $R_I = B_I = 1$, $\rho_0^2 = 1$, $\alpha_{12} = 2$, with $\sigma_1 = -1$, $\sigma_2 = 1$.

$$f = 1 + \exp\eta_1, \tag{3.20a}$$

$$h = 1 + \exp(\eta_1 + i\phi_1), \tag{3.20b}$$

which gives an envelope-hole type soliton solution,

$$|A|^2 = \rho_0^2[1 - \sin^2(\phi_1/2)\mathrm{sech}^2(\eta_1/2)]. \tag{3.21}$$

For constructing the two-soliton solution, we start with

$$f^{(1)} = \exp\eta_1 + \exp\eta_2, \tag{3.22a}$$

$$h^{(1)} = \exp(\eta_1 + i\phi_1) + \exp(\eta_2 + i\phi_2). \tag{3.22b}$$

These solutions satisfy (3.11) if

$$\widehat{\Omega}_i = \pm(\sigma_1 P_i^2 - Q_i^2)\sqrt{2\sigma_2\rho_0^2/(\sigma_1 P_i^2 + Q_i^2) - 1} \tag{3.23}$$

and

$$\sin^2(\phi_i/2) = (\sigma_1 P_i^2 + Q_i^2)/2\sigma_2\rho_0^2. \tag{3.24}$$

Then, from (3.12) we have

$$f^{(2)} = D_{12}\exp(\eta_1 + \eta_2), \tag{3.25a}$$

$$h^{(2)} = D_{12}\exp(\eta_1 + \eta_2 + i\phi_1 + i\phi_2), \tag{3.25b}$$

where

$$D_{ij} = \{-2\sigma_2\rho_0^2\sin(\phi_i/2)\sin(\phi_j/2)\cos[(\phi_i - \phi_j)/2] + (\sigma_1 P_i P_j + Q_i Q_j)\}\{-2\sigma_2\rho_0^2\sin(\phi_i/2)\sin(\phi_j/2) \times\cos[(\phi_i + \phi_j)/2] + (\sigma_1 P_i P_j + Q_i Q_j)\}^{-1}. \tag{3.26}$$

Substituting (3.22) and (3.25) into (3.13) and (3.14), we find all higher terms can be chosen to be zero and we have a two-envelope-hole solution,

$$f = 1 + \exp\eta_1 + \exp\eta_2 + D_{12}\exp(\eta_1 + \eta_2), \tag{3.27a}$$

$$f = 1 + \exp(\eta_1 + i\phi_1) + \exp(\eta_2 + i\phi_2) + D_{12}\exp(\eta_1 + \eta_2 + i\phi_1 + i\phi_2). \tag{3.27b}$$

It is to be noted that the reality of f demands: 1) $\exp\eta_1$, $\exp\eta_2$, ϕ_1, ϕ_2 are real, or 2) $\exp\eta_2 = \exp\eta_1^*$ and $\phi_2 = \phi_1^*$. The form of (3.27) suggests that the N-soliton solution can be expressed as

$$f \equiv f_N = \sum_{\mu=0,1}\exp\left(\sum_{i<j}^{(N)}\mu_i\mu_j A_{ij} + \sum_{i=1}^{N}\mu_i\eta_i\right), \tag{3.28a}$$

$$h \equiv h_N = \sum_{\mu=0,1}\exp\left(\sum_{i<j}^{(N)}\mu_i\mu_j A_{ij} + \sum_{i=1}^{N}\mu_i(\eta_i + i\phi_i)\right), \tag{3.28b}$$

where

$$\exp A_{ij} = D_{ij} \tag{3.29}$$

[the proof that this solution satisfies (3.3) is given in the Appendix].

In order to construct rational solutions from soliton solutions we consider $P_i \ll 1$ and $R_i (= Q_i/P_i)$ are $O(1)$ for $i = 1,2,\ldots,N$. It is readily seen that the following expansions in terms of P_i hold

$$\exp\eta_i = 1 + P_i\theta_i + O(P_i^2), \tag{3.30a}$$

$$\exp(i\phi_i) = 1 + 2iP_i B_i + O(P_i^2), \tag{3.30b}$$

$$D_{ij} = 1 + P_i P_j\alpha_{ij} + O(P_i^2), \tag{3.30c}$$

where

$$\theta_i = x + R_i y - [-2\sigma_1 k + 2lR_i + (\sigma_1 - R_i^2)/B_i]t, \tag{3.31}$$

$$\alpha_{ij} = \frac{4B_i^2 B_j^2}{2B_i B_j - (\sigma_1 + R_i R_j)/\sigma_2\rho_0^2} = \frac{(\sigma_1 + R_i^2)(\sigma_1 + R_j^2)}{\sigma_2\rho_0^2(\epsilon_i\epsilon_j\sqrt{\sigma_1 + R_i^2}\sqrt{\sigma_1 + R_j^2} - \sigma_1 - R_i R_j)}, \tag{3.32}$$

and

$$B_i = \epsilon_i\sqrt{(\sigma_1 + R_i^2)/2\sigma_2\rho_0^2} \quad (\epsilon_i = \pm 1). \tag{3.33}$$

Taking $\exp(\eta_1^{(0)}) = -1$ and $P_1 \to 0$, we have from the one-soliton solution (3.20),

$$f = 1 - \exp\eta_1 = 1 - (1 + P_1\theta_1) + O(P_1^2) = -P_1[\theta_1 + O(P_1)], \tag{3.34a}$$

$$h = 1 - \exp(\eta_1 + i\phi_1) = -P_1[(\theta_1 + 2iB_1) + O(P_1)], \tag{3.34b}$$

where we have used the notation $\eta_i = P_i x + Q_i y - \Omega_i t$ instead of (3.17) (we may add another constant to this phase factor). Thus in the limit we obtain a singular solution,

$$A = \rho_0 e^{i\xi}(1 + 2iB_1/\theta_1). \tag{3.35}$$

Next we take $\exp(\eta_i^{(0)}) = -1$ and $P_i \to 0$ with $P_1/P_2 = O(1)$ in the two-soliton solution (3.27). Then we have

$$f = 1 - \exp\eta_1 - \exp\eta_2 + D_{12}\exp(\eta_1 + \eta_2) = P_1P_2[(\theta_1\theta_2 + \alpha_{12}) + O(\mathbf{P})]. \tag{3.36}$$

The long wave limit of h is obtained simply by taking $\theta_i \to \theta_i + 2iB_i$ in f,

$$h = P_1P_2[(\theta_1 + 2iB_1)(\theta_2 + 2iB_2) + \alpha_{12} + O(\mathbf{P})]. \tag{3.37}$$

In this limit, (3.36) and (3.37) yields the solution

$$A = \rho_0 e^{i\xi}\frac{(\theta_1 + 2iB_1)(\theta_2 + 2iB_2) + \alpha_{12}}{\theta_1\theta_2 + \alpha_{12}}. \tag{3.38}$$

This solution is generally singular. However, if we choose $\theta_2 = \theta_1^*$ and $\alpha_{12} > 0$, it gives a nonsingular lump solution. These conditions are realized when the parameters satisfy $R_2 = R_1^*$ and $\sigma_1\sigma_2 = -1$. It is to be noted that this choice of parameters guarantees the reality of f.

The envelope of the solution is rational. It is written as

$$|A|^2 = \rho_0^2\left[1 + 4\left(\frac{B_1\theta_1^* + B_1^*\theta_1}{\theta_1\theta_1^* + \alpha_{12}}\right)^2\right]. \tag{3.39}$$

To see the shape of this envelope lump, we rewrite the parameters as

$$R_1 = R_\mathrm{R} + iR_\mathrm{I}, \tag{3.40a}$$

$$B_1 = B_\mathrm{R} + iB_\mathrm{I}, \tag{3.40b}$$

and

$$(\sigma_1 - R_1^2)/B_1 = S_\mathrm{R} + iS_\mathrm{I}. \tag{3.40c}$$

Then introducing

$$x' = x + [2\sigma_1 k + (S_\mathrm{I}R_\mathrm{R} - S_\mathrm{R}R_\mathrm{I})/R_\mathrm{I}]t, \tag{3.41a}$$

$$y' = y - (2l + S_I/R_I)t, \tag{3.41b}$$

(3.39) is rewritten as

$$|A|^2 = \rho_0^2\left[1 + 16\left(\frac{B_R x' + (B_R R_R - B_I R_I)y'}{(x' + R_R y')^2 + R_I^2 y'^2 + \alpha_{12}}\right)^2\right]. \tag{3.42}$$

The second term of the right-hand side of (3.42) has its maximum value $4(B_R^2 + B_I^2)/\alpha_{12}$ at the two points, $x' = x_0 \equiv (B_R + B_I B_R/R_I)\sqrt{\alpha_{12}/(B_R^2 + B_I^2)}$, $y' = y_0 \equiv -B_I/R_I(\alpha_{12}/(B_R^2 + B_I^2))^{1/2}$ and $x' = -x_0 y' = -y_{0'}$ and it becomes zero on the line $B_R x' + (B_R R_R - B_I R_I)y' = 0$. Moreover, it decays as $O(1/x^2, 1/y^2)$ for $|x|, |y| \to \infty$. Thus, (3.42) expresses a permanent lump with two humps on a plane wave field. Its velocity is given by

$$v_x = -2\sigma_1 k - (S_I R_R - S_R R_I)/R_I, \tag{3.43a}$$

$$v_y = 2l + S_I/R_I. \tag{3.43b}$$

In Fig. 2, the envelope (3.42) is drawn for a particular choice of the parameters.

We have taken the limit of $P_i \to 0$ in order to obtain rational solutions from soliton solutions. Even if we take another limit of $Q_i \to 0$ and P_i/Q_i finite, the argument is almost parallel and a lump solution similar to (3.42) except some minor change of parameters is obtained under the same condition $\sigma_1\sigma_2 = -1$.

Finally, we consider the long wave limit of the N-soliton solution (3.28). Taking $\exp(\eta_i^{(0)}) = -1$ for $i = 1,2,\ldots,N$, f_N may be written as

$$f_N = \sum_{\mu=0,1} \prod_{i=1}^{N} (-1)^{\mu_i} \exp(\mu_i \eta_i) \prod_{i<j}^{(N)} (D_{ij})^{\mu_i \mu_j}. \tag{3.44}$$

It can be shown by the same procedure as the K-P equation that f_N is factorized by $\Pi_{i=1}^N P_i$. Expanding f_N in terms of P_i, we have

$$f_N = \sum_{\mu=0,1} \prod_{i=1}^{N} (-1)^{\mu_i} (1 + \mu_i P_i \theta_i)$$

$$\times \prod_{i<j}^{(N)} (1 + \mu_i P_i \mu_j P_j \alpha_{ij}) + O(\mathbf{P}^{N+1}). \tag{3.45}$$

The leading terms of f_N are given by those of order $\Pi_{i=1}^N P_i$ in $\Pi_{i=1}^N(1 + P_i\theta_i)\Pi_{i<j}^{(N)}(1 + P_i P_j \alpha_{ij})$. Similarly, if we expand h_N in terms of P_i, the leading terms are given by those of order $\Pi_{i=1}^N P_i$ in $\Pi_{i=1}^N[1 + P_i(\theta_i + 2iB_i)]\Pi_{i<j}^{(N)}(1 + P_i P_j \alpha_{ij})$. Thus, the long wave limit of the N-soliton solution is expressed as

$$A = \rho_0 e^{i\xi} \hat{h}_N / \hat{f}_N, \tag{3.46}$$

where

$$\hat{f}_N = \prod_{i=1}^{N} \theta_i + \frac{1}{2} \sum_{i,j}^{(N)} \alpha_{ij} \prod_{l \neq i,j}^{N} \theta_l + \cdots + \frac{1}{M!2^M}$$

$$\times \sum_{i,j,\ldots,m,n}^{(N)} \overbrace{\alpha_{ij}\alpha_{kl}\cdots\alpha_{mn}}^{M} \prod_{p \neq i,j,\ldots,m,n}^{N} \theta_p + \cdots, \tag{3.47}$$

and $\hat{h}_N$ is given by changing θ_i to $\theta_i + 2iB_i$ in $\hat{f}_N$. This gives a solution of the 2DNLS equation whose envelope is rational. Since the functional form of f_N is the same as (2.26), (3.46) becomes nonsingular if we choose $\sigma_1\sigma_2 = -1$ and $R_{M+i} = R_i^*$ ($i = 1,2,\ldots,M$) for $N = 2M$. As for the asymptotic state, the same argument is possible as that of the K–P equation and hence we deduce that the nonsingular solution describes multiple collisions of M lumps each of which has the functional form (3.42) and which has no phase shift when they collide with each other.

ACKNOWLEDGMENT

This research was partially sponsored by the Air Force Office of Scientific Research, Air Force Systems Command, USAF, under Grant No. AFOSR—The United States government is authorized to reproduce and distribute reprints for Governmental purposes notwithstanding any copyright notation hereon.

APPENDIX

In this appendix we show that the N-soliton solution (3.28) satisfies

$$F(f,h) = [i(D_t - 2\sigma_1 k D_x + 2l D_y) - \sigma_1 D_x^2 + D_y^2] h \cdot f = 0, \tag{A1}$$

$$G(f,h) = (\sigma_1 D_x^2 + D_y^2 - \sigma_2\rho_0^2) f \cdot f + \sigma_2\rho_0^2 hh^* = 0. \tag{A2}$$

Substituting (3.28) into (A1) and (A2), using (3.23) and (3.24), we have

$$F(f,h) = \sum_{\mu=0,1} \sum_{\nu=0,1} \Bigg[-i \sum_{i=1}^{N} (\mu_i - \nu_i)(\sigma_1 P_i^2 - Q_i^2) \cot(\phi_i/2)$$

$$- \sigma_1\left(\sum_{i=1}^{N} (\mu_i - \nu_i) P_i\right)^2 + \left(\sum_{i=1}^{N} (\mu_i - \nu_i) Q_i\right)^2\Bigg]$$

$$\times \exp\left(\sum_{i=1}^{N} (\mu_i + \nu_i)\eta_i + \sum_{i=1}^{N} i\mu_i\phi_i + \sum_{i<j}^{N} (\mu_i\mu_j + \nu_i\nu_j)A_{ij}\right), \tag{A3}$$

$$G(f,h) = \sum_{\mu=0,1} \sum_{\nu=0,1} \Bigg[\sigma_1\left(\sum_{i=1}^{N} (\mu_i - \nu_i) P_i\right)^2$$

$$+ \left(\sum_{i=1}^{N} (\mu_i - \nu_i) Q_i\right)^2 - \sigma_2\rho_0^2\left(1 - \exp\left(\sum_{i=1}^{N} i(\mu_i - \nu_i)\phi_i\right)\right)\Bigg]$$

$$\times \exp\left(\sum_{i=1}^{N} (\mu_i + \nu_i)\eta_i + \sum_{i<j}^{(N)} (\mu_i\mu_j + \nu_i\nu_j)A_{ij}\right). \tag{A4}$$

Let us call the coefficients of the factors $\exp(\Sigma_{i=1}^n \eta_i + \Sigma_{i=n+1}^m 2\eta_i)$ of F and G $\bar{F}(1,2,\ldots,n;n+1,\ldots,m)$ and $\bar{G}(1,2,\ldots,n;n+1,\ldots,m)$ respectively. Then we have

$$\bar{F} = \sum_{\mu=0,1} \sum_{\nu=0,1} \mathrm{cond}(\mu,\nu) \Bigg[-i \sum_{i=1}^{N} (\mu_i - \nu_i)(\sigma_1 P_i^2 - Q_i^2)$$

$$\times \cot(\phi_i/2) - \sigma_1\left(\sum_{i=1}^{N} (\mu_i - \nu_i) P_i\right)^2 + \left(\sum_{i=1}^{N} (\mu_i - \nu_i) Q_i\right)^2\Bigg]$$

$$\times \exp\left(\sum_{i=1}^{N} i\mu_i\phi_i + \sum_{i<j}^{N} (\mu_i\mu_j + \nu_i\nu_j)A_{ij}\right), \tag{A5}$$

$$\tilde{G} = \sum_{\mu=0,1} \sum_{\nu=0,1} \mathrm{cond}(\mu,\nu)\left[\sigma_1\left(\sum_{i=1}^{N} (\mu_i - \nu_i)P_i\right)^2 + \left(\sum_{i=1}^{N} (\mu_i - \nu_i)Q_i\right)^2 - \sigma_2\rho_0^2\left(1 - \exp\left(\sum_{i=1}^{N} i(\mu_i - \nu_i)\phi_i\right)\right)\right]$$
$$\times \exp\left(\sum_{i<j}^{(N)} (\mu_i\mu_j + \nu_i\nu_j)A_{ij}\right), \tag{A6}$$

where $\mathrm{cond}(\mu,\nu)$ means that the summation over μ, ν are performed under the conditions

$$\mu_i + \nu_i = 1, \quad \text{for } i = 1,2,\ldots,n,$$
$$\mu_i = \nu_i = 1, \quad \text{for } i = n+1,n+2,\ldots,m,$$
$$\mu_i = \nu_i = 0, \quad \text{for } i = m+1,m+2,\ldots,N.$$

Introducing notations $\lambda_i = \mu_i - \nu_i$ for $i = 1,2,\ldots,n$ and substituting (3.29) with (3.26), (A5) and (A6) are reduced to

$$\tilde{F} = \mathrm{const} \sum_{\lambda=\pm 1} \left(-i \sum_{i=1}^{n} (\sigma_1 P_i^2 - Q_i^2)\cot(\lambda_i\phi_i/2) - \sigma_1\left(\sum_{i=1}^{n} \lambda_i P_i\right)^2 + \left(\sum_{i=1}^{n} \lambda_i Q_i\right)^2\right) \sum_{i=1}^{n} (\sin(\lambda_i\phi_i/2)$$
$$- i\cos(\lambda_i\phi_i/2)) \prod_{i<j}^{(n)}(-2\sigma_2\rho_0^2 \cos[(\lambda_i\phi_i - \lambda_j\phi_j)/2]$$
$$+ (\sigma_1 P_iP_j + Q_iQ_j)/\sin(\phi_i/2)\sin(\phi_j/2)), \tag{A7}$$

$$\tilde{G} = \mathrm{const} \sum_{\lambda=\pm 1} \left(\sigma_1\left(\sum_{i=1}^{n} \lambda_i P_i\right)^2 + \left(\sum_{i=1}^{n} \lambda_i Q_i\right)^2 - \sigma_2\rho_0^2\left[1 - (-1)^n \prod_{i=1}^{n} (\sin(\lambda_i\phi_i/2) - i\cos(\lambda_i\phi_i/2))^2\right]\right)$$
$$\times \prod_{i<j}^{(n)}(-2\sigma_2\rho_0^2 \cos[(\lambda_i\phi_i - \lambda_j\phi_j)/2]$$
$$+ (\sigma_1 P_iP_j + Q_iQ_j)/\sin(\phi_i/2)\sin(\phi_j/2)), \tag{A8}$$

where the constants do not depend on λ_i. Rewritting the variables in (A7) and (A8) as

$$x_i = -i\sin(\phi_i/2), \tag{A9}$$

$$y_i = \sqrt{(\sigma_1 P_i^2 - Q_i^2)/(\sigma_1 P_i^2 + Q_i^2)}, \tag{A10}$$

we find

$$\tilde{F} = \mathrm{const}\hat{F}(\lambda_1x_1,\lambda_2x_2,\ldots,\lambda_nx_n;y_1^2,y_2^2\ldots,y_n^2), \tag{A11}$$

$$\tilde{G} = \mathrm{const}\hat{G}(\lambda_1x_1,\lambda_2x_2,\ldots,\lambda_nx_n;y_1^2,y_2^2\ldots,y_n^2), \tag{A12}$$

where

$$\hat{F} = \sum_{\lambda=\pm 1} \left[2 \sum_{i=1}^{n} \lambda_ix_i\sqrt{1+x_i^2}\,y_i^2 + \left(\sum_{i=1}^{n} \lambda_ix_i\sqrt{1+y_i^2}\right)^2 - \left(\sum_{i=1}^{n} \lambda_ix_i\sqrt{1-y_i^2}\right)^2\right]$$
$$\times \prod_{i=1}^{n} (\lambda_ix_i - \sqrt{1+x_i^2}) \prod_{i<j}^{(n)} (2\lambda_i\lambda_jx_ix_j - 2\sqrt{1+x_i^2}\sqrt{1+x_j^2} + \sqrt{1+y_i^2}\sqrt{1+y_j^2}$$
$$+ \sqrt{1-y_i^2}\sqrt{1-y_j^2}), \tag{A13}$$

$$\hat{G} = \sum_{\lambda=\pm 1} \left[\left(\sum_{i=1}^{n} \lambda_ix_i\sqrt{1+y_i^2}\right)^2 + \left(\sum_{i=1}^{n} \lambda_ix_i\sqrt{1-y_i^2}\right)^2 + 1 - \prod_{i=1}^{n} (\lambda_ix_i - \sqrt{1+x_i^2})^2\right] \prod_{i<j}^{(n)} (2\lambda_i\lambda_jx_ix_j$$
$$- 2\sqrt{1+x_i^2}\sqrt{1+x_j^2} + \sqrt{1+y_i^2}\sqrt{1+y_j^2}$$
$$+ \sqrt{1-y_i^2}\sqrt{1-y_j^2}). \tag{A14}$$

Thus f and h are solutions of (A1) and (A2) if the identities,

$$\hat{F}(\lambda_1x_1,\lambda_2x_2,\ldots,\lambda_nx_n;y_1^2,y_2^2\ldots,y_n^2) = 0, \tag{A15}$$

$$\hat{G}(\lambda_1x_1,\lambda_2x_2,\ldots,\lambda_nx_n;y_1^2,y_2^2\ldots,y_n^2) = 0, \tag{A16}$$

hold for $n = 1,2,\ldots,N$.

These identities can be verified by mathematical induction. The verification is easy for $n = 1$ and 2. We assume that the identities hold for $n-1$ and $n-2$. $\hat{F}$ and $\hat{G}$ have the following properties:

(a): $\hat{F}$ and $\hat{G}$ are even functions of $x_1,x_2,\ldots,x_n$, and $y_1,y_2,\ldots,y_n$.

(b): $\hat{F}$ and $\hat{G}$ are unchanged by the replacement x_i and y_i with x_j and y_j for arbitrary i and j.

(c):

$$\hat{F}(\lambda_1x_1,\lambda_2x_2,\ldots,\lambda_nx_n;y_1^2,y_2^2\ldots,y_n^2)|_{x_1=0} = -\prod_{i=2}^{n} (-2\sqrt{1+x_i^2} + \sqrt{1+y_1^2}\sqrt{1+y_i^2} + \sqrt{1-y_1^2}\sqrt{1-y_i^2}$$
$$\times \hat{F}(\lambda_2x_2,\lambda_3x_3,\ldots,\lambda_nx_n;y_2^2,y_3^2\ldots,y_n^2),$$

$$\hat{G}(\lambda_1x_1,\lambda_2x_2,\ldots,\lambda_nx_n;y_1^2,y_2^2\ldots,y_n^2)|_{x_1=0} = \prod_{i=2}^{n} (-2\sqrt{1+x_i^2} + \sqrt{1+y_1^2}\sqrt{1+y_i^2} + \sqrt{1-y_1^2}\sqrt{1-y_i^2})$$
$$\times \hat{G}(\lambda_2x_2,\lambda_3x_3,\ldots,\lambda_nx_n;y_2^2,y_3^2\ldots,y_n^2).$$

(d):

$$\hat{F}(\lambda_1x_1,\lambda_2x_2,\ldots,\lambda_nx_n;y_1^2,y_2^2\ldots,y_n^2)|_{x_1=x_2 \text{ and } y_1=y_2} = -8x_1^2 \prod_{i=3}^{n} [-4x_1^2x_i^2 + (-2\sqrt{1+x_1^2}\sqrt{1+x_i^2}$$
$$+ \sqrt{1+y_1^2}\sqrt{1+y_i^2} + \sqrt{1-y_1^2}\sqrt{1-y_i^2})^2]$$

$$\times \hat{F}(\lambda_3 x_3, \lambda_4 x_4, \ldots, \lambda_n x_n; y_3^2, y_4^2 \ldots, y_n^2),$$

$$\hat{G}(\lambda_1 x_1, \lambda_2 x_2, \ldots, \lambda_n x_n; y_1^2, y_2^2 \ldots, y_n^2)|_{x_1 = x_2 \text{ and } y_1 = y_2}$$

$$= -8x_1^2 \prod_{i=3}^{n} [-4x_1^2 x_i^2 + (-2\sqrt{1+x_1^2}\sqrt{1+x_i^2} + \sqrt{1+y_1^2}\sqrt{1+y_i^2} + \sqrt{1-y_1^2}\sqrt{1-y_i^2})^2]$$

$$\times \hat{G}(\lambda_3 x_3, \lambda_4 x_4, \ldots, \lambda_n x_n; y_3^2, y_4^2 \ldots, y_n^2).$$

We may consider $\hat{F}$ and $\hat{G}$ to be functions of $u_1, u_2, \ldots, u_n$ and $v_1^\pm, v_2^\pm, \ldots, v_n^\pm$ with $u_n = (1 + x_i^2)^{1/2}$ and $v_i^\pm = (1 \pm y_i^2)^{1/2}$. Then property (b) implies

(b)′: $\hat{F}$ and $\hat{G}$ are unchanged by the replacement u_i and $v_i^\pm$ with u_j and $v_j^\pm$ for arbitrary i and j.

The properties (b)′ and (c) means $\hat{F}$ and $\hat{G}$ are factorized by $\Pi_{i=1}^{n}(u_i^2 - 1)$. The properties (b)′ and (d) show that $\hat{F}$ and $\hat{G}$ are written as

$$(u_1 - u_2)^2 J_1 + (u_1 - u_2)(v_1 - v_2) J_2 + (v_1 - v_2)^2 J_3,$$

where J_i $(i = 1,2,3)$ are certain polynomial of $u_1, u_2, \ldots, u_n$ and $v_1, v_2, \ldots, v_n$, and v_i stands for v_i^+ or v_i^-. Using property (b)′, we see that $\hat{F}$ and $\hat{G}$ are written as

$$\prod_{i<j}^{(n)} (u_i - u_j)^2 K_1 + (u_1 - u_2)(v_1 - v_2) \prod_{i<j}^{(n)}{}' (u_i - u_j)^2 K_2 + \cdots + \prod_{i<j}^{(n)} (v_i - v_j)^2 K_m,$$

where $K_1, K_2, \ldots, K_m$ are polynomials of $u_1, u_2, \ldots, u_n$ and $v_1, v_2, \ldots, v_n$ and the prime attached to $\Pi_{i<j}^{(n)}$ means product over all i, j except $i = 1$ and $j = 2$. Thus, $\hat{F}$ and $\hat{G}$ are expressed as

$$\prod_{i=1}^{n} (u_i^2 - 1)\left(\prod_{i<j}^{(n)} (u_i - u_j)^2 K_1 + \cdots \prod_{i<j}^{(n)} (v_i - v_j)^2 K_m \right),$$

if they would not be identically zero. We examine the degree of $\hat{F}$ and $\hat{G}$ with respect to u_1 and v_1. The above expression implies $\hat{F}$ and $\hat{G}$ are at least of degree $2n$. On the other hand, (A13) and (A14) show $\hat{F}$ and $\hat{G}$ are at most of degree $n + 2$ and $2n - 1$, respectively. Thus $\hat{F}$ and $\hat{G}$ must be zero for $n > 2$ and (A15) and (A16) have been proved.

[1] See for example, "Nonlinear Dispersive Waves," *Series of Selected Papers in Physics*, **59**, edited by N. Yajima, and N. Kakutani, (Phys. Soc. Japan, 1975).

[2] An account of IST is found in M.J. Ablowitz, D.J. Kaup, A.C. Newell, and H. Segur, Stud. Appl. Math. **53**, 249 (1973).

[3] B. Kadomtsev, and V. Petviashvili, Sov. Phys. Dokl. **15**, 539 (1970); M. Oikawa, J. Satsuma, and N. Yajima, J. Phys. Soc. Jpn **37**, 511 (1974); D.J. Benney, and G.J. Roskes, Stud. Appl. Math. **1**, 377 (1969); A. Davey, and K. Stewartson, Proc. R. Soc. **388**, 191 (1974); N.C. Freeman and A. Davey, Proc. R. Soc. London, Ser. A **344**, 427 (1975); V.D. Djovdjevic, and L.G. Redekopp, J. Fluid Mech. **79**, 703 (1977); M.J. Ablowitz and H. Segur (preprint, 1978).

[4] V.E. Zakharov and A.B. Shabat, Funct. Anal. Appl. **8**, 226 (1974).

[5] M.J. Ablowitz and R. Haberman, Phys. Rev. Lett. **35**, 1185 (1975).

[6] J. Satsuma, J. Phys. Soc. Jpn. **40**, 286 (1976).

[7] D. Anker and N.C. Freeman, Proc. R. Soc. London, Ser. A **360**, 529 (1978).

[8] M.J. Ablowitz and J. Satsuma, J. Math. Phys. **19**, 2180 (1978).

[9] S.V. Manakov, V.E. Zakharov, L.A. Bordag, A.R. Its, and V.B. Matveev, Phys. Lett. A **63**, 205 (1977).

[10] R. Hirota, *Bäcklund Transformations, the Inverse Scattering Method, Solitons, and Their Applications*, edited by R.M. Miura (Springer, New York, 1976), p. 40.

Transport scattering coefficients from reflection and transmission measurements

N. J. McCormick
Department of Nuclear Engineering, University of Washington, Seattle, Washington 98195

(Received 26 April 1978; revised manuscript received 19 January 1979)

Angular moments of the distributions on the surfaces of a homogeneous slab are manipulated to provide two complementary sets of linear equations for calculating all the one-speed angular expansion coefficients of the scattering kernel of the medium. Use of the results requires that at least one incident distribution be dependent upon the azimuthal angle.

I. INTRODUCTION

For the simplest inverse transport problem, corresponding to an infinite medium containing a localized azimuthally-symmetric plane source (i.e., the Green's function problem), a method equivalent to the "method of moments" has been utilized to extract the scattering coefficients in terms of angular and spatial moments of the angular flux throughout the infinite medium.[1,2] Solutions of the azimuthally-independent inverse problem also have been worked out for the multi-energy-group[3] and time-dependent cases.[4] An elegant solution recently has been developed by Cacuci and Goldstein[5] who obtained, as a by-product of their analysis of the slowing down problem, the spatial moments of the angle-integrated one-speed flux in terms of a series solution [see Eqs. (40) and (48) of Ref. 5].

The infinite-medium inverse problem with an azimuthally asymmetric source also has been solved, where it has been shown that a single angular and spatial moment of the one-speed azimuth-dependent Green's function distribution can be related to a single scattering coefficient.[6] Both these moments and those for the one-speed azimuth-independent problems[5] are special cases of a generalized family of moments which may be expressed in a series solution.[7]

The inverse problem for a finite slab has been solved[8] for the multigroup case by an extension of the method used for the infinite-medium problem;[3] the desired expansion coefficients were expressed implicitly in terms of angular and spatial moments of the interior distributions and angular moments of the distributions on the slab surfaces. From the point-of-view of possible experimental application of the equations, avoiding the need for measurement of the interior distribution eliminates the need for moving the detector within the medium and the possibility of inducing perturbations due to the detector. An explicit evaluation of the coefficients for isotropic and linearly anisotropic scattering using only angular moments of the surface distributions has been completed for the one-speed case.[9]

The previous slab-geometry analyses utilized an azimuthally-independent incident source distribution.[8,9] Use of an azimuthally-dependent incident source distribution, however, enlarges the number of possible moments which can be combined in a judicious manner to yield the expansion coefficients. It is the purpose of this paper to generalize the innovative work of Siewert[9,10] by using angular moments of the surface distributions to develop two complementary sets of linear equations for calculating all the expansion coefficients for a medium with arbitrary anisotropic scattering.

After an introduction to the notation (Sec. II), the equations for calculating all the coefficients of a general scattering law are presented in Sec. III. The relationship of this work to that with azimuthally-symmetric moments[9,10] also is shown in Sec. III for the special case of linearly anisotropic scattering of the medium. Comments on experimental applications of the equations are contained in Sec. IV.

II. BASIC DEFINITIONS AND NOTATION

For a plane source, the radiation intensity (or neutron angular flux) $I(\tau,\mu,\phi)$ depends upon one coordinate (τ), as measured in mean-free-paths, on the cosine of the polar angle with respect to the positive τ-axis (μ), and on the azimuth (ϕ). The equation of transfer may be written as[11]

$$\mu \frac{\partial I(\tau,\mu,\phi)}{\partial \tau} + I(\tau,\mu,\phi) = \frac{1}{4\pi}\int_{-1}^{1} d\mu' \int_0^{2\pi} d\phi'\, p(\cos\delta) I(\tau,\mu',\phi'), \quad 0 \leqslant \tau \leqslant \tau_0, \tag{1}$$

where anisotropic scattering of finite order N is admitted,

$$p(\cos\delta) = \sum_{n=0}^{N} \omega_n P_n(\cos\delta), \tag{2}$$

and where $0 < \omega_0 < 1$. The azimuthally-dependent incoming distributions on the slab surfaces at $\tau = 0$ and $\tau = \tau_0$ are allowed to be arbitrary, although use of the equations is slightly simplified if only a collimated distribution is incident upon one surface, so that

$$I(0,\mu,\phi) = \delta(\mu - \mu_0)\delta(\phi), \quad 0 \leqslant \mu \leqslant 1 \tag{3}$$

and

$$I(\tau_0,\mu,\phi) = 0, \quad -1 \leqslant \mu \leqslant 0. \tag{4}$$

By an established procedure[11,12] the ϕ-dependence in Eq. (1) can be eliminated by a finite Fourier expansion

$$I(\tau,\mu,\phi) = \sum_{m=0}^{N} (2 - \delta_{m0}) I^m(\tau,\mu)(1-\mu^2)^{m/2} \cos m\phi + I_u(\tau,\mu,\phi). \tag{5}$$

Here $I_u(\tau,\mu,\phi)$ is a portion of the uncollided distribution given, when Eq. (3) is valid, by

 0022-2488/79/071504-04$01.00

$$I_u(\tau,\mu,\phi) = \delta(\mu-\mu_0)e^{-\tau/\mu}\Big(\delta(\phi) - (2\pi)^{-1} \times \sum_{m=0}^{N}(2-\delta_{m0})\cos m\phi\Big), \quad 0\leqslant\mu\leqslant1, \tag{6}$$
$$= 0, \quad \text{otherwise.}$$

The resulting $(N+1)$ independent transport equations are

$$\mu\frac{\partial I^m(\tau,\mu)}{\partial\tau} + I^m(\tau,\mu) = \frac{1}{2}\int_{-1}^{1} p^m(\mu,\mu')I^m(\tau,\mu')dm(\mu'), \tag{7}$$

while the boundary conditions corresponding to Eqs. (3) and (4) are

$$I^m(0,\mu) = \frac{\delta(\mu-\mu_0)}{2\pi(1-\mu_0^2)^{m/2}}, \quad 0\leqslant\mu\leqslant1, \tag{8}$$

$$I^m(\tau_0,\mu) = 0, \quad -1\leqslant\mu\leqslant0, \tag{9}$$

for $m = 0$ to N. In Eq. (7),

$$p^m(\mu,\mu') = \sum_{k=m}^{N} c_k^m p_k^m(\mu)p_k^m(\mu'), \tag{10}$$

where

$$p_k^m(\mu) = \frac{d^m P_k(\mu)}{d\mu^m} = (1-\mu^2)^{-m/2}P_k^m(\mu), \tag{11}$$

$$c_k^m = \omega_k\frac{(k-m)!}{(k+m)!} \tag{12}$$

and, for brevity,

$$dm(\mu) = (1-\mu^2)^m\, d\mu. \tag{13}$$

It may be noted that

$$p_k^m(-\mu) = (-1)^{k-m}p_k^m(\mu). \tag{14}$$

The desired equations for calculating the ω_k, $k = 0$ to N, using moments of the surface intensities, require the definition of the spatially-dependent moments

$$i_{jk}^m(\tau) = (2\pi)^{-1}\int_0^{2\pi} d\phi\,\cos m\phi\int_{-1}^{1} d\mu\,\mu^j P_k^m(\mu)I(\tau,\mu,\phi) \tag{15}$$

for $j = 0,1,\cdots$. After use of Eqs. (5), (11), and (13), these moments may be written as

$$i_{jk}^m(\tau) = \int_{-1}^{1}\mu^j p_k^m(\mu)I^m(\tau,\mu)\,\mathrm{d}m(\mu). \tag{16}$$

On the surfaces of the slab, these integrals are evaluated by use of the boundary conditions; in the special case of a collimated incident distribution, use of Eqs. (3) and (4) or, equivalently, Eqs. (8) and (9), gives

$$i_{jk}^m(0) = (2\pi)^{-1}\mu_0^j P_k^m(\mu_0) + (-1)^{k-m+j} \times\int_0^1\mu^j p_k^m(\mu)I^m(0,-\mu)\,dm(\mu), \tag{17}$$

$$i_{jk}^m(\tau_0) = \int_0^1\mu^j p_k^m(\mu)I^m(\tau,\mu)\,dm(\mu), \tag{18}$$

for $m = 0$ to N.

For convenience in future notation, the superscript m with $I^m(x,\mu)$ and other functions will be suppressed unless it is required for clarity.

A general equation for the moments in Eq. (16) follows by multiplying Eq. (7) by $\mu^j p_k(\mu)$ and integrating over $dm(\mu)$ to obtain

$$i'_{j+1,k}(\tau) + i_{jk}(\tau) = \frac{1}{2}\sum_{n=m}^{N} c_n i_{0n}(\tau) \times\int_{-1}^{1}\mu^j p_k(\mu)p_n(\mu)\,dm(\mu), \tag{19}$$

where i' denotes $di/d\tau$. For $j = 0$, Eq. (19) takes the convenient form

$$i_{0k}(\tau) = -[(2k+1)/h_k]i'_{1k}(\tau) \tag{20}$$

after use of orthogonality and Eq. (12). Here h_k is defined as

$$h_k = 2k+1-\omega_k \tag{21}$$

and is independent of m.

A second set of moments which are needed is defined as

$$S_n(\tau) = 4\int_0^1 d\mu\,\mu^{2n}\Big((2\pi)^{-1}\int_0^{2\pi} d\phi\,\cos(m\phi)I(\tau,-\mu,\phi)\Big) \times\Big((2\pi)^{-1}\int_0^{2\pi} d\phi\,\cos(m\phi)I(\tau,\mu,\phi)\Big) \tag{22}$$

for $n = 0, 1,\cdots$. After use of Eqs. (5) and (13), the moments may be written as

$$S_n(\tau) = 4\int_0^1\mu^{2n}I(\tau,-\mu)I(\tau,\mu)\,dm(\mu)$$
$$= 2\int_{-1}^{1}\mu^{2n}I(\tau,-\mu)I(\tau,\mu)\,dm(\mu). \tag{23}$$

On the surfaces of the slab, these integrals are evaluated by use of boundary conditions; use of Eqs. (8) and (9) leads to the simplest possible results.

III. THE GENERAL SOLUTION

The first general set of equations for the ω_k values is derived by first multiplying Eq. (7) by $2\mu^{2n}I^m(\tau,-\mu)$ and integrating over $dm(\mu)$ to obtain

$$2\int_{-1}^{1}\mu^{2n}I(\tau,-\mu)\mu\frac{\partial I(\tau,\mu)}{\partial\tau}dm(\mu) + S_n(\tau) = \sum_{k=m}^{N}(-1)^{k-m}c_k\, i_{0k}(\tau)i_{2n,k}(\tau), \tag{24}$$

after using Eqs. (14), (16), and (23). Differentiation of Eq. (24) leads to the result

$$2\int_{-1}^{1}\mu^{2n}I(\tau,-\mu)\frac{\partial}{\partial\tau}\Big(\mu\frac{\partial I(\tau,\mu)}{\partial\tau}\Big)dm(\mu) + S'_n(\tau) = \sum_{k=m}^{N}(-1)^{k-m}c_k\, i_{0k}(\tau)i'_{2n,k}(\tau) + \sum_{k=m}^{N}(-1)^{k-m}c_k\, i'_{0k}(\tau)i_{2n,k}(\tau) \tag{25}$$

after recognizing that a term on the left-hand side (lhs) vanishes because it is an integral of an odd function over symmetric limits. Substituting Eq. (7) for $\mu\partial I/\partial\tau$ in the first term on the lhs leads to a simplified expression,

$$-2\int_{-1}^{1}\mu^{2n}I(\tau,-\mu)\frac{\partial I(\tau,\mu)}{\partial\tau}dm(\mu) + S'_n(\tau)$$

$$= \sum_{k=m}^{N} (-1)^{k-m} c_k\, i_{0k}(\tau) i'_{2n,k}(\tau). \tag{26}$$

By symmetry, however,

$$\int_{-1}^{1} \mu^{2n} I(\tau, -\mu) \frac{\partial I(\tau,\mu)}{\partial \tau} dm(\mu) = \tfrac{1}{4} S'_n(\tau), \tag{27}$$

so therefore

$$\tfrac{1}{2} S'_n(\tau) = \sum_{k=m}^{N} (-1)^{k-m} c_k\, i_{0k}(\tau) i'_{2n,k}(\tau). \tag{28}$$

For $n = 0$, Eq. (28) becomes

$$\frac{d}{d\tau} S_0(\tau) = \frac{d}{d\tau} \sum_{k=m}^{N} (-1)^{k-m} c_k [i_{0k}(\tau)]^2. \tag{29}$$

Integration of Eq. (29) over $0 \leqslant \tau \leqslant \tau_0$ and use of Eq. (12) leads to the final result

$$S_0^m = \sum_{k=m}^{N} \omega_k s_{0k}^m, \tag{30}$$

where the required surface integrals are defined as

$$S_0^m = S_0^m(0) - S_0^m(\tau_0) \tag{31}$$

and

$$s_{0k}^m = (-1)^{k-m}[(k-m)!/(k+m)!] \times \{[i_{0k}^m(0)]^2 - [i_{0k}^m(\tau_0)]^2\}. \tag{32}$$

The first general set of equations for calculating the ω_k values follows from Eq. (30) by taking $m = 0,1,2,\ldots, N$ to obtain the "upper triangular" set of equations

$$\begin{aligned} S_0^0 &= \omega_0 s_{00}^0 + \omega_1 s_{01}^0 + \cdots + \omega_N s_{0N}^0 \\ S_0^1 &= \omega_1 s_{01}^1 + \cdots + \omega_N s_{0N}^1 \\ &\vdots \\ S_0^N &= \omega_N s_{0N}^N \end{aligned} \tag{33}$$

This set of $(N+1)$ equations may be immediately inverted to give the desired ω values from the S and s values.

The second general set of equations for the ω_k values is derived by first multiplying Eq. (7) by $2\mu^{2n-1} I^m(\tau, -\mu)$ for $n = 1,2,\cdots$ and integrating over $dm(\mu)$ to obtain

$$\tfrac{1}{2} S'_n(\tau) = -\sum_{k=m}^{N} (-1)^{k-m} c_k i_{0k}(\tau) i_{2n-1,k}(\tau). \tag{34}$$

On the lhs, Eq. (27) has been used, along with the fact that the integeral of an odd function over symmetric limits vanishes; on the rhs, Eqs. (14) and (16) have been used. For $n = 1$, Eq. (34) becomes

$$\frac{d}{d\tau} S_1(\tau) = \frac{d}{d\tau} \sum_{k=m}^{N} (-1)^{k-m} c_k \left(\frac{2k+1}{h_k}\right) [i_{1k}(\tau)]^2 \tag{35}$$

after use of Eq. (20). Integration of Eq. (35) over $0 \leqslant \tau \leqslant \tau_0$ and use of Eq. (12) leads to the final result

$$S_1^m = \sum_{k=m}^{N} (\omega_k/h_k) s_{1k}^m, \tag{36}$$

where the required surface integrals are defined as

$$S_1^m = S_1^m(0) - S_1^m(\tau_0) \tag{37}$$

and

$$s_{1k}^m = (-1)^{k-m}(2k+1)[(k-m)!/(k+m)!] \times \{[i_{1k}^m(0)]^2 - [i_{1k}^m(\tau_0)]^2\}. \tag{38}$$

In a manner similar to Eq. (33), the $(N+1)$ equations obtained by setting $m = 0,1,\ldots, N$ in Eq. (36) immediately lead to the values of ω_k/h_k, and hence to the ω values after use of Eq. (21).

Various mutations of Eqs. (30) and (36) also may be written down by selecting, for example, the even-m equations from Eq. (30) and the odd-m equations from Eq. (36). In such cases, the equations for the ω_k values are nonlinear, but remain upper triangular. Of course, if both Eqs. (30) and (36) are used for a particular value of m, then the equations also are no longer upper triangular.

For the case where $m = 0$, Eqs. (30) and (36) give the two equations

$$S_0^0 = \sum_{k=0}^{N} \omega_k s_{0k}^0, \tag{39}$$

$$S_1^0 = \sum_{k=0}^{N} (\omega_k/h_k) s_{1k}^0. \tag{40}$$

In the event that scattering is at most linearly anisotropic so that $N = 1$, these two equations are sufficient to determine ω_0 and ω_1 and are precisely the results derived earlier by Siewert;[9,10] the evaluation of the ω_0 and ω_1 from Eqs. (39) and (40) is somewhat more involved algebraically, however, than the use of either Eq. (30) or (36) since the equations are nonlinear.

It is tempting to seek additional equations for $m = 0$ by using equations for S_2, S_3, etc., as obtained from Eqs. (28) and (34). Complications arise, however, since the right-hand sides are no longer perfect differentials, although a solution has been completed for $N = 2$.[10]

IV. COMMENTS

Equations (30) and (36) are two schemes for expressing all the desired ω_k values in terms of Eqs. (15) and (22) for moments of the surface intensities. These equations are valid for any (known or unknown) slab thickness and for any finite order of anisotropic scattering N. The value of N required is easily ascertained since the moments defined in Eqs. (31), (32), (37), and (38) vanish for $m > N$. Thus, a major concern in the use of Eqs. (30) and (36) is that there be sufficient azimuthal asymmetry in at least one incident distribution so that the moments do not vanish at a prematurely low value, i.e., for $N' < N$. Of course, use of a collimated incident distribution like that in Eq. (3), with $\mu_0 \ll 1$, allays any such fears.

Of the two schemes, that in Eq. (36) may be somewhat more prone to numerical inaccuracies if the medium is very weakly absorbing, so that h_0 becomes very small. In such a case, the $m = 0$ equation of the set causes a problem because i_{10}^0, the net current, is a constant for a nonabsorbing medium.

Numerical accuracy of the ω_k values, obtained by measuring the emerging distributions from the slab for a given incident distribution, always may be checked by repeating the experiment and analysis for a different incident distribution.

ACKNOWLEDGMENT

This paper never would have been completed in this form without the suggestions of C.E. Siewert.[10]

[1]K.M. Case and P.F. Zweifel, *Linear Transport Theory* (Addison-Wesley, Reading, Massachusetts, 1967), pp. 99ff.
[2]K.M. Case, Phys. Fluids **16**, 1607 (1973); **18**, 927 (1975); **20**, 2031 (1977).
[3]C.E. Siewert, M.N. Özişik, and Y. Yener, Nucl. Sci. Eng. **63**, 95 (1977).
[4]E. Canfield, Nucl. Sci. Eng. **53**, 137 (1974).
[5]D.G. Cacuci and H. Goldstein, J. Math. Phys. **18**, 2436 (1977).
[6]N.J. McCormick and I. Kuščer, J. Math. Phys. **15**, 926 (1974).
[7]N.J. McCormick and J.A.R. Veeder, J. Math. Phys. **19**, 994 (1978); **20**, 216 (1979).
[8]C.E. Siewert, Nucl. Sci. Eng. **67**, 259 (1978).
[9]C.E. Siewert, "On establishing a two-term scattering law in the theory of radiative transfer," Z. Angew. Math. Phys. (to be published).
[10]C.E. Siewert, private communication.
[11]S. Chandrasekhar, *Radiative Transfer* (Oxford U.P., London and New York, 1950; Dover, New York, 1960).
[12]N.J. McCormick and I. Kuščer, J. Math. Phys. **7**, 2036 (1966).

Shear-free gravitational collapse

E. N. Glass

Physics Department, University of Windsor, Windsor, Ontario N9B 3P4
(Received 20 November 1978)

Spherically symmetric perfect fluids are studied under the restriction of shear-free motion. All solutions of the field equations are found by solving a single second order nonlinear equation containing an arbitrary function. It is shown that this arbitrary function is a geometric invariant, E, which measures the gravitational field energy, and it is shown that $E = \text{const}$ generates all the homogeneous density solutions. An improved proof is given for the nonexistence of any one-parameter equation of state. A number of exact solutions are presented and discussed.

1. INTRODUCTION AND DISCUSSION

Spherical shear-free collapse of a perfect fluid is interesting to study for two reasons: For a given spherical matter configuration, Raychaudhuri's equation[1] shows that it is the slowest collapse possible; It is simple to obtain many inequivalent analytic solutions of the Einstein field equations, since the number of unknown metric functions can be reduced from three to one.

It was recognized some time ago[2] that shear-free solutions of the Einstein equations could be generated by solving a single second order nonlinear equation containing an arbitrary function of the radial coordinate. The simplest form of this equation was first given by Faulkes.[3] In this work, it is shown that the arbitrary time-independent function is E, where E measures the purely gravitational field energy inside spheres of radius R. Each choice of E specifies a class of shear-free solutions. In particular, it is shown that $E = \text{const}$ generates all the homogeneous density solutions.

It has recently been proved by Mansouri[4] that an equation of state $p = p(w)$ can never hold for spherical perfect fluids in shear-free collapse with closed $p = 0$ surfaces. Mansouri's proof has an invalid assumption which is not used in the version given here.

The argument showing the lack of an equation of state holds equally well for Newtonian systems in shear-free collapse, and it is interesting to ask whether this lack is related to a lack of local thermodynamic equilibrium (LTE)? There appears to be a relation since the (rate of) shear-free condition does not allow the fluid to relax in response to the (increasing) tidal force. There are two characteristic scales relevant to the question.

Measured values of the Riemann tensor provide (in principle) a scale L for the tidal force where magnitude $(R_{\mu\nu\rho\sigma}) \sim L^{-2}$. In order to maintain the validity of the fluid picture, there must be a minimum particle number density and hence an upper bound $D_{\max}$ for D, the interparticle distance. If we consider the neighborhood of any point in the fluid, the shear-free condition will play the role of a net external force (and thus not allow statistical equilibrium) when $D \sim L$. This will occur in two situations: When the radius of curvature L is very large and $D > D_{\max}$ is equally large. In this case the system will be too diffuse for the fluid picture to be valid. The other situation occurs when both L and D are equally small (compared with the Schwarzschild radius of the entire system). This happens in the neighborhood of a curvature singularity where it is unreasonable to expect LTE to hold.

The situations considered here (away from singularities) have $D \ll L$. (Among all isolated collapsing systems for which the fluid picture is valid, the closest correspondence between D and L that is observable occurs when the core of a typical globular cluster falls through its Schwarzschild radius with $L \sim 10D$.) For $D \ll L$ the effects of tidal forces can be neglected locally and (assuming the nongravitational interactions will allow) LTE will hold.

All the explicit collapse solutions examined in this work have trapped surfaces inside of which singularities occur. It remains to be shown in general that, if a curvature singularity exists, it lies inside a trapped surface.

The paper is organized as follows: The field equations are formulated in Sec. 2, where it is shown that the system can be completely described by one function $F(x,t)$, which is the solution of a nonlinear second order equation containing an arbitrary function, and the boundary conditions. In Sec. 3, the arbitrary function is identified with E and related to the total energy scalar for the system. The equations of motion are presented in Sec. 4, and a proof is given for the lack of an equation of state. The condition for trapped surfaces is formulated in Sec. 5. Sections 6 and 7 treat all the conformally flat interiors. In Sec. 8, it is shown that all homogeneous density solutions are generated by $E = \text{const.}$ and they all contain trapped surfaces. Nonuniform density solutions are reviewed in Secs. 9 and 10, and adiabatic flow is treated in Sec. 11.

2. FIELD EQUATIONS

We consider the motion of a spherical perfect fluid with energy–momentum tensor

$$T^{\mu\nu} = w u^{\mu} u^{\nu} - p \gamma^{\mu\nu}, \tag{1}$$

where u^{μ} is the unit 4-velocity of matter, w is the mass–energy density, p is the isotropic pressure, and $\gamma^{\mu\nu} = g^{\mu\nu} - u^{\mu} u^{\nu}$ is the induced metric on the 3-surfaces orthogonal to

 0022-2488/79/071508-06$01.00

u^μ. The spacetime is divided into an exterior vacuum region covered by the Schwarzschild metric and an interior region with metric

$$g_{\mu\nu}\,dx^\mu dx^\nu = A^2 dt^2 - B^2\,dr^2 - R^2\,d\Omega^2, \qquad (2)$$

where

$$A = A(r,t), \quad B = B(r,t), \quad R = R(r,t).$$

R is the usual scalar field constructed from the three Killing vectors of the rotation group, and t is a comoving time parameter with $u^\mu = A^{-1}\delta^\mu_t$.

The fluid moves according to

$$u_{\mu;\nu} = a_\mu u_\nu + \sigma_{\mu\nu} + \tfrac{1}{3}\theta\gamma_{\mu\nu}, \qquad (3a)$$

$$a_\mu = -(A'/A)\,\delta^r_\mu, \qquad (3b)$$

$$\theta = A^{-1}(\dot B/B + 2\dot R/R), \qquad (3c)$$

$$\sigma^\mu_\nu = A^{-1}(\dot R/R - \dot B/B)(\tfrac{1}{3}\gamma^\mu{}_\nu - \delta^\mu{}_r\delta^r{}_\nu), \qquad (3d)$$

where primes denote $\partial/\partial r$ and dots denote $\partial/\partial t$. The shear-free condition is

$$\dot R/R = \dot B/B, \quad \text{or} \quad R = f(r)B. \qquad (4)$$

Choosing $f(r) = r$ with no loss of generality, the metric is henceforth restricted to[5]

$$g_{\mu\nu}\,dx^\mu dx^\nu = A^2\,dt^2 - B^2(dr^2 + r^2\,d\Omega^2). \qquad (5)$$

For the field equations, zero mass-energy flux yields

$$\dot B = AB\,H(t), \qquad (6)$$

where $H(t)$ is an arbitrary function of integration.

The equation due to pressure isotropy is

$$A''/A + B''/B - (2B'/B + 1/r)(A'/A + B'/B) = 0. \qquad (7)$$

Considerable simplification is achieved by using Eq. (6) to eliminate A from Eq. (7):

$$F_{xx} + J(x)F^2 = 0, \qquad (8)$$

where $x{:} = r^2$, $F{:} = B^{-1}$, *and* $J(x)$ is an arbitrary function of integration. The density and pressure are given by

$$8\pi w = 3H^2 + 4F^2[2x(F_{xx}/F) + 3(F_x/F) - 3x(F_x/F)^2], \qquad (9)$$

$$8\pi p = (F/\dot F)(H^2)^{\cdot} - 3H^2 + 4F^2\{[\log(\dot F/F^2)]_x - x[(\log F)_x]\,[\log(\dot F^2/F^3)]_x\}, \qquad (10)$$

where Eq. (6) has again been used to eliminate A.

When $\dot w_x \neq 0$, Eqs. (9) and (10) yield

$$p(x,t) = -w + \dot w w_x/\dot w_X$$

or

$$p(r,t) = -w + \dot w w'/\dot w'. \qquad (11)$$

The system of equations (8), (9), and (10) is complete[6] for the three unknowns p, w, and F. Only $J(x)$ can be specified arbitrarily since $H(t)$ must satisfy the first order time equation $p(x_b,t) = 0$ at the boundary x_b of the fluid where the interior solution is matched to exterior Schwarzschild. At the $p = 0$ boundary, Eq. (10) is an exact time derivative yielding

$$H^2 = K_1F_b^3 - 4F_b(F_x)_b + 4x_b(F_x)_b^2, \qquad (12)$$

where K_1 is a constant of integration.

3. INTERPRETATION OF *H(t)* AND *J(x)*

Comparing Eqs. (3c) and (6) immediately yields

$$\theta = 3H. \qquad (13)$$

Thus $H(t)$ gives the rate of expansion of the fluid and can be called the "Hubble parameter" for the system.

$J(x)$ turns out to label the conformal curvature of the fluid system. With the definition

$$\psi_2{:} = \tfrac{1}{2}C_{\mu\nu\rho\sigma}\theta^\mu\varphi^\nu\theta^\rho\varphi^\sigma/\theta^\alpha\theta_\alpha\varphi^\beta\varphi_\beta$$

for the vector fields $\theta^\mu = \delta^\mu{}_\theta$ and $\varphi^\mu = \delta^\mu{}_\varphi$, it follows from the decomposition of the Riemann tensor into its Weyl and Ricci parts that[7]

$$\tfrac{4}{3}x^{5/2}J(x) = -R^3\psi_2. \qquad (14)$$

ψ_2 (identical with the Newman–Penrose scalar for spherical systems) labels the Weyl tensor invariant:

$$C^{\mu\nu\rho\sigma}C_{\mu\nu\rho\sigma} = 48\psi_2^2.$$

A total energy scalar[8] can be invariantly defined in terms of the sectional curvature of $r = \text{const.}$, $t = \text{const.}$ surfaces:

$$m{:} = (\tfrac{1}{2}R^3)R_{\mu\nu\rho\sigma}\theta^\mu\varphi^\nu\theta^\rho\varphi^\sigma/\theta^\alpha\theta_\alpha\varphi^\beta\varphi_\beta,$$

which yields the formula (valid for the general case with shear)

$$m = \tfrac{4}{3}\pi R^3 w + E, \qquad (15)$$

where $E{:} = -R^3\psi_2$. E can be interpreted as purely gravitational field energy (not binding energy) within spheres of surface radius R.

Equation (10) with $p = 0$ no longer determines x_b since, when H^2 is eliminated using Eq. (12), (10) is satisfied identically at x_b. At the boundary, the exterior Schwarzschild mass M must equal the total energy parameter

$$M = m(x_b,t),$$

and so K_1 is determined from Eq. (15) with the use of Eqs. (9), (12), and (14):

$$K_1x_b^{3/2} = 2M, \quad K_1 > 0.$$

The Bianchi identities can be written in terms of the total energy parameter as

$$m_x = \tfrac{4}{3}\pi w(R^3)_x,$$

$$\dot m = -\tfrac{4}{3}\pi p(R^3)^{\cdot},$$

and with Eq. (15) in terms of the E function as

$$E_x = -\tfrac{4}{3}\pi R^3 w_x, \qquad (16a)$$

$$\dot E = -\tfrac{4}{3}\pi R^3(p + w)(\dot R/R - \dot B/B). \qquad (16b)$$

It is clear from Eq. (16b) that the shear-free case corresponds to $E = E(x)$, and that all conformally flat solutions must be shear-free.

4. EQUATION OF STATE

The equations of motion $T^{\mu\nu}{}_{;\nu} = 0$ have components

$$u^{\mu}\nabla_{\mu} w + (p + w)\theta = 0, \qquad (17a)$$

$$(p + w)a_{\mu} = p_{,\nu}\gamma^{\nu}{}_{\mu}. \qquad (17b)$$

These can be rewritten with the aid of Eqs. (3a), (6) and (13) as

$$\dot{w} + 3(p + w)\tau = 0, \qquad (18a)$$

$$p' + (p + w)(\log\tau)' = 0, \qquad (18b)$$

where $\tau := \dot{B}/B$. Eliminating τ between (18a) and (18b) yields Eq. (11). The case $\dot{w}' = 0$ can now be considered: The solution of $\dot{w}' = 0$ is

$$w = h(t) + s(r),$$

and upon substituting into Eq. (18) one obtains $s' = 0$ or $w = h(t)$ only. Thus one can use Eq. (11) whenever $w = w(r,t)$.

Mansouri[4] has proved that spherical perfect fluids bounded by closed surfaces and in shear-free flow never admit equations of state $p = p(w)$. An improved proof is given here: Assume an equation of state $p = p(w)$ exists for the shear-free system. At the fluid boundary $p(x_b, t) = p(w_b) = 0$, so w must satisfy

$$w(x_b, t) = w_b \text{ const.}$$

(otherwise $p = p(t)$ at x_b). Using Eqs. (9) and (12) to evaluate w at x_b yields

$$8\pi w_b = (3K_1 - 8x_b J_b)F_b^3.$$

Since $R = x^{1/2}/F$, F_b is constant only for static interiors. Therefore, w_b const. implies $w_b = 0$, and also $J_b = \frac{3}{4}Mx_b^{-5/2}$, in agreement with Eqs. (14) and (15).

In addition, there exists a function $N(w)$ such that

$$3dN/N = dw/(p + w), \qquad (19)$$

which along with Eq. (17a) yields

$$(N^3 u^{\mu})_{;\mu} = 0.$$

This conserved current, together with $(-g)^{1/2} = AB^3 r^2 \sin\theta$ from metric (5) and $u^{\mu} = A^{-1}\delta^{\mu}_t$, implies

$$N^3 B^3 = h(r)$$

for an arbitrary function $h(r)$, or

$$N^3(x,t) = F^3\beta^3(x), \quad \beta \text{ arbitrary}. \qquad (20)$$

At x_b $N_b(w_b) = N_b(0)$ const., and since $F_b = F_b(t)$ it must true be that $\beta(x_b) = \beta_b = 0$, and that $N_b = 0$.

The condition of functional dependence for p and w is

$$p_x \dot{w} - \dot{p} w_x = 0,$$

which applied to Eq. (11) yields

$$(\log w_x)_x = (\log\dot{w})_x. \qquad (21)$$

The general solution of Eq. (21) is

$$w = w(\varphi), \quad \varphi = \alpha(x)\,\gamma(t)$$

for arbitrary functions α and γ. Since $w_b = 0$, it follows that $\alpha_b = \varphi_b = 0$. From Eq. (20) and $N = N(w) = N(\varphi)$

$$F = N(\varphi)/\beta(x). \qquad (22)$$

[Mansouri makes the invalid assumption that $\beta(x) = f(x)\,\alpha(x)$ where $f(x_b) \neq 0$. See Eqs. (34) and (64) of Ref. 4.]

Equations (3b) and (17b) can be integrated using Eq. (19):

$$N^3 = (p + w)A\delta(t), \qquad (23)$$

where the integration function $\delta(t)$ can be set to unity by choice of boundary condition and t coordinate normalization. $A = -\dot{F}/FH$ and Eq. (22) establish

$$A = -(\varphi\hat{N}/N)(\dot{\gamma}/\gamma/H),$$

where overhead carets denote $d/d\varphi$. Since $A = A(\varphi)$, it follows

$$H = \text{const.}(\dot{\gamma}/\gamma) = -\dot{\gamma}/\gamma,$$

$$A = \varphi\hat{N}/N, \qquad (24a)$$

and at x_b

$$A_b = (\varphi\hat{N}/N)_b = 1, \qquad (24b)$$

by normalizing the t coordinate.

Having established the necessary boundary conditions, the consistency of Eq. (22) is examined by substituting for F in field Eq. (8):

$$(\alpha_x/\alpha)^2 = k_1(J/\beta), \qquad (25a)$$

$$(\alpha_{xx}/\alpha) - 2(\alpha_x/\alpha)(\beta_x/\beta) = k_2(J/\beta), \qquad (25b)$$

$$2(\beta_x/\beta)^2 - (\beta_{xx}/\beta) = k_3(J/\beta), \qquad (25c)$$

$$k_1\varphi^2(d\hat{N}/d\varphi) + k_2\varphi\hat{N} + k_3 N + N^2 = 0, \qquad (25d)$$

with separation constants k_1, k_2, k_3. Equations (25a) and (25b) can be integrated:

$$\alpha_x/\alpha^{(k_2/k_1)} = (k_1 k_4)^{1/2}\beta^2, \qquad (26a)$$

$$J = k_4\beta^5\alpha^{(k_2/k_1)-2}. \qquad (26b)$$

Equation (25d) and boundary condition (24b) establish (using the fact that $\hat{A}$ is bounded at x_b)

$$k_2 + k_3 = 0,$$

which allows Eqs. (25b) and (25c) to be integrated:

$$\beta = c_2\alpha/(c_1 - x). \qquad (27)$$

Substituting (27) into (26b) yields

$$J(x) = k_4 c_2^5 \alpha^{3 + 2(k_2/k_1)}(c_1 - x)^{-5}. \qquad (28)$$

There are two cases: (i) $c_1 \neq x_b$, in which case J_b = nonzero const implies $\alpha_b \neq 0$. But $\alpha_b = 0$, so this implies $k_2/k_1 = -3/2$. (ii) $c_1 = x_b$, *where* J_b = const. implies

$$\alpha(x) = (x_b - x)^{5/[3 + 2(k_2/k_1)]}\nu(x), \quad \nu(x_b) \neq 0. \qquad (29)$$

Substituting Eq. (27) into (26a) and integrating produces

$$\alpha^{1 + (k_2/k_1)} = (c_1 - x)/[k_5(c_1 - x) + k_6]. \qquad (30)$$

Comparison with Eq. (29) establishes $k_2/k_1 = -2/3$, and thus Eq. (30) has two cases: (i) $c_1 \neq x_b$ and $k_2/k_1 = -3/2$. (ii) $c_1 = x_b$ and $k_2/k_1 = -2/3$. Equation (28) yields

$$J(x) = k_4 c_2^5 (c_1 - x)^{-5}, \quad k_2/k_1 = -3/2, \qquad (31a)$$

$$= k_7[k_5(x_b - x) + k_6]^{-5}, \quad k_2/k_1 = -2/3. \quad (31b)$$

Solutions (27), (30), and (31) can now be tested for consistency by using the Bianchi identities. With $\varphi \hat{w} = 3N^3$ from Eqs. (19), (23), and (24), it follows that

$$w_x = 3N^3(\alpha_x/\alpha).$$

Equation (16a) can be written, using $R = x^{1/2}\beta/N$, as

$$\tfrac{5}{2}J + xJ_x = -3\pi\beta^3(\alpha_x/\alpha). \quad (32)$$

Upon substitution, one sees that solutions (27), (30), and (31) are inconsistent with Eq. (32). Therefore, shear-free flow is inconsistent with a single one-parameter equation of state.

5. TRAPPED SURFACES

A null tetrad is chosen for metric (2):

$$l^\mu = \frac{1}{\sqrt{2}}(A^{-1}\delta^\mu_t + B^{-1}\delta^\mu_r),$$

$$n^\mu = \frac{1}{\sqrt{2}}(A^{-1}\delta^\mu_t - B^{-1}\delta^\mu_r),$$

$$m^\mu = \frac{1}{\sqrt{2}}R^{-1}\left(\delta^\mu{}_\theta + \frac{i}{\sin\theta}\delta^\mu{}_\varphi\right),$$

where l^μ and n^μ are outgoing and incoming principal null vectors with respective expansion rates given by[9]

$$\rho = -l^\alpha\nabla_\alpha \log R, \quad \mu = n^\alpha\nabla_\alpha \log R. \quad (33)$$

The Newman–Penrose field equation[10]

$$\delta\alpha + \bar{\delta}\alpha - 4\alpha^2 - \mu\rho = -\psi_2 + \Lambda + \Phi_{11}$$

yields

$$\tfrac{1}{2}R^{-2} - \mu\rho = \tfrac{4}{3}\pi w - \psi_2. \quad (34)$$

The total energy parameter m of Eq. (15) can now be written as

$$2m = R - 2\mu\rho R^3.$$

A trapped surface will exist when both μ and ρ are positive, with the marginally trapped surface given by $\rho = 0$. The time history of the marginally trapped surface constitutes the apparent horizon with equation

$$2m = R \quad \text{at } r = r_{\text{ah}}. \quad (35)$$

To test a solution of Eqs. (8), (9), and (10) for the existence of a trapped surface, one sets $\rho = 0$ in Eq. (33):

$$F + x^{1/2}H - 2xF_x = 0. \quad (36)$$

For a solution of Eq. (8) given as $F(x,t)$, a trapped surface exists if Eq. (36) has a real solution $x^{1/2}{}_{\text{ah}}$ corresponding to a positive R_{ah}. The contributions of the density and conformal factor to x_{ah} can be shown explicitly by rewriting Eq. (36) with the use of Eqs. (8) and (9):

$$x^2(8JF^3) + x(8\pi w) - 3F^2 = 0, \quad \text{at } x = x_{\text{ah}}.$$

6. CONFORMALLY FLAT INTERIORS

All the conformally flat solutions are found upon setting $J(x) = 0$. The complete solution of Eq. (8) is then

$$F = xh_1(t) + h_2(t). \quad (37)$$

Just as in the static case, all conformally flat Schwarzschild interiors have uniform density. Substituting (37) into Eq. (9) yields

$$8\pi w = 3(H^2 + 4h_1h_2). \quad (38)$$

The solutions of Bonnor and Faulkes,[11] and Bondi,[12] belong to this class. The pressure, from Eq. (10), is given by

$$8\pi p = \frac{(xh_1 + h_2)}{(x\dot{h}_1 + \dot{h}_2)}[(H^2)^{\cdot} + 4\dot{h}_1(h_2 - xh_1)]$$

$$- 3H^2 + 4h_1(xh_1 - 2h_2). \quad (39)$$

The condition $t_0\dot{h}_1 = h_1$, and $t_0\dot{h}_2 = h_2$, leads to geodesic flow with $p = p(t)$ and a Robertson–Walker metric so that case is excluded in this work.

At the $p = 0$ boundary, Eq. (39) is an exact time derivative yielding [from Eq. (12)]

$$H^2 = -4h_1h_2 + K_1(x_bh_1 + h_2)^3, \quad (40)$$

where K_1 is a positive constant of integration. The time dependence of the boundary is given by

$$R_b = (2M/K_1)^{1/3}/(x_bh_1 + h_2), \quad (41)$$

and the apparent horizon is located from Eq. (36):

$$x^{1/2}_{\text{ah}} = [H \pm (H^2 + 4h_1h_2)^{1/2}]/2h_1.$$

Use of Eqs. (40) and (41) yields the relation

$$R^3_b = 2MR^2_{\text{ah}}. \quad (42)$$

It is now clear that all the conformally flat solutions with $p = 0$ boundary have an apparent horizon. With the density

$$8\pi w = 3K_1(x_bh_1 + h_2)^3, \quad (43)$$

Eqs. (39)–(43) give a complete description of all conformally flat interior solutions in terms of two arbitrary functions $h_1(t)$ and $h_2(t)$.

7. SOME CONFORMALLY FLAT EXACT SOLUTIONS

(a) *Collapse*: Upon making the simple choice

$$h_2 = \alpha h_1 + \beta, \quad \alpha \text{ and } \beta \text{ const.},$$

Eqs. (39)–(43) yield

$$H^2 = K_1[(x_b + \alpha)h_1 + \beta]^3 - 4h_1(\alpha h_1 + \beta), \quad (44a)$$

$$8\pi w = 3K_1[(x_b + \alpha)h_1 + \beta]^3, \quad (44b)$$

$$p/w = [h_1 + \beta/(x + \alpha)][h_1 + \beta/(x_b + \alpha)]^{-1} - 1, \quad (44c)$$

$$R_b = (4\pi w/3M)^{-1/3}, \quad (44d)$$

$$R_{\text{ah}} = (8\pi w/3)^{-1/2}, \quad (44e)$$

with $x_b = (2M/K_1)^{2/3}$.

Choose h_1 to be an increasing function of time such as $\exp(t/t_0)$. Choose α positive (to avoid an additional pressure singularity at points other than $R = 0$). K_1 must be positive, and choose β positive. Adjust the constants so that H^2 is always positive. The negative square root of H^2 gives the collapse interpretation, and H becomes increasingly negative as time advances. R_b will decrease faster than $R_{\rm ah}$ so at some finite time the boundary will fall through the apparent horizon and an absolute Schwarzschild horizon will remain.

(b) *Explosion*: Choose h_2 as above, and h_1 to be decreasing function of time such as $\exp(-t/t_0)$. Choose the constants as above. The positive square root of H^2 gives the expansion interpretation, and H decreases in time in response to the gravitational attraction. At $t = 0$, the constants can be chosen so that $R_b > R_{\rm ah}$, but $R_{\rm ah}$ will increase faster than R_b so that at some finite time the exploding matter will disappear behind a Schwarzschild horizon.

(c) *Oscillation*: See Bondi's[13] discussion of choices for h_1 and h_2.

8. UNIFORM DENSITY

Is conformal flatness a necessary condition for uniform density? To answer this question (with answer no) one sets $w_x = 0$ in Eq. (16a). All uniform density solutions are generated from $E =$ const. or $J(x) = kx^{-5/2}$.

$$F_{xx} + kx^{-5/2}F^2 = 0. \tag{45}$$

The previous section treated $k = 0$, and so here $k \neq 0$ is considered. The double transformation

$$F = x^{1/2}\Gamma, \quad x = e^y,$$

gives rise to

$$\Gamma_{yy} - \tfrac{1}{4}\Gamma + k\Gamma^2 = 0. \tag{46}$$

The general solution of Eq. (46) is a Weierstrass P function which satisfies the standard equation

$$P_{yyy} = 12PP_y.$$

The first integral of Eq. (46) allows a special solution in terms of elementary functions:

$$\tfrac{1}{2}\Gamma_y^2 + \tfrac{1}{3}k\Gamma^3 - \tfrac{1}{8}\Gamma^2 + c_1(t) = 0,$$

where choosing $c_1(t) = 0$ yields the solution

$$F = (3/2k)xc_2(1 - x^{1/2}c_2)^{-2}, \tag{47}$$

with $c_2(t)$ the second function of integration.

There is an apparent horizon, since Eq. (36) yields

$$y^3 + y^2(\alpha - 3) + y(\alpha + 3) - 1 = 0,$$

where $\alpha = 3/(2kH)$, and $y = x_{\rm ah}^{1/2}c_2$.

The cubic has one real root when $\alpha^2 < 108$, and three real unequal roots when $\alpha^2 > 108$.

9. $J(x)$ CONSTANT

The simplest solutions with nonuniform density arise in the case $J(x) = K_2$ nonzero:

$$F_{xx} + K_2F^2 = 0. \tag{48}$$

The general solution of Eq. (48) is a Weierstrass P function but, as in Sec. 8, the first integral of (48) allows special solutions in terms of elementary functions:

$$\tfrac{1}{2}F_x^2 + \tfrac{1}{3}K_2F^3 + h_1(t) = 0.$$

Choosing $h_1(t) = 0$ yields the solution

$$F = (\alpha x + h_2)^{-2}, \tag{49}$$

where $h_2(t)$ is the second function of integration, and $K_2 = -6\alpha^2$. K_2 negative is necessary in order that R be positive. This solution was examined by Faulkes,[3] but some of the possibilities were overlooked in that work.

Equations (9), (10), and (12) yield

$$\tfrac{8}{3}\pi w = H^2 - 8\alpha(\alpha x + h_2)^{-5}, \tag{50a}$$

$$\tfrac{8}{3}\pi p = 8\alpha(\beta + h_2)(\alpha x + h_2)^{-6} - H^2, \tag{50b}$$

$$H^2 = 8\alpha(\beta + h_2)(\alpha x_b + h_2)^{-6}, \tag{50c}$$

with $\beta := 3\alpha x_b + K_1/8\alpha, \quad x_b = (2M/K_1)^{2/3}$.

An interesting difference between solution (50) and the ones above is that an apparent horizon is not always present for this solution. With the use of Eq. (36), one obtains

$$y^7 + 3y^5 + 3y^3 + 5h_3y^2 + y + h_3 = 0, \tag{51a}$$

where $y := (\alpha x_{\rm ah}/h_2)^{1/2}$, and $h_3 := (\alpha/h_2^5)^{1/2}/H$. In this case α and h_2 are both positive (or both negative). y must be positive in order that $R_{\rm ah}$ be positive. Descartes' rule of signs requires at least one sign change in Eq. (51a) for the possibility of a positive root. Thus h_3 and hence H must be negative, and so there is no apparent horizon for an expanding solution.

When there is a relative sign difference between α and h_2, then H^2 positive requires $3\alpha^2x_b + K_1/8 + \alpha h_2 > 0$. (This is the case Faulkes discussed.) Equation (36) yields

$$y^7 + 3y^5 + 3y^3 + 5h_3y^2 - y - h_3 = 0, \tag{51b}$$

where $y := (-\alpha/h_2)^{1/2}x_{\rm ah}^{1/2}$,
and $h_3 := (-\alpha/h_2^5)^{1/2}/H$. Equation (51b) can have positive roots for H both positive and negative.

10. COLLAPSING STAR CLUSTER MODEL

The solution of Glass and Mashhloon[14] (which includes the solutions of McVittie,[15] and of Nariai[16] as special cases) has

$$F = (\lambda/\lambda_0)(\gamma x + \delta)(\alpha x + \beta)[(\alpha x + \beta)^{1/2} + (\lambda/2)(\gamma x + \delta)^{1/2}]^{-2},$$

where $\lambda = \lambda(t)$, $\alpha, \beta, \gamma, \delta, \lambda_0$ are constants. Equation (8) yields[17]

$$J(x) = (3\lambda_0/4)\Delta^2(\alpha x + \beta)^{-5/2}(\gamma x + \delta)^{-5/2},$$

with $\Delta = \alpha\delta - \beta\gamma$.

The collapsing star cluster is modeled for parameter ranges $\alpha > 0$, $\beta \geqslant 0$, $\gamma < 0$, and $\delta > 0$. The static case ($\lambda = \lambda_0$, $\alpha > 0, \beta > 0, \gamma = 0, \delta = 1$) is the relativistic Plummer model[18] for a globular cluster.

There is a pressure singularity (and hence a Riemann tensor singularity) at

$$x_s = [\delta(\lambda/2)^2 - \beta][\alpha - \gamma(\lambda/2)^2]^{-1}.$$

The radial coordinate is restricted to the range $x_s < x \leqslant x_b$ with x_s always inside a trapped surface.

11. ADIABATIC FLOW

Assuming local thermodynamic equilibrium,

$$Tds = du + pd(1/\rho), \tag{52}$$

where s is the specific entropy, u the specific internal energy, and ρ the proper mass density. u and ρ are related to the mass–energy density by

$$w = \rho(1 + u).$$

Upon rewriting Eq. (52) in comoving coordinates as

$$T\dot{s} = [\dot{w}/(w + p) - \dot{\rho}/\rho](w + p)/\rho,$$

comparison with Eq. (17a) shows that adiabatic motion implies proper mass conservation

$$(\rho u^\mu)_{;\mu} = 0$$

and vice versa. Since a conserved 4-current for the system is

$$(B^{-3}u^\mu)_{;\mu} = (R^{-3}u^\mu)_{;\mu} = 0, \tag{53}$$

we are led to identify

$$\rho(r,t) = \rho_0 R^{-3}, \tag{54}$$

where $\rho_0 = \rho_0(r)$, thus maintaining adiabatic flow as a feature of the systems considered here.

ACKNOWLEDGMENTS

I have profited from discussions with Dr. G. Szamosi and correspondence with Dr. B. Mashhoon.

[1] A. Raychaudhuri, Z. Astrophys. **43**, 161 (1957).
[2] M. Wyman, Phys. Rev. **70**, 396 (1946).
[3] M.C. Faulkes, Prog. Theor. Phys. **42**, 1139 (1969).
[4] R. Mansouri, Ann. Inst. H. Poincaré **27**, 175 (1977).
[5] The first author to recognize that metric (5) (up to change of r) is necessary and sufficient for the shear-free condition was H. Nariai, Prog. Theor. Phys. **40**, 1013 (1968).
[6] In order that the metric be admissable at $R = 0$ the condition $\lim_{r\to 0^\cdot}(rR'/R) = 1$ or $\lim_{r\to 0^\cdot}(xF_x/F) = 0$ is required.
[7] A less tedious method of verifying (14) is to use Eqs. (8), (9), (33), and (34).
[8] m is the commonly used energy parameter, cf., C.W. Misner and D.H. Sharp, Phys. Rev. **136**, B 571 (1964).
[9] E. Newman and R. Penrose, J. Math. Phys. **3**, 566 (1962).
[10] For the tetrad chosen here, spherical symmetry yields six nonzero real spin coefficients $\mu, \rho, \alpha, \beta, \epsilon$, and γ, with $\alpha = \quad -\beta = -(2\sqrt{2}R)^{-1}\cot\theta$.
[11] W.B. Bonnor and M.C. Faulkes, Mon. Not. R. Astron. Soc. **137**, 239 (1967).
[12] H. Bondi, Mon. Not. R. Astron. Soc. **142**, 333 (1969).
[13] Equation (22) of Ref. 12 identifies the conformally flat class.
[14] E.N. Glass and B. Mashhoon, Astrophys. J. **205**, 570 (1976).
[15] G.C. McVittie, Mon. Not. R. Astron. Soc. **93**, 325 (1933).
[16] H. Nariai, Prog. Theor. Phys. **38**, 92 (1967).
[17] The scale freedom of Eq. (8) is $J \to k_0 J$, $F \to k_0^{-1}F$ for k_0 const. This scale freedom always allows the choice $w_b = 0$, which is made for the collapsing star cluster model.
[18] E.D. Fackerell, Astrophys. J. **165**, 489 (1971).

Boost matrix elements of the homogeneous Lorentz group

M. A. Rashid

Department of Mathematics, Ahmadu Bello University, Zaria, Nigeria
(Received 7 July 1978)

The boost matrix elements of the homogeneous Lorentz group are expressed in terms of analytic continuation of the Clebsch–Gordan coefficients of the three-dimensional rotation group.

I. INTRODUCTION

The matrix elements of finite hyperbolic rotation (boost) of the homogeneous Lorentz group for irreducible unitary representations belonging to the principal series have been computed by several authors using both the global and the local methods.[1-4] Usually these matrix elements are expressed as a *finite* double sum of the product of some gamma functions and an ${}_2F_1$, with an argument which involves the boost.[4,5] However, when we examine the unitary irreducible representations (finite dimensional) of the O(4) group, on account of its decomposition as $O(3) \otimes O(3)$, we can express the corresponding matrix elements as a single finite sum of a product of two Clebsch–Gordan (CG) coefficients and an exponential in terms of the corresponding rotation.[6] We wish to rewrite the boost matrix elements of the homogeneous Lorentz group essentially in the same form involving a product of two CG coefficients and an exponential in the boost. One expects such an expression to exist on account of two reasons. First, dynamically, it has been shown that the scattering amplitude expressions at zero momentum transfer in terms of the matrix elements of the O(4) and O(3,1) groups are continuable into one another.[7] Secondly, though the boost matrix elements of the homogeneous Lorentz group are usually expressed using the argument $(1 - e^{-2a})$, we can nevertheless express them also in terms of e^{-2a} using, e.g., some analytic continuation. This results in an expression for the boost matrix D in terms of two E's.[8] We show that when we attempt this exercise, we obtain an expression for the D matrix in terms of sum of two single infinite sums of products of CG coefficients (which are continuations of the CG coefficients of the three-dimensional rotation group with some angular momenta and their third components as complex) and an exponential in the boost. We also show that we can obtain the matrix elements of the corresponding O(4) rotation by giving special values to the variables appearing in the expression for the Lorentz group boost matrix elements.

This paper is organized as follows. In Sec. II, we attempt to rewrite the known expression for the D matrix in terms of sum of two infinite series in the argument e^{-2a}. In Sec. III, we show that this expression can indeed be represented using analytic continuation of the CG coefficients of the three-dimensional rotation group. In Sec. IV, we show how the matrix elements of the corresponding O(4) rotation are obtained.

II. BOOST MATRIX ELEMENTS OF THE HOMOGENEOUS LORENTZ GROUP

The irreducible representations of the homogeneous Lorentz group belonging to the principal series are characterized by two numbers ν and ρ where $|\nu| = 0, \frac{1}{2}, 1 \cdots$ and $-\infty \leqslant \rho \leqslant \infty$. The matrix element $D^{\nu\rho}_{Jm,J'm'}(a)$ of the hyperbolic rotation characterized by the parameter a where $-\infty \leqslant a \leqslant \infty$ is given by[4,5]

$$D^{\nu\rho}_{Jm,J'm'}(a) = \delta_{mm'} D^{\nu\rho}_{JJ'm}(a), \tag{1}$$

where

(i) $J - |\nu|$, $J' - |\nu|$, $J \pm m$, $J' \pm m$ are nonnegative integers,

(ii) $-\min(J,J') \leqslant m \leqslant \min(J,J')$,

and

$$\begin{aligned} D^{\nu\rho}_{JJ'm}(a) = {} & \alpha^{\nu\rho *}_{J} \alpha^{\nu\rho}_{J'} [(J + J' + 1)!]^{-1} \\ & \times [(2J+1)(J+\nu)!(J-\nu)!(J+m)!(J-m)!(2J'+1)(J'+\nu)!(J'-\nu)!(J'+m)!(J'-m)!]^{1/2} \\ & \times \sum (-1)^{d+d'} \frac{(\nu+m+d+d')!(J+J'-\nu-m-d-d')! \exp[-a(2d'+\nu+m+1-\frac{1}{2}i\rho)]}{d!(J-\nu-d)!(J-m-d)!(\nu+m+d)!d'!(J'-\nu-d')!(J'-m-d')!(\nu+m+d')!} \\ & \times {}_2F_1(J'+1-\tfrac{1}{2}i\rho, \nu+m+1+d+d', J+J'+2; 1-e^{-2a}). \end{aligned} \tag{2}$$

In the above expression, the phase $\alpha^{\nu\rho}_{J}$ is given by

 0022-2488/79/071514-06$01.00

$$\alpha_J^{\nu\rho} = \prod_{s=|\nu|}^{J} \frac{(-2s+i\rho)}{(4s^2+\rho^2)^{1/2}} = \frac{\Gamma(-|\nu|+\frac{1}{2}i\rho+1)}{|\Gamma(|-\nu|+\frac{1}{2}i\rho+1)|} \times \frac{|\Gamma(-J+\frac{1}{2}i\rho)|}{\Gamma(-J+\frac{1}{2}i\rho)}, \tag{3}$$

i.e.,

$$\alpha_J^{\nu\rho *}\alpha_{J'}^{\nu\rho} = \left(\frac{\Gamma(-J+\frac{1}{2}i\rho)\Gamma(-J'-\frac{1}{2}i\rho)}{\Gamma(-J-\frac{1}{2}i\rho)\Gamma(-J'+\frac{1}{2}i\rho)}\right)^{1/2}. \tag{4}$$

In Eq. (2) and in the following, we shall omit the range of summation *for those indices* for which the range is dictated by the nonnegativity of the arguments of the real factorials. Also we shall generally use the notation $x!$ for $\Gamma(x+1)$ even when x is not necessarily a nonnegative integer. Again since using

$${}_2F_1(a,b,c;z) = (1-z)^{-a}\ {}_2F_1(a,c-b,c;z/(z-1)), \tag{5}$$

we see that

$$D_{JJ'm}^{\nu\rho}(a) = D_{JJ'm}^{\nu,-\rho}(a), \tag{6}$$

we shall, in the following, concentrate our attention on the case

$$0 \leqslant a < \infty.$$

Next we replace the hypergeometric function in Eq. (2) by its (infinite) series expansion

$$\frac{(J+J'+1)!}{(J'-\frac{1}{2}i\rho)!(\nu+m+d+d')!} \sum_{n=0}^{\infty} \frac{(J'-\frac{1}{2}i\rho+n)!(\nu+m+d+d'+n)!}{n!(J+J'+1+n)!}(1-e^{-2a})^n,$$

which is valid whenever $0 \leqslant a < \infty$ (note that when $a \to \infty$, the argument $1-e^{-2a} \to 1$, and we require some condition on the parameters in the hypergeometric function for its expansion to be valid), and we use

$$\frac{(\nu+m+d+d'+n)!(J+J'-\nu-\lambda-d-d')!}{(J+J'+1+n)!(J-\nu-d)!(J'-m-d')!} = \sum_t (-1)^{J-\nu-d-t} \frac{(J+m+d'+n-t)!}{t!(J+J'+1+n-t)!(J-\nu-d-t)!}, \tag{7}$$

which express $D_{JJ'm}^{\nu\rho}(a)$ in a form in which the d summation can be performed utilizing

$$\sum_d \frac{1}{d!(J-m-d)!(J-\nu-t-d)!(\nu+m+d)!} = \frac{(2J-t)!}{(J-m)!(J+\nu)!(J-\nu-t)!(J+m-t)!}. \tag{8}$$

This results in (on replacing t by d again)

$$D_{JJ'm}^{\nu\rho}(a) = \alpha_J^{\nu\rho *}\alpha_{J'}^{\nu\rho} \frac{(-1)^{J'-\nu}}{(J'-\frac{1}{2}i\rho)!}\left(\frac{(2J+1)(J-\nu)!(J+m)!(2J'+1)(J'-\nu)!(J'+\nu)!(J'-m)!(J'+m)!}{(J+\nu)!(J-m)!}\right)^{1/2}$$

$$\times \sum_{dd'}\sum_{n=0}^{\infty}(-1)^{d+d'} \frac{(2J-d)!(J'-\frac{1}{2}i\rho+n)!(J+m-d-d'+n)!\exp[-a(2d'+\nu+m+1-\frac{1}{2}i\rho)](1-e^{-2a})^n}{d!(J-\nu-d)!(J+m-d)!d'!(J'-\nu-d')!(\nu+m+d')!(J+J'+1-d+n)!n!}. \tag{9}$$

Now

$$\sum_{n=0}^{\infty} \frac{(J'-\frac{1}{2}i\rho+n)!(J+m-d+d'+n)!}{n!(J+J'+1-d+n)!}(1-e^{-2a})^n$$

$$= \frac{(J'-\frac{1}{2}i\rho)!(J+m-d+d')!}{(J+J'+1-d)!}\ {}_2F_1(J'+1-\tfrac{1}{2}i\rho, J+m+1-d+d', J+J'+2-d; 1-e^{-2a})$$

$$= \frac{(J'-\frac{1}{2}i\rho)!(J+m-d+d')!}{(J+J'+1-d)!}\exp[-2a(-m+\tfrac{1}{2}i\rho-d')]$$

$$\times {}_2F_1(J+1+\tfrac{1}{2}i\rho-d, J'-m+1-d', J+J'+2-d; 1-e^{-2a}) \tag{10}$$

on making use of

$${}_2F_1(a,b,c;z) = (1-z)^{c-a-b}\ {}_2F_1(c-a,c-b,c;z).$$

Thus

$$\sum_{n=0}^{\infty} \frac{(J'-\frac{1}{2}i\rho+n)!(J+m-d+d'+n)!}{n!(J+J'+1-d+n)!}(1-e^{-2a})^n$$

$$= \frac{(J'-\frac{1}{2}i\rho)!(J+m-d+d')!}{(J+\frac{1}{2}i\rho-d)!(J'-m-d')!}\exp[-2a(-m+\tfrac{1}{2}i\rho-d')]$$

$$\times \sum_{n=0}^{\infty} \frac{(J+\frac{1}{2}i\rho-d+n)!(J'-m-d'+n)!}{n!(J+J'+1-d-n)!}(1-e^{-2a})^n. \tag{11}$$

Combining Eqs. (9) and (11) we obtain

$$D\mathscr{Y}^{\nu\rho}_{J'm}(a)=\alpha\mathscr{Y}^{\nu\rho*}\alpha\mathscr{Y}^{\nu\rho}(-1)^{J'-\nu}\left(\frac{(2J+1)(J-\nu)!(J+m)!(2J'+1)(J'-\nu)!(J'+\nu)!(J'-m)!(J'+m)!}{(J+\nu)!(J-m)!}\right)^{1/2}$$

$$\times\sum_{dd'}\sum_{n=0}^{\infty}(-1)^{d+d'}\frac{(2J-d)!}{d!(J-\nu-d)!(J+m-d)!(J+\frac{1}{2}i\rho-d)!}\exp[-a(\nu-m+1+\tfrac{1}{2}i\rho)]$$

$$\times\frac{(J+m-d+d')!(J+\frac{1}{2}i\rho-d+n)!(J'-m-d'+n)!}{d'!(J'-\nu-d')!(J'-m-d')!(\nu+m+d')!n!(J+J'+J'-d+n)!}(1-e^{-2a})^n. \tag{12}$$

Note that the above steps have resulted in making the boost-dependent factors in any term to be independent of the summation index d. Next we transform the d' summation using

$$\frac{(J+m-d+d')!(J'-m-d'+n)!}{n!(J'-m-d')!(J+J'+1-d+n)!}=\sum_t(-1)^{J'-m-d'-t}\frac{(J+J'-d-t)!}{t!(J+J'+1-d+n-t)!(J'-m-d'-t)!}. \tag{13}$$

Now the d' summation can be performed. Indeed

$$\sum_{d'}\frac{1}{d'!(J'-\nu-d')!(J'-m-t-d')!(\nu+m+d')!}=\frac{(2J'-t)!}{(J'-\nu)!(J'+m)!(J'+\nu-t)!(J'-m-t)!}. \tag{14}$$

From Eqs. (12)–(14) and replacing t by d', we arrive at

$$D\mathscr{Y}^{\nu\rho}_{J'm}(a)=\alpha\mathscr{Y}^{\nu\rho*}\alpha\mathscr{Y}^{\nu\rho}(-1)^{m-\nu}\left(\frac{(2J+1)(J-\nu)!(J+m)!(2J'+1)(J'+\nu)!(J'-m)!}{(J+\nu)!(J-m)!(J'-\nu)!(J'+m)!}\right)^{1/2}$$

$$\times\sum_{dd'}\sum_{n=0}^{\infty}\frac{(2J-d)!}{d!(J-\nu-d)!(J+m-d)!(J+\frac{1}{2}i\rho-d)!}\exp[-a(\nu-m+1+\tfrac{1}{2}i\rho)]$$

$$\times\frac{(2J'-d')!(J+J'-d-d')!(J+\frac{1}{2}i\rho-d+n)!}{d'!(J'+\nu-d')!(J'-m-d')!(J+J'+1-d-d'+n)!}(1-e^{-2a})^n. \tag{15}$$

Also

$$\sum_{n=0}^{\infty}\frac{(J+\frac{1}{2}i\rho-d+n)!}{(J+J'+1-d-d'+n)!}(1-e^{-2a})^n$$

$$=\sum_{n=J+J'+1-d-d'}^{\infty}\frac{(-J'-1+\frac{1}{2}i\rho+d'+n)!}{n!}(1-e^{-2a})^{-J-J'-1+d+d'+n}$$

$$=\sum_{n=0}^{\infty}\frac{(-J'-1+\frac{1}{2}i\rho+d'+n)!}{n!}(1-e^{-2a})^{-J-J'-1+d+d'+n}-\sum_{n=0}^{J+J'-d-d'}$$

$$\times\frac{(-J'-1+\frac{1}{2}i\rho+d'+n)!}{n!}(1-e^{-2a})^{-J-J'-1+d+d'+n}$$

$$=E_1-E_2. \tag{16}$$

In the expression for E_1 above, the n summation consists of an infinite binomial series which can be summed and results in

$$E_1=(-J'-1+\tfrac{1}{2}i\rho+d')!\exp[-2a(J'-\tfrac{1}{2}i\rho-d')](1-e^{-2a})^{-J-J'-1+d+d'}$$

$$=(-J'-1+\tfrac{1}{2}i\rho+d')!\exp[-2a(J'-\tfrac{1}{2}i\rho-d')]\sum_{r=0}^{\infty}\frac{(J+J'-d-d'+r)!}{r!(J+J'-d-d')!}e^{-2ar}. \tag{17}$$

Further

$$E_2=\sum_{r=0}^{\infty}\sum_n\frac{(-J'-1+\frac{1}{2}i\rho+d'+n)!(J+J'-d-d'-n+r)!}{r!n!(J+J'-d-d'+n)!}e^{-2ar}. \tag{18}$$

The n summation in the above equation *being finite* can be computed and results in

$$E_2 = \sum_{r=0}^{\infty} \frac{(-J'-1+\frac{1}{2}i\rho+d')!(J+\frac{1}{2}i\rho-d+r)!}{(J+J'-d-d')!(-J'+\frac{1}{2}i\rho+d'+r)!} e^{-2ar}. \tag{19}$$

Substituting for E_1, E_2 in Eq. (16) from Eqs. (17), (19), the Eq. (15) takes the form

$$D^{\nu\rho}_{JJ'm}(a) = \alpha^{\nu\rho*}_{J}\alpha^{\nu\rho}_{J'}(-1)^{m-\nu}\left(\frac{(2J+1)(J-\nu)!(J+m)!(2J'+1)(J'+\nu)!(J'-m)!}{(J+\nu)!(J-m)!(J'-\nu)!(J'+m)!}\right)^{1/2}$$

$$\times \sum_{dd'} (-1)^{d+d'} \frac{(2J-d)!(2J'-d')!(-J'-1+\frac{1}{2}i\rho+d')!}{d!(J-\nu-d)!(J+m-d)!(J+\frac{1}{2}i\rho-d)!d'!(J'+\nu-d')!(J'-m-d')!}$$

$$\times \exp[-a(\nu-m+1+\tfrac{1}{2}i\rho)]\left(\sum_{r=0}^{\infty}\frac{(J+J'-d-d'+r)!}{r!}\exp[-2a(J'-\tfrac{1}{2}i\rho-d'+r)]\right.$$

$$\left.-\sum_{r=0}^{\infty}\frac{(J+\frac{1}{2}i\rho-d+r)!}{(-J'+\frac{1}{2}i\rho+d'+r)!}e^{-2ar}\right). \tag{20}$$

Our aim is to express $D^{\nu\rho}_{JJ'm}(a)$ in terms of sums of products of two CG coefficients and an exponential in the boost. In the 2nd term above, the d and d' summations are split, the exponential e^{-2ar} does not involve the indices d and d' and we can trivially change the order of summations to do the r summation at the end. In the first term, in order to make the exponential free of d and d', let us use the summation index r in place of $J'+\nu-d'+r$. (Note that $J'+\nu$ is an integer.) Then the r summation is over the range $J'+\nu-d'$ to ∞. On changing orders, the r summation will be form $\min(J'+v-d') = \max(\nu+m,0)$ to ∞. Thus we arrive at

$$D^{\nu\rho}_{JJ'm}(a) = \alpha^{\nu\rho*}_{J}\alpha^{\nu\rho}_{J'}(-1)^{m-\nu}\left(\frac{(2J+1)(J-\nu)!(J+m)!(2J'+1)(J'+\nu)!(J'-m)!}{(J+\nu)!(J-m)!(J'-\nu)!(J'+m)!}\right)^{1/2}$$

$$\times \sum_{dd'r} (-1)^{d+d'} \frac{(2J-d)!(2J'-d')!(-J'-1+\frac{1}{2}i\rho+d')!}{d!(J-\nu-d)!(J+m-d)!(J+\frac{1}{2}i\rho-d)!d'!(J'+\nu-d')!(J'-m-d')!}\left(\frac{(J-\nu-d+r)!}{(-J'-\nu+d'+r)!}\right.$$

$$\left.\times\exp[-a(-\nu-m+1-\tfrac{1}{2}i\rho+2r)] - \frac{(J+\frac{1}{2}i\rho-d+r)!}{(-J'+\frac{1}{2}i\rho+d'+r)!}\exp[-a(\nu-m+1+\tfrac{1}{2}i\rho+2r)]\right). \tag{21}$$

where the r summation in the first term is from $\max(\nu+m,0)$ to ∞ whereas in the second term it is from 0 to ∞.

In the next section, we shall replace the *finite* d and d' summations in the above equation by analytically continued CG coefficients. However, before proceeding, we wish to remark that *none* of the two terms above can vanish. This can be seen on examining Eq. (16). [Indeed, the two terms in Eq. (21) above, correspond to the two terms in Eq. (16).] Both terms in Eq. (16) are singular for $a=0$, whereas the matrix element $D^{\nu\rho}_{JJ'm}(a)$ is just $\delta_{JJ'}$ for $a=0$. Furthermore in Eq. (21), when we consider $a=0$, we see that the resulting r summation in each of the two terms is divergent since $J+J'-d-d'\geqslant 0$.

We also note that the first term in Eq. (21) can be formally obtained from the second by replacing r by $-\nu-\frac{1}{2}i\rho+r$, i.e.,

$$D^{\nu\rho}_{JJ'm}(a) = \sum_{r=\max(\nu+m,0)}^{\infty} F(-\nu-\tfrac{1}{2}i\rho+r) - \sum_{r=0}^{\infty}F(r), \tag{22}$$

where

$$F(r) = \alpha^{\nu\rho*}_{J}\alpha^{\nu\rho}_{J'}\left(\frac{(2J+1)(J-\nu)!(J+m)!(2J'+1)(J'+\nu)!(J'-m)!}{(J+\nu)!(J-m)!(J'-\nu)!(J'+m)!}\right)^{1/2}$$

$$\times\sum_{dd'}(-1)^{d+d'}\frac{(2J-d)!(J+\frac{1}{2}i\rho-d+r)!}{d!(J-\nu-d)!(J+m-d)!(J+\frac{1}{2}i\rho-d)!}$$

$$\times\frac{(2J'-d)!(-J'-1+\frac{1}{2}i\rho+d')!}{d'!(J'+\nu-d')!(J'-m-d')!(-J'+\frac{1}{2}i\rho+d'+r)!}\exp[-a(\nu-m+1+\tfrac{1}{2}i\rho+2r)]. \tag{23}$$

III. THE BOOST MATRIX ELEMENTS IN TERMS OF THE ANALYTICALLY CONTINUED CLEBSCH-GORDAN COEFFICIENTS

We shall use the definition

$$\langle j_1,m_1;j_2,m_2|j,m\rangle = \left(\frac{(2j+1)(j_1+j_2-j)!(j_1-m_1)!(j_2-m_2)!(j+m)!(j-m)!}{(j_1+j_2+j+1)!(j_1-j_2+j)!(-j_1+j_2+j)!(j_1+m_1)!(j_2+m_2)!}\right)^{1/2}$$

$$\times \sum_s (-1)^s \frac{(2j_1-s)!(-j_1+j_2+j+s)!}{s!(j_1-m_1-s)!(-j_1+j-m_2+s)!(j_1+j_2-j-s)!} \tag{24}$$

of the analytically continued CG coefficient of the three-dimensional rotation group for the case where j_2, m_2, j, m may be complex. Note that in such a case, in Eq. (24), we indeed have a finite summation over s since $(j_1 - m_1)$ is a nonnegative integer. Comparing the d and d' summations in Eq. (23) with the above definition, and using

$$\Gamma(z)\Gamma(1-z) = (z-1)!(-z)! = \frac{\pi}{\sin\pi z} \tag{25}$$

repeatedly, we arrive at

$$D^{\nu\rho}_{JJ'm}(a) = (-1)^{(J+J'+2m)/2}(\nu - \tfrac{1}{2}i\rho)^{-1}[(2J+1)(2J'+1)]^{1/2}$$

$$\times\Bigg(\sum_{r=\max(\nu+m,0)}^{\infty} \langle J,-m;-j_1-1,m_1|-j_2-1,m_1-m\rangle\langle J',-m;j_1,m_1|j_2,m_1-m\rangle e^{-a(2m_1-m)}$$

$$-\sum_{r=0}^{\infty}\langle J,-m;-j_1-1,m_1'|-j_2-1,m_1'-m\rangle\langle J',-m;j_1,m_1'|j_2,m_1'-m\rangle e^{-a(2m_1'-m)}\Bigg), \tag{26}$$

where

$$j_1 = \tfrac{1}{2}\nu + \tfrac{1}{4}i\rho - \tfrac{1}{2},\quad j_2 = -\tfrac{1}{2}\nu + \tfrac{1}{4}i\rho - \tfrac{1}{2},\quad m_1 = -j_1 + r,\quad m_1' = j_1 + 1 + r. \tag{27}$$

We shall use the following two symmetry relations of CG coefficients (for the case where the phase is independent of ρ):

$$\langle j_1,m_1;j_2,m_2|j,m\rangle = (-1)^{j-j_1-m_2}\langle j_1,m_1;-j_2-1,m_2|j,-m\rangle \tag{28a}$$

$$= (-1)^{j_2+m_2}\left(\frac{2j+1}{2j_1+1}\right)^{1/2}\langle j_2,-m_2;j,m|j_1,m_1\rangle \tag{28b}$$

to recast Eq. (26) in the form[9]

$$D^{\nu\rho}_{JJ'm}(a) = (-1)^{(J+J'-1-2\nu)/2}\Bigg(\sum_{r=\max(\nu+m,0)}^{\infty}\langle J_1,-m_1;-j_2,m_1-m|J,-m\rangle$$

$$\times\langle J_1,-m_1;j_2,m_1-m|J',-m\rangle\exp[-a(2m_1-m)]$$

$$-\sum_{r=0}^{\infty}\langle j_1,-m_1';j_2,m_1'-m|J,-m\rangle\langle j_1,-m_1';j_2,m_1'-m|J',-m\rangle\exp[-a(2m_1'-m)]\Bigg). \tag{29}$$

IV. MATRIX ELEMENT OF THE O(4) ROTATIONS IN FINITE DIMENSIONAL IRREDUCIBLE (UNITARY) REPRESENTATIONS

In the last section, we expressed the matrix elements of the Lorentz group boost matrix elements for hyperbolic rotations in terms of the analytically continued CG coefficients. Now we examine how we can obtain the matrix elements of the corresponding O(4) rotation in any irreducible (finite-dimensional) spinor representation. To achieve it, we shall use the correspondence[10]

$$-\nu + \tfrac{1}{2}i\rho - 1 \leftrightarrow 2j_1, \tag{30a}$$

$$-\nu + \tfrac{1}{2}i\rho - 1 \leftrightarrow 2j_2, \tag{30b}$$

$$a \leftrightarrow i\delta, \tag{30c}$$

and use j_1, j_2 to characterize the irreducible representations under consideration in place of ν, ρ. Note that $2j_1$, $2j_2$ take the values 0,1,2,.... Using the symmetry relation[11]

$$\langle j_1,m_1;j_2,m_2|j,m\rangle = (-1)^{j_1+j_2+j}\langle j_1,-m_1;j_2,-m_2|j,-m\rangle \tag{28c}$$

and the correspondences in Eq. (30) above, the Eq. (29) takes the form

$$D^{j_1j_2}_{JJ'm}(\delta)$$

$$= (-1)^{(J+J'+1-2\nu)/2}\Bigg(\sum_{r=\max(j_1-j_2+m,0)}^{\infty}\langle j_1,m_1;j_2,-m_1+m|J,m\rangle\langle J_1,m_1;j_2,-m_1+m|J',m\rangle\exp[-i\delta(2m_1-m)]$$

$$-\sum_{r=0}^{\infty}\langle j_1,m_1';j_2,-m_1'+m|J,m\rangle\langle j_1,m_1';j_2,-m_1'+m|J',m\rangle\exp[-i\delta(2m_1'-m)]\Bigg). \tag{31}$$

In the first term we require

$$-j_1\leqslant m_1\leqslant j_1,\quad -j_2\leqslant -m_1+m\leqslant j_2,$$

which on using Eq. (27) become

$$0\leqslant r\leqslant 2j_1,\quad j_1-j_2+m\leqslant r\leqslant j_1+j_2+m$$

and shows that this term in Eq. (31) is essentially what appears as the O(4) "boost" matrix element in Freedman and Wong. Again from Eq. (27), the second term in Eq. (27) required

$$-j_1\leqslant m_1'\leqslant j_1+1+r\leqslant j_1$$

which is impossible for any r in the range 0 to ∞. Thus the second term does not contribute to the O(4) boost matrix elements. Removing the second term, Eq. (31) takes the form

$$D^{j_1j_2}_{JJ'm}(a)=(-1)^{-(J+J'+1-2j_1+2j_2)/2}\sum_{m_1}\langle j_1,m_1;j_2,-m_1+m|J,m\rangle\langle j_1,m_1;j_2,-m_1+m|J',m\rangle\exp[-i\delta(2m_1-m)] \tag{32}$$

where, for convenience, we have replaced the r summation by a summation over the magnetic quantum number m_1 which takes either integral or half-integral values.

V. DISCUSSION

The major result of our paper is contained in Eq. (29) which expresses the boost matrix elements of the homogeneous Lorentz group in terms of analytically continued CG coefficients of the rotation group. We have also shown that the corresponding result for the O(4) group could be trivially derived by replacing the variables in this equation by their corresponding values. Starting from Smorodinskii and Shepelev's integral representation for the boost matrix elements,[12] Wong and Yeh[13] have tried to arrive at an expression which is essentially the same as ours. However, they concluded, using a transformation of ${}_3F_2$, that one of the terms in this expression vanishes except for the case $J=J'=\nu=m=0$. We have started from an explicit representation of the boost matrix elements as the sum of hypergeometric functions ${}_2F_1$ in the argument $1-e^{-2a}$ and examined how the split into the two terms occurs. Following our Eq. (21) we have already remarked that the two terms into which the boost matrix elements split are both singular as $a\to 0$ though the matrix elements are just $\delta_{JJ'}$. Thus none of the two terms can vanish. Indeed we have checked for the simple case of $D^{0\rho}_{100}(a)$ for which we can easily compute the two terms and verified our conclusions that the two terms are indeed nonvanishing and singular for $a\to 0$. We believe that Wong and Yeh have not drawn a correct conclusion from the relationship

$$\Gamma(a)\ {}_3F_2(a,b,c;e,f)=\frac{\Gamma(s)\Gamma(e)\Gamma(f)}{\Gamma(s+b)\Gamma(s+c)}\ {}_3F_2(s,e-a,f-a;s+b,s+c),$$

where

$$s=e+f-a-b-c.$$

Using their notation, on the left hand, we have a terminating ${}_3F_2$ and $a=\sigma-\nu+1-t$; hence the left-hand side contains poles in t at the positions where a is a nonpositive integer. On the right-hand side, all of $s,e,f,s+b,s+c,e-a,f-a$ are nonsingular for these values of t which led Wong and Yeh to conclude that the above-mentioned poles have disappeared. However, when we examine the quantity ${}_3F_2(s,e-a,f-a;s+b,s+c)$ on the right-hand side, we note that the difference of the parameters in the numerator and the denominator is

$$-a=\nu-\sigma-1+t,$$

which is nonnegative at the position of these poles, i.e., the ${}_3F_2$ on the right-hand side is ∞ which is consistent with the presence of the poles on the left-hand side. Thus the set of poles do not disappear as is also confirmed by our results.

[1]S. Ström, Ark. Fys. **29**, 467 (1965).

[2]D.W. Duc and N.V. Hieu, Dokl. Akad. Nauk SSSR **172**, 1281 (1967) [Sov. Phys. Dokl. **12**, 312 (1967)].

[3]A. Sciarrino and M. Toller, J. Math. Phys. **8**, 1252 (1967).

[4]R.L. Anderson, R. Rącska, M.A. Rashid, and P. Winternitz, J. Math. Phys. **11**, 1050 (1970).

[5]We shall utilize the expression given in Ref. 4 and its notation.

[6]D.Z. Freedman and J.M. Wong, Phys. Rev. **160**, 1560 (1967).

[7]D.A. Akyeampong, J.F. Boyce, and M.A. Rashid, J. Math. Phys. **11**, 706 (1970).

[8]See for example Ref. (3) above.

[9]Note that this has been achieved since $\frac{1}{2}i\rho-\nu$ in Eq. (26) is just $2j_2+1$.

[10]R.L. Anderson, R. Rączka, M.A. Rashid, and P. Winteritz, J. Math. Phys. **11**, 1059 (1970). Note that in this reference the symbols k and n are used in place of $2j_1$ and $2j_2$.

[11]We could not make use of this symmetry relation in Eq. (29) for the Lorentz group boost matrix elements since $j_1+j_2-j=\frac{1}{2}i\rho-1-j$ is not an integer.

[12]Ya.A. Smorodinskii and G.I. Shepelev, Yad. Fiz. **13**, 441 (1971) [Sov. J. Nucl. Phys. **13**, 248 (1971)].

[13]M.K.F. Wong and H.Y. Yeh, J. Math. Phys. **18**, 1768 (1978).

New nonlinear realizations of SL(2,*C*)

B. J. Dalton

Ames Laboratory, U. S. Department of Energy and Department of Physics, Iowa State University, Ames, Iowa 50011

(Received 23 August 1978; revised manuscript received 18 January 1979)

In this paper we present a study of some of the possible realizations of SL(2, C) as a transformation group on a complex three-dimensional space **S**. Several inequivalent categories of nonlinear realizations are found. We demonstrate that complex vector spaces can be found which transform under certain of these nonlinear realizations and for which the Poincaré mass and spin operators are diagonal. Within these new categories we obtain one spin-0, two spin-1/2, and several spin-1 realizations. In addition we obtain one new realization for each spin greater than 1. Certain of these realizations are unusual in that for both integer and half-odd integer spins the rotational period is 4π. For each new realization we derive the invariant metric on the space **S** and obtain a covariant wave equation. The development in this study suggests the interesting possibility that different categories of elementary particles could correspond to different categories of SL(2, C) realizations.

I. INTRODUCTION

Physical particles with spin are traditionally associated with *linear* realizations (i.e., representations) of the Poincaré group. This group is the semidirect product of the translation group in space–time with the diagonal subgroup of the direct product of two copies of SL(2,C),[1] one generated by the space–time orbital generators $l_{\mu\gamma}$ and the other by the intrinsic spin operators $S_{\mu\gamma}$. For linear realizations (generated by $S_{\mu\gamma}$) one can form the direct product because the partial derivatives (∂_μ) in $l_{\mu\gamma}$ commute with the $S_{\mu\gamma}$. This feature is not necessarily true for nonlinear realizations.

The primary goal of this paper is to exhibit several categories of nonlinear realizations of SL(2,C) and to demonstrate that for some of them the realizations have the above mentioned desired property that the intrinsic generators commute with the space–time partial derivatives.

These nonlinear realizations are important because with them one can diagonalize the Poincaré Casimir invariants and thereby describe particles with definite mass and intrinsic spin. Since these nonlinear realizations differ substantially from the linear ones, they offer some new mathematical categories for the possible classification of elementary particles.

In the past work, Hinds[2] has studied the O(3,1)/O(3) coset realizations of the Lorentz group. This coset space can be parametrized by three variables which transform nonlinearly under the group. In a later paper Hopkinson and Reya[3] made a more extensive study of these coset realizations, constructed the invariant metric on the coset space and derived a covariant wave equation.

In a more recent article[4] on nonlinear relizations involving the direct product of two Lorentz groups, the present author has obtained, although indirectly, another category of nonlinear infinitesimal realizations of SL(2,C). These realizations, which were defined as transformations on a complex three-dimensional space **S**, are not equivalent to the above mentioned coset realizations.

In the present paper we use direct algebraic techniques to explore some possible infinitesimal realizations of SL(2,C) on a three-dimensional complex space **S**. We discuss these realizations in terms of the various inequivalent categories found. The linear category corresponds to the Lorentz transformations on the electromagnetic field via $\mathbf{S} = \mathbf{B} + i\mathbf{E}$. In another category we find, as expected, the above mentioned coset realizations. For another category, the realizations have the same form as the special de Sitter group realizations studied by Melvin,[5] Takabayasi,[6] and later by Philips and Wigner.[7] In addition to the categories just mentioned, we find two categories of the type obtained in the author's recent work.[4] In these two categories, as well as for the special de Sitter type just mentioned, it is possible to find realizations for which $\partial_\mu S_i$ transforms like the product $p_\mu S_i$.

For these latter realizations we can diagonalize the Poincaré mass and spin invariants. We obtain one new realization with Poincaré spin 0, two new realizations with Poincaré spin $\frac{1}{2}$ and several new realizations with Poincaré spin 1. In addition, we obtain one new realization for all Poincaré spin values greater than one.

In Sec. II we establish the notation and define the group generators. These are used in Sec. III where we give exact infinitismal realizations of an SU(2) subgroup. In Sec. IV we use these SU(2) realizations to exhibit several inequivalent categories of SL(2,C) realizations. In Sec. V we discuss the Poincaré group invariants for these realizations. Using the invariant metric on **S** we also drive covariant wave equations. For certain of these realizations we show that it is possible to have covariant nonlinear wave equations which

 0022-2488/79/071520-10$01.00

can have soliton solutions.[8-11] We have developed this paper such that for each of the SL(2,C) realizations considered, we can easily obtain a corresponding realization of O(4) by simple parameter complexification, as explained in Sec. II.

II. NOTATION AND GROUP ALGEBRA

In this section we establish the notation used in this paper and describe a convenient basis for the Lie algebra of SL(2,C) or O(4) together with expressions for the Casimir invariants. For convenience in manipulation, the Pauli metric convention of letting the fourth space–time component be imaginary is used (that is, $x_4 = ict$), together with the convention of summing from 1–4 (from 1–3) over repeated Greek (Latin) indices.

We start with a basis $\{M_{\mu\nu}\}$ for an arbitrary realization of the Lie algebra of SL(2,C) such that with the above convention, the generators $M_{\mu\nu}$ satisfy the following commutation relations

$$[M_{\mu\nu},M_{\rho\alpha}] = i(\delta_{\mu\rho}M_{\nu\alpha} - \delta_{\mu\alpha}M_{\nu\rho} + \delta_{\nu\alpha}M_{\mu\rho} - \delta_{\nu\rho}M_{\mu\alpha}). \tag{2.1}$$

Corresponding to the $\{M_{\mu\nu}\}$ we define a vector basis $\{J_i,K_i\}$ as follows[12]:

$$J_i = i\epsilon_{ijk}M_{jk}/2, \quad K_i = M_{4i}. \tag{2.2}$$

Here ϵ_{ijk} is the usual total antisymmetric tensor with normalization $\epsilon_{123} = +1$. For These generators the commutation relations corresponding to (2.1) are given as follows:

$$[J_i,J_j] = -\epsilon_{ijk}J_k, \tag{2.3}$$

$$[K_i,K_j] = +\epsilon_{ijk}J_k, \tag{2.4}$$

$$[J_i,K_j] = -\epsilon_{ijk}K_k. \tag{2.5}$$

To describe the various types of nonlinear realizations we find it more convenient to use another basis with operator components T_i and Z_i defined as follows

$$T_i = (J_i + iK_i)/2, \quad Z_i = (J_i - iK_i)/2. \tag{2.6}$$

For this basis we have the following commutation relations:

$$\begin{aligned}[T_i,T_j] &= -\epsilon_{ijk}T_k,\\ [Z_i,Z_j] &= -\epsilon_{ijk}Z_k,\end{aligned} \tag{2.7}$$

$$[T_i,Z_j] = 0. \tag{2.8}$$

This basis corresponds to a direct product decomposition which we indicate by $T \times Z$.

We will indicate a general element (g) for a realization of the group in the ususal exponential form.

$$\begin{aligned} g &= \exp(\boldsymbol{\omega}\cdot\mathbf{J} + \boldsymbol{\nu}\cdot\mathbf{K}) = \exp(\boldsymbol{\alpha}\cdot\mathbf{T})\exp(\boldsymbol{\beta}\cdot\mathbf{Z}),\\ \boldsymbol{\beta} &= \boldsymbol{\omega} + i\boldsymbol{\nu}, \quad \boldsymbol{\alpha} = \boldsymbol{\omega} - i\boldsymbol{\nu}. \end{aligned} \tag{2.9}$$

The equality of the two forms for g in (2.9) follows from using (2.8) in the Baker–Cambell–Hausdorff formula.[13] With the choice of basis used here, both $\boldsymbol{\omega}$ and $\boldsymbol{\nu}$ are real. If we make the substitution $\boldsymbol{\nu} \to i\boldsymbol{\nu}$ the realization indicated in (2.9) will be a realization of the group O(4), rather than SL(2,C).[14] Although the emphasis in this paper is on the SL(2,C) group it should be clear that for each SL(2,C) direct product realization we will also have a realization of O(4) via the above mentioned complexification. We emphasize, however, that the nature of the space will generally differ in the two cases (see Sec. IV for illustrations).

A particular realization of this group on a component S_l of a space S is indicated by g: $S_l \to S'_l(\boldsymbol{\omega},\boldsymbol{\nu},S)$ where $\boldsymbol{\omega}$ and $\boldsymbol{\nu}$ are the real group parameters discussed above. With this we define the commutation relation, "action," of the generators J_i and K_i with a component S_l by the expressions

$$[J_i,S_l] = \left.\frac{\partial S'_l}{\partial \omega_i}\right|_{\omega_i=0}, \quad [K_i,S_l] = \left.\frac{\partial S'_l}{\partial \nu_i}\right|_{\nu_i=0} \tag{2.10}$$

The Jacobi identities imposed by the group properties are discussed and used in the following sections to help determine the possible forms for the action $[T_i,S_l] = ([J_i,S_l] + i[K_i,S_l])/2$.

To help characterize the particular realizations studied in the following sections we need expressions for the Casimir invariants of SL(2,C). These are defined here as follows

$$\begin{aligned} F_1 &= M_{\mu\nu}M_{\mu\nu}/4 = (\mathbf{K}\cdot\mathbf{K} - \mathbf{J}\cdot\mathbf{J})/2 = -[\mathbf{Z}\cdot\mathbf{Z} + \mathbf{T}\cdot\mathbf{T}],\\ F_2 &= M_{\mu\nu}\widetilde{M}_{\mu\nu}/4 = i\mathbf{J}\cdot\mathbf{K} = -[\mathbf{Z}\cdot\mathbf{Z} - \mathbf{T}\cdot\mathbf{T}], \end{aligned} \tag{2.11}$$

where $\widetilde{M}_{\mu\nu} \equiv \epsilon_{\mu\nu\alpha\beta}M_{\alpha\beta}/2$ is the usual dual of $M_{\mu\nu}$. We also have the alternate set of Casimir invariants,

$$C_1 = (F_1 + F_2)/2 = -Z\cdot Z, \quad C_2 = (F_1 - F_2)/2 = -T\cdot T. \tag{2.12}$$

For a given component S_l of a space S the eigenvalue equations for these Casimir invariants are given by

$$C_1S_l \equiv -[Z_i,[Z_i,S_l]] = j_1(j_1+1)S_l, \tag{2.13}$$

$$C_2S_l \equiv -[T_i,[T_i,S_l]] = j_2(j_2+1)S_l. \tag{2.14}$$

For linear realizations (representations) j_1 and j_2 may be complex but for finite-dimensional nonunitary representations j_1 and j_2 together are integer or half-odd integer. In the following sections we will consider some particular finite dimensional *nonlinear* realizations for which the above Casimir eigenvalue equations are satisfied with j_1 and j_2 both taking on either integer or half-odd integer values. We will also obtain some *nonlinear* realizations for which j_1 and j_2 are not restricted to discrete values.

III. REALIZATIONS OF THE LIE ALGEBRA OF T

In this section we discuss the possible realizations of the Lie algebra for the subgroup T, generated by the operators $T_i \equiv (J_i + iK_i)/2$, on a complex three-dimensional space **S**. We include both linear and nonlinear transformations and explore the different types of realizations on **S** consistent with the commutation relations (2.7) for the Lie algebra of T. The extension of these realizations to the full lie algebra of SL(2,C) is made in Sec. IV, to which we relegate most of our comparisons with previous work.

We are interested here in the commutation relations, "action," of the generators $\{T_i\}$ with the components of **S**. For this action we first write down the following form

$$[T_i, S_j] = F\delta_{ij} + hS_iS_j + G\epsilon_{ijk}S_k. \tag{3.1}$$

The functions F, h and G, are to be determined. In this expression and throughout this paper the subscript on the component S_i is the same Cartesian index appearing on the generator T_i or Z_i. With this restriction, (3.1) is the most general form one can write consistent with the indices, assuming of course that F, h and G are arbitrary functions of the components S_i. In this paper we study only those realizations for which F, h and G depend upon S only through the function $S^2 = S_iS_i$. With the possible nonlinearity, as expressed by (3.1), the inner product S^2 will not in general be an SL(2,C) invariant. Using (3.1) we obtain the relation

$$[T_i,S^2] = 2[F + hS^2]S_i, \tag{3.2}$$

from which it follows that S^2 will be invariant only if the quantity $F + hS^2$ is identically zero. The functions F, h and G in (3.1) cannot be arbitrary since (3.1) must be consistent with the commutation relations of the Lie algebra. To find the allowed values of F, h and G, or restrictions on them, we first impose the following Jacobi identity.

$$[T_i,[T_j,S_k]] - [T_j,[T_i,S_k]] = [[T_i,T_j],S_k]. \tag{3.3}$$

If we substitute (3.1) into (3.3) and carry through the algebra we obtain the following equation,

$$\begin{aligned}&[2F'(F + hS^2) - hF - G^2 - G](S_i\delta_{jk} - S_j\delta_{ik})\\&\quad + [(2G+1)h - 2G'(F + hS^2)]\epsilon_{ijl}S_lS_k\\&\quad + [(2G+1)F + 2G'(F + hS^2)S^2]\epsilon_{ijk} = 0\end{aligned} \tag{3.4}$$

Here, the prime denotes differentiation with respect to S^2. To rearrange terms in the derivation of (3.4) we used the following identity:

$$S^2\epsilon_{ijk} = \epsilon_{ijl}S_lS_k + \epsilon_{jkl}S_lS_i + \epsilon_{kil}S_lS_j. \tag{3.5}$$

For (3.4) to hold, the coefficients of each of the three indexed expressions must separately be zero, yielding the following equations:

$$2F'(F + hS^2) - hF - G^2 - G = 0, \tag{3.6}$$

$$(2G+1)h - 2G'(F + hS^2) = 0, \tag{3.7}$$

$$(2G+1)F + 2G'(F + hS^2)S^2 = 0. \tag{3.8}$$

From these equations it is easy to show that G can have only a constant value of either 0, $-\frac{1}{2}$, or -1. If G has either the value of 0 or -1 both F and h must be zero.

As we show below there is an infinite number of possible solutions of (3.6)–(3.8). In earlier work on nonlinear realizations,[4,15] those solutions which could be obtained one from the other by a one-to-one coordinate redefinition of the form

$$S_i \to \widetilde{S}_i = \phi(S^2)S_i \tag{3.9}$$

were called equivalent. Letting $\widetilde{F}$, $\widetilde{h}$ and $\widetilde{G}$ be solutions of (3.6)–(3.8) in the ($\widetilde{\ }$) coordinate system, equivalence via (3.9) means that ϕ must satisfy the following conditions

$$\widetilde{F} = \phi F, \quad \widetilde{h} = \phi^{-2}[2\phi'(F + hS^2) + \phi h], \quad \widetilde{G} = G. \tag{3.10}$$

With (3.9) we have $\widetilde{S}^2 = \phi^2(S)S^2$ so that if S^2 is invariant then $\widetilde{S}^2$ will likewise be invariant. We may therefore catagorize the different solutions into two groups, one for which S^2 is invariant and another for which S^2 is not invariant.

We emphasize here that although a function $\phi(S^2)$ may be found which satisfies (3.10) for two different solutions of (3.6)–(3.8), the corresponding spaces $\widetilde{\mathbf{S}}$ and $\mathbf{S}$ may not be equivalent via (3.9) with respect to other properties of physical interest. In particular, since S^2 is, in general, not an invariant, the tranformation (3.9) could connect realizations with different values of the Casimir invariants. Since the Casimir invariants are of physical interest we shall in this paper call equivalent only those realizations which can be obtained one from the other via (3.9) and for which the values of the Casimir invariants are the same. For the subgroup T we have the Casimir invariant C_2 defined in (2.14). In terms of the functions F, h, and G, C_2 acting on S_j takes on the following form,

$$\begin{aligned}C_2S_j &= [2(F' + h'S^2)(F + hS^2) + 4hF\\&\quad + 2h^2S^2 - 2G^2]S_j.\end{aligned} \tag{3.11}$$

With the above definition of equivalence the space $\mathbf{S}$ and $\widetilde{\mathbf{S}}$ are equivalent provided a ϕ can be found which satisfies (3.10) and the condition $\widetilde{C}_2 = C_2$ where $\widetilde{C}_2$ has the form of (3.11) in the $\sim$ coordinate system. For the present work we are interested in those realizations for which C_2 has a definite eigenvalue $[+j_2(j_2+1)]$.

Together with each realization discussed below we give the corresponding invariant quadratic metric form from which covariant field equations can be obtained. To construct an invariant metric, we first assume the following form

$$-d\eta^2 = g_{ij}dS_idS_j \tag{3.12}$$

and look at the conditions on g_{ij} imposed by invariance of $d\eta^2$ under the general transformation (3.1). This metric form is invariant if the metric tensors satisfy the following equations:

$$\left\{[T_i,g_{jk}] + 2g_{jl}\frac{\partial}{\partial S_k}[T_i,S_l]\right\}dS_kdS_j = 0. \tag{3.13}$$

Since g_{ij} is symmetric we assume the following form for it.

$$g_{ij} = A(S^2)\delta_{ij} + B(S^2)S_iS_j, \tag{3.14}$$

where the functions A and B are to be determined. Use of (3.14) in (3.13) with a bit of algebra leads directly to the following condition,

$$\begin{aligned}&[A(2F' + h) + B(F + hS^2)]dS_i(S_kdS_k)\\&\quad + [(A + BS^2)2h' + 2B(F' + h) + B'(F + hS^2)]S_i\\&\quad \times(S_jdS_j)^2 + [A'(F + hS^2) + Ah]S_idS_jdS_j = 0.\end{aligned} \tag{3.15}$$

Here as before the prime indicates differentiation with respect to S^2. Equation (3.15) holds true provided the coefficient of each indexed expression vanishes. This condition of invariance imposes the following equations:

$$A(2F' + h) + B(F + hS^2) = 0, \tag{3.16}$$

$$A'(F + hS^2) + Ah = 0. \tag{3.17}$$

These two equations together with (3.6)–(3.8) can be used to show that the coefficient of the middle term in (3.15) automatically vanishes. We also note here that both (3.16) and (3.17) are invariant under the simultaneous transformation $F \to -F$ and $h \to -h$. Because of this fact, the above metric form will also be invariant under the extension of the group to the full SL(2,C) or O(4), which we discuss in the following section.

Equations (3.16) and (3.17) will differ for each type of realization. For this reason we give the specific solutions and metric form as we discuss each separate realization below.

We now consider some solutions in detail, discussing first those solutions for which S^2 is invariant. From (3.2) we see that S^2 is invariant provided the expression $(F + hS^2)$ is identically zero. There are three solutions of (3.6)–(3.8) consistent with this condition. The first is the trivial null solution with $F = h = G = 0$. A second solutions is $F = h = 0$ and $G = -1$. This solution is the only linear realization found. For $F + hS^2 = 0$ we also have the solution $G = -1/2, F = \pm iS/2$ and $h = \mp i/2S$. Here, S is one of the roots of $(S^2)^{1/2}$. For later reference we label and discuss below these three cases. We give for each the Casimir invariant C_2.

$$TA\text{: Null case,} \quad [T_i, S_j] = 0, \quad C_2 = 0. \tag{3.18}$$

The metric is arbitrary in this case.

$$TB\text{:} \quad [T_i, S_j] = -\epsilon_{ijk} S_k, \quad C_2 = 2. \tag{3.19}$$

Again for this case there are no restrictions on the metric.

$$TC\text{:} \quad [T_i, S_j] = [\pm iS\delta_{ij} \mp iS_i S_j/S - \epsilon_{ijk} S_k]/2. \tag{3.20}$$

For this case we have $C_2 = 0$ and the invariant metric is given as follows.

$$-d\eta^2 = B(S^2)(S_i dS_i)^2. \tag{3.21}$$

Here B can be any function of the invariant S^2. For this case Eq. (3.17) requires $A = 0$, which is as should be, since $dS_i dS_i$ is not invariant under the action expressed in (3.20). With $C_2 = j_2(j_2 + 1)$ the quantum number j_2 for the realizations TA, TB, and TC takes the values 0, 1, and 0 respectively.

We next consider those realizations for which S^2 is not invariant. From (3.7) and (3.8) we have $G = -\frac{1}{2}$ as the only possible value if $F + hS^2$ is not zero. Using $G = -\frac{1}{2}$ in (3.6) we obtain the expression

$$(2F'F + 1/4) + h(2F'S^2 - F) = 0. \tag{3.22}$$

There are two separate categories of solutions of this equation corresponding to whether or not the expression in the second parenthesis is zero. We first consider the category for which this quantity is zero, that is, $2F'S^2 - F = 0$. With this condition we find the solutions $F = \pm iS/2$ and h can be any arbitrary function of S^2. With this form for F, the Casimir invariant C_2 in (3.11) will have a definite eigenvalue only if the function h has the form $\mp ib/2S$ where b is a constant which we shall relate below to the value of the Casimir invariant $j_2(j_2 + 1)$. With h restricted in this way we have the following solutions for this category:

$$F = \pm iS/2, \quad h = \mp ib/2S. \tag{3.23}$$

The special solutions for $b = 1$ have already been included in the TC realization discussed above.

If we substitute F and h given in (3.23) into (3.11) for C_2 we obtain the following relation

$$b = 2 \pm [1 + 4j_2(j_2 + 1)]^{1/2} = 2 \pm (2j_2 + 1). \tag{3.24}$$

The signs in this expression, which arise in solving a quadratic equation for b, are uncorrelated with the signs in (3.23). From (3.24) it is clear that for any value of j_2 we have a nonlinear infinitesimal realization of the subgroups T with the following explicit action which we label by TD,

$$TD\text{:} \quad [T_i, S_j] = \{\pm iS\delta_{ij} \mp ibS_i S_j/S - \epsilon_{ijk} S_k\}/2. \tag{3.25}$$

We may rewrite this expression in the following form,

$$[T_i, S_j] = \{\pm iS\delta_{ij} \mp iS_i S_j/S - \epsilon_{ijk} S_k\}/2 + \{\mp 1 \pm [1 + 4j_2(j_2 + 1)]^{1/2}\} iS_i S_j/2S. \tag{3.26}$$

The first term in brackets is identical to the action for the above mentioned TC class under which S^2 is invariant. It is the last term in (3.26), the magnitude of which is determined by the Casimir invariant C_2, that leads to the noninvariance of S^2. The reader should notice here that the number j_2 above is not restricted to discrete values.

For this realization we obtain from (3.16) and (3.17) the following solutions for A and B,

$$A = \lambda S^{2\alpha}, \quad B = -\lambda S^{2\alpha - 2} = -AS^{-2}. \tag{3.27}$$

Here, λ is an arbitrary constant and α is given by

$$\alpha = b(1 - b)^{-1}, \tag{3.28}$$

where b is given in (3.24). The invariant metric for this realization has the following form,

$$-d\eta^2 = \lambda S^{2\alpha}[dS_i dS_i - S^{-2}(S_i dS_i)^2]. \tag{3.29}$$

This metric is not valid for $b = 1$ which corresponds to the TC case previously discussed.

We next consider the second category of solutions of (3.22) mentioned above. For this category we have $2F'S^2 - F \neq 0$ so that the function h is given by the expression

$$h = (2F'F + 1/4)/(2F'S^2 - F). \tag{3.30}$$

For any function F subject to the condition $2F'S^2 - F \neq 0$ the function h can be computed from (3.30). We consider here the following solution of (3.30)

$$F = \pm(D^2 - S^2)^{1/2}/2, \quad D \neq 0, \quad h = 0. \tag{3.31}$$

This solution is such that $2F'F + 1/4 = 0$ but $2F'S^2 - F \neq 0$. Here, D is a group invariant, that is $[T_i, D] = 0$. For this realization which we label by TE we have the following action,

$$TE\text{:} \quad [T_i, S_j] = \{\pm(D^2 - S^2)^{1/2}\delta_{ij} - \epsilon_{ijk} S_k\}/2. \tag{3.32}$$

Using (3.31) in (3.11) we obtain $C_2 = 3/4$, corresponding to $j_2 = \frac{1}{2}$. The presence of the nonzero arbitrary invariant D, which distinguishes this realization from the previous ones, represents, in essence, two additional degrees of freedom (since D may be complex).

For the class of solutions given by (3.30) the functions A and B which solve (3.16) and (3.17) are given by

$$A = (F^2 + S^2/4)^{-1},$$
$$B = -(F^2 + S^2/4)^{-1}(2F' + h)(F + hS^2)^{-1}. \tag{3.33}$$

With these solutions we have the following invariant metric form,

$$-d\eta^2 = (F^2 + S^2/4)^{-1}[\delta_{ij} + (2F' + h)(F + hS^2)^{-1}S_i S_j] \, dS_i \, dS_j. \tag{3.34}$$

This particular form has been obtained in previous work[15] in which the nonlinear action corresponded to the present TE case, or an equivalent form thereof. The solutions for A and B in (3.33) are not valid for the cases TC and TD. In the TD realization we have $F^2 + S^2/4 = 0$ and in the TC realization we have $F + hS^2 = 0$. The solution (3.33) is singular for these cases.

The realization (3.31) can be linearized by embedding in a space with one more dimension. This can easily be seen from (3.31) if we set $S_4 = (D^2 - S^2)^{1/2}$. Using this and the action on S_4 which follows from (3.2) we have the following equations:

$$[T_i, S_j] = \{ \pm S_4\delta_{ij} - \epsilon_{ijk}S_k \}/2, \tag{3.35}$$

$$[T_i, S_4] = \mp S_i/2. \tag{3.36}$$

These equations represent linear transformations in the space $(\mathbf{S}, S_4)$ which have $D^2 = S^2 + S_4^2$ invariant.

IV. EXTENSIONS TO $T\times Z$

Here we extend the infinitesimal realizations developed in the previous section to the Lie algebra of SL(2,C) or O(4) decomposed as the direct product $T\times Z$. For the action of the generators Z_i on $\mathbf{S}$ we write down the general form:

$$[Z_i, S_j] = \bar{F}\delta_{ij} + \bar{h}S_i S_j + \bar{G}\epsilon_{ijk}S_k. \tag{4.1}$$

Since the generators $\{Z_i\}$ satisfy the same commutation relations as the generators $\{T_i\}$ the Jacobi identity corresponding to (2.7) leads to conditions identical to (3.6)–(3.8) for $\bar{F}$, $\bar{h}$, and $\bar{G}$. For reference we label the solutions for $\bar{F}$, $\bar{h}$, and $\bar{G}$ corresponding to the classes TA, TB–TE discussed in the previous section by ZA, ZB,-ZC, respectively.

In order for the realizations of T and Z on $\mathbf{S}$ to be realizations of the algebra of $T\times Z$, the following Jacobi identity

$$[T_i,[Z_j,S_k]] - [Z_j,[T_i,S_k]] = 0 \tag{4.2}$$

corresponding to (2.8) must also be satisfied. The use of (3.1) and (4.1) in (4.2) leads, after a little algebra, to the following equation,

$$\begin{aligned}
&[2\bar{F}'(F + hS^2) - h\bar{F} - \bar{G}G]\delta_{jk}S_i + (\bar{h}F - h\bar{F})\delta_{ij}S_k \\
&\quad - [2F'(\bar{F} + \bar{h}S^2) - \bar{h}F - \bar{G}G]\delta_{ki}S_j \\
&\quad + (\bar{h}G + h\bar{G})\epsilon_{ijl}S_l S_k + [2\bar{h}'(F + hS^2) \\
&\quad - 2h'(\bar{F} + \bar{h}S^2)]S_j S_k S_i + (\bar{G}F + G\bar{F})\epsilon_{ijk} = 0.
\end{aligned} \tag{4.3}$$

In the derivation of this expression we used the fact that G and $\bar{G}$ must both be constant.

For (4.3) to hold the coefficients of each of the six indexed expressions must be identically zero. With (3.6) to (3.8), the resulting six equations can be satisfied in two ways. The first is with any realization TA–TE with the null realization ZA, or any realization ZA–ZE with the null realization TA. The second way is with $\bar{G} = G = -\frac{1}{2}$ together with $\bar{F} = -F$ and $\bar{h} = -h$. As we mentioned earlier, Eqs. (3.16) and (3.17) are invariant under the simultaneous transformation $F\to -F$ and $h\to -h$. Because of this property the invariant metrics discussed in the previous sections will also be invariant under the action of Z.

We now consider the various realizations of the Lie algebra of $T\times Z$ on $\mathbf{S}$. For each we give the pair of quantum numbers (j_1, j_2). We first discuss the linear case and the TE case which can be linearized in four dimensions.

$TB\times ZA$: *Linear realization*:

$$[T_i, S_j] = -\epsilon_{ijk}S_k, \quad [Z_i, S_j] = 0, \tag{4.4}$$

$$(j_1, j_2) = (0,1). \tag{4.5}$$

Suppose the complex conjugate of $\mathbf{S}$ transforms like $TA\times ZB$, that is,

$$[T_i, S_j^*] = 0, \quad [Z_i, S_j^*] = -\epsilon_{ijk}S_k^*, \tag{4.6}$$

$$(j_1, j_2) = (1,0). \tag{4.7}$$

The combined space $(\mathbf{S},\mathbf{S}^*)$ transforms as a $(0,1) + (1,0)$ representation. If we write $\mathbf{S} = \mathbf{B} + i\mathbf{E}$ and identify $\mathbf{B}$ and $\mathbf{E}$ with the magnetic and electric fields, then (4.4) and (4.6) together generate the usual Lorentz transformation on these fields.

$TE\times ZE$: 4-*vector realization*:

From (3.35) and (3.36) with $S_4 = (D^2 - S^2)^{1/2}$ we have the following action:

$$[T_i, S_j] = \{ \pm S_4\delta_{ij} - \epsilon_{ijk}S_k \}/2, \tag{4.8}$$

$$[T_i, S_4] = \mp S_i/2, \tag{4.9}$$

$$[Z_i, S_j] = \{ \mp S_4\delta_{ij} - \epsilon_{ijk}S_k \}/2, \tag{4.10}$$

$$[Z_i, S_4] = \pm S_i/2, \tag{4.11}$$

$$(j_1, j_2) = (1/2, 1/2). \tag{4.12}$$

The nature of this realization becomes more apparent if we use the alternate basis $\{\mathbf{J},\mathbf{K}\}$ the components of which are given by the following inverse of (2.6):

$$J_i = T_i + Z_i, \quad K_i = (-i)(T_i - Z_i). \tag{4.13}$$

In terms of these generators the above action has the following form:

$$[J_i, S_j] = -\epsilon_{ijk}S_k, \quad [J_i, S_4] = 0, \tag{4.14}$$

$$[K_i, S_j] = \mp iS_4\delta_{ij}, \quad [K_i, S_4] = \pm iS_i. \tag{4.15}$$

From these expressions it is clear that this case represents the usual linear transformation of the Lorentz group [or O(4)] on four dimensions. The realizations for the two solutions indicated by the two signs can be obtained one from the other by changing the sign on the group parameters. The reader should recall that we are dealing with the Lorentz group if the parameter $\mathbf{v}$ in (2.9) is real and the O(4) group if $\mathbf{v}$ is pure imaginary. The invariant metric for this case has the following form:

$$-d\eta^2 = 4[dS_i\,dS_i + (S_i\,dS_i)^2/(D^2 - S^2)]/D^2. \tag{4.16}$$

For the metric in (4.16) the function D is not required to be a constant.

Several realizations that are equivalent via (3.9) to this 4-vector realization have appeared in the literature. These include the O(3,1)/O(3) coset realizations of the Lorentz group studied by Hind[2] and later by Hopkinson and Reya[3] as well as the nonlinear realizations of the chirality group $SU(2)\times SU(2)$ on the pion field studied by Weinberg.[15] The specific transformations between the O(3,1)/O(3) coset realizations and the transformation $TE\times ZE$ given here can be found in the recent work by the present author.[4]

$TE\times ZA$: This case is much like the electromagnetic realization considered above in the sense that the Z action is null. We have the following action:

$$[T_i,S_j] = \{\pm S_4\delta_{ij} - \epsilon_{ijk}S_k\}/2, \qquad (4.17)$$

$$[T_i,S_4] = \mp S_i/2, \quad [Z_i,S_j] = 0, \qquad (4.18)$$

where $S_4 = (D^2 - S^2)^{1/2}$ is one of the roots of $(D^2 - S^2)^{1/2}$. For the angular momentum and boost generators we have the following equations:

$$[J_i,S_j] = \{\pm S_4\delta_{ij} - \epsilon_{ijk}S_k\}/2, \qquad (4.19)$$

$$[K_i,S_j] = (-i)[J_i,S_j]. \qquad (4.20)$$

We have $(j_1j_2) = (0,1/2)$ and the invariant metric is given by (4.16). We also have the following eigenvalue equation:

$$-[J_i,[J_i,S_j]] = (3/4)S_j. \qquad (4.21)$$

Equations (4.21) suggests that we are dealing with a type of spin 1/2 field. We are not dealing with a Dirac 4-spinor, however. This can be seen from considering Eq. (4.20). For a Dirac spinor we have the following relations:

$$M_{\mu\nu} = -i[\gamma_\mu,\gamma_\nu]/4, \quad \gamma_\mu\gamma_\nu + \gamma_\nu\gamma_\mu = 2\delta_{\mu\nu}. \qquad (4.22)$$

This particular definition of $M_{\mu\nu}$ satisfies the commutation relations (2.1). From this and (2.2) we have

$$J_i = \tfrac{1}{4}\epsilon_{ijk}\gamma_j\gamma_k, \quad K_i = (-i)\tfrac{1}{2}\gamma_4\gamma_i. \qquad (4.23)$$

Using (4.20) with J_i and K_i defined by (4.22) we arrive at $\gamma_4^2 = -1$ which is inconsistent with (4.22) in which $\gamma_4^2 = +1$.

Another interesting point about this realization is that we do not have a diagonal angular momentum projection operator. If one had $S_4 = i|\mathbf{S}|$, then from (4.18) it is clear that we could diagonalize J_3, but this condition is not possible since we have $D^2 \equiv S_4^2 + S^2 \neq 0$ for this realilzation.

We may directly integrate, using (4.19), to obtain the following finite transformation for real ω_1 (for generator J_1):

$$S_1' = S_1\cos\theta_1 \pm S_4\sin\theta_1, \quad S_4' = S_4\cos\theta_1 \mp S_1\sin\theta_1, \qquad (4.24)$$

$$S_2' = S_2\cos\theta_1 - S_3\sin\theta_1, \quad S_3' = S_3\cos\theta_1 + S_2\sin\theta_1. \qquad (4.25)$$

Since we have $\theta_1 = \omega_1/2$ the period of rotation is 4π. Unlike the 4-vector realizations just discussed, the realizations corresponding to the two signs cannot be obtained from each other by changing the sign on the group parameters because the two signs do not occur in all parts of the generator (4.19). A space $\mathbf{S}$ either transforms under one sign or the other. For the real boost parameter ν_1 we can directly integrate, using (4.20), to obtain the following finite transformation generated by K_1:

$$S_1' = S_1\cosh\phi_1 \mp iS_4\sinh\phi_1, \quad S_4' = S_4\cosh\phi_1 \pm S_1\sinh\phi_1, \qquad (4.26)$$

$$S_2' = S_2\cosh\phi_1 + iS_3\sinh\phi_1, \quad S_3' = S_3\cosh\phi_1 - iS_2\sinh\phi_1. \qquad (4.27)$$

where

$$\phi_1 = \nu_1/2.$$

The invariant Lagrangian for this case is given by

$$\mathscr{L} = [\partial_\mu S_j\partial_\mu S_j + (S_i\partial_\mu S_i)/(D^2 - S^2)]/D^2. \qquad (4.28)$$

If we have two spaces, say A_μ and S_μ, which transform under this realization, then the quantity $A_\mu S_\mu$ is invariant. We again emphasize that this realization is not the usual Dirac realization for spinors.

$TC\times ZC$: *S^2 is invariant*:

$$[T_i,S_j] = [\pm iS\delta_{ij} \mp iS_iS_j/S - \epsilon_{ijk}S_k]/2, \qquad (4.29)$$

$$[Z_i,S_j] = [\mp iS\delta_{ij} \pm iS_iS_j/S - \epsilon_{ijk}S_k]/2, \qquad (4.30)$$

$$(j_1,j_2) = (0,0). \qquad (4.31)$$

For the angular momentum and boost generators we have the action

$$[J_i,S_j] = -\epsilon_{ijk}S_k, \qquad (4.32)$$

$$[K_i,S_j] = \pm S\delta_{ij} \mp S_iS_j/S. \qquad (4.33)$$

The angular momentum action generates the usual linear vector rotation. For the boost generators, a direct integration using (4.33) gives the following finite transformation for the real parameter ν_1:

$$S_1' = (S_1\cosh\nu_1 \pm S\sinh\nu_1)/Q, \quad S_2' = S_2/Q, \qquad (4.34)$$

$$Q = (S\cosh\nu_1 \pm S_1\sinh\nu_1)/S, \quad S_3' = S_3/Q. \qquad (4.35)$$

We have here a nonlinear realization which leaves S^2 invariant. It is obvious from (4.29) that this realization is not defined for $S^2 = 0$. Suppose we write $\mathbf{S} = \mathbf{R} + i\mathbf{I}$ where $\mathbf{R}$ and $\mathbf{I}$ represent the real and imaginary components. We have two invariants $R^2 - I^2$ and $\mathbf{R}\cdot\mathbf{I}$. One or the other but not both of these invariants could be zero for this realization.

Equation (4.32) suggests that we are dealing with a vector type field in this case. In contrast to the previous realization we can diagonalize an angular momentum projection operator, $[J_3,S_3] = 0$.

This particular type of nonlinear realization has been studied previously in conjunction with extensions of the space–time Lorentz group to the de Sitter group. An excellent treatment and review can be found in the work by Philips and Wigner.[7] Following this earlier work one can extend the above K_i generators by adding a term of form $\lambda S_i/S$, that is,

$$K_i(TC\times ZC) \to K_i(TC\times ZC) + \lambda S_i/S, \qquad (4.36)$$

where λ is a constant which can be related to $j_2(j_2+1)$ by evaluating (3.11). Under the action of these extended generators, S^2 is still invariant because the added nondifferential term commutes with S^2.

$TC\times ZA$: *S^2 is invariant*:

$$[J_i,S_j] = [\pm iS\delta_{ij} \mp iS_iS_j/S - \epsilon_{ijk}S_k/2], \qquad (4.37)$$

$$[K_i,S_j] = (-i)[J_i,S_j], \qquad (4.38)$$

$$(j_1,j_2) = (0,1/2). \qquad (4.39)$$

For this realization we have the invariant eigenvalue equation

$$- [J_i,[J_i,S_j]] = 0. \tag{4.40}$$

This equation suggests that we are dealing with a special type of scalar field for which the group action is not the usual scalar field *null* action. Even though we have the nonlinear form in (4.37) we can find a "point" in the space S for which J_3 is diagonal. Equation (4.37) with $S_1 = S_2 = 0$ gives $[J_3,S_3] = 0$.

The finite transformation generated by J_1, with real ω_1, is obtained by integration using (4.37):

$$S'_1 = (S_1\cos\theta_1 \pm iS\sin\theta_1)/Q, \tag{4.41}$$

$$S'_2 = (S_2\cos\theta_1 - S_3\sin\theta_1)/Q, \tag{4.42}$$

$$S'_3 = (S_3\cos\theta_1 + S_2\sin\theta_1)/Q, \tag{4.43}$$

$$Q = (S\cos\theta_1 \pm iS_1\sin\theta_1)/S. \tag{4.44}$$

Here we have $\theta_1 = \omega_1/2$. In a like manner the finite transformations generated by K_1 with real parameter ν_1 are obtained:

$$S'_1 = (S_1\cosh\phi_1 \pm S\sinh\phi_1)/Q, \tag{4.45}$$

$$S'_2 = (S_2\cosh\phi_1 + iS_3\sinh\phi_1)/Q, \tag{4.46}$$

$$S'_3 = (S_3\cosh\phi_1 - iS_2\sinh\phi_1)/Q, \tag{4.47}$$

$$Q = (S\cosh\phi_1 \pm S_1\sinh\phi_1)/S. \tag{4.48}$$

Here we have used $\phi_1 = \nu_1/2$.

$TD \times ZD$: *Arbitrary* $j_1 = j_2$:

$$[T_i,S_j] = \{ \pm iS\delta_{ij} \mp ibS_iS_j/S - \epsilon_{ijk}S_k \}/2 \tag{4.49}$$

$$[Z_i,S_j] = \{ \mp iS\delta_{ij} \pm ibS_iS_j/S - \epsilon_{ijk}S_k \}/2, \tag{4.50}$$

$$[J_i,S_j] = -\epsilon_{ijk}S_k, \tag{4.51}$$

$$[K_i,S_j] = \pm S\delta_{ij} \mp bS_iS_j/S. \tag{4.52}$$

The relation between the constant b and $j_1 = j_2$ is given in Eq. (3.24). The angular momentum subgroup has the usual linear form. For real parameter ν_1 we can integrate, using (4.52) to obtain the following finite transformation:

$$S'_1 = (S_1\cosh\nu_1 \pm S\sinh\nu_1)/Q, \tag{4.53}$$

$$S'_2 = S_2/Q, \quad S'_3 = S_3/Q, \tag{4.54}$$

$$Q = (\cosh\nu_1 \pm (S_1/S)\sinh\nu_1)^b. \tag{4.55}$$

$$b = 2 \pm [1 + 4j_2(j_2+1)]^{1/2}. \tag{4.56}$$

The generator K_i for the action in (4.52) can be written as the corresponding generator for the $TC \times ZC$ realization plus an added differential term,

$$K_i(TD \times ZD) = K_i(TC \times ZC) \mp i(b-1)\frac{S_i}{2S}S_l\frac{\partial}{\partial S_l}. \tag{4.57}$$

In contrast to the extended $TC \times ZC$ boost generators indicated in (4.36) this operator does not commute with S^2.

We remind the reader that although the $TD \times ZD$ realization does reduce to the $TC \times ZC$ case for $j_2 = 0$ with the negative sign in b, we have treated the latter case separately because we wish to associate with each class of realizations an invariant Lagrangian form. The metric (3.29) is not valid for the $TC \times ZC$ realization, for which $F + hS^2 = 0$ so that we cannot use this metric to construct an invariant Lagrangian for the $TC \times ZC$ case.

$TD \times ZA$:$(j_1,j_2) = (0,j_2)$:

$$[J_i,S_j] = \{ \pm iS\delta_{ij} \mp ibS_iS_j/S - \epsilon_{ijk}S_k \}/2, \tag{4.58}$$

$$[K_i,S_j] = (-i)[J_i,S_j]. \tag{4.59}$$

For this case we have the following invariant eigenvalue equation:

$$-[J_i,[J_i,S_j]] = j_2(j_2+1)S_j \tag{4.60}$$

We can integrate, using (4.58) to obtain the following finite transformations for real parameter ω_1:

$$S'_1 = (S_1\cos\theta_1 \pm iS\sin\theta_1)/Q, \tag{4.61}$$

$$S'_2 = (S_2\cos\theta_1 - S_3\sin\theta_1)/Q, \tag{4.62}$$

$$S'_3 = (S_3\cos\theta_1 + S_2\sin\theta_1)/Q, \tag{4.63}$$

$$Q = (\cos\theta_1 \pm i(S_1/S)\sin\theta_1)^b. \tag{4.64}$$

Here, $\theta_1 = \omega_1/2$ and b is given by (4.56). For the real parameter ν_1 we have, using (4.59), the following finite transformations:

$$S'_1 = (S_1\cosh\phi_1 \pm S\sinh\phi_1)/Q, \tag{4.65}$$

$$S'_2 = (S_2\cosh\phi_1 + iS_3\sinh\phi_1)/Q, \tag{4.66}$$

$$S'_3 = (S_3\cosh\phi_1 - iS_2\sinh\phi_1)/Q, \tag{4.67}$$

$$Q = (\cosh\phi_1 \pm (S_1/S)\sinh\phi_1)^b.. \tag{4.68}$$

For the case where $j_2 = 1/2$ and the negative sign in b this realization becomes linear. We can in this case diagonalize a projection operator, say J_3, at the point $S_2 = S_1 = 0$,

$$i[J_3,S_3] = \mp\tfrac{1}{2}S_3. \tag{4.69}$$

Because of (4.59), this case, like the $TE \times ZA$ case, does not correspond to a Dirac spinor. One important characteristic of this realization is that for the angular momentum subgroup the period for ω_1 is 4π for *all* nonzero spin values. This feature clearly distinguishes this class of realizations from the above mentioned $TD \times ZD$ realizations, and from the usual integer spin representations.

For the $TD \times TD$ and $TD \times ZA$ realizations the invariant metric is given by (3.29). From this invariant form we can construct the following invariant Lagrangian density:

$$\mathscr{L} = S^{2\alpha}[(\partial_\mu S_i)(\partial_\mu S_i) - S^{-2}(S_i\partial_\mu S_i)^2]. \tag{4.70}$$

Here, α is given by (3.28). For the special spinor case mentioned above with $j_2 = \frac{1}{2}$ we have $\alpha = 0$. We remind the reader that this Lagrangian form is not invariant for the case where $j_2 = 0$.

Before leaving this section we point out that with the particular basis used here for the Lie algebra we can easily obtain from each of the above realizations of SL(2,C) a realization of O(4) by the substution $\mathbf{v} \to i\mathbf{v}$. Although the nature of the space will generally change, the above forms for the invariant metrics, or Lagrangian densities, are still valid.

V. WAVE EQUATIONS, MASS AND SPIN

In this section we discuss covariant wave equations for the new realizations presented in the preceeding sections with the particular aim of demonstrating that complex vector spaces can be found which transform under certain of these nonlinear realizations and for which the mass and spin operators are diagonal.

Traditionally, physical particles with spin have been associated with spaces which transform linearly (i.e, representations) under the group SL(2,C). The homogeneous group of interest is the diagonal subgroup of the direct product of two copies of SL(2,C), one generated by the space–time orbital generators $l_{\mu\gamma}$ and the other by the intrinsic spin generators $S_{\mu\gamma}$. The generators for this diagonal subgroup are given by

$$M_{\mu\gamma} = l_{\mu\gamma} + S_{\mu\gamma}, \quad l_{\mu\gamma} = -i(\chi_\mu\partial_\gamma - \chi_\gamma\partial_\mu) \tag{5.1}$$

and satisfy the commutation relations (2.1). The group of interest here is the Poincaré group which is the semidirect product of the above diagonal Lorentz subgroup with the space–time translation group. The generators P_μ for the latter group are defined by

$$P_\mu = -i\partial_\mu. \tag{5.2}$$

In addition to (2.1) the Lie algebra for this group has the following commutation relations.

$$[M_{\mu\gamma}, P_\rho] = i[\delta_{\mu\rho}P_\gamma - \delta_{\gamma\rho}P_\mu], \tag{5.3}$$

$$[P_\mu, P_\rho] = 0. \tag{5.4}$$

In the linear realizations the physical mass and spin of particles are identified with the Casimir invariants of this group, that is

$$[P_\mu, [P_\mu, S_j]] = -M^2 S_j, \tag{5.5}$$

$$[W_\mu, [W_\mu, S_j]] = M^2 j(j+1)S_j, \tag{5.6}$$

where W_μ is a component of the Pauli–Lubanski 4-vector given by

$$W_\mu = -\epsilon_{\mu\nu\alpha\rho}M_{\nu\alpha}P_\rho/2. \tag{5.7}$$

With (5.1) one can satisfy the commutation relations (2.1) and (5.3) because the intrinsic spin generators $S_{\mu\gamma}$ commute with the partial derivatives appearing in $l_{\mu\gamma}$ and P_μ.

We are concerned here with the interesting question, "Are there complex vector spaces transforming nonlinearly under SL(2,C) for which the generators also commute with the operators $l_{\mu\nu}$ and P_μ?" The intrinsic generators T_i acting on S will commute with $l_{\mu\nu}$ and P_μ if the following equation is satisfied.

$$[\partial_\mu, [T_i, S_j]] = [T_i, [\partial_\mu, S_j]]. \tag{5.8}$$

To demonstrate that (5.8) can be satisfied for some of the nonlinear realizations considered in Sec. IV. we consider spaces which can be factored in the following way:

$$S_j = A_j\Psi, \tag{5.9}$$

where $\partial_\mu A_j = 0$ and Ψ is a Lorentz scalar. Using (5.9) in (5.8) with (3.1) we arrive at the following equation for $\Psi \neq 0$:

$$[2F'S^2\delta_{ij} + 2(h'S^2 + h)S_iS_j + G\epsilon_{ijk}S_k]\Psi^{-1}\partial_\mu\Psi$$
$$= [F\delta_{ij} + hS_iS_j + G\epsilon_{ijk}S_k]\Psi^{-1}\partial_\mu\Psi. \tag{5.10}$$

In deriving this expression we have assumed that the space–time dependence of F and h arise only through the space–time dependence of S^2. Comparing the two sides of (5.10) we must conclude that either $F = h = 0$, corresponding to either the null or linear realization, or that these functions satisfy the following conditions:

$$2F'S^2 = F, \quad 2(h'S^2 + h) = h. \tag{5.11}$$

Equations (3.22) and (5.11) can both be satisfied with $F = \pm iS/2$ and $h = \mp ib/2S$ which, if the reader recalls, are just the solutions for the $TD \times ZD$ and $TD \times ZA$ realizations; as well as the $TC \times ZC$ and $TC \times ZA$ realizations when $b = 1$. The equation corresponding to (5.8) for the generator Z_i leads to the same conditions given in (5.11) so that we need not consider it separately. With Eq. (5.11) the transformed space S′ will also have a factored form:

$$S_j' = A_j'\Psi. \tag{5.12}$$

With (5.9) the Jacobi identity

$$[P_\mu, [P_\gamma, S_j]] - [P_\gamma, [P_\mu, S_j]] = 0 \tag{5.13}$$

corresponding to (5.4) is satisfied. The commutation relations for the Lie algebra of the Poincaré group are satisfied so that we may choose to diagonalize the mass and spin invariants as in (5.5) and (5.6). With (5.9) and (5.5) the function Ψ will satisfy the Klein–Gordon equation

$$\Box^2\Psi = M^2\Psi. \tag{5.14}$$

With (5.9) we have a complex vector space S with each component having the same space–time dependence which may be a solution of (5.14).

We next consider some extremum equations obtained from invariant Lagrangian densities for the above realizations and show that with (5.9) these equations are consistent with the Klein–Gordon equation. For the $TC \times ZC$ and $TC \times ZA$ realizations S^2 is invariant. Using the invariant metric given by (3.21) we can construct an invariant differential form $(S_i\partial_\mu S_i)(S_j\partial_\mu S_j)$. We may choose several combinations for the invariant Lagrangian density. Consider for example the following choice:

$$\mathscr{L} = S^{-2}(S_i\partial_\mu S_i)(S_j\partial_\mu S_j) + M^2R(S^2), \tag{5.15}$$

where R is some function of S^2, and M^2 is a constant. The Euler–Lagrange equations for this $\mathscr{L}$ reduce to the following form:

$$-\frac{1}{S^4}(S_i\partial_\mu S_i)^2 + \frac{1}{S^2}\partial_\mu(S_i\partial_\mu S_i) = M^2\frac{\partial R}{\partial S^2}. \tag{5.16}$$

With the special choice of $R = S^2$ and using (5.9), Eq. (5.16) reduces to the Klein–Gordon equation (5.14) for Ψ. We can thus choose an invariant Lagrangian for the $TC \times ZC$ and $TC \times ZA$ realizations for which the extremum equations reduce to the Klein–Gordon equation for a space which satisfies (5.9). With different choices for the function R in (5.15), one can obtain a number of interesting equations for Ψ. For instance, with the choice $R = -2\cos(S^2)^{1/2}$ and $A^2 \equiv A_iA_i = 1$, Eq. (5.15) reduces to the covariant sine-Gordon equation,

$$\Box^2\Psi = \sin\Psi.$$

The point of interest here is that with the $TC \times ZC$ and $TC \times ZA$ realizations it is possible also to have an invariant wave equation that admits soliton solutions,[8–11] at least in two dimensions.

For the $TD \times ZD$ and $TD \times ZA$ realizations S^2 is not an invariant so that for these realizations we do not have as much freedom in choosing the invariant Lagrangian as in the previous cases. The extremum equations for the invariant differential expression in (4.70) simplify to the following form:

$$[\Box^2 S_i - S^{-2} S_i (S_l \Box^2 S_l)] + 2\alpha S^{-2} (S_l \partial_\mu S_l) \cdot [\partial_\mu S_i - S^{-2} (S_l \partial_\mu S_l) S_i] = 0. \tag{5.17}$$

For the factored space indicated in (5.9) the expressions in both brackets in (5.17) are each identically zero. This simply means that the factoring of the space as in (5.9) is sufficient to solve (5.17). With (5.9) then, the extremum equations impose no additional conditions on the space so that we can choose Ψ to be a solution of the Klein–Gordon equation.

In the above we have demonstrated that for certain of these nonlinear realizations we can find at least some complex vector spaces for which the mass and spin operators are diagonal. With (5.9) and (5.14) these particular spaces are somewhat analogous to a classical electromagnetic field picture of the photon in which each component of the field has the same space–time propagation, (e.g, plane waves).

To demonstate the possible intrinsic spin features for these spaces we consider (5.9) and plane wave solutions $\Psi = \exp(i\chi_\mu P_\mu)$. For nonzero mass the spin invariant (5.6) in the rest frame $\mathbf{p} = 0$ is proportional to the Casimir invariant for the angular momentum subgroup, that is

$$\begin{aligned}[][W_\mu,[W_\mu,S_k]] &= -M^2[J_i,[J_i,S_k]] \\ &= M^2 j(j+1) S_k.\end{aligned} \tag{5.18}$$

For both the $TC \times ZC$ and $TD \times ZD$ realizations we have $j = 1$ as the only possible value. Since $j_1 = j_2$ is not fixed, we have in the $TD \times ZD$ case an unlimited number of realizations which have the same physical (Poincaré) spin $j = 1$ and rotational period of 2π. In the TD realization all of the solutions for different j_2 are equivalent under the change of variables

$$\widetilde{S}_i = (S^2)^\beta S_i, \quad \beta = (b - \widetilde{b})/2(1 - b), \tag{5.19}$$

where b and $\widetilde{b}$ depend on j_2 and $\widetilde{j}_2$ as in (3.24). For the $TD \times ZD$ realizations the physical spin ($j = 1$) will not distinguish between these realizations. However, because of the nonlinearity in (5.19) the masses will differ. To see this consider (5.9) and plane wave solutions for Ψ. With (5.19) we arrive at the following relationship between the masses.

$$\widetilde{M} = \frac{(1 - \widetilde{b})}{(1 - b)} M. \tag{5.20}$$

In this $TD \times ZD$ realization we can have many vector ($j = 1$) solutions but each with a different mass.

For the $TC \times ZA$ realization we have $j = 0$ as the only possibility. The rotational period for this case is 4π. This is most interesting! We have in this case a nontrivial realization with zero spin, rotational period of 4π, and the mass operator can be diagonal. This is possible with nonlinear realizations.

In the $TD \times ZA$ case we have a realization for all values of the spin j. In each of these realizations the rotational period is also 4π. The $j = 1$ realization in this case differs from the $j = 1$ realizations in the $TD \times ZD$ and $TC \times ZC$ cases. In these realizations we have $j = j_2$ so that the different j_2 realizations have different Poincaré spins. In this case we also have many complex vector spaces that are equivalent via (5.19) with mass relation (5.20). In this case, however, we have $j_2 = j$ so that the physical spin distinguishes between the masses. With (3.24) in (5.20) we have the relation:

$$\widetilde{M} = \frac{[-1 \pm (2\widetilde{j} + 1)]M}{[-1 \pm (2j + 1)]}. \tag{5.21}$$

We can have $j = 0$ for both the $TC \times ZA$ and $TD \times ZA$ realizations. We have with these realizations two possible categories for the spin zero mesons. The only other possible category for these mesons is the usual trivial or null realization. In both cases the components can be eigenfunctions of the Klein–Gordon wave equation but in the nonlinear case we have three wave components which mix under $SL(2,C)$ and transform under rotations with a period of 4π.

For nonzero mass we have a Poincaré spin of $\frac{1}{2}$ for the $TE \times ZA$ realization. We also have a Poincaré spin of $\frac{1}{2}$, among other spins, in the $TD \times ZA$ realization. These two, along with the Dirac spinors give three possible categories for spin $\frac{1}{2}$ fields.

For both the $TC \times ZC$ and $TD \times ZD$ realizations we have $j = 1$. The $TD \times ZA$ category also contains a $j = 1$ realization. With these and the usual vector realizations of electrodynamics we have more different types of realizations with spin 1 than any other spin. With this number and variety of spin 1 realizations it does not seem surprising to find in nature several different types of spin 1 particles as evidenced by the photons, vector mesons and the recently discovered J/ψ particle.

The new realizations discussed above significantly enlarge the number of known mathematical realizations of the physical space–time group. The fact that these new realizations exist mathematically does not mean that they exist in nature as transformations on physical fields. However, the possibility that physical fields could exist which transform under these (or other) nonlinear realizations of $SL(2,C)$ is a very interesting conjecture, especially in light of the fact that we have in nature several physically inequivalent categories of particles such as the leptons, mesons, and hadrons. If, for instance, physical particles with spin $\frac{1}{2}$ exist which transform under one of the above new spin $\frac{1}{2}$ realizations, then the use of Dirac's equation to describe the properties of these particles would be invalid.

ACKNOWLEDGMENTS

The author would like to thank Professor Derek Pursey who took the time to independently derive and check the mathematics in this paper. Professor Stan Williams and Professor Brian de Facio are acknowledged for reading the manuscript and making several helpful suggestions. This work was supported by the U.S. Department of Energy, Office of Basic Sciences, Nuclear Physics Division.

[1] D.L. Pursey, Ann. Phys. (N.Y.) **32**, 157 (1965). This article gives references to several classical papers on this subject.
[2] J.D. Hind, Nuovo Cimento A **4**, 71 (1971).
[3] J.F.L. Hopkinson and E. Reya, Phys. Rev. D **10**, 342 (1972).
[4] B.J. Dalton, J. Math. Phys. **19**, 1335 (1978). These new realizations refer to solutions given by (3.17) of this reference.
[5] M.A. Melvin, Bull. Am. Phys. Soc. **7**, 493 (1962); **8**, 356 (1963).
[6] T. Takabayasi, Prog. Theor. Phys. **36**, 1074 (1966).
[7] T.O. Philips and E.P. Wigner, in *Group Theory and Its Applications*, edited by Ernest M. Loebl (Academic, New York, 1968), pp 331–76.
[8] This possibility has been independently suggested to the author by Professor J.P. Davidson and Professor C.L. Hammer.
[9] R. Jackiw, Rev. Mod. Phys. **49**, 681 (1977).
[10] G.B. Whitham, *Linear and Non-Linear Waves* (Wiley, New York, 1974).
[11] R. Miura, editor, *Bäcklund Transformations*, Springer Lecture Notes in Mathematics No. 515 (Springer, New York, 1976).
[12] We follow the work of Itzhak Bars and Feza Gürsey, J. Math. Phys. **13**, 131 (1972), except that our generator J_i and K_i are related to those in this reference via $J_i = iJ_i(BG)$ and $K_i = -iK_i(BG)$.
[13] See for instance, N. Jacobson, *Lie Algebras* (Interscience, New York, 1966).
[14] It is more common to use real parameters with an alternate basis along with a change of sign in (2.4). For our particular choice of bases we easily go from SL(2,C) to O(4) realizations with only $\mathbf{v} \rightarrow i\mathbf{v}$.
[15] S. Weinberg, Phys. Rev. **166**, 1568 (1968).

Nth-order multifrequency coherence functions: A functional path integral approach[a)]

Conrad M. Rose

Naval Surface Weapons Center, Dahlgren, Virginia 22448

Ioannis M. Besieris

Virginia Polytechnic Institute and State University, Blacksburg, Virginia 24061

(Received 21 August 1978)

A functional (or path) integral applicable to a broad class of randomly perturbed media is constructed for the nth-order multifrequency coherence function (a quantity intimately linked to nth-order pulse statistics). This path integral is subsequently carried out explicitly in the case of a nondispersive, deterministically homogeneous medium, with a simplified (quadratic) Kolmogorov spectrum, and a series of new results are derived. Special cases dealing with the two-frequency mutual coherence function for plane and beam pulsed waves are considered, and comparisons are made with previously reported findings.

1. INTRODUCTION

There has been recently a mounting interest in the subject of propagation of pulsed signals through randomly perturbed media. Proposed high data rate communication systems at millimeter and optical frequencies, remote sensing schemes, low-frequency underwater sound signaling and detection, interpretation of signals emitted by extraterrestrial radio sources such as pulsars, all require a quantitative assessment of stochastic pulse broadening. The latter leads to an irreversible degradation of a signal, in contradistinction to dispersive pulse spreading, which is a reversible phenomenon. (The receiver is usually equipped with "built-in dispersion" in order to make optimal use of the additional signal bandwidth).

Earlier contributions in this area (cf., e.g., Refs. 1–3) were confined to weakly turbulent media and/or short propagation paths, for which methods such as the Born and Rytov approximations were adequate. More recent theoretical analyses which account for multiple scattering, large-scale inhomogeneities and long propagation distances are based for the most part on the parabolic equation for the complex field amplitude and the Markov random process approximation.[4] Within the framework of this formalism, it has been recognized that complete information about transient signal propagation in random media requires the solutions of moment equations for the wave field at different frequencies and different positions. Although a complete set of such equations can be derived (cf. e.g., Ref. 5), solutions are available only for the two-frequency mutual coherence function, and these results (both analytical and numerical) are generally restricted to plane wave calculations.[6–15]

There exist physical situations (e.g., laser beam propagation in the atmosphere and in lightguides, and underwater sound wave propagation) where the restriction to spatially planar sources must be lifted. The problem of propagation of pulsed beam waves in randomly perturbed environment has been investigated by the method of "temporal moments."[16] Since this technique is based on the two-frequency mutual coherence function, it yields only information at the level of second-order pulse statistics (e.g., mean arrival time, mean square pulse width, etc.). Pulse shapes cannot be obtained directly by this method; however, they can be synthesized by superimposing individual temporals moments. An alternative technique which, at least in principle, enables one to compute approximately multifrequency coherence functions for arbitrary beam waves propagating in a random medium was proposed recently by Fante.[17] It is based on the phase-screen approximation along with the extended Huygens–Fresnel principle. So far, only results pertaining to the two-frequency mutual coherence function have been reported in the literature.[18]

One of the reasons why the aforementioned approaches (with the exception, perhaps, of the one suggested by Fante[17]) have not yielded sufficient information in connection with multifrequency coherence functions for beams in a random medium is that the study of the asymptotic (or even exact in special cases) behavior of these functions based on their governing local equations is a nontrivial problem. In contradistinction, recently formulated methods based on functional path integration (cf. Refs. 19–21) have the distinct advantage that they work on a *global* rather than a *local* level, thus making the algorithmic derivation of asymptotic solutions to higher-order moments easier.

It is our specific intent in this exposition to use the functional path integration approach in order to evaluate nth-order coherence functions for arbitrary source distributions. These evaluations will be performed within the confines of a simplified (quadratic) Kolmogorov spectrum. (Similar re-

a)Research supported in part by the Office of Naval Research under contract No. N00014-76-C-0056 and the Graduate and Internal Research Programs of the Naval Surface Weapons Center.

sults, for a single frequency and a specific source distribution, have been determined by Furutsu[22] using a different technique.)

The structure of the paper can be outlined as follows: Several preliminary concepts pertaining to pulse propagation within the domain of validity of the quasioptics (or parabolic) approximation are developed in Sec. 2. A functional path integral applicable to a broad class of random media is constructed at the beginning of Sec. 3 for the nth-order multifrequency coherence function (a quantity intimately linked to nth-order pulse statistics). This path integral is subsequently carried out explicitly in the case of a nondispersive, deterministically homogeneous medium, with a simplified (quadratic) Kolmogorov spectrum, and a series of new results are derived. Special cases dealing with the two-frequency mutual coherence function for plane and beam pulsed waves are considered in Sec. 4, where comparisons are also made with previously reported findings. Finally, the possibility of asymptotic expansions of our general exact results in the partially and fully saturated regimes are briefly discussed in Sec. 5.

2. PRELIMINARY CONCEPTS

The preliminary analysis, as well as the notation, in this section will be specialized to the case of electromagnetic wave propagation in a random channel. It should be emphasized, however, that the resulting stochastic complex parabolic equation is equally applicable in other physical areas (e.g., random underwater sound propagation).

A. The quasioptics approximation

Ignoring depolarization effects, the transverse, complex, electric field of radiation is governed by the stochastic Helmholtz equation

$$\nabla^2 E(\mathbf{r},\omega;\alpha) + (\omega/c)^2\epsilon_r(\mathbf{r},\omega;\alpha)E(\mathbf{r},\omega;\alpha) = 0, \quad \mathbf{r}\in R^3. \quad (2.1)$$

Here, ω is the angular frequency, c is the speed of light in *vacuo*, and $\epsilon_r(\mathbf{r},\omega;\alpha)$—the relative permittivity provided the medium is nonmagnetic—is a dimensionless, scalar, random function depending on a parameter $\alpha\in A$, (A,F,P) being an underlying probability measure space. If, in addition to dispersion, the medium is characterized by either gain or loss, the relative permittivity $\epsilon_r(\mathbf{r},\omega;\alpha)$ is complex. In the sequel, we shall restrict our attention to physical situations where this quantity is real.

Let $E\{\epsilon_r(\mathbf{r},\omega;\alpha)\}$ and $\delta\epsilon_r(\mathbf{r},\omega;\alpha)$ denote respectively the average and fluctuating parts of the relative permittivity. Let, furthermore, $\epsilon_0(\omega)$ be a convenient "reference" quantity. It may, for example, coincide with the average relative permittivity if the latter is independent of the position variable $\mathbf{r}$. We introduce, next, three new quantities as follows:

$$k^2(\omega) = (\omega/c)^2\epsilon_0(\omega), \quad (2.2)$$

$$\epsilon_1(\mathbf{r},\omega;\alpha) = \delta\epsilon_r(\mathbf{r},\omega;\alpha)/\epsilon_0(\omega), \quad (2.3)$$

$$\epsilon_2(\mathbf{r},\omega) = [E(\epsilon_r(\mathbf{r},\omega;\alpha)) - \epsilon_0(\omega)]/\epsilon_0(\omega). \quad (2.4)$$

With these definitions, (2.1) assumes the form

$$\nabla^2 E(\mathbf{r},\omega;\alpha) + k^2(\omega)[1 + \epsilon_2(\mathbf{r},\omega) + \epsilon_1(\mathbf{r},\omega;\alpha)]E(\mathbf{r},\omega;\alpha) = 0. \quad (2.5)$$

It is clear that $\epsilon_2(\mathbf{r},\omega)$ accounts for deterministic inhomogeneities in the medium and, in light of a statement made earlier, it vanishes in the absence of such background profiles. On the other hand, $\epsilon_1(\mathbf{r},\omega;\alpha)$, a zero-mean random function, is directly associated with the superimposed random effects.

For plane and beam propagation in the z-direction, it is convenient to resort to the transformation

$$E(\mathbf{r},\omega;\alpha) = \Psi(\mathbf{x},z,\omega;\alpha)\exp(ikz); \mathbf{r} = (\mathbf{x},z), k = k(\omega). \quad (2.6)$$

In the quasioptical description, the slowly varying complex random amplitude function $\Psi(\mathbf{x},z,\omega;\alpha)$ is described exceedingly well by the stochastic complex parabolic equation

$$\frac{i}{k}\frac{\partial}{\partial z}\Psi(\mathbf{x},z,\omega;\alpha) = -\frac{1}{2k^2}\nabla_{\mathbf{x}}^2\Psi(\mathbf{x},z,\omega;\alpha) - \tfrac{1}{2}[\epsilon_2(\mathbf{x},z,\omega) + \epsilon_1(\mathbf{x},z,\omega;\alpha)]\Psi(\mathbf{x},z,\omega;\alpha), \quad z>0. \quad (2.7a)$$

In the presence of a deterministic profile ($\epsilon_2\neq 0$), the parabolic equation (2.7a) constitutes a valid approximation to (2.5) if the normals to the wavefronts in the unperturbed problem, where $\epsilon_1 = 0$, remain close to the z axis.

Corresponding to the boundary condition $E(\mathbf{x},0,\omega;\alpha)\equiv E_0(\mathbf{x},\omega;\alpha)$ for (2.5), one has the "initial" condition

$$\Psi(\mathbf{x},0,\omega;\alpha)\equiv\Psi_0(\mathbf{x},\omega;\alpha) = E_0(\mathbf{x},\omega;\alpha), \quad (2.7b)$$

which incorporates all the information concerning the temporal frequency spectrum and the spatial distribution of the source at the initial plane $z = 0$.

In our subsequent formulation based on the functional path integral technique we shall require the fundamental solution (referred to alternatively as the propagator or Green's function) of (2.7). This quantity, denoted here by $G(\mathbf{x},\mathbf{x}',z,\omega;\alpha)$, provides a link between the wavefunction $\Psi(\mathbf{x},z,\omega;\alpha), z>0$, and the boundary condition $\Psi_0(\mathbf{x},\omega;\alpha)$, viz.,

$$\Psi(\mathbf{x},z,\omega;\alpha) = \int_{R^2} d\mathbf{x}' G(\mathbf{x},\mathbf{x}',z,\omega;\alpha)\Psi_0(\mathbf{x}',\omega;\alpha), \quad (2.8)$$

and satisfies the equation

$$\frac{i}{k}\frac{\partial}{\partial z}G(\mathbf{x},\mathbf{x}',z,\omega;\alpha) = -\frac{1}{2k^2}\nabla_{\mathbf{x}}^2 G(\mathbf{x},\mathbf{x}',z,\omega;\alpha) - \tfrac{1}{2}[\epsilon_2(\mathbf{x},z,\omega) + \epsilon_1(\mathbf{x},z,\omega;\alpha)]G(\mathbf{x},\mathbf{x}',z,\omega;\alpha), \quad z>0, \quad (2.9a)$$

$$G(\mathbf{x},\mathbf{x}',0,\omega;\alpha) = \delta(\mathbf{x}-\mathbf{x}'). \quad (2.9b)$$

B. Nth-order pulse statistics

Consider next the situation where a receiver at range z is characterized by a temporal spectrum $F_r(\omega)$—usually a bandpass function of frequency. It follows, then, from our work so far together with the existing linearity, that the wavefunction of interest at the receiver site is $E_r(\mathbf{x},z,\omega;\alpha)\equiv E(\mathbf{x},z,\omega;\alpha)F_r(\omega)$. The corresponding time-dependent, real, random signal can be expressed as

$$e_r(\mathbf{x},z,t;\alpha) = \frac{1}{2\pi}\int_{-\infty}^{\infty} d\omega \int_{R^2} d\mathbf{x}'\, G(x,\mathbf{x}',z,\omega;\alpha) \times F_r(\omega)E_0(\mathbf{x}',\omega;\alpha)\exp\{-i[\omega t - k(\omega)z]\}, \tag{2.10}$$

provided that the parabolic approximation is valid. The signal $e_r(\mathbf{x},z,t;\alpha)$ itself is not an observable quantity. However, a substantial amount of information associated with physically measurable pulse statistics is contained in the nth-order moments

$$E\left(\prod_{p=1}^{n} e_r(\mathbf{x}_p,z,t_p;\alpha)\right) = \frac{1}{(2\pi)^n}\int_{R^n} d\boldsymbol{\omega}\int_{R^{2n}} d\,\mathbf{X}' E\{G^{(n)}(\mathbf{X},\mathbf{X}',z,\boldsymbol{\omega};\alpha)\} \times F_r^{(n)}(\boldsymbol{\omega})E\{E_0^{(n)}(\mathbf{X}',\boldsymbol{\omega};\alpha)\} \times \exp\left(\sum_{p=1}^{n}(-i)\xi_p[\omega_p t_p - k(\omega_p)z]\right), \tag{2.11}$$

where n is assumed to be an even integer; $\xi_p = 1, p$ odd, $\xi_p = -1, p$ even, and the following notation is used:

$$\boldsymbol{\omega} = (\omega_1,\omega_2,\ldots,\omega_n)\in R^n, \tag{2.12a}$$

$$\mathbf{X} = (\mathbf{x}_1,\mathbf{x}_2,\ldots,x_n)\in R^{2n}, \tag{2.12b}$$

$$\mathbf{X}' = (\mathbf{x}_1'\mathbf{x}_2',\ldots,\mathbf{x}_n')\in R^{2n}; \tag{2.12c}$$

$$F_r^{(n)}(\boldsymbol{\omega}) = \prod_{p=1}^{n/2} F_r^*(\omega_{2p})F_r(\omega_{2p-1}), \tag{2.12d}$$

$$E_0^{(n)}(\mathbf{X}',\boldsymbol{\omega};\alpha) = \prod_{p=1}^{n/2} E_0^*(\mathbf{x}_{2p}',\omega_{2p};\alpha)E_0(\mathbf{x}_{2p-1}',\omega_{2p-1};\alpha), \tag{2.12e}$$

$$G^{(n)}(\mathbf{X},\mathbf{X}',z,\boldsymbol{\omega};\alpha) = \prod_{p=1}^{n/2} G^*(\mathbf{x}_{2p},\mathbf{x}_{2p}',z,\omega_{2p};\alpha) \times G(\mathbf{x}_{2p-1},\mathbf{x}_{2p-1}',z,\omega_{2p-1};\alpha). \tag{2.12f}$$

Several specific remarks are in order: (1) The derivation of (2.11) presupposes statistical independence between source incoherencies and random fluctuations in the medium; moreover, the receiver is assumed to be coherent; (2) The choice of n even is made on the strength of physical evidence that moments $E\{G^{(n)}\}$ with n odd—unequal number of conjugated and unconjugated terms in (2.12f)—decay relatively fast with increasing range z; (3) At the receiver site, further processing of the nth-order moments given in (2.11), such as averaging over the coordinates $\mathbf{x}_p$, may be necessary.

C. General remarks

It is clear from the foregoing discussion that the study of pulse propagation in a random medium requires knowledge of the nth-order coherence functions $E\{G^{(n)}(\mathbf{X},\mathbf{X}',z,\boldsymbol{\omega};\alpha)\}$ at different frequencies and different transverse (with respect to z) coordinates.

Transport equations for the moments $E\{G^{(n)}(\mathbf{X},\mathbf{X}',z,\boldsymbol{\omega};\alpha)\}$ can be obtained using the Markovian random process approximation, even for dispersive media. In the absence of a deterministic profile ($\epsilon_2 = 0$), Lee,[5] for example, has derived such a set of transport equations for the special case of a randomly perturbed cold plasma.

The solution of these transport equations is difficult, in general. Nevertheless, as already mentioned in the introduction, some progress has been made in connection with the 2-frequency mutual coherence function $\Gamma(\mathbf{x}_2,\mathbf{x}_1,z,\omega_2,\omega_1) \equiv E\{\Psi^*(\mathbf{x}_2,z,\omega_2;\alpha)\Psi(\mathbf{x}_1,z,\omega_1;\alpha)\}$, albeit for planar source distributions, viz., $\Gamma(\mathbf{x}_2,\mathbf{x}_1,0,\omega_2,\omega_1) = \Gamma_0(\omega_2,\omega_1)$. This assumption gives rise to a great deal of simplification since, upon resorting to the center of mass and difference coordinates $\mathbf{x} = (\mathbf{x}_1 + \mathbf{x}_2)/2$ and $\mathbf{y} = \mathbf{x}_2 - \mathbf{x}_1$, one recognizes that Γ depends on $\mathbf{y}$ but not on $\mathbf{x}$, and the corresponding equation for $\Gamma(\mathbf{y},z,\omega_2,\omega_1)$ can be solved—analytically in the case of a simplified (quadratic) Kolmogorov spectrum (cf. Ref. 12), or numerically under more relaxed assumptions regarding the spectrum of the random inhomogeneities (cf. Refs. 7, 9, and 10).

The procedure outlined above in connection with the computation of $\Gamma(\mathbf{x}_2,\mathbf{x}_1,z,\omega_2,\omega_1)$ in the case of spatially planar source distributions cannot be extended easily to higher-order multifrequency moments. The complexity of the required generalizations is compounded in physical situations where nonplanar source distributions must be considered.

The aforementioned difficulties can be alleviated to some extent by using the functional path integral method. This formalism can be briefly outlined as follows: The propagator $G(\mathbf{x},\mathbf{x}',z,\omega;\alpha)$ and, in turn, $G^{(n)}(\mathbf{X},\mathbf{X}',z,\boldsymbol{\omega};\alpha)$, are first expressed as (continuous) functional path integrals. Upon ensemble averaging $E\{G^{(n)}(\mathbf{X},\mathbf{X}',z,\boldsymbol{\omega};\alpha)\}$, the required quantity, is also expressed as a path integral. The latter is finally evaluated—exactly in special cases, or asymptotically in general. Since this evaluation is based on a *global* expression for $E\{G^{(n)}(\mathbf{X},\mathbf{X}',z,\boldsymbol{\omega};\alpha)\}$, the underlying analysis (exact or asymptotic) is invariably simpler than the one required for the direct solution of the *local* transport equation for $E\{G^{(n)}(\mathbf{X},\mathbf{X}',z,\boldsymbol{\omega};\alpha)\}$.

3. *N*th-ORDER MULTIFREQUENCY COHERENCE FUNCTIONS: A FUNCTIONAL PATH INTEGRAL APPROACH

A. The Feynman path integral

The solution of the stochastic complex parabolic Eq. (2.9) for the propagator $G(\mathbf{x},\mathbf{x}',z,\omega;\alpha)$ can be expressed as a continuous functional path integral (cf. Refs. 23 and 24; see, also Ref. 25). Specifically,

$$G(\mathbf{x},\mathbf{x}',z,\omega;\alpha) = \int d[\mathbf{x}(\zeta)]\exp\left\{ik\int_0^z d\zeta\,\tfrac{1}{2}\left[\dot{\mathbf{x}}^2(\zeta) + \epsilon_2[\mathbf{x}(\zeta),\zeta,\omega] + \epsilon_1[\mathbf{x}(\zeta),\zeta,\omega;\alpha]\right]\right\}, \tag{3.1}$$

where $d[\mathbf{x}(\zeta)]$ is the usual Feynman path differential measure, and the integration is over "all" paths $\mathbf{x}(\zeta)$ subject to the boundary conditions $\mathbf{x}(0) = \mathbf{x}', \mathbf{x}(z) = \mathbf{x}$. The dot over $\mathbf{x}(\zeta)$ designates a derivative with respect to the argument ζ. As mentioned earlier, k is an abbreviation for $k(\omega)$. Finally $\dot{\mathbf{x}}^2(\zeta)$ denotes the square norm of the vector-valued function $\dot{\mathbf{x}}(\zeta)$.

Equation (3.1) can then be used as a basis for constructing a path integral represenation for the nth-order quantity $G^{(n)}(\mathbf{X},\mathbf{X}',z,\omega;\alpha)$ [cf. Eq. (2.12f)]. Specifically,

$$G^{(n)}(\mathbf{X},\mathbf{X}',z,\omega;\alpha) = \int d[\mathbf{X}(\zeta)]\exp\Big(\frac{i}{2}\sum_{p=1}^{n}\xi_p k_p\int_0^z d\zeta\Big[\dot{\mathbf{x}}_p^2(\zeta) + \epsilon_2[\mathbf{x}_p(\zeta),\zeta,\omega_p] + \epsilon_1[\mathbf{x}_p(\zeta),\zeta,\omega_p;\alpha]\Big]\Big), \quad (3.2)$$

where $d[\mathbf{X}(\zeta)] = d[\mathbf{x}_1(\zeta)]d[\mathbf{x}_2(\zeta)]\cdots d[\mathbf{x}_n(\zeta)]$; $\xi_p = 1$, p odd, $\xi_p = -1$, p even; $k_p = k(\omega_p)$; and the integration is over "all" paths $\mathbf{x}_p(\zeta)$, $p = 1,2,\ldots,n$, subject to the boundary conditions $\mathbf{x}_p(0) = \mathbf{x}_p', \mathbf{x}_p(z) = \mathbf{x}_p$.

B. Statistical analysis

Ensemble averaging (3.2) over the statistical realizations α results in the expression

$$E\{G^{(n)}(\mathbf{X},\mathbf{X}',z,\omega;\alpha)\} = \int d[\mathbf{X}(\zeta)]\exp\Big(\frac{i}{2}\sum_{p=1}^{n}\xi_p k_p\int_0^z d\zeta\Big[\dot{x}_p^2(\zeta) + \epsilon_2[\mathbf{x}_p(\zeta),\zeta,\omega_p]\Big]\Big)E\Big\{\exp\Big(\frac{i}{2}\sum_{p=1}^{n}\xi_p k_p\int_0^z d\zeta \times\epsilon_1[\mathbf{x}_p(\zeta),\zeta,\omega_p;\alpha]\Big)\Big\}. \quad (3.3)$$

To proceed further, we need to specify the structure of $\epsilon_1[\mathbf{x}_p(\zeta),\zeta,\omega_p;\alpha]$. We assume, first, that the dependence of the function ϵ_1 on ω_p (arising from the dispersive properties of the medium) enters multiplicatively,[26] viz.,

$$\epsilon_1[\mathbf{x}_p(\zeta),\zeta,\omega_p;\alpha] = \nu(\omega_p)\mu[\mathbf{x}_p(\zeta),\zeta;\alpha]. \quad (3.4)$$

All the information about the random fluctuations in the medium is now contained in the quantity $\mu[\mathbf{x}_p(\zeta),\zeta;\alpha]$. If the latter is assumed to be a Gaussian random process, the statistical averaging appearing in (3.3) can be carried out explicitly, with the result

$$I_1 \equiv E\Big\{\exp\Big(\frac{i}{2}\sum_{p=1}^{n}\xi_p k_p\int_0^z d\zeta\,\epsilon_1[\mathbf{x}_p(\zeta),\zeta,\omega_p;\alpha]\Big)\Big\} = \exp\Big(-\frac{1}{8}\sum_{p=1}^{n}\sum_{q=1}^{n}\xi_p\xi_q k_p k_q \nu_p\nu_q\int_0^z d\zeta \times\int_0^z d\zeta'\gamma[\mathbf{x}_p(\zeta),\mathbf{x}_q(\zeta'),\zeta,\zeta']\Big), \quad (3.5)$$

where $\nu_p = \nu(\omega_p)$ and γ is the correlation function of the random process μ, viz.,

$$\gamma[\mathbf{x}_p(\zeta),\mathbf{x}_q(\zeta'),\zeta,\zeta'] = E\{\mu[\mathbf{x}_p(\zeta),\zeta;\alpha]\mu[\mathbf{x}_q(\zeta'),\zeta';\alpha]\}. \quad (3.6)$$

It should be noted that the Gaussian assumption invoked in deriving (3.5) can be relaxed somewhat by using the theory of cumulants (cf. Ref. 27; see, also, Ref. 20).

We resort, next, to the usual Markovian approximation, i.e., we assume that the process μ is δ correlated along the longitudinal direction of propagation. We have, then, in the place of (3.6)

$$\gamma[\mathbf{x}_p(\zeta),\mathbf{x}_q(\zeta'),\zeta,\zeta'] = A[\mathbf{x}_p(\zeta),\mathbf{x}_q(\zeta')]\delta(\zeta-\zeta'). \quad (3.7)$$

With this simplification, the integration over ζ' in (3.5) can be performed trivially. The resulting expression for I_1 is given as follows:

$$I_1 = \exp\Big(-\frac{1}{8}\sum_{p=1}^{n}\sum_{q=1}^{n}\xi_p\xi_q k_p k_q\nu_p\nu_q \times\int_0^z d\zeta A[\mathbf{x}_p(\zeta),\mathbf{x}_p(\zeta)]\Big). \quad (3.8)$$

In many cases of physical interest, the "transverse" correlation $A[\mathbf{x}_p(\zeta),\mathbf{x}_q(\zeta)]$ is homogeneous and isotropic,[28] viz., $A[\mathbf{x}_p(\zeta),\mathbf{x}_q(\zeta)] = A[|\mathbf{x}_p(\zeta)-\mathbf{x}_q(\zeta)|]$, and of a power-law type, viz.,

$$A[|\mathbf{x}_p(\zeta)-\mathbf{x}_q(\zeta)|] = A(0)\Big\{1-\frac{1}{2}\Big[\frac{1}{L_0}|\mathbf{x}_p(\zeta)-\mathbf{x}_q(\zeta)|\Big]^\beta\Big\}. \quad (3.9)$$

Here, L_0 is a characteristic length, and the parameter β is usually within the range $1<\beta<4$ (cf. Refs. 9 and 20). For optical propagation through a turbulent medium such as the atmosphere, one has $\beta\leqslant 2$. This range includes the Gaussian spectrum and the Kolmogorov spectrum.

Introducing (3.9) into (3.8) and, in turn, the resulting expression for I_1 into (3.3), we obtain

$$E\{G^{(n)}(\mathbf{X},\mathbf{X}',z,\omega;\alpha)\} = \exp\Big[-\tfrac{1}{8}A(0)\Big(\sum_{p=1}^{n}\sum_{q=1}^{n}\xi_p\xi_q k_p k_q\nu_p\nu_q\Big)z\Big] \times\int d[\mathbf{X}(\zeta)]\exp\Big\{\frac{i}{2}\sum_{p=1}^{n}\xi_p k_p\int_0^z d\zeta\Big(\dot{\mathbf{x}}_p^2(\zeta) + \epsilon_2[\mathbf{x}_p(\zeta),\zeta,\omega_p] - \frac{i}{16}A(0)\sum_{q=1}^{n}\xi_q k_q\nu_p\nu_q \times\Big[\frac{1}{L_0}|\mathbf{x}_p(\zeta)-\mathbf{x}_q(\zeta)|\Big]^\beta\Big)\Big\}. \quad (3.10)$$

This expression is the starting point for all asymptotic evaluations of nth-order multifrequency coherence functions in the presence of dispersion, a deterministic inhomogeneous profile, and for a wide class of fluctuation spectra (cf. Ref. 20).

The versatility of the functional path integration technique will be illustrated below for a simple (but physically nontrivial) setting: a nondispersive, deterministically flat medium, characterized by a simplified (quadratic) Kolmogorov spectrum. In this case, the path integral in (3.10) can be carried out explicitly, yielding a series of presently unavailable results.

C. Specialization to a nondispersive, deterministically homogeneous medium, with a simplified (quadratic) Kolmogorov spectrum

In (3.10), let $\epsilon_2 = 0$, $\nu_p = 1, p = 1,2,...,n$, and $\beta = 2$. The corresponding path integral representation assumes the simpler form

$$E\{G^{(n)}(\mathbf{X},\mathbf{X}',z,\omega;\alpha)\} = \exp\left[-\frac{1}{8}A(0)\left(\sum_{p=1}^{n}\sum_{q=1}^{n}\xi_p\xi_q k_p k_q\right)z\right] \times \int d[\mathbf{X}(\zeta)]\exp(iS), \tag{3.11a}$$

$$S = \int_0^z d\zeta L[\dot{\mathbf{x}}_p(\zeta),\mathbf{x}_p(\zeta)], \tag{3.11b}$$

$$L[\dot{\mathbf{x}}_p(\zeta),x_p(\zeta)] = \frac{1}{2}\sum_{p=1}^{n}\xi_p k_p\left\{\dot{\mathbf{x}}_p^2(\zeta) - \frac{i}{4}D \times \sum_{q=1}^{n}\xi_q k_q[\mathbf{x}_p(\zeta) - \mathbf{x}_q(\zeta)]^2\right\}, \tag{3.11c}$$

$$D = A(0)/2L_0^2. \tag{3.11d}$$

Borrowing terminology from quantum mechanics (in connection with which the functional path integration method was originally developed by Feynman), we shall refer to L and S in (3.11) as the Lagrangian and action, respectively, and to $\mathbf{x}_p$ and $\dot{\mathbf{x}}_p$ as the coordinates and velocities (or momenta with respect to a "mass" normalized to unity), respectively.

For the problem under consideration here, S [cf. Eq. (3.11b)] is a quadratic action functional. It is well known in this case (cf. Ref. 24) that the path integral in (3.11a) becomes

$$I_2 \equiv \int d[\mathbf{X}(\zeta)]\exp(iS) = N(z)\exp(S_c). \tag{3.12}$$

Here, S_c is the "classical" action, i.e., the action S evaluated along the "classical" paths $\mathbf{x}_{cp}(\zeta)$ satisfying the Euler–Lagrange equations

$$\frac{d}{d\zeta}[\partial L/\partial\dot{\mathbf{x}}_{cp}(\zeta)] - \partial L/\partial\mathbf{x}_{cp}(\zeta) = 0, \tag{3.13a}$$

with the boundary conditions

$$\mathbf{x}_{cp}(0) = \mathbf{x}'_p, \mathbf{x}_{cp}(z) = \mathbf{x}_p, \quad p = 1,2,...,n. \tag{3.13b}$$

In general, S_c is a function of $\mathbf{x}'_p$, $\mathbf{x}_p$, and z. The normalization quantity N in (3.12), however, depends only on z, and is related to the classical action as follows:

$$N(z) = \left(\det(i/2\pi)\left[\frac{\partial^2 S_c}{\partial\mathbf{x}_p\partial\mathbf{x}'_q}\right]\right)^{1/2}, \quad p,q,1,2,...,n. \tag{3.14}$$

The $2n\times 2n$ matrtix within the square brackets is referred to in the literature as the Van Vleck–Morette matrix (or "Hessian of the action").

Our attention will be directed next to the evaluation of the classical action S_c and the normalization factor $N(z)$ required in (3.12).

D. Evaluation of S_c

The Euler–Lagrange equations (3.13) corresponding to the Lagrangian $L[\dot{\mathbf{x}}_{cp}(\zeta),\mathbf{x}_{cp}(\zeta)]$ in (3.11c) yield the following equations for the classical paths $x_{cp}(\zeta)$:

$$\ddot{\mathbf{x}}_{cp}(\zeta) + g^2\sum_{q=1}^{n}\xi_q k_q[\mathbf{x}_{cp}(\zeta) - \mathbf{x}_{cq}(\zeta)] = 0, \tag{3.15a}$$

$$\mathbf{x}_{cp}(0) = \mathbf{x}'_p, \mathbf{x}_{cp}(z) = \mathbf{x}_p, \tag{3.15b}$$

$$g^2 = \frac{i}{2}D\sum_{p=1}^{n}\xi_p k_p. \tag{3.15c}$$

We consider next the quadratic action functional S [cf. Eq. (3.11b)]. If the momentum-dependent term is integrated by parts, and the resulting expression for S is evaluated along the classical paths (3.15), we obtain

$$S_c = \frac{1}{2}\sum_{p=1}^{n}\xi_p k_p \mathbf{x}_{cp}(\zeta)\cdot[\dot{\mathbf{x}}_{cp}(\zeta)]_0^z. \tag{3.16}$$

The classical action given in the last equation can be manipulated into a form which is more suitable for further analysis. Specifically,

$$S_c = \tfrac{1}{2}\left(\sum_{p=1}^{n}\xi_p k_p\right)^{-1}\left[\left(\mathbf{v}(\zeta)\cdot\dot{\mathbf{v}}(\zeta) + \frac{1}{2}\sum_{p=1}^{n} \times \sum_{q=1}^{n}\xi_p\xi_q k_p k_q \mathbf{u}_{pq}(\zeta)\cdot\dot{\mathbf{u}}_{pq}(\zeta)\right)\right]_0^z, \tag{3.17}$$

where

$$\mathbf{v}(\zeta) = \sum_{p=1}^{n}\xi_p k_p \mathbf{x}_{cp}(\zeta), \tag{3.18}$$

and

$$\mathbf{u}_{pq}(\zeta) = \mathbf{x}_{cp}(\zeta) - \mathbf{x}_{cq}(\zeta). \tag{3.19}$$

We shall determine next all the quantities required in (3.17); that is, $\mathbf{v}(\zeta)$, $\mathbf{u}_{pq}(\zeta)$, and their derivatives. Toward this goal, we note that a relatively simple manipulation of Eqs. (3.15) for the classical paths yields[29]

$$\ddot{\mathbf{u}}_{pq}(\zeta) + g^2\mathbf{u}_{pq}(\zeta) = 0, \quad p\neq q, \tag{3.20a}$$

$$\mathbf{u}_{pq}(0) = \mathbf{u}'_{pq}, \quad \mathbf{u}_{pq}(z) = \mathbf{u}_{pq}, \tag{3.20b}$$

and

$$\ddot{\mathbf{v}}(\zeta) = 0, \tag{3.21a}$$

$$\mathbf{v}(0) = \mathbf{v}', \mathbf{v}(z) = \mathbf{v}. \tag{3.21b}$$

The solutions to the two-point boundary value problems

(3.20) and (3.21) can be found in a straightforward manner—say, by a Laplace transformation. They are given as follows:

$$\mathbf{u}_{pq}(\zeta) = [\mathbf{u}_{pq}\operatorname{sing}\zeta - \mathbf{u}'_{pq}\operatorname{sing}(\zeta - z)]/\operatorname{sing}z, \tag{3.22}$$

$$\mathbf{v}(\zeta) = (\mathbf{v} - \mathbf{v}')(\zeta/z) + \mathbf{v}'. \tag{3.23}$$

The desired expression for the classical action can be found by using the solutions (3.22) and (3.23) in conjunction with (3.17). One has, finally,

$$S_c = \frac{1}{2}\Big(\sum_{p=1}^{n} \xi_p k_p\Big)^{-1}\Big[\frac{1}{z}\Big(\sum_{p=1}^{n} \xi_p k_p(\mathbf{x}_p - \mathbf{x}'_p)\Big)^2 - \Big(\frac{g}{\operatorname{sing}z}\Big)\sum_{p=1}^{n}\sum_{q=1}^{n} \xi_p\xi_q k_p k_q(\mathbf{x}_p - \mathbf{x}_q)\cdot(\mathbf{x}'_p - \mathbf{x}'_q) + \tfrac{1}{2}g\operatorname{cotg}z \sum_{p=1}^{n}\sum_{q=1}^{n} \xi_p\xi_q k_p k_q[(\mathbf{x}_p - \mathbf{x}_q)^2 + (\mathbf{x}'_p - \mathbf{x}'_q)^2]\Big\}. \tag{3.24}$$

E. Evaluation of $N(z)$

The normalization factor $N(z)$ defined in (3.14) is rewritten as

$$N(z) = (i/2\pi)^n\Big(\det\Big[\frac{\partial^2 S_c}{\partial\mathbf{X}\partial\mathbf{X}'}\Big]\Big)^{1/2}, \tag{3.25}$$

where $\mathbf{X}$ and $\mathbf{X}'$ are the $2n$ vectors given in (2.12b) and (2.12c), respectively. The $2n\times 2n$ Van Vleck–Morette matrix appearing within the square brackets in (3.25) can be determined easily using the expression for the classical action which was derived in the previous subsection [cf. Eq. (3.24)]. Unfortunately, a direct evaluation of the determinant of this matrix is rather difficult for large n. In the following, we shall pursue an alternative procedure.

Consider the linear transformations,

$$T\mathbf{X} = \mathrm{R},\quad T\mathbf{X}' = \mathbf{R}', \tag{3.26}$$

where T is a $2n\times 2n$ matrix and $\mathbf{R}$, $\mathbf{R}'$ are $2n$-vectors. Since the Van Vleck–Morette matrix is of the Hessian form, we have[30]

$$\Big[\frac{\partial^2 S_c}{\partial\mathbf{X}\partial\mathbf{X}'}\Big] = \overline{T}\Big[\frac{\partial^2 S_c}{\partial\mathbf{R}\partial\mathbf{R}'}\Big]T, \tag{3.27}$$

where the overbar denotes the transpose of the matrix T. It follows, then, from (3.27) that

$$\det\Big[\frac{\partial^2 S_c}{\partial\mathbf{X}\partial\mathbf{X}'}\Big] = (\det T)^2\det\Big[\frac{\partial^2 S_c}{\partial\mathbf{R}\partial\mathbf{R}'}\Big]. \tag{3.28}$$

A convenient change of variables is the following:

$$\mathbf{x}_1 - \mathbf{x}_r = \mathbf{u}_{1r},\quad r = 2,3,\ldots,n, \tag{3.29a}$$

$$\sum_{p=1}^{n} \xi_p k_p \mathbf{x}_p = \mathbf{v}, \tag{3.29b}$$

and similar relations for the primed coordinates. With this change of variables, the linear transformation $T\mathbf{X} = \mathbf{R}$ [cf. Eq. (3.26)] has the realization

$$\begin{bmatrix} 1 & -1 & 0 & 0 & \cdots & 0 & 0 \\ 1 & 0 & -1 & 0 & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 1 & 0 & 0 & 0 & \cdots & 0 & -1 \\ k_1 & -k_2 & k_3 & -k_4 & \cdots & k_{n-1} & -k_n \end{bmatrix} \times \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \\ \vdots \\ \mathbf{x}_{n-1} \\ \mathbf{x}_n \end{bmatrix} = \begin{bmatrix} \mathbf{u}_{12} \\ \mathbf{u}_{13} \\ \vdots \\ \mathbf{u}_{1n} \\ \mathbf{v} \end{bmatrix}. \tag{3.30}$$

A similar realization holds for the linear transformation $T\mathbf{X}' = \mathbf{R}'$. Because of the partitioning in the column matrices for $\mathbf{X}$ and $\mathbf{R}$, each entry of the $2n\times 2n$ matrix T in (3.30) must be understood as a 2×2 diagonal matrix.

The determinant of the matrix T can be easily evaluated from the realization of T shown in (3.30). Specifically,

$$\det T = \Big(\sum_{p=1}^{n} \xi_p k_p\Big)^2. \tag{3.31}$$

Substituting this result into (3.28), we obtain

$$\det\Big[\frac{\partial^2 S_c}{\partial\mathbf{X}\partial\mathbf{X}'}\Big] = \Big(\sum_{p=1}^{n} \xi_p k_p\Big)^4\det\Big[\frac{\partial^2 S_c}{\partial\mathbf{U}\partial\mathbf{U}'}\Big]\det\Big[\frac{\partial^2 S_c}{\partial\mathbf{v}\partial\mathbf{v}'}\Big], \tag{3.32}$$

where we have resorted to the obvious partitioning

$$\Big[\frac{\partial^2 S_c}{\partial\mathbf{R}\partial\mathbf{R}'}\Big] = \left[\begin{array}{c|c} \Big[\dfrac{\partial^2 S_c}{\partial\mathbf{U}\partial\mathbf{U}'}\Big] & 0 \\ \hline 0 & \Big[\dfrac{\partial^2 S_c}{\partial\mathbf{v}\partial\mathbf{v}'}\Big] \end{array}\right], \tag{3.33}$$

with $\mathbf{U} = (\mathbf{u}_{12},\mathbf{u}_{13},\ldots,\mathbf{u}_{1n})$ and $\mathbf{U}' = (\mathbf{u}'_{12},\mathbf{u}'_{13},\ldots,\mathbf{u}'_{1n})$.

To proceed further, we shall have to express the classical action given in (3.24) in terms of the new coordinates $\mathbf{U}$, $\mathbf{U}'$, $\mathbf{v}$, and $\mathbf{v}'$. Omitting intermediate steps, we present below the final result:

$$S_c = \frac{1}{2}\Big(\sum_{p=1}^{n} \xi_p k_p\Big)^{-1}\Big[\frac{1}{z}(\mathbf{v} - \mathbf{v}')^2 - \Big(\frac{g}{\operatorname{sing}z}\Big)\times\Big(2\xi_1 k_1\sum_{r=2}^{n} \xi_r k_r \mathbf{u}_{1r}\cdot\mathbf{u}'_{1r} + \sum_{r=2}^{n}\sum_{s=2}^{n}\times\xi_r\xi_s k_r k_s(\mathbf{u}_{1r} - \mathbf{u}_{1s})\cdot(\mathbf{u}'_{1r} - \mathbf{u}'_{1s})\Big) + \tfrac{1}{2}g\operatorname{cotg}z\times\Big(2\xi_1 k_1\sum_{r=2}^{n} \xi_r k_r(\mathbf{u}^2_{1r} + \mathbf{u}'^2_{1r}) + \sum_{r=2}^{n}\sum_{s=2}^{n}\times\xi_r\xi_s k_r k_s[(\mathbf{u}_{1r} - \mathbf{u}_{1s})^2 + (\mathbf{u}'_{1r} - \mathbf{u}'_{1s})^2]\Big)\Big]. \tag{3.34}$$

The individual matrices needed in (3.33) can now be determined by straightforward differentiation:

$$\left[\frac{\partial^2 S_c}{\partial \mathbf{U}\partial \mathbf{U}'}\right] = \left[\frac{\partial^2 S_c}{\partial u^i_{1r}\partial u^{i'}_{1r'}}\right]$$

$$= -\left(\frac{g}{\sin gz}\right)\left(\sum_{p=1}^{n}\xi_p k_p\right)^{-1}\Big[\xi_1 k_1 \xi_r k_r \delta_{rr'} + \xi_r k_r \times\left(\sum_{s=2}^{n}\xi_s k_s\right)\delta_{rr'} - \xi_r k_r \xi_{r'} k_{r'}\Big]\delta_{ii'}, \qquad (3.35)$$

for all $r, r' = 2,3,\ldots,n$; $i,i' = 1,2$, and

$$\left[\frac{\partial^2 S_c}{\partial \mathbf{v}\partial \mathbf{v}'}\right] = \left[\frac{\partial^2 S_c}{\partial v_i \partial v'_{i'}}\right] = -\left(z\sum_{p=1}^{n}\xi_p k_p\right)^{-1}\delta_{ii'}, \qquad (3.36)$$

for $i, i' = 1,2$.

The determinants of the matrices given in (3.35) and (3.36) can be computed without difficulty. Omitting again intermediate steps, we present below the final results:

$$\det\left[\frac{\partial^2 S_c}{\partial \mathbf{U}\partial \mathbf{U}'}\right] = \left(\frac{g}{\sin gz}\right)^{2(n-1)}\left(\prod_{p=1}^{n}\xi_p k_p \Big/ \sum_{p=1}^{n}\xi_p k_p\right)^2, \qquad (3.37)$$

$$\det\left[\frac{\partial^2 S_c}{\partial \mathbf{v}\partial \mathbf{v}'}\right] = \left(z\sum_{p=1}^{n}\xi_p k_p\right)^{-2}. \qquad (3.38)$$

Introducing (3.37) and (3.38) into (3.28), and the resulting expression into (3.25), we find that the desired normalization factor $N(z)$ is given by

$$N(z) = (i/2)^n z^{-1}(g/\sin gz)^{n-1}\left(\sum_{p=1}^{n}\xi_p k_p\right). \qquad (3.39)$$

The solution to our problem is now complete. Using (3.39), (3.24), and (3.12), together with (3.11), the final expression for the nth-order multifrequency coherence function is given as follows:

$$E\{G^{(n)}(\mathbf{X},\mathbf{X}',z,\boldsymbol{\omega};\alpha)\}$$
$$= (i/2\pi)^n z^{-1}\left(\sum_{p=1}^{n}\xi_p k_p\right)(g/\sin gz)^{n-1}$$
$$\times\exp\left[-\tfrac{1}{8}A(0)\left(\sum_{p=1}^{n}\sum_{q=1}^{n}\xi_p\xi_q k_p k_q\right)z\right]$$
$$\times\exp\left\{\frac{i}{2}\left(\sum_{p=1}^{n}\xi_p k_p\right)^{-1}\left[\frac{1}{z}\left(\sum_{p=1}^{n}\xi_p k_p(\mathbf{x}_p - \mathbf{x}'_p)\right)^2\right.\right.$$
$$-\left(\frac{g}{\sin gz}\right)\sum_{p=1}^{n}\sum_{q=1}^{n}\xi_p\xi_q k_p k_q(\mathbf{x}_p - \mathbf{x}'_q)\cdot(\mathbf{x}'_p - \mathbf{x}'_q)$$
$$+\tfrac{1}{2}g\cot gz\sum_{p=1}^{n}\sum_{q=1}^{n}\xi_p\xi_q k_p k_q[(\mathbf{x}_p - \mathbf{x}_q)^2$$
$$\left.\left. + (\mathbf{x}'_p - \mathbf{x}'_q)^2]\right]\right\}. \qquad (3.40)$$

It should be noted that this result is exact under the restrictions specified earlier in this section. When used in conjunction with (2.10), nth order pulse statistics can be studied. A special case ($n = 2$) of (3.40) will be examined in Sec. 4, and comparisons will be made with previously reported results.

4. SPECIAL CASES

We consider (3.40) in the special case where $n = 2$. We resort, also, to the following center of mass and difference variables: $\mathbf{x} = (\mathbf{x}_1 + \mathbf{x}_2)/2, \mathbf{y} = \mathbf{x}_2 - \mathbf{x}_1$; $\mathbf{x}' = (\mathbf{x}'_1 + \mathbf{x}'_2)/2$, $\mathbf{y}' = \mathbf{x}'_2 - \mathbf{x}'_1$; $k_s = (k_1 + k_2)/2, k_d = k_2 - k_1$; $\omega_s = (\omega_1 + \omega_2)/2, \omega_d = \omega_2 - \omega_1$. (For a nondispersive medium, we have, in general, the relationship $k = \omega/v$, where v is a characteristic reference velocity; hence $k_s = \omega_s/v$ and $k_d = \omega_d/v$ for the sum and difference quantities.) With these specifications, (3.40) simplifies to

$$E\{G^{(2)}(\mathbf{x},\mathbf{y},\mathbf{x}',\mathbf{y}',z,\omega_s,\omega_d;\alpha)\}$$
$$= (2\pi)^{-2}\lambda z^{-1}(g/\sin gz)\exp[-\tfrac{1}{8}A(0)k_d^2 z]$$
$$\times\exp\left\{-\frac{i}{2}\frac{1}{k_d}\left[\frac{1}{z}[k_d(\mathbf{x}' - \mathbf{x}) + k_s(\mathbf{y}' - \mathbf{y})]^2\right.\right.$$
$$\left.\left. + 2\lambda\left(\frac{g}{\sin gz}\right)\mathbf{y}\cdot\mathbf{y}' - \lambda g\cot gz(\mathbf{y}^2 + \mathbf{y}'^2)\right]\right\}, \qquad (4.1)$$

where $\lambda = k_1 k_2 = k_s^2 - (k_d^2/4)$ and $g^2 = -(iDk_d)/2$ [cf. Eq. (3.15c)].

For a planar source distribution, the boundary condition $E_0(\mathbf{x},\omega;\alpha)$ [cf. Eq. (2.7d)] has the form

$$E_0(\mathbf{x},\omega;\alpha) = F_s(\omega), \qquad (4.2)$$

if we ignore source incoherencies. This initial distribution is introduced next into (2.11)—the latter must be specialized to $n = 2$ and (4.1) must be taken into consideration. In this case, the spatial integrations in (2.11) can be carried out explicitly, with the result

$$E\{e_r(\mathbf{x} + \tfrac{1}{2}\mathbf{y}, t + \tfrac{1}{2}\tau;\alpha)e_r(\mathbf{x} - \tfrac{1}{2}\mathbf{y}, t - \tfrac{1}{2}\tau;\alpha)\}$$
$$= (2\pi)^{-2}\int_{-\infty}^{\infty}d\omega_s\int_{-\infty}^{\infty}d\omega_d F_r^{(2)}(\omega_s,\omega_d)F_s^{(2)}(\omega_s,\omega_d)$$
$$\times\sec gz\,\exp\left(-\frac{1}{8v^2}A(0)\omega_d^2 z\right)\exp\left(-i\frac{\lambda}{2\omega_d}(g\tan gz)\mathbf{y}^2\right)$$
$$\times\exp[i\omega_s\tau + i\omega_d(t - z/v)], \qquad (4.3)$$

where $t = (t_1 + t_2)/2$, $\tau = t_2 - t_1$; $F_{r,s}^{(2)}(\omega_s,\omega_d) = F^*_{r,s}[\omega_s + (\omega_d/2)]F_{r,s}[\omega_s - (\omega_d/2)]$; $\lambda = [\omega_s^2 - (\omega_d^2/4)]/v^2$; and $g^2 = -(iD\omega_d)/2v$. As expected from physical considerations, the second-order moment in (4.3) is independent of the center of mass coordinate $\mathbf{x}$.

In the special case where $\mathbf{x}_2 = \mathbf{x}_1$ ($\mathbf{y} = 0$) and $t_2 = t_1$ ($\tau = 0$), Eq. (4.3) simplifies even further:

$$E\{e_r^2(z,t;\alpha)\}$$
$$= (2\pi)^{-2}\int_{-\infty}^{\infty}d\omega_s\int_{-\infty}^{\infty}d\omega_d F_r^{(2)}(\omega_s,\omega_d)F_s^{(2)}(\omega_s,\omega_d)$$

$$\times \sec gz \exp\left(-\frac{1}{8v^2}A(0)\omega_d^2 z\right)\exp\left[i\omega_d\left(t-\frac{z}{v}\right)\right]. \quad (4.4)$$

This is essentially the expression for the average pulse intensity reported by Sreenivasiah *et al.* (cf. Ref. 12). For a broadband receiver, i.e., $F_r^{(2)}(\omega)\simeq 1$, and an impulsive source intensity, the integrations over the sum and difference frequencies in (4.4) can be carried out. The resulting expression for the mean pulse intensity [cf., e.g., Eqs. (19)–(22) of Ref. 12] exhibits a smearing effect (broadening) caused by the random inhomogeneities in the medium.

Consider next a boundary condition $E_0(\mathbf{x},\omega;\alpha)$ of the form

$$E_0(\mathbf{x},\omega;\alpha) = F_s(\omega)\exp(-\mathbf{x}^2/2\sigma_0^2). \quad (4.5)$$

This initial distribution may represent, for example, the field of a coherent, pulsed, collimated, Gaussian laser beam having an aperture radius equal to σ_0. It turns out in this case that the spatial integrations in (2.11) can be performed exactly. As a specific illustration of such an operation, we present below the ensemble averaged pulsed intensity evaluated on the beam axis:

$$E\{e_r^2(z,t;\alpha)\} = (2\pi)^{-2}\int_{-\infty}^{\infty} d\omega_s \int_{-\infty}^{\infty} d\omega_d F_r^{(2)}(\omega_s,\omega_d)F_s^{(2)}(\omega_s,\omega_d)$$
$$\times H(\omega_s,\omega_d)\exp\left[-\frac{1}{8v^2}A(0)\omega_d^2 z\right]\exp\left[i\omega_d\left(t-\frac{z}{v}\right)\right]; \quad (4.6a)$$

$$H(\omega_s,\omega_d) = (\lambda/4z)(g/\sin gz)[\sigma_0^{-2} + i(\omega_d/2vz)]^{-1}[(2\sigma_0)^{-2} + (\omega_s/2vz)^2][\sigma_0^{-2} + i(\omega_d/2vz)]^{-1} + i(\omega_s^2/2vz\omega_d) - i(\lambda vg\cot gz/\omega_d)^{-1}. \quad (4.6b)$$

In contrast to the plane wave case [cf. Eq. (4.4)], the integrations over the sum and difference frequencies in (4.6) must be evaluated asymptotically (e.g., by the method of steepest descent), even for a broadband receiver and an impulsive source intensity.

5. CONCLUDING REMARKS

The two-frequency mutual coherence function [cf. Eq. (3.40) in the special case $n = 2$, or Eq. (4.1)] has also been derived by the authors independently by solving the transport equation for $E\{G^{(2)}\}$. The effort involved is considerable, and is expected to be even more substantial if nth-order multifrequency coherence functions are to be obtained by solving directly the associated local transport equations. On the other hand, it has been demonstrated in this paper that the path integration technique yields these results in a simple and straightforward manner.

The simplified (quadratic) Kolmogorov spectrum implicit in the derivation of (3.40) is valid in many physical situations (e.g., optical propagation through a turbulent atmosphere) provided propagation distances are large.[31] Under this assumption, the main results in this paper incorporated in (3.40) are exact, that is, *all* path contributions have been accounted for. However, one can proceed along the lines suggested by Dashen (cf. Ref. 20) in order to compute $E\{G^{(n)}(\mathbf{X},\mathbf{X}',z,\omega;\alpha)\}$ asymptotically with respect to a small parameter "a" whose order of magnitude is roughly $a \sim [6(L_0/z)^{3/2}/E\{\mu^2\}]^{1/2}$ in terms of the scale size, the range, and the rms fluctuations. Two regions are delineated: the fully and partially saturated regimes. To $O(a)$, the statistics of $G(\mathbf{x},\mathbf{x}',z,\omega;\alpha)$ in the fully saturated region is Gaussian. The situation is a little more complicated in the partially saturated regime. In both cases, however, the ensuing expressions for $E\{G^{(n)}(\mathbf{X},\mathbf{X}',z,\omega;\alpha)\}$ are expressible in terms of two-frequency mutual coherence functions. This simplification would facilitate considerably the computation of nth-order pulse statistics [cf. Eq. (2.11)] in physical situations where terms of $O(a)$ are negligible.

[1]D. Mintzer, J. Acoust. Soc. Am. **25**, 922 (1953).
[2]H.H. Su and M.A. Plonus, J. Opt. Soc. Am. **61**, 256 (1971).
[3]L.M. Erukhimov, I.G. Zarnitsyna, and P.I. Kirsch, Izv. VUZ, Radiofiz. **16**, 573 (1973).
[4]V.I. Tatarskii, *Propagation of Waves in a Turbulent Atmosphere* (Nauka, Moskow, 1967).
[5]L.C. Lee, J. Math. Phys. **15**, 1431 (1974).
[6]C.H. Liu, A.W. Wernik, and K.C. Yeh, IEEE Trans. Ant. Prop. **AP-22**, 624 (1974).
[7]V.I. Shishov, Sov. Astron. AJ. **17**, 598 (1974).
[8]C.H. Liu and K.C. Yeh, Radio Sci. **10**, 1055 (1975).
[9]L.C. Lee and J.R. Jokipii, Astrophys. J. **201**, 532 (1975).
[10]A. Ishimaru and S.T. Hong, Radio Sci. **10**, 637 (1975).
[11]S.T. Hong and A. Ishimaru, Radio Sci. **11**, 551 (1976).
[12]I. Sreenivasiah, A. Ishimaru, and S.T. Hong, Radio Sci. **11**, 775 (1976).
[13]K.C. Yeh and C.H. Liu, Radio Sci. **12**, 671 (1977).
[14]B.J. Rickett, Ann. Rev. Astron. Astrophys. **15**, 479 (1977). This is a comprehensive review article, with particular emphasis on scattering and scintillation of radio waves by interstellar matter.
[15]B.S. Uscinski, *The Elements of Wave Propagation in Random Media* (McGraw–Hill, New York, 1978).
[16]C.H. Liu and K.C. Yeh, J. Opt. Soc. Am. **67**, 1261 (1977).
[17]R.L. Fante, IEEE Trans. Ant. Prop. **AP-26**, 621 (1978).
[18]The brief literature cited here in connection with pulse propagation in random media is based direclty on the parabolic equation and the Markov random process approximation. It should be pointed out, however, that there are available alternative techniques which transcend several limitations of the aforementioned "pure" Markovian approximation. For example, a wave kinetic approach to second-order pulse statistics in dispersive and/or absorptive random media can be found in I.M. Besieris and F.D. Tappert, J. Math. Phys. **14**, 704 (1973); J. Appl. Phys. **44**, 2119 (1973).
[19]P.L. Chow, J. Math. Phys. **13**, 1224 (1972); J. Stat. Phys. **12**, 93 (1975); Indiana U. Math. J. **25**, 609 (1976).
[20]R. Dashen, "Path Integrals for Waves in Random Media," Stanford Research Inst. Tech. Report, JSR-76-1, May 1977.
[21]Extensive applications of Dashen's work (cf. Ref. 20) in the area of underwater sound wave propagation are included in S.M. Flatté, R. Dashen, W.H. Munk, and F. Zachariasen, "Sound Transmission through a Fluctuating Ocean," Stanford Research Inst. Tech. Report JSR-76-39, May 1977.
[22]K. Furutsu, J. Math. Phys. **17**, 1252 (1976).
[23]R.P. Feynman, Rev. Mod. Phys. **20**, 267 (1948).
[24]R.P. Feynman and A.R. Hibbs, *Quantum Mechanics and Path Integrals* (McGraw-Hill, New York, 1965).
[25]J.B. Keller and D.W. McLaughlin, Am. Math. Monthly **82**, 451 (1975).

[26]This multiplicative decomposition appears in many physical systems. A specific example is the case of a randomly perturbed cold plasma (cf. Ref. 5).

[27]R. Kubo, J. Math. Phys. **4**, 174 (1963).

[28]This assumption is not valid in a realistic ocean; in general, random fluctuations due to internal wave activity are inhomogeneous and anisotropic.

[29]In the language of quantum mechanics, Eqs. (3.20) and (3.21) govern the classical motion of uncoupled harmonically bound and free particles, respectively.

[30]A.C. Aitken, *Determinants and Matrices* (Interscience, New York, 1959).

[31]A. Ishimaru, *Wave Propagation and Scattering in Random Media* (Academic, New York, 1978), Vol. 2.

Unitary declension of dynamical symmetries for the time-dependent harmonic oscillator

Neil J. Günther

Department of Physics, The University of Southampton, England, U.K. S09 5NH
(Received 5 December 1977; revised manuscript received 18 April 1978)

Elaboration is given to a previous discussion concerning dynamical symmetries of the Lewis–Riesenfeld time-dependent harmonic oscillator. A boson operator formalism is used to define a generator of impicitly time-dependent unitary transformations. Bilinear combinations of the transformed boson operators are shown to span the symplectic algebra $\mathrm{Sp}(2n,\mathbb{R})$ regarded as a larger noninvariance group for the Lewis–Riesenfeld Hamiltonian. The multiplet structure of the embeddings, $\mathrm{Sp}(2n)\downarrow\mathrm{SU}(n)$, and $\mathrm{SU}(n+1)\downarrow\mathrm{SU}(n)$, is determined for the case $n=3$, using established branching rules. Boson operator realizations are presented for the multiplet structures in each $\mathrm{SU}(n)$ declension of these dynamical groups.

1. INTRODUCTION

In a previous paper[1] the three-dimensional time-dependent harmonic oscillator $3H(t)$ in our notation of the Lewis–Riesenfeld type[2] was shown to preserve only the rotational (geometrical) symmetry as an invariance of the time-dependent Hamiltonian. The usual dynamical symmetry associated with the extra (accidental) degeneracy of states in the time-independent Hamiltonian is SU(3).[3,4] This symmetry was shown to be a *noninvariance*[5] of the $3H(t)$ Hamiltonian provided one defined a rank-3 symmetric tensor invariant, I_{ij}, the elements of which contain certain implicit functions of time.[1] Although this tensor does not commute with the $3H(t)$ Hamiltonian, its total time derivative vanishes so confirming it as a bona fide dynamical invariant. An algebra, isomorphic to SU(3), was shown to close under commutation with the trace I_{ij} instead of the original $3H(t)$ Hamiltonian. It is in this sense that the Lewis–Riesenfeld Hamiltonian admits exact solutions analogous to those of the $3H$ Hamiltonian, so that recourse to perturbative methods is rendered unnecessary.[6–8]

If p_i and q_i form a conjugate canonical realization with commutation relations

$$[q_i, q_j] = 0 = [p_i, p_j], \quad [q_i, p_j] = i\delta_{ij},$$

then the Lewis–Riesenfeld $3H(t)$ Hamiltonian is

$$H = \tfrac{1}{2}p_i p_i + \tfrac{1}{2}\mu^2(t) q_i q_i,$$

where $\mu(t)$ is an arbitrary real or complex function of time representing the time-dependent frequency. The symmetric tensor invariant

$$I_{ij}(\rho,\dot\rho) = (1/2\rho^2) q_i q_j + \tfrac{1}{2}(\rho p_i - \dot\rho q_i)(\rho p_j - \dot\rho q_j)$$

developed in Ref. 1 is an exact invariant provided the auxiliary function $\rho(t)$ satisfies the nonlinear equation

$$\rho(t) + \mu^2(t)\rho(t) - \nu\rho^{-3} = 0, \tag{1}$$

where ν is a real-valued constant associated with the angular momentum of auxiliary motion in a four-dimensional hyperspace.[1,9] $I_{ij}(\rho,\dot\rho)$ falls into a class of functions, which up to an arbitrary multiplicative constant, are the most general homogeneous bilinear forms that can be prescribed for a $3H(t)$ Hamiltonian.

In this paper we given further perspective to our previous results concerning questions of symmetry by considering larger dynamical symmetry groups for the $3H(t)$. The conventional notion of a symmetry group, in quantum mechanics, relates to an exact invariance of the Hamiltonian which thereby incorporates conservation laws and selection rules. As an example recall the canonical realization of the SU(3) symmetry of $3H$ oscillator.[1] (Table I). In this case there exists a 3×3 symmetric matrix invariant which we have previously referred to as the Fradkin tensor A_{ij},[1] such that

$$[H, A_{ij}] = 0.$$

The 28 commutation relations of the SU(3) algebra may be written

$$[T_{ij}, H] = [L_i, H] = 0,$$
$$[L_i, L_j] = i\epsilon_{ijk} L_k,$$
$$[L_i, T_{jk}] = i\epsilon_{ijl} T_{lk} + i\epsilon_{ikl} T_{jl},$$
$$[T_{ij}, T_{kl}] = i(\epsilon_{jlm}\delta_{ik} + \epsilon_{jkm}\delta_{il} + \epsilon_{ilm}\delta_{jk} + \epsilon_{ikm}\delta_{jl}) L_m,$$

where $T_{ij} = A_{ij} - \frac{1}{3}\delta_{ij}A_{kk}$ is a traceless form of A_{ij} and L_{ij} is the angular momentum. Hence SU(3) is an exact dynamical symmetry for the $3H$ oscillator.

For the $3H(t)$ Hamiltonian, the 3×3 symmetric invariant is a constant of the motion but does *not* commute with the Hamiltonian, viz.,

TABLE I.

$3H$ oscillator	$3H(t)$ oscillator
A. Significant properties of SU(3) generators	
$[H, L_i] = 0$	$[H(t), L_i] = 0$
$[H, A_{ij}] = 0$	$[H(t), I_{ij}] \neq 0$
$\mathrm{Tr}\, A_{ij} = H$	$\mathrm{Tr}\, I_{ij} \neq H(t)$
$\mathrm{Tr}\, A_{ij}$ is SU(3) invariant	$\mathrm{Tr}\, I_{ij}$ is SU(3) invariant
H is a constant of the motion	$H(t)$ is not a constant of the motion
B. Symmetries	
Exact: SO(3); SU(3)	SO(3)
Approx: SU(4); Sp(6)	SU(3); SU(4); Sp(6)

$$[I_{ij}, H(t)] = \frac{\partial Iij}{\partial t}.$$

Furthermore, since

$$\mathrm{Tr}\, I_{ij} \neq H(t)$$

we refer[1] to SU(3) as an *approximate* symmetry of this system[5] (see Table I).

The reduction of invariance symmetries, SU(3)↓SO(3), for the $3H$ case is presented in Sec. 3. By defining a suitable unitary transformation for the boson operators ξ and η we show that bilinear combinations of the transformed variables ξ' and η', span the symplectic algebra Sp(6). Since the generators of this symmetry do not themselves commute with the original $3H(t)$ Hamiltonian we refer to this group as a "higher dimensional-approximate-symmetry" (H.A.S.). Then the noninvariance symmetry SU(3) forms a maximal compact unimodular subgroup. Other higher order nonivariance groups, e.g., SU(4) are also considered and the respective multiplet structure of SU(3) embeddings are determined using the branching rules outlined in the Appendix. General remarks bearing on the relationship of the $nH(t)$ to other problems of physical interest are made in conclusion to this study.

2. NOTATION

We continue to abide by the notation developed in Ref. 1, ie.,:

n: n-dimensional, $n = 1,2,3,\cdots$,

(t): time-dependent; suppressed argument implies time-independence,

H: isotropic harmonic oscillator,

A: anisotropic harmonic oscillator.

Then, e.g., $nH(t)$ represents the phrase "n-dimensional time-dependent isotropic harmonic oscillator". Units are chosen such that $m = h = 1$. The conventional summation over explicitly repeated indices applies unless otherwise stated. Roman subscripts take the values 1,2,3, in general. Other notation, relevant to the branching rules outlined in the Appendix, is defined there.

3. INVARIANCE SYMMETRIES

In order to establish a context for the approach adopted in the rest of this paper, we briefly reconsider and augment the essential results of Ref. 1 concerning the dynamical symmetry of the $3H(t)$ oscillator. As mentioned in Sec. 1, SU(3) is the *minimal dynamical* symmetry of the $3H$ Hamiltonian while SO(3)[10] is the *maximal invariance* symmetry of the $3H(t)$ Hamiltonian. The multiplet structure of SU(3) appearing in SO(3) may be obtained in the following way.

Using the notation of the Appendix, the regular (octet) representation of SU(3) is given by $D_n\{\lambda\} = 8$ which corresponds to the Young tableau

3	4
2	

In our notation this representation is written $\{2,1\}$. From the branching rules given in the Appendix (ii) we have:

$$\{\lambda\} \supset \sum_{\{\delta\}} \oplus \left[\frac{2,1}{\delta}\right]$$

$$\supset [2,1] \oplus [1].$$

In evaluating the dimensionality of the first terms of this direct sum of irreducible representations it is important to note that for the group SO($2k+1$) the number of partitions in the tableau p must satisfy $p \leqslant k$. This is not the case for [2,1] in SO(3). Consequently it is necessary to invoke the modification rules developed in Ref. 11(b). This results in a tableau $\boxed{2}\boxed{5}$ corresponding to rep. [2] having dimensionality $D_n[\lambda] = 5$. Then SU(3)↓SO(3):$\mathbf{8} \supset \mathbf{5} \oplus \mathbf{3}$. The branching rules provide a vindication of our previous identification of these multiplets with a canonical realization.[1] In the case where $\rho(t) = \mu^{-1/2} =$ constant, the tensor invariant $I_{ij}(\rho,\dot\rho,\xi,\eta)$ reduces to a 3×3 symmetric matrix which we have previously referred to as the Fradkin tensor.[1,3] The traceless version of this invariant (quadrupole tensor) has five independent components associated with the coset space[12] SU(3)/SO(3) and together with the three generators of SO(3) it furnishes the regular representation of SU(3) for the $3H$ Hamiltonian. For the $3H(t)$ Hamiltonian only the SO(3) generators commute with the Hamiltonian, making the maximal compact invariance symmetry. SU(3) is "restored" only as an approximate symmetry under commutation with the trace of $I_{ij}(\rho,\dot\rho,\xi,\eta)$ which is the singlet representation for the $3H(t)$. This result is tantamount to the statement that the energy-level degeneracy (including the "accidental" degeneracy) present in the $3H$ is unaltered in the $3H(t)$ case. The only difference in two cases is easily understood by considering the temporal development of the elliptic orbit in phase space. As the phase-space ellipse undergoes its peristaltic temporal development in accordance with $\rho(t)$ satisfying (1), the relative separation of each degenerate set is changed but the overall degeneracy amongst the energy levels remains unlifted. From this description of the $3H(t)$ in phase space, we are able to understand Lewis's result [2a] for the expectation value of the $1H(t)$ Hamiltonian in a state $|k\rangle$, viz.,

$$\langle k|H(t)|k\rangle = \tfrac{1}{2}(\rho^{-2} + \mu^2(t)\rho^2 + \dot\rho^2)(k+\tfrac{1}{2})\langle k|k\rangle. \tag{2}$$

In a canonical realization the $\mathbf{5} \oplus \mathbf{3}$ decomposition is associated with the angular momentum.[1,4] SO(3) is maintained as an invariance symmetry since the relative orientation of the phase-space orbit is preserved during its temporal development. The tensor invariant satisfies the eigenvalue equation

$$\sum_j^3 I_{ij}(\rho,\dot\rho,\xi,\eta)|\lambda_j\rangle = \lambda_j|\lambda_j\rangle \tag{3}$$

with eigenvalues

$$\lambda_j = \lambda_{\pm;0} = \tfrac{1}{2}(I_{jj} \pm \sqrt{I_{jj}^2 - D_j^2});0 \tag{4}$$

and normalized eigenvectors

$$\begin{bmatrix} 4\lambda_+ I_{31} + D_3 D_1 \\ 4\lambda_+ I_{32} + D_3 D_2 \\ 4\lambda_+ I_{33} - D^2 + D_3^2 \end{bmatrix}, \quad \begin{bmatrix} 4\lambda_- I_{31} + D_3 D_1 \\ 4\lambda_- I_{32} + D_3 D_2 \\ 4\lambda_- I_{33} - D^2 + D_3^2 \end{bmatrix}, \quad \begin{bmatrix} D_1 \\ D_2 \\ D_3 \end{bmatrix}, \tag{5}$$

where $D_{ij} = \frac{1}{2}(\xi_i \eta_j - \eta_i \xi_j)$. The relative orientation of the eigenvectors to the principal axes is constant while the eigenvalues are implicit functions of time.[1]

4. APPROXIMATE SYMMETRIES

A. Sp (6, $\mathbb{R}$) as an H.A.S.

The most significant of the higher dimensional noninvariance symmetries (H.A.S.) of nH is Sp($2n$, $\mathbb{R}$). Our purpose here is to show that Sp (6,$\mathbb{R}$) is an H.A.S. for the $3H(t)$. The usual boson operators ξ_i and η_i span a nonsemisimple algebra given by the commutation relations

$$[\xi_i, \xi_j] = [\eta_i, \eta_j] = 0, \quad [\xi_i, \eta_j] = \delta_{ij} \tag{6}$$

such that

$$[\xi_i, H_0] = \xi_i \quad \text{and} \quad [\eta_i, H_0] = -\eta_i, \tag{7}$$

where $H_0 = \frac{1}{2}(\xi_i \eta_i + \xi_i \eta_i)$ is the $3H$ Hamiltonian. It has been shown [12,13] that H_0 is a suitable realization for the Sp($2n$, $\mathbb{R}$) singlet. If α_{ij} and β_{ij} are infinitesimal generators spanning Sp($2n$, $\mathbb{R}$) then to first order, the symplectic matrix S which preserves the metric

$$\left[\begin{array}{c|c} 0 & \mathbb{1}n \\ \hline -\mathbb{1}n & 0 \end{array}\right]$$

may be written as

$$S = \mathbb{1} + \sum_{i,j}^{n} A_{ij}\alpha_{ij} + \sum_{i,j}^{n} (B_{ij}\beta_{ij} + C_{ij}\beta_{ij}^{+}), \tag{8}$$

where A_{ij}, B_{ij}, and C_{ij} are arbitrary complex coefficients. In terms of the boson operators (true for any bilinear forms)

$$\alpha_{ij} = \eta_j \xi_i, \quad \beta_{ij} = \xi_i \xi_j, \quad \beta_{ij}^{+} = \eta_i \eta_j. \tag{9}$$

These generators close under the following commutation relations:

$$\begin{aligned}
&[(\xi_i\xi_j), (\xi_k\xi_l)] = [(\eta_i\eta_j), (\eta_k\eta_l)] = 0, \\
&[(\xi_i\xi_j), (\eta_k\eta_l)] = \delta_{ik}\eta_l\xi_j + \delta_{il}\eta_k\xi_j + \delta_{jk}\eta_l\xi_i + \delta_{jl}\eta_i\xi_k, \\
&[(\xi_i\xi_j), (\eta_k\xi_l)] = \delta_{ik}\xi_l\xi_j + \delta_{jk}\xi_i\xi_l, \\
&[(\eta_i\eta_j), (\eta_k\xi_l)] = -(\delta_{il}\eta_k\eta_j + \delta_{jl}\eta_i\eta_k), \\
&[(\eta_i\xi_j), (\eta_k\xi_l)] = -\delta_{il}\eta_k\xi_j + \delta_{jk}\eta_i\xi_l
\end{aligned} \tag{10}$$

In terms of the basis vectors $\pm\boldsymbol{\lambda}_i \pm \boldsymbol{\lambda}_j$, $\pm 2\boldsymbol{\lambda}_i$, $\boldsymbol{\lambda}_i \cdot \boldsymbol{\lambda}_j = \delta_{ij}$, the standard form

$$\begin{aligned}
&\xi_i\eta_i = H_i, \\
&\eta_i\xi_j = E_{\lambda_i - \lambda_j}, \\
&\eta_i\eta_j = E_{\lambda_i + \lambda_j}, \quad (i \neq j), \\
&\xi_i\xi_j = E_{-\lambda_i - \lambda_j} \quad (i \neq j), \\
&\eta_i\eta_i = E_{2\lambda_i}, \\
&\xi_i\xi_i = E_{-2\lambda_i},
\end{aligned} \tag{11}$$

can be shown to span the root space C_n.

We now define a unitary transformation which takes the operators in the $3H$ case, ξ_i and η_i, to a form involving $\rho(t)$ and $\dot\rho(t)$ as real valued coefficients such that the original boson algebra is preserved. Evidently this must be a symplectic transformation having the effect

$$\begin{aligned}
U\xi_i U^{-1} &= \xi_i' = \xi_i'(\xi_i \eta_i, \rho, \dot\rho), \\
U\eta_i U^{-1} &= \eta_i' = \eta_i'(\xi_i, \eta_i, \rho, \dot\rho).
\end{aligned} \tag{12}$$

Let

$$G(\rho, \dot\rho, \xi_i', \eta_i) = \tfrac{1}{2}\rho(t)\{(\eta_i\xi_i' + \xi_i'\eta_i) - \dot\rho(t)(\xi_i'\xi_i' + \eta_i\eta_i)\} \tag{13}$$

be an Hermitian operator generating the unitary transformation specified in (12). This definition is evidently a special case of the transformation matrix S in (8). If F is an arbitrary polynomial in ξ,η then the commutation identities

$$\begin{aligned}
[\xi_i, F(\eta_j)] &= \delta_{ij}\frac{\partial F}{\partial \eta_j}(\eta_j) \\
[\eta_i, F(\xi_j)] &= \delta_{ij}\frac{\partial F}{\partial \xi_j}(\xi_j)
\end{aligned} \tag{14}$$

allow us to evaluate the result of a finite unitary transformation of ξ_i and η_i. The transformed operators

$$\begin{aligned}
\xi_i' &= \rho^{-1}\xi_i + \dot\rho\eta_i, \\
\eta_i' &= \rho^{-1}\eta_i - \dot\rho\xi_i + \rho\dot\rho^2\eta_i,
\end{aligned} \tag{15}$$

clearly satisfy the original boson algebra. The H.A.S. algebra Sp (6, $\mathbb{R}$) is now spanned by the bilinear generators:

$$\begin{aligned}
\xi_i'\xi_j' &= \rho^{-2}\beta_{ij} + \dot\rho^2\beta_{ij}^{+} + \rho^{-1}\dot\rho(\alpha_{ij} + \alpha_{ij}^{+}), \\
\eta_i'\eta_j' &= \rho^2\beta_{ij}^{+} + \dot\rho^2\beta_{ij} + \rho^2\dot\rho^2\beta_{ij}^{+} - \rho\dot\rho(\alpha_{ij} + \alpha_{ij}^{+}) \\
&\quad - \rho\dot\rho^3(\alpha_{ij}^{+} + \alpha_{ij}) + 2\rho^2\dot\rho^2\beta_{ij}^{+}, \\
\eta_i'\xi_j' &= \alpha_{ij} - \rho^{-1}\dot\rho\beta_{ij} + \rho\dot\rho\alpha_{ij} + \rho\dot\rho\beta_{ij}^{+} - \dot\rho^2\alpha_{ij} + \rho\dot\rho^3\beta_{ij}^{+}.
\end{aligned} \tag{16}$$

The invariant tensor I_{ij} $(\rho, \dot\rho, \xi, \eta)$ transforms as

$$U I_{ii} U^{-1} = \tfrac{1}{2}(\xi_i'\eta_i' + \eta_i'\xi_i') \equiv H_0'. \tag{17}$$

The following linear combinations of $\alpha_{ij}' = \eta_i'\xi_j'$ produce the usual subgroup structure (Table I),

$$\begin{aligned}
K_{ij}' &= \alpha_{ij}' + \alpha_{ji}' : \mathrm{U}(3), \\
J_{ij}' &= \alpha_{ij}' - \alpha_{ji}' : \mathrm{SO}(3), \\
T_{ij}' &= K_{ij}' - \tfrac{2}{3}\delta_{ij}\alpha_{kk}' : \mathrm{SU}(3)/\mathrm{SO}(3).
\end{aligned} \tag{18}$$

B. Other H.A.S.'s and SU(3) embeddings

Several SU(3) multiplets are embedded in a single irreducible representation of Sp(6) and for the regular representation these may be identified using the branching rules given in the Appendix.

(a) Sp(6)↓SU(3):

The regular representation of Sp(6) is $\langle 2 \rangle$ and has dimension $D_\eta\langle\lambda\rangle = 21$. The rule (A3) yields the following multiplets:

$$
\begin{aligned}
\langle \lambda \rangle &\supset \sum_{\{\xi,\delta\}} \oplus \left\{\bar{\xi}, \frac{2}{\xi,\delta}\right\} \\
&\supset \left\{0, \frac{2}{0}\right\} \oplus \left\{\bar{1}, \frac{2}{1}\right\} \oplus \left\{\bar{1}^2, \frac{2}{1^2}\right\} \oplus \left\{\bar{2}, \frac{2}{2}\right\} \qquad (19) \\
&\supset \mathbf{6} \oplus (\mathbf{8} \oplus \mathbf{1}) \oplus \bar{\mathbf{6}}
\end{aligned}
$$

in agreement with Ref. 14 and the boson realization discussed in Sec. 4A. (see Table II).

(b) SU(4)↓SU(3):

Introducing the higher order symmetry SU(4),[15] the symmetric tensor representations of rank l of SU(4) contain the first l-levels of the $3H(t)$ uniquely while the remainder are accomodated in one infinite dimensional unitary representation of the noncompact subgroup SU(3,1) of Sl(4,C). For example, ,the symmetric tensor reps. of SU(3) are:

$\boxed{3}$	dimension $= \mathbf{3}$	rank $l = 1$
$\boxed{3}\boxed{4}$	$= \mathbf{6}$	$= 2$
$\boxed{3}\boxed{4}\boxed{5}$	$= \mathbf{10}$	$= 3$
$\boxed{3}\boxed{4}\boxed{5}\boxed{6}$	$= \mathbf{15}$	$= 4$
$\cdot$	$= \cdot$	$= \cdot$
$\cdot$	$= \cdot$	$= \cdot$
$\cdot$	$= \cdot$	$= \cdot$

Now the **35**-dimensional symmetric rep. of SU(4) contains the following multiplets:

$$
\begin{aligned}
\boxed{4}\boxed{5}\boxed{6}\boxed{7} &\supset \sum_{\{m\}} \oplus \left\{\frac{4}{m}\right\} \\
&\supset \{4\} \oplus \{3\} \oplus \{2\} \oplus \{1\} \oplus \{0\} \\
&\supset \mathbf{15} \oplus \mathbf{10} \oplus \mathbf{6} \oplus \mathbf{3} \qquad (20)
\end{aligned}
$$

which are exactly the first symmetric reps. of SU(3). Decomposition of the regular representation permits us to identify the SU(4) generators. The regular rep. is $\{2,1^2\}$ with $D_n\{\lambda\} = 15$. Then from (A1) we find

$$
\begin{aligned}
\{\lambda\} &\supset \{2,1^2\} \oplus (\{1^3\} \oplus \{2,1\}) \oplus \{1^2\} \\
&\supset \mathbf{3} \oplus (\mathbf{1} \oplus \mathbf{8}) \oplus \bar{\mathbf{3}}. \qquad (21)
\end{aligned}
$$

A realization is given in Table II.

5. DISCUSSION

In this paper several points have been made regarding the higher approximate symmetries of the $nH(t)$.

Firstly it should be emphasized that the symmetries discussed, viz., SU(3), SU(4), and Sp(6) are all approximate symmetries of the Hamiltonian in the sense of Sec. 1, i.e., not all the generators of the corresponding Lie algebra commute with the $3H(t)$ Hamiltonian. One could take the view that there is some underlying "breaking" of the respective symmetry of the Hamiltonian and we consider this approach in Ref. 16. Since the Hamiltonian can be written as a function of the SU(3) generators the effect of this form of symmetry breaking is to split the SU(3) multiplets without mixing them. This splitting of levels is time-dependent in the case of th $3H(t)$ Hamiltonian as described in Sec. 3.

Next we have considered two questions:

1. Do the known dynamical symmetries of the $3H$ Hamiltonian hold for the $3H(t)$ case?

As an example, in Sec. 4A we have shown in terms of the usual boson operators that the there exists a unitary canonical transformations which takes them into those which describe the $3H(t)$. In this way it was demonstrated that Sp(6) is also an approximate symmetry of the $3H(t)$ oscillator. Generally we associate Sp($2n$, ℝ) with $nH(t)$.

2. How are the U(3) multiplets (corresponding to the nine generators) contained in these larger approximate symmetry groups?

To answer this question we used established branching rules for the reduction of a group to its subgroup (Appendix A). In particular we have considered the chain Sp(6)⊃SU(4)⊃SU(3)⊃SO(3) and looked at the resulting UIR's for the U(3) representation, viz., the octet and singlet. The paradigm for this approach was presented in Sec. 3, by examining the octet reduction in SU(3)↓SO(3). The quintet and triplet reps. arrived at in this way supported the dynamical assignment given in our earlier paper where the Hamiltonian was to be associated with the singlet.

The same method was used for the regular reps. of the groups, Sp(6) [see (19)] and SU(4) [see (21)] since it is these reps. which are directly related to the number of generators.

TABLE II.

Group	Representation	Subgroup	Multiplets	Boson realization
SU(3)	[2,1]	SO(3)	**5**	$\xi_i\eta_j + \eta_i\xi_j - \frac{2\delta_{ij}}{3}H_0$
			3	$\xi_i\eta_j - \eta_i\xi_j$
SU(4)	$\{2,1^2\}$	SU(3)	$\mathbf{8}\oplus\mathbf{1}$	$\xi_i\eta_j + \eta_i\xi_j$
			3	ξ_i
			3	η_j
Sp(6)	$\langle 2\rangle$	SU(3)	$\mathbf{8}\oplus\mathbf{1}$	$\xi_i\eta_j + \eta_i\xi_j$
			6	$\xi_i\xi_j$
			6	$\eta_i\eta_j$

The embedded U(3) multiplets are immediately identifiable and their respective Boson realizations are presented in Table II. The other multiplets are also assigned a realization amongst the remaining bilinear operators.

6. GENERAL REMARKS

To conclude this study we make some comments, of a more speculative nature, on the relevance of these and our previous results to more diverse problems associated with a Hamiltonian of the $nH(t)$ type.

A. Infinite dimensional Hamiltonian

The relevance of the $nH(t)$ to the understanding of field-theoretic problems may be of some interest. The problems of mathematical consistency concerning infinite component unitary transformations on nH Hamiltonians ($n\to\infty$) and the boundedness of the conventionally defined Hilbert space have been discussed, for example, by Segal.[17] These difficulties, presumably pertain to the $nH(t)$ *mutatis mutandis*.

Apart from questions of mathematical consistency, the $nH(t)$ Hamiltonian could be viewed as representing the discrete case of a continuous boson field theory. For example, if $\varphi(k)$ is a boson field, then the total Hamiltonian (continuum analog) in momentum space,

$$H[\varphi(k)] = \frac{1}{2}\int d^4k\,[(\partial_\mu\varphi)^2 + m^2(k)\varphi^2],$$

would have a momentum dependent mass term (as in naïve lattice-type models) reflecting the time dependent frequency of the Lewis–Riesenfeld Hamiltonian. A subtle difference, however, is that the above Hamiltonian is translation invariant in k whilst the Lewis–Riesenfeld Hamiltonian is not translation invariant in t. Furthermore, if specific symmetries are known for the m = constant case then in view of our previous discussion they may be found to hold for the momentum dependent case $m = m(k)$.

As an example of this consider a field theory described by the total Hamiltonian

$$H_f[\psi(k)] = \int d^4k\,[(\bar\psi \not{D}_\mu\psi) + m(k)\bar\psi\psi].$$

Here $\psi(k)$ is a fermion field and $\not{D}_\mu$ is a covariant derivative with $\mu = 1,2,3,4$. By analogy, one may be led to suspect that those invariance symmetries of $H_f(m = \text{const})$, e.g., SU(6) type in a theory of strong interactions with fundamental fermions fields $\psi(k)$, hold as symmetries for $H_f[m = m(k)]$ (quark binding). Further investigation along these lines, with the Lewis–Riesenfeld Hamiltonian providing an example, may furnish some insight into the problems of constructing a field theory for the strong interactions.

B. Anisotropic oscillator

The $3A$ has long been a subject of interest.[4,18] Remarks concerning the algebra of the $3A(t)$, in a canonical realization have been made elsewhere.[1]

As another example, we note that the high order behavior of the $1A$ Hamiltonian, $H = \frac{1}{2}[\dot x^2 - (mx)^2] + (g/4!)x^4$ has been investigated.[19] This Hamiltonian is known to model degenerate vacua arising from spontaneous symmetry breaking in certain scalar field theories. The existence of a Lewis type invariant, viz., $\frac{1}{2}[\rho^{-2}x^2 + (\rho x - \dot\rho x)^2]$ (Sec. 1) would be useful for certain calculations. Our investigations suggest that such a *quadratic* invariant cannot be prescribed for this Hamiltonian.

Whilst the above remarks lie somewhat outside the mainstream of ideas contained in the earlier sections of this paper, the necessity for them is evidenced by the connection between the symmetry of a particular problem and its possible solutions.[20] The problems outlined in Sec. 6 have the symmetry properties of the $nH(t)$ Hamiltonian in common.

APPENDIX: BRANCHING RULES[11,21]

The problem of determining those irreducible representations of multiplets in a given subgroup H of G, and let λ be an irreducible representation of H. In general $G\downarrow H$ for the λ representation yields a set of matrices $\{\lambda(h):h\epsilon H\}$ which are reducible so that

$$G\downarrow H \sim \lambda\downarrow\sum_\mu \oplus B^\mu{}_\lambda\mu,$$

where the $B^\mu{}_\lambda$ are the branching coefficients. We list the branching rules for the groups discussed.

(i) $\mathrm{U}(n,\mathbb{C})\downarrow\mathrm{U}(n-1,\mathbb{C})$: The dimensionality $D_n\{\lambda\}$ is given by the quotient $N_n\{\lambda\}/H(\lambda)$ where

$$N_n\{\lambda\} = \prod_{(i,j)}(n-i+j)$$

and

$$H(\lambda) = \prod_{(i,j)} h_{ij} \quad \text{(Hook length).}$$

Here λ labels the irreducible representation and (i,j) refers to the box in the ith row and jth column of the $\{\lambda\}$ tableau. Then the branching rule

$$\{\lambda\}\supset\sum_{\{m\}}\oplus\left\{\frac{\lambda}{m}\right\} \tag{A1}$$

is specified by a partition, m, of all one row tableaux given by $\{m\} = \{0\}, \{1\}, \{2\}, \{3\},\cdots$.

(ii) $\mathrm{U}(n,\mathbb{C})\downarrow\mathrm{O}(n,\mathbb{R})$: Here the dimensionality $D_n[\lambda]$ is defined by

$$N_n[\lambda] = \prod_{(i,j)}[n-i+j+\delta_{ij\lambda}a_k - (1-\delta_{ij\lambda})(1+b_k)],$$

where

$$\delta_{ij\lambda} = \begin{cases}1, & \text{if diagonal through } (i,j)\cap \text{ box at end of row,}\\ 0, & \text{otherwise.}\end{cases}$$

a_k is the arm length and b_k is the leg length measured from the (i,j) box. Then

$$\{\lambda\}\supset\sum_{\{\delta\}}\oplus\left\{\frac{\lambda}{\delta}\right\} \tag{A2}$$

is specified by $\{\delta\} = \{0\}, \{2\}, \{4\}, \{2^2\}, \{6\}, \{4,2\},\cdots$.

(iii) $\mathrm{Sp}(2n,\mathbb{R})\downarrow\mathrm{U}(n,\mathbb{C})$: The dimensionality $D_n\langle\lambda\rangle$ is defined by

$$N_n\langle\lambda\rangle = \prod_{(i,j)}|n-i+j-\delta_{ij\tilde\lambda}b_k + (1-\delta_{ij\tilde\lambda})(1+a_k)|,$$

where

$$\delta_{ij\lambda} \doteq \begin{cases} 1, & \text{if diag. through } (i,j) \cap \text{ box bottom of column,} \\ 0, & \text{otherwise.} \end{cases}$$

Then

$$\langle \lambda \rangle \supset \sum_{\{\xi,\delta\}} \oplus \left\{ \bar{\xi}; \frac{\lambda}{\delta} \right\} \tag{A3}$$

is specified by $\{\xi\} = \{0\}, \{1\}, \{1^2\}, \{2\}, \cdots$.

(iv) $U(n, \mathbb{C}) \downarrow Sp(n, \mathbb{R})$:

$$\{\lambda\} \supset \sum_{\{\beta\}} \oplus \left\langle \frac{\lambda}{\beta} \right\rangle, \tag{A4}$$

where $\{\beta\} = \{0\}, \{1^2\}, \{1^4\}, \{2^2\}, \cdots$.

Note added in proof: Several papers by M. E. Major [J. Math. Phys. **18**, (1) 1938, (2) 1944, (3) 1952 (1977)] have come to my attention. In paper (3) there is a demonstration that the semidirect product of group $SU(n) \odot W_G(n)$ is an appropriate candidate for a dynamical symmetry of the nH oscillator in the sense of Sec. 4B(b). Here $W_G(n)$ is the little group of the Poincaré or conformal groups and should not be confused with the group S_n of reflections in $SU(n)$ root space. The implications of the this result for the $H(t)$ oscillator will be discussed elsewhere.[16]

ACKNOWLEDGMENTS

With pleasure I take this opportunity to express my gratitude to Dr. P. D. Jarvis for conversations which initiated this work and for his critical reading of the manuscript. Thanks are also due to Dr. R. C. King for a helpful discussion, and the University of Southampton for the award of a Research Grant.

[1]N.J. Günther and P.G.L. Leach, J. Math. Phys. **18**, 572 (1977).

[2](a) H.R. Lewis, Phys. Rev. Lett. **18**, 510, 636 (1967); (b) H.R. Lewis, J. Math. Phys. **9**, 1976 (1968); (c) H.R. Lewis and W.B. Riesenfeld, J. Math. Phys. **10**, 1458 (1969).

[3]D.M. Fradkin, Am. J. Phys. **33**, 207 (1965); D.M. Fradkin, Prog. Theor. Phys. **37**, 798 (1967).

[4]J.M. Jauch and E.L. Hill, Phys. Rev. **57**, 641 (1940).

[5]By the term "noninvariance" symmetry, we shall mean a group whose generators do not close under commutation with the Hamiltonian of the system, either time-dependent or time-independent.

[6]M. Kruskal, J. Math. Phys. **3**, 806 (1962).

[7]S. Solimeno, P. Di Porto, and B. Crosignani, J. Math. Phys. **10**, 1922 (1969); C.L. Hammer and T.A. Weber, ibid.**6**, 1591 (1965).

[8]In view of these remarks it would seem that the phrase "time-dependent oscillator" is so encompassing as to be obscure and that a more specific term of reference would be suitable, e.g., Lewis–Riesenfeld–Kruskal (LRK) oscillator.

[9]C.J. Eliezer and A. Gray, S.I.A.M. **30**, 463 (1976).

[10]We point out that the group structure preserved by the $3H(t)$ is SO(3) and *not* SU(2). This can be seen most evidently by examining the octet reductions, viz., $SU(3)\downarrow SU(2) \supset \mathbf{3} \oplus \mathbf{2} \oplus \mathbf{2} \oplus \mathbf{1}$, while $SU(3)\downarrow SO(3) \supset \mathbf{5} \oplus \mathbf{3}$. Obviously it is the SO(3) representations which we are considering.

[11](a) R.C. King, J. Phys. A **8**, 429 (1975); (b) R.C. King, J. Math. Phys. **12**, 1588 (1971); (c) Unpublished lecture notes on Group Theory, Southampton University (1977).

[12]G. Gilmore, *Lie Groups, Lie Algebras and Some of Their Applications* (Wiley, New York, 1974); B.G. Wybourne, *Classical Groups for Physicists* (Wiley, New York, 1974).

[13]M. Moshinsky and C. Quesne, J. Math. Phys. **12**, 1772 (1971).

[14]H. Bacry, J. Nuyts, and L. Van Hove, Nuovo Cimento **35**, 510 (1965).

[15]N. Mukunda, L. O'Raifeartaigh, and E.C.G. Sudarshan, Phys. Rev. Lett. **15**, 1041 (1965).

[16]N.J. Günther, "Dynamical Symmetry Breaking in the Lewis–Riesenfeld Oscillator," THEP preprint, University of Southampton (1978).

[17]I.E. Segal, *Mathematical Problems of Relativistic Physics* (Am. Math. Soc., Providence, Rhode Island, 1963).

[18]M. Moshinsky, J. Patera, and P. Winternitz, J. Math. Phys. **16**, 82 (1975) and References cited thererin.

[19]C.M. Bender and T.T. Wu, Phys. Rev. D **7**, 1620 (1973) and cited refs.

[20]Y.S. Kim and M.E. Noz, Phys. Rev. **D15**, 335 (1977); R.P. Feynman, *Photon-Hadron Interactions* (Benjamin, New York, 1972).

[21]M.L. Whippman, J. Math. Phys. **6**, 1534 (1965).

Projective representations of $SU(2)_\Lambda R^4$

N. B. Backhouse

Department of Applied Mathematics and Theoretical Physics, Liverpool University, P. O. Box 147, Liverpool L69 3BX, United Kingdom

J. W. B. Hughes

Department of Applied Mathematics, Queen Mary College, Mile End Road, London El 4NS, United Kingdom

(Received 23 May 1978)

The Mackey little group procedure is used to classify the projective irreducible unitary representations of the group $SU(2)_\Lambda R^4$. Whereas an arbitrary ordinary representation of this group plays the role of both spectral and dynamical group for the three-dimensional hydrogen atom, it is found that one of its projective representations acts as spectral group for the two-dimensional isotropic oscillator. The projective representations of the Lie algebra of $SU(2)_\Lambda R^4$ are treated as ordinary representations of the central extension of this Lie algebra, and shift operator techniques used to obtain the matrix elements of its representatives. The connection with the isotropic oscillator is shown to be related to the realizability of the extended Lie algebra by two-boson creation and annihilation operators.

1. INTRODUCTION

The Lie algebra $SU(2)_A(T_2 \times T_2^\#)$, which is the subject of this paper, was first considered by Hughes and Yadegar[1] under the name $O(3)_A(T_2 \times T_2^\#)$. This Lie algebra, obtained by putting together the SU(2) algebra with basis $\{L_0, L_\pm\}$, and two mutually commuting Abelian algebras T_2 and $T_2^\#$ with bases $\{Q_{\pm 1/2}\}$ and $\{Q^\#_{\pm 1/2}\}$, which are modules for the spin-$\frac{1}{2}$ representation of SU(2), has nonzero commutation relations

$$\begin{aligned} &[L_0, L_\pm] = \pm L_\pm, \quad [L_+, L_-] = 2L_0, \\ &[L_0, Q^{(\#)}_{\pm 1/2}] = \pm \tfrac{1}{2} Q^{(\#)}_{\pm 1/2}, \\ &[L_\pm, Q^{(\#)}_{\mp 1/2}] = Q^{(\#)}_{\pm 1/2}. \end{aligned} \tag{1.1}$$

In their paper, Hughes and Yadegar determined the equivalence classes of irreducible unitary representations (IUR) of this algebra by a shift operator method, and noted that on reduction to the SU(2) algebra the generic IUR had the correct bound state multiplicities for the hydrogen atom. The SU(2) algebra was not, however, the SO(3) algebra physically realized by the angular momentum operators whose representations correspond to integral angular momentum. On restriction to the SU(2) generated by the above $\{L_0, L_\pm\}$, the IUR of $SU(2)_A(T_2 \times T_2^\#)$ yield both integral and half-integral representations, and it is for this reason that we choose in this paper to replace the term $O(3)_A(T_2 \times T_2^\#)$ by the more appropriate $SU(2)_A(T_2 \times T_2^\#)$. In addition to showing that this Lie algebra acts as a spectrum generating algebra for the hydrogen atom alternative to the algebra O(4,1) usually considered, they showed that from the representatives of elements of the enveloping algebra of $SU(2)_A(T_2 \times T_2^\#)$ in this generic IUR, one could realize precisely that IUR of the O(4,2) algebra which was employed by Barut and Kleinert[2] to deal with dynamical properties of the hydrogen atom. The seven-dimensional $SU(2)_A(T_2 \times T_2^\#)$ was thereby shown to serve equally well as a dynamical algebra for the hydrogen atom as the fifteen-dimensional algebra of O(4,2).

Subsequently, Backhouse[3] rederived the results of Hughes and Yadegar from a group theoretical point of view. It was shown that the complexification of the Lie algebra of the group $SU(2)_A R^4$ has a real form isomorphic to $SU(2)_A(T_2 \times T_2^\#)$. In the construction of this group, the action of SU(2) on R^4 comes about through the standard homomorphism of $SU(2)\times SU(2)$ onto SO(4). The IUR of $SU(2)_A R^4$ were obtained using Mackey's method of induced representations, and were naturally realized by an action on $L^2(S^3)$—exactly the space of bound states of the hydrogen atom.[4] Thus, just as $SU(2)_A(T_2 \times T_2^\#)$ acts as a spectrum generating algebra, the corresponding group $SU(2)_A R^4$ was shown to be a spectrum generating group, for the hydrogen atom. Unlike the group SO(4,1) usually considered in this context, however, $SU(2)_A R^4$ does not contain the symmetry group SO(4), nor the angular momentum group SO(3), as subgroups; rather, they are generated by certain elements in the enveloping algebra of $SU(2)_A(T_2 \times T_2^\#)$. The IUR of SO(4) which arise are constrained to be the diagonal representations, as in the usual group theoretical treatment of the hydrogen atom.[4,5]

In this paper we explore further the representation theory of $SU(2)_A R^4$ by computing its projective irreducible unitary representations (PIUR), showing incidentally that one particular PIUR, realizable on $L^2(R^2)$, makes $SU(2)_A R^4$ a spectrum generating group of the two-dimensional isotropic harmonic oscillator, containing the symmetry, or degeneracy, group SU(2) as a subgroup. Other applications of projective representations are mentioned by Backhouse.[6]

An alternative approach to the construction of PIUR of a group is by the construction of ordinary IUR of central extensions of the group. It is shown that, modulo group automorphisms, $SU(2)_A R^4$ has essentially just one central extension. At the Lie algebra level this is the central extension $C(SU(2)_A(T_2 \times T_2^{(\#)}))$ generated by L_0, $L_\pm$, and $Q^{(\#)}_{\pm 1/2}$, which still satisfy the commutation relations (1.1), but appended by the identity operator **1** and where $Q^\#_{\pm 1/2}$ and $Q^\#_{\mp 1/2}$, instead of mutually commuting, have commutators

 0022-2488/79/071545-10$01.00

$$[Q_{\pm 1/2}, Q^{\#}_{\mp 1/2}] = \mp \mathbf{1}. \quad (1.2)$$

The IUR of $C(\mathrm{SU}(2)_\Lambda(T_2\times T_2^{\#}))$ can be classified by SU(2) shift operator methods developed by Hughes and Yadegar[7] and used by them for the case of ordinary IUR of $\mathrm{SU}(2)_\Lambda(T_2\times T_2^{\#})$.[1] It is found that the IUR of $C(\mathrm{SU}(2)_\Lambda(T_2\times T_2^{\#}))$ obtained by these methods are in exact one to one correspondence with the PIUR of $\mathrm{SU}(2)_\Lambda R^4$. This Lie algebra has a single invariant I_4, of fourth order in the basis elements, whose eigenvalues provide a unique label for the IUR. In addition, as in the case of SU(3) in an SU(2) basis,[8] it possesses three independent SU(2) scalar operators, any one of which provides a state label to distinguish states corresponding to the same value $l(l+1)$ of the SU(2) Casimir L^2. In addition to classifying the IUR, the shift operator methods also enable us to write down matrix elements for the $Q^{(\#)}_{\pm 1/2}$ in an arbitrary one of these IUR, thus completing their analysis. Just one of these IUR contains no l-degeneracies; it is also shown to be the only one which can be realized in terms of two-boson creation and annihilation operators. This provides further explanation of the fact that it generates the spectrum of the two-dimensional harmonic oscillator.

Independently, as part of a more general study, Major[9] showed that the spectrum generating group of the two-dimensional isotropic harmonic oscillator is $\mathrm{SU}(2)_\Lambda N$, where N is the so-called Weyl group. This group, a central extension of $\mathrm{SU}(2)_\Lambda R^4$, has ordinary representations which are obtained by lifting the projective representations found in the present paper. It is worth pointing out that we prove Major's group to be essentially the only central extension of $\mathrm{SU}(2)_\Lambda R^4$ by the circle group.

We begin in Sec. 2 by reviewing the definitions and elementary theory of projective representations of groups and algebras. As a prerequisite to finding the projective representations of a given group one needs explicit knowledge of its multipliers, or factor systems. In Sec. 3 we recall the Mackey decomposition for analyzing the multipliers of semidirect product groups and apply it to the group $\mathrm{SU}(2)_\Lambda R^4$. The PIUR of $\mathrm{SU}(2)_\Lambda R^4$ are obtained in Sec. 4 using the Mackey little group method. It is here that we recognize $\mathrm{SU}(2)_\Lambda R^4$ as the spectral group of the two-dimensional isotropic harmonic oscillator by explicitly performing the reduction of one particular PIUR, denoted π^0, to SU(2). The relationship with earlier accounts of the oscillator is shown up by computing, as differential operators, the action of the generators of SU(2) on the carrier space $L^2(R^2)$ of π^0. In Sec. 5 we consider the Lie algebra $C(\mathrm{SU}(2)_\Lambda(T_2\times T_2^{\#}))$ and construct the invariant I_4, the SU(2) scalar operators, and the shift operators $A^{(\#)\pm 1/2}$ needed for an analysis of its IUR. The actual analysis is given in Sec. 6, where the relation to the two-dimensional harmonic oscillator is also further discussed. We conclude in Sec. 7 with a short discussion of related groups and algebras to which the techniques of this paper might be applied.

2. PROJECTIVE REPRESENTATIONS

Algebraically a projective unitary representation of a group G is a mapping $g\to U(g)$ of G into the unitary operators on some Hilbert space, together with a circle-valued function ω on $G\times G$, which satisfy, for all elements g_1, g_2, g_3 of G, both of the relations

$$U(g_1)U(g_2) = \omega(g_1,g_2)U(g_1g_2), \quad (2.1)$$

$$\omega(g_1,g_2)\omega(g_1g_2,g_3) = \omega(g_1,g_2g_3)\omega(g_2,g_3). \quad (2.2)$$

ω is called the multiplier, or factor system, of the representation. ω is called trivial if there exists a circle-valued function δ on G such that $\omega(g_1,g_2) = \delta(g_1)\delta(g_2)/\delta(g_1g_2)$ for all $g_1,g_2\in G$. Triviality establishes an equivalence relation on multipliers and their associated projective representations over and above the usual unitary equivalence. In particular we say that projective unitary representations with trivial multipliers are projectively equivalent to ordinary unitary representations. Also, we normally choose equivalence class representatives for multipliers such that $\omega(e,g) = \omega(g,e) = 1$ for all $g\in G$, where e is the identity of G.

Given a multiplier ω for G we can form the central extension G^ω of G by the circle group T^1 as follows: $G^\omega = \{(g,z) : g\in G, z\in \mathbb{C}, \text{ and } |z| = 1\}$ with multiplication $(g_1,z_1)(g_2,z_2) = (g_1g_2, \omega(g_1,g_2)z_1z_2)$ for all $g_1,g_2\in G$ and $z_1,z_2\in T^1$. We shall identify the set $\{(e,z) : z\in T^1\}$ with T^1, and note that G^ω/T^1 is isomorphic to G. The significance of G^ω is that a projective representation U of G which multiplier ω gives rise to an ordinary representation V of G^ω defined by $V(g,z) = zU(g)$. Conversely, an ordinary representation of G^ω for which $V(e,z)$ acts as scalar multiplication by z yields a projective representation U of G by defining $U(g) = V(g,1)$.

We have made no mention of topological or measure theoretic conditions; these are dealt with thoroughly by Mackey,[10] Bargmann,[11] and Parthasarathy.[12] We mention here only that we require our representations to be continuous, and that if G is a connected Lie group, then representatives ω for the equivalence classes of multipliers can be so chosen that the extension groups G^ω are also connected Lie groups. In particular ω is analytic near (e,e) which enables us to relate the projective representations of G with projective representations of its Lie algebra LG as follows.

First we can check that if $g\to U(g)$ is a projective representation of G, then

$$U(g_1)U(g_2)U(g_1)^{-1}U(g_2)^{-1} = \mu(g_1,g_2)U(g_1g_2g_1^{-1}g_2^{-1}), \quad (2.3)$$

where

$$\mu(g_1,g_2) = \frac{\omega(g_1,g_2)\omega(g_1^{-1},g_2^{-1})\omega(g_1g_2,g_1^{-1}g_2^{-1})}{\omega(g_1,g_1^{-1})\omega(g_2,g_2^{-1})} \quad (2.4)$$

for all $g_1,g_2\in G$. Now if $t\to\exp(tX)$, $X\in LG$, is a one-parameter subgroup of G, then Stone's theorem provides an operator $U(X)$, defined on some domain, such that, for all t, $U[\exp(tX)] = \exp[tU(X)]$. It is known that, for any $X\in LG$, $U(X)$ is skew-Hermitian on the dense subspace of so-called analytic vectors. Then the infinitesimal version of the commutator relation (2.3) is

$$[U(X),U(Y)] = U([X,Y]) + F(X,Y)\mathbf{1} \quad (2.5)$$

for all $X,Y\in LG$, where $\mathbf{1}$ is the identity operator and F is a bilinear skew-symmetric function on LG with purely imaginary values. F is the infinitesimal counterpart of the function μ defined by (2.4). Equation (2.5) can be alternatively viewed as giving an ordinary representation of a central extension $C(LG)$ of LG by the real numbers, in which a real number r is represented by $ir\mathbf{1}$. This approach is taken up in Sec. 5 and the following sections.

3. MULTIPLIERS

Let $G = P_A N$ be a semidirect product of the groups P and N, where N is a normal subgroup of G, and denote the general element of G by $g = np$, where $n \in N$ and $p \in P$. Then it is a well established fact[10,12,13] that equivalence classes of multipliers of G contain representatives which can be decomposed in the form

$$\omega(n_1p_1,n_2p_2) = \gamma(n_1,p_1n_2p_1^{-1})\beta(p_1,n_2)\alpha(p_1,p_2), \qquad (3.1)$$

where

$$\beta(p,n_1n_2) = \beta(p_1,n_1)\beta(p,n_2)\frac{\gamma(pn_1p^{-1},pn_2p^{-1})}{\gamma(n_1,n_2)} \qquad (3.2)$$

and

$$\beta(p_1p_2,n) = \beta(p_1,p_2np_2^{-1})\beta(p_2,n) \qquad (3.3)$$

for all values of the elements. Here α and γ are multipliers on P and N, respectively, and β is a circle-valued function on $P \times N$. It can be further shown[12] that if N is a real vector group, such as R^4, and P is a simply connected and connected semisimple Lie group, such as SU(2), then α and β can be reduced to unity. This leaves

$$\omega(n_1p_1,n_2p_2) = \gamma(n_1,p_1n_2p_1^{-1}), \qquad (3.4)$$

where

$$\gamma(pn_1p^{-1},pn_2p^{-1}) = \gamma(n_1,n_2) \qquad (3.5)$$

for all values of the group elements. Finally, we know that each class of multipliers for a real vector group contains a unique representative of the form defined by $\gamma(n_1,n_2) = \exp[iA(n_1,n_2)]$, where A is a real skew-symmetric bilinear form on N.

We now apply the above general results to the particular case where $G = \mathrm{SU}(2)_A R^4$, first of all summarizing the realization of this group employed by Backhouse.[3] G may be identified with the set of (4×4) matrices

$$\begin{pmatrix} v & m \\ 0 & I \end{pmatrix}, \text{ where } v \in \mathrm{SU}(2),\ m = \begin{pmatrix} \alpha & -\beta^* \\ \beta & \alpha^* \end{pmatrix},$$
$$\alpha = \tau + i\sigma, \text{ and } \beta = \nu + i\mu.$$

We relate m to a point in R^4 by defining $t = (\mu,\nu,\sigma,\tau)^T$; it is also convenient to define $z = (\alpha,\alpha^*,\beta,\beta^*)^T \in C^4$. This complex form is most useful for defining the action of SU(2) on R^4: if

$$v = \begin{pmatrix} \phi & -\psi^* \\ \psi & \phi^* \end{pmatrix}, \text{ where } \phi\phi^* + \psi\psi^* = 1,$$

then the conjugate

$$\begin{pmatrix} v & 0 \\ 0 & I \end{pmatrix}\begin{pmatrix} I & m \\ 0 & I \end{pmatrix}\begin{pmatrix} v & 0 \\ 0 & I \end{pmatrix}^{-1} = \begin{pmatrix} I & vm \\ 0 & I \end{pmatrix}$$

is determined by the action on $(\alpha,\alpha^*,\beta,\beta^*)^T$ of the matrix

$$\Phi = \begin{pmatrix} \phi & 0 & -\psi^* & 0 \\ 0 & \phi^* & 0 & -\psi \\ \psi & 0 & \phi^* & 0 \\ 0 & \psi^* & 0 & \phi \end{pmatrix} \qquad (3.6)$$

We can express the typical multiplier for R^4 in either of the forms $\gamma(t_1,t_2) = \exp(it_1^TAt_2)$ or $\gamma'(z_1,z_2) = \exp(iz_1^TBz_2)$, where $\gamma(t_1,t_2) \equiv \gamma'(z_1,z_2)$ and A is a real skew-symmetric matrix. If γ is to be lifted to G it must satisfy Eq. (3.5). By using the above matrix Φ one may easily check that A and B have to have the forms

$$A = \begin{pmatrix} 0 & -2a & q & r \\ 2a & 0 & r & -q \\ -q & -r & 0 & -2a \\ -r & q & 2a & 0 \end{pmatrix},$$
$$B = \begin{pmatrix} 0 & ia & c & 0 \\ -ia & 0 & 0 & c^* \\ -c & 0 & 0 & ia \\ 0 & -c^* & -ia & 0 \end{pmatrix} \qquad (3.7)$$

where a, q, and r are real, and $c = \frac{1}{2}(q + ir)$. thus $\mathrm{SU}(2)_A R^4$ has a three-parameter family of multipliers.

Using these results we can obtain the function F of Eq. (2.5). Now Eq. (3.4) implies that $F(X,Y)$ vanishes unless X and Y belong to the Lie algebra of R^4. Thus if T_μ, T_ν, T_σ, and T_τ denote the generators of translations along the coordinate axes, we find

$$F(T_\mu,T_\nu) = -4ai, \quad F(T_\mu,T_\sigma) = 2iq, \quad F(T_\mu,T_\tau) = 2ir, \qquad (3.8)$$
$$F(T_\nu,T_\sigma) = 2ir, \quad F(T_\nu,T_\tau) = -2iq, \quad F(T_\sigma,T_\tau) = -4ai.$$

Alternatively, in terms of the $Q^{(\#)}_{\pm 1/2}$ of $\mathrm{SU}(2)_A(T_2\times T_2^{\#})$[1,3] defined by

$$Q_{1/2} = \tfrac{1}{2}(T_\tau - iT_\sigma), \quad Q_{-1/2} = \tfrac{1}{2}(T_\nu - iT_\mu), \qquad (3.9)$$
$$Q^{\#}_{1/2} = -\tfrac{1}{2}(T_\nu + iT_\mu), \quad Q^{\#}_{-1/2} = \tfrac{1}{2}(T_\tau + iT_\sigma),$$

we can extend F by linearity to the complexification to obtain the nonzero values

$$F(Q_{1/2},Q_{-1/2}) = iq - r, \quad F(Q^{\#}_{1/2},Q^{\#}_{-1/2}) = iq + r, \qquad (3.10)$$
$$F(Q_{1/2},Q^{\#}_{-1/2}) = -2a, \quad F(Q_{-1/2},Q^{\#}_{1/2}) = 2a.$$

We remark here that F could be determined purely at the Lie algebra level by the requirement that the central extension of LG by the real numbers, essentially defined by Eq. (2.5), is a Lie algebra, i.e., that F be compatible with the Jacobi identity for the commutators. One then still has the problem of lifting F to the whole group G to form the multiplier ω. This approach is in general inadequate since it is possible for G to have a nontrivial central extension even though LG does not. For instance, although the Lie algebra of SO(3) has no nontrivial central extension, the group SO(3) itself does have a nontrivial central extension, namely SU(2), and the corresponding projective representations of SO(3) are precisely the spin representations. The group SU(2), of course, does not possess a nontrivial central extension.

We conclude this section by showing that it is possible to make a definite choice of the parameters, a, q, r without losing any essential information. The general idea is that if σ is an automorphism of the group G, then for every multiplier ω we can define a new multiplier ω_σ by $\omega_\sigma(g_1,g_2) = \omega(\sigma(g_1),\sigma(g_2))$ for all $g_1,g_2 \in G$. This gives us a new equivalence relation on multipliers which reduces the number of equivalence classes that we need to consider. For some groups, inner automorphisms can be of use. However, in the present case, because R^4 is Abelian and because of Eq. (3.4), multipliers of $SU(2)_\Lambda R^4$ are invariant under inner automorphisms; therefore we look to outer automorphisms. We recall[1] that $SU(2)_\Lambda R^4$ is a subgroup of $[SU(2)\times SU(2)]_\Lambda R^4$, in which the two SU(2) groups commute but act differently on R^4. The second SU(2) group produces a group of outer automorphisms of $SU(2)_\Lambda R^4$ which fixes the first factor, but which acts on R^4 and therefore also on the multiplier. The action on R^4 is defined by $m \to mv^{-1}$, for $m \in R^4$ and $v \in SU(2)$. By these means we are able to reduce to zero the parameters q and r defining the multiplier and which appear in Eqs. (3.7), (3.8), and (3.10). Finally, the parameter a can be reduced to the value $\frac{1}{2}$ by using those outer automorphisms which fix SU(2) but scale R^4. Thus the matrix A takes the form

$$J \oplus J, \quad \text{where } J = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

One may see this at the Lie algebra level as follows: Suppose that the function F has the values for $Q^{(\#)}_{\pm 1/2}$ given in Eq. (3.10). Then if one chooses an alternative basis for the $T_2 \times T_2^{\#}$ algebra by

$$Q'_{\pm 1/2} = \alpha Q_{\pm 1/2} + \beta Q^{\#}_{\pm 1/2},$$
$$Q^{\#\prime}_{\pm 1/2} = -\beta^* Q_{\pm 1/2} + \alpha Q^{\#}_{\pm 1/2}, \tag{3.11}$$

where

$$\alpha = \frac{(r^2+q^2)^{1/2}}{\{2(4a^2+r^2+q^2)[(4a^2+r^2+q^2)^{1/2}-2a]\}^{1/2}} \tag{3.12}$$

and

$$\beta = \frac{-(r-iq)}{\{2(4a^2+r^2+q^2)[(4a^2+r^2+q^2)^{1/2}+2a]\}^{1/2}}, \tag{3.13}$$

then it is easy to check that

$$F(Q'_{1/2},Q'_{-1/2}) = 0, \quad F(Q^{\#\prime}_{1/2},Q^{\#\prime}_{-1/2}) = 0, \tag{3.14}$$
$$F(Q'_{1/2},Q^{\#\prime}_{-1/2}) = -1, \quad F(Q'_{-1/2},Q^{\#\prime}_{1/2}) = 1.$$

The outer automorphisms defined above for the group R^4 clearly correspond to the outer automorphism $Q^{(\#)}_{\pm 1/2} \to Q^{(\#)\prime}_{\pm 1/2}$ of the Lie algebra $T_2 \times T_2^{\#}$.

4. PROJECTIVE REPRESENTATIONS OF $SU(2)_\Lambda R^4$

In order to obtain the PIUR of $SU(2)_\Lambda R^4$ we must first have an understanding of the PIUR of R^4. We use the generalization of the results of Backhouse and Bradley[14,15] as given by Baggett and Kleppner.[16] The multiplier γ, defined by the matrix A above, is nondegenerate in the sense that if $\gamma(t_1,t_2) = 1$ for all $t_2 \in R^4$, then $t_1 = 0$. For such a multiplier we have the nice result that up to unitary equivalence there exists a unique PIUR of R^4, π say, with that multiplier. This is a generalization of the Stone–von Neumann theorem for the uniqueness of the canonical commutation relations in the Weyl exponential form.

From the results of Baggett and Kleppner, in order to construct π we must decompose R^4 into two complementary subspaces w_+ and w_-, both isotropic with respect to ω—ω is unity on w_+ and w_-. An obvious choice in our case is $w_+ = \{(0,\nu,0,\tau) : (\nu,\tau) \in R^2\}$ and $w_- = \{(\mu,0,\sigma,0) : (\mu,\sigma) \in R^2\}$. Then π is obtained by ω-inducing the trivial representation of w_+ up to R^4. π acts on $L^2(R^4/w_+) \equiv L^2(w_-) \equiv L^2(R^2)$ and is given by

$$(\pi(\mu,\nu,\sigma,\tau)f)(x,y) = \exp[-2i(x\nu + y\tau)] \times \exp[-i(\mu\nu + \sigma\tau)] f(x+\mu, y+\sigma) \tag{4.1}$$

for $(x,y) \in R^2$ and $f \in L^2(R^2)$. At the Lie algebra level we quickly check that T_ν, T_τ are represented by the operations of multiplication by $-2ix$ and $-2iy$, respectively, and T_μ, T_σ are represented by $(\partial/\partial x)$ and $(\partial/\partial y)$, respectively. From these we see that $Q^{(\#)}_{\pm 1/2}$ are represented by the operators

$$q_{1/2} = -\tfrac{1}{2}i\left(2y + \frac{\partial}{\partial y}\right), \quad q_{-1/2} = -\tfrac{1}{2}i\left(2x + \frac{\partial}{\partial x}\right), \tag{4.2}$$
$$q^{\#}_{1/2} = -\tfrac{1}{2}i\left(-2x + \frac{\partial}{\partial x}\right), \quad q^{\#}_{-1/2} = -\tfrac{1}{2}i\left(2y - \frac{\partial}{\partial y}\right).$$

Note that the element $X = Q_{1/2}Q^{\#}_{-1/2} - Q_{-1/2}Q^{\#}_{1/2}$ from the center of the enveloping algebra of $SU(2)_\Lambda(T_2 \times T_2^{\#})$ is now represented by the noninvariant operator

$$X = \frac{1}{4}\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\right) - (x^2+y^2) - \mathbf{1}, \tag{4.3}$$

so that $X + \mathbf{1}$ is just the negative of the Hamiltonian for the two-dimensional isotropic harmonic oscillator written in certain appropriate units: $X + \mathbf{1} = -H$.

We are now in a position to follow the Mackey procedure[10] for finding PIUR of semidirect product groups $P_\Lambda N$:

(i) For each PIUR, π, of N, find the little cogroup $P_\pi = \{p \in P : \to \pi_p(n) = \pi(pnp^{-1})$ is equivalent to $\pi\}$.

(ii) If $p \in P_\pi$, there exists a unitary operator U_p such that $\pi(pnp^{-1}) = U_p \pi(n) U_p^{-1}$ for all $n \in N$. Then $p \to U_p$ gives a projective unitary representation of P_π with multiplier ω', say.

(iii) Find PIUR of P_π, V say, with multiplier ω'', where $\omega'\omega'' = \omega$ on P_π. Then $np \to [\pi(n)U_p] \otimes V_p$ is a PIUR of $P_{\pi\Lambda}N$ with multiplier ω.

(iv) Induce this representation up to $P_\Lambda N$. The resulting representations are irreducible, and for certain classes of groups the procedure is exhaustive.

Our case is rather special; in the first place R^4 has, up to unitary equivalence, a unique PIUR with multiplier ω. It follows that π_v must be equivalent to π for all $v\in SU(2)$, hence $P_\pi = P = SU(2)$. This means that the induction step (iv) is not needed. In the second place, the group SU(2) has only trivial multipliers, which means that we can arrange for the multipliers ω' and ω'' to be unity. The IUR of SU(2) are, of course, well known, so it remains to discover the intertwining representation $v\to U_v$ defined in (ii) above. As before it is convenient to work with the complex formalism. Thus we first compute the representation π, given in (4.1), in terms of $z=(\alpha,\alpha^*,\beta,\beta^*)$, by

$$(\pi(z)f)(x,y)=\exp\{-i[x(\beta+\beta^*)+y(\alpha+\alpha^*)]\}$$
$$\times\exp\{-\tfrac{1}{4}[(\beta^2-\beta^{*2})+(\alpha^2-\alpha^{*2})]\}$$
$$\times f\Big(x+\frac{1}{2i}(\beta-\beta^*),y+\frac{1}{2i}(\alpha-\alpha^*)\Big), \tag{4.4}$$

acting on $f\in L^2(R^2)$.

Let

$$v=\begin{pmatrix}\phi & -\psi^*\\ \psi & \phi^*\end{pmatrix},\quad \phi\phi^*+\psi\psi^*=1,$$

be an element of SU(2). Then, using the matrix Φ given in Eq. (3.6), we find

$$\begin{aligned}(\pi_v(z)f)(x,y)
&=\exp\{-i[x(\psi\alpha+(\psi\alpha)^*+\phi^*\beta+\phi\beta^*)\\
&\quad+y(\phi\alpha+(\phi\alpha)^*-\psi^*\beta-\psi\beta^*)]\}\exp\{-\tfrac{1}{4}[\psi^2\alpha^2\\
&\quad+\phi^{*2}\beta^2+2\psi\phi^*\alpha\beta-\phi^2\beta^{*2}-\psi^{*2}\alpha^{*2}-2\\
&\quad\times\phi(\psi\alpha\beta)^*+\phi^2\alpha^2+\psi^{*2}\beta^2\\
&\quad-2\phi\psi^*\alpha\beta-\psi^2\beta^{*2}-\phi^{*2}\alpha^{*2}+2\psi(\phi\alpha\beta)^*]\}\\
&\quad\times f\Big(x+\frac{1}{2i}[\psi\alpha-(\psi\alpha)^*+\phi^*\beta-\phi\beta^*],y\\
&\quad+\frac{1}{2i}[\phi\alpha-(\phi\alpha)^*+\psi\beta^*-\psi^*\beta]\Big).\end{aligned} \tag{4.5}$$

This very complicated action at the R^4 group level implies the following operator relations at the level of the Lie algebra $T_2\times T_2^{\#}$:

$$\begin{aligned}\pi_v(Q_{1/2}^{(\#)})&=\phi q_{1/2}^{(\#)}+\psi q_{-1/2}^{(\#)},\\
\pi_v(Q_{-1/2}^{(\#)})&=-\psi^*q_{1/2}^{(\#)}+\phi^{\#}q_{-1/2}^{(\#)},\end{aligned} \tag{4.6}$$

where $q_{1/2}^{(\#)}=\pi(Q_{\pm 1/2}^{(\#)})$ are given by Eqs. (4.2), and $\pi_v(Q_{\pm 1/2}^{(\#)})$ means $U_v q_{\pm 1/2}^{(\#)}(U_v)^{-1}$. Thus the q's transform as spinors under the representation $v\to U_v$ of SU(2). These relations allow us to find commutation relations for the generators of SU(2) in the representation U. If A_1, A_2, A_3 are the generators of SU(2) as given by Backhouse,[3] that is

$$A_1=\begin{pmatrix}0 & -\frac{1}{2}i\\ -\frac{1}{2}i & 0\end{pmatrix},\quad A_2=\begin{pmatrix}0 & \frac{1}{2}\\ -\frac{1}{2} & 0\end{pmatrix},$$
$$A_3=\begin{pmatrix}\frac{1}{2}i & 0\\ 0 & -\frac{1}{2}i\end{pmatrix}, \tag{4.7}$$

so that

$$L_\pm=\pm A_2+iA_1,\quad L_0=-iA_3, \tag{4.8}$$

then we find from Eqs. (4.6) that their representatives satisfy

$$\begin{aligned}[U(A_1),q_{\pm 1/2}^{(\#)}]&=-\tfrac{1}{2}iq_{\mp 1/2}^{(\#)},\\
[U(A_2),q_{\pm 1/2}^{(\#)}]&=\mp\tfrac{1}{2}iq_{\mp 1/2}^{(\#)},\\
[U(A_3),q_{\pm 1/2}^{(\#)}]&=\pm\tfrac{1}{2}iq_{\pm 1/2}^{(\#)}.\end{aligned} \tag{4.9}$$

Using the irreducibility of the representation π we find the following unique solution to (4.9):

$$\begin{aligned}U(A_1)&=i\Big(xy-\frac{1}{4}\frac{\partial^2}{\partial x\partial y}\Big),\\
U(A_2)&=-\frac{1}{2}\Big(x\frac{\partial}{\partial y}-y\frac{\partial}{\partial x}\Big),\\
U(A_3)&=\frac{i}{2}\Big[\frac{1}{4}\Big(\frac{\partial^2}{\partial y^2}-\frac{\partial^2}{\partial x^2}\Big)-(y^2-x^2)\Big],\end{aligned} \tag{4.10}$$

acting on analytic vectors in $L^2(R^2)$. In terms of $\{L_0,L_\pm\}$, Eqs. (4.10) become

$$\begin{aligned}U(L_0)&=\frac{1}{8}\Big(\frac{\partial^2}{\partial y^2}-\frac{\partial^2}{\partial x^2}\Big)-\tfrac{1}{2}(y^2-x^2),\\
U(L_\pm)&=\Big(\frac{1}{4}\frac{\partial^2}{\partial x\partial y}-xy\Big)\pm\frac{1}{2}\Big(y\frac{\partial}{\partial x}-x\frac{\partial}{\partial y}\Big).\end{aligned} \tag{4.11}$$

Note that Eq. (4.10) are precisely what are obtained when SU(2) is realized as the symmetry group of the two-dimensional isotropic harmonic oscillator–see Jauch and Hill,[17] Major.[9] Observe further that Eqs. (4.10) and (4.11) can in fact be constructed as bilinear products of the $q^{(*)}$'s; for instance:

$$\begin{aligned}&U(L_0)=\tfrac{1}{2}(q_{1/2}q_{-1/2}^{\#}+q_{-1/2}q_{1/2}^{\#}),\quad U(L_+)=-q_{1/2}q_{1/2}^{\#},\\
&U(L_-)=q_{-1/2}q_{-1/2}^{\#}.\end{aligned} \tag{4.12}$$

The explicit decomposition into irreducibles of the intertwining representation of SU(2) was achieved algebraically by Major,[9] using well known properties of Hermite functions. The representation decomposes as

$$\bigoplus_{l=0,1/2,1,3/2,\cdots} D^l,$$

where D^l is the $(2l+1)$-dimensional IUR of SU(2). In view of Major's work,[9] we restrict the exposition of our independent group theoretical derivation of this decomposition to a statement of our main results. We found that the decomposition is most easily obtained using certain functions which crop up rather naturally: define, for $(a,b)\in R^2$, a function $G_{a,b}$ by

$$G_{a,b}(x,y) = \exp[-(a^2+b^2)]\exp[-(x^2+y^2)] \times \exp[2\sqrt{2}(ax+by)], \quad (4.13)$$

for $(x,y)\in R^2$. $G_{a,b}$ is a simultaneous eigenfunction of $q_{\pm 1/2}$. Then, for

$$v = \begin{pmatrix} \phi & -\psi^* \\ \psi & \phi^* \end{pmatrix} \in \mathrm{SU}(2),$$

we find

$$U_v G_{a,b} = G_{\phi a+\psi^* b,\, -\psi a+\phi^* b}. \quad (4.14)$$

Now if we put $H_{m,n}(x,y) = \exp[-(x^2+y^2)]H_m(\sqrt{2}x) \times H_n(\sqrt{2}y)$, where H_i is the ith Hermite polynomial, then there is the well known expansion

$$G_{a,b} = \sum_{m,n=0}^{\infty} \frac{1}{m!}\frac{1}{n!} H_{m,n} a^m b^n. \quad (4.15)$$

These results can easily be put to work to show that, if $W_N = \{\text{linear span of } H_{m,n} \text{ for } m+n=N\}$, then W_N is an invariant subspace for U and, moreover, it carries the representation $D^{N/2}$ of SU(2), $N\geqslant 0$. This is the last step in the determination of the equivalence classes of PIUR of $\mathrm{SU}(2)_\Lambda R^4$: The classes are labelled by an integer p, and as representatives we choose the representations π^p defined by

$$\begin{pmatrix} v & m \\ 0 & I \end{pmatrix} \to (\pi(m)U_v) \otimes D^{p/2}(v) \quad (4.16)$$

acting on $L^2(R^2)\otimes C^{p+1}$, where π is defined by the equivalent formula (4.1), $U = \bigoplus_{l=0,1/2,1,\cdots} D^l$.

It is interesting to determine the SU(2) content of π^p. In fact

$$\pi^p \downarrow \mathrm{SU}(2) = \Big(\bigoplus_{l=0,1/2,1,\cdots} D^l \Big) \otimes D^{p/2}$$

$$= D^0 \oplus 2D^{1/2} \oplus 3D^1 \oplus \cdots \oplus pD^{(p-1)/2} \oplus (p+1) \times \Big(\bigoplus_{k\geqslant p/2} D^k \Big). \quad (4.17)$$

As we shall see in Sec. 6, this agrees with the results obtained by applying shift operator techniques to the Lie algebra $C(\mathrm{SU}(2)_\Lambda(T_2\times T_2^{\#}))$.

The representation π^0 makes $\mathrm{SU}(2)_\Lambda R^4$ a spectrum generating group for the two-dimensional isotropic harmonic oscillator. This follows because: (1) we know from elementary quantum mechanics that W_N is an energy eigenspace for the oscillator; (2) the commutation relations imply that $q^{(\#)}_{\pm 1/2}$ map W_N into $W_{N-1}+W_{N+1}$, for $N\geqslant 0$, where $W_{-1} \equiv \{0\}$. Finally, it is worth pointing out that the properties of Hermite polynomials used by Major[9] are consequences of our approach.

5. FURTHER PROPERTIES OF THE LIE ALGEBRA $C(\mathrm{SU}(2)_\Lambda(T_2\times T_2^{\#}))$

We now examine the Lie algebra $C(\mathrm{SU}(2)_\Lambda(T_2\times T_2^{\#}))$ in more detail, writing down its SU(2) scalar operators, invariant, and shift operators, in preparation for an analysis of the IUR obtained in the previous section.

In the following we shall denote the representatives in any representation π of L_0, $L_\pm$, $Q^{(\#)}_{\pm 1/2}$ by l_0, $l_\pm$, $q^{(\#)}_{\pm 1/2}$, respectively, and retain the same symbols as were used in previous sections for the SU(2) Casimir L^2 and SU(2) scalar operator X. For convenience, we summarize here the nonvanishing commutation relations for $C(\mathrm{SU}(2)_\Lambda(T_2\times T_2^{\#}))$, which are

$$[l_0,l_\pm] = \pm l_\pm, \quad [l_+,l_-] = 2l_0, \quad [q_{\pm 1/2}, q^{\#}_{\mp 1/2}] = \mp \mathbf{1},$$
$$[l_0, q^{(\#)}_{\pm 1/2}] = \pm \tfrac{1}{2} q^{(\#)}_{\pm 1/2}, \quad [l_\pm, q^{(\#)}_{\mp 1/2}] = q^{(\#)}_{\pm 1/2}. \quad (5.1)$$

We choose the same hermiticity relations as were satisfied by the above operators in the realization of the representation π^0 by the operators given in Eqs. (4.2) and (4.11) acting on $L^2(R^2)$, namely

$$l_0^\dagger = l_0, \quad l_\pm^\dagger = l_\mp, \quad q^\dagger_{\pm 1/2} = \mp q^{\#}_{\mp 1/2}. \quad (5.2)$$

Note that the alternative choice, $q^\dagger_{\pm 1/2} = \pm q^{\#}_{\mp 1/2}$, would lead to exactly the same Lie algebra but with $q_{\pm 1/2}$ and $q^{\#}_{\pm 1/2}$ interchanged, so we are losing no generality by considering only the hermiticity relations (5.2).

Just as for the case of $\mathrm{SU}(2)_\Lambda(T_2\times T_2^{\#})$[1], the following operators are SU(2) scalars:

$$X = q_{1/2}q^{\#}_{-1/2} - q_{-1/2}q^{\#}_{1/2}, \quad (5.3)$$

$$Y_0 = q_{1/2}q^{\#}_{-1/2}l_0 + q_{-1/2}q^{\#}_{1/2}l_0 - q_{1/2}q^{\#}_{1/2}l_- + q_{-1/2}q^{\#}_{-1/2}l_+, \quad (5.4)$$

$$Y_+ = -(2q_{1/2}q_{-1/2}l_0 + q_{-1/2}q_{-1/2}l_+ - q_{1/2}q_{1/2}l_-), \quad (5.5)$$

$$Y_- = 2q^{\#}_{1/2}q^{\#}_{-1/2}l_0 + q^{\#}_{-1/2}q^{\#}_{-1/2}l_+ - q^{\#}_{1/2}q^{\#}_{1/2}l_-, \quad (5.6)$$

but unlike the case of $\mathrm{SU}(2)_\Lambda(T_2\times T_2^{\#})$, Y_0 and $Y_\pm$ do not satisfy the commutation relations of an extra SU(2) algebra. Instead, we have

$$[Y_0,Y_+] = XY_+ = Y_+(X+2), \quad (5.7)$$

$$[Y_0,Y_-] = -(X+2)Y_- = -Y_-X, \quad (5.8)$$

$$[Y_+,Y_-] = 2(X+1)(Y_0 - 2L^2). \quad (5.9)$$

one may also easily check that

$$[X,Y_0] = 0, \quad [X,Y_\pm] = \pm 2Y_\pm, \quad (5.10)$$

so $Y_\pm$ raise the eigenvalue of X by ± 2.

In addition to these operators, $C(\mathrm{SU}(2)_\Lambda(T_2\times T_2^{\#}))$ possesses a single independent invariant I_4 given by

$$I_4 = X^2 + 2X - 4Y_0 + 4L^2. \quad (5.11)$$

Thus Y_0 and X are diagonal together. Finally, one may easily show that $Y^\dagger_\pm = Y_\mp$, whereas I_4, Y_0, and X are Hermitian operators, and $(X+2\mathbf{1})$ is also negative semidefinite.

The shift operators we shall need in order to analyze the IUR of $C(\mathrm{SU}(2)_A(T_2\times T_2^{\#}))$ are particular cases of more general operators derived and discussed by Hughes and Yadegar.[8] These shift the eigenvalues l and m of R and l_0 by ± 1, where $R(R+1)\equiv L^2$, when acting to the right on eigenstates of these operators, and are given by

$$0^{(\#)\,1/2,1/2} = q_{1/2}^{(\#)}(R+l_0+1) + q_{-1/2}^{(\#)}l_+, \tag{5.12}$$

$$0^{(\#)\,-1/2,-1/2} = -q_{-1/2}^{(\#)}(R+l_0) + q_{1/2}^{(\#)}l_-. \tag{5.13}$$

We shall find it more convenient to consider the normalized operators

$$A_l^{(\#)\,1/2} = (l+m+1)^{-1/2}0_{l,m}^{(\#)\,1/2,1/2},$$
$$A_l^{(\#)\,-1/2} = (l+m)^{-1/2}0_{l,m}^{(\#)\,-1/2,-1/2}, \tag{5.14}$$

in which the internal structure of IUR of SU(2) has effectively been divided out.

Various L^2-commuting products of these operators, which we shall need in the following section, are

$$A_{l-1/2}^{1/2}A_l^{-1/2} = A_{l+1/2}^{-1/2}A_l^{1/2} = Y_+,$$
$$A_{l-1/2}^{\#\,1/2}A_l^{\#\,-1/2} = A_{l+1/2}^{\#\,-1/2}A_l^{\#\,1/2} = -Y_-, \tag{5.15}$$
$$A_{l-1/2}^{1/2}A_l^{-1/2} = -Y_0 - lX,$$
$$A_{l-1/2}^{\#\,1/2}A_l^{-1/2} = -Y_0 + l(X+2), \tag{5.16}$$
$$A_{l+1/2}^{-1/2}A_l^{\#\,1/2} = -Y_0 + (l+1)X,$$
$$A_{l+1/2}^{\#\,-1/2}A_l^{1/2} = -Y_0 - (l+1)(X+2), \tag{5.17}$$

where it is assumed that all operators act to the right on eigenstates of R. They also satisfy the commutation relations

$$[X,A_l^{\pm 1/2}] = A_l^{\pm 1/2}, \quad [X,A_l^{\#\,\pm 1/2}] = -A_l^{\#\,\pm 1/2}, \tag{5.18}$$

and so shift the eigenvalues of X by ± 1.

Finally, using methods similar to those employed in the general case by Hughes and Yadegar,[8] one may deduce from (5.2) the following hermiticity relations of the shift operators:

$$\langle\gamma',l\pm\tfrac{1}{2}|A_l^{\pm 1/2}|\gamma,l\rangle = \pm\frac{(2l+1)}{2(l+\frac{1}{2}\pm\frac{1}{2})}\langle\gamma',l\pm\tfrac{1}{2}|(A_{l\pm 1/2}^{\#\,\mp 1/2})^\dagger|\gamma,l\rangle, \tag{5.19}$$

$$\sum_{\gamma'}|\langle\gamma',l\pm\tfrac{1}{2}|A_l^{\pm 1/2}|\gamma,l\rangle|^2 = \pm\frac{(2l+1)}{2(l+\frac{1}{2}\pm\frac{1}{2})}\langle\gamma,l|A_{l\pm 1/2}^{\#\,\mp 1/2}A_l^{\pm 1/2}|\gamma,l\rangle, \tag{5.20}$$

$$\sum_{\gamma'}|\langle\gamma',l\pm\tfrac{1}{2}|A_l^{\#\,\pm 1/2}|\gamma,l\rangle|^2 = \mp\frac{(2l+1)}{2(l+\frac{1}{2}\pm\frac{1}{2})}\langle\gamma,l|A_{l\pm 1/2}^{\mp 1/2}A_l^{\#\,\pm 1/2}|\gamma,l\rangle, \tag{5.21}$$

where γ denotes an additional set of state labelling parameters, which will be specified in the following section, and the m-values of the states have been suppressed. Thus $A^{\#\,-1/2}A^{1/2}$ and $A^{1/2}A^{\#\,-1/2}$ are positive semidefinite, and $A^{\#1/2}A^{-1/2}$, $A^{-1/2}A^{\#1/2}$ negative semidefinite operators. This completes the algebraic apparatus that will be needed in the next section for the analysis of the IUR of $C(\mathrm{SU}(2)_A(T_2\times T_2^{\#}))$.

6. ANALYSIS OF THE IUR OF $C(\mathrm{SU}(2)_\Lambda(T_2\times T_2^{\#}))$

Since $C(\mathrm{SU}(2)_A(T_2\times T_2^{\#}))$ has, apart from the trivial invariant **1**, a single independent invariant I_4 and one further independent SU(2) scalar X, the state label γ will consist of a pair (p,s), where p labels the IUR and s serves to distinguish between states within an IUR corresponding to the same degenerate l-value. We define $-(p+2)$ to be the eigenvalue of X for the state of minimum l-value, $l=\underline{l}$, say; since $(X+2\mathbf{1})$ is a Hermitian negative semidefinite operator, p must be real and nonnegative. We can define s uniquely by

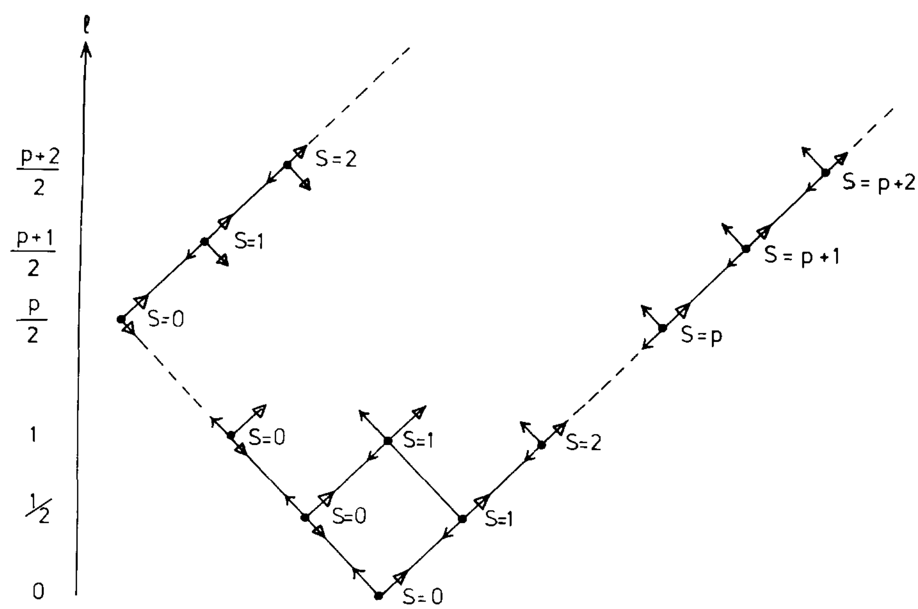

FIG. 1. States of the IUR π^p of $C(\mathrm{SU}(2)_A(T_2\times T_2^{\#}))$. Ordinary arrows indicate the actions of $A^{\pm 1/2}$, solid arrows indicate the actions of $A^{\#\,\pm 1/2}$.

requiring $s = 0$ for the $l = \underline{l}$ state, $A^{1/2}$ and $A^{\# -1/2}$ to leave s unchanged, $A^{\# 1/2}$ to raise the value of s by 1, and $A^{-1/2}$ to lower it by 1, as depicted in Fig. 1. From this it is clear that s must lie somewhere in the range $0 \leqslant s \leqslant 2(l - \underline{l})$ for any state with given l value.

We start the analysis by considering the $l = \underline{l}$ state, which must be annihilated by $A^{(\#)-1/2}$. Thus we have

$$A^{\# 1/2}_{\underline{l}-1/2} A^{-1/2}_{\underline{l}} \mid p;\underline{l},0\rangle = A^{1/2}_{\underline{l}-1/2} A^{\# -1/2}_{\underline{l}} \mid p;\underline{l},0\rangle = 0$$

which, together with

$$X \mid p;\underline{l},0\rangle = -(p+2) \mid p;\underline{l},0\rangle,$$

yields, on substitution into Eqs. (5.16),

$$Y_0 \mid p;\underline{l},0\rangle = \underline{l}(p+2) \mid p;\underline{l},0\rangle = -\underline{l}p \mid p;\underline{l},0\rangle.$$

Since $p \geqslant 0$, the only possibility is that $\underline{l} = 0$, in agreement with Eq. (4.17). Using (5.11) we now get

$$I_4 \mid p;0,0\rangle = p(p+2) \mid p;0,0\rangle; \tag{6.1}$$

thus $p(p+2)$ is the eigenvalue of I_4 for all states of the IUR labelled by p.

Now consider the general $|p;l,s\rangle$ state. From the definition of s and the fact that $0 \leqslant s \leqslant 2l$, we see that $|p;l,s\rangle \propto (A^{1/2})^{(2l-s)}(A^{\# 1/2})^s|p;0,0\rangle$. From the commutation relations (5.18), we see that $(A^{1/2})^{(2l-s)}(A^{\# 1/2})^s$ raises the eigenvalue of X by $(2l - 2s)$, so

$$X \mid p;l,s\rangle = -(p - 2l + 2s + 2) \mid p;l,s\rangle. \tag{6.2}$$

Since $X + 2\mathbf{1}$ is negative semidefinite, we must have $(p - 2l + 2s) \geqslant 0$. Since $s \geqslant 0$, this is always satisfied if $l \leqslant \frac{1}{2}p$, but if $l > \frac{1}{2}p$, it will not be satisfied by the $|p;l,0\rangle \propto (A^{1/2})^{2l}$ $|p;0,0\rangle$ state, for instance. Thus, in order to preserve the Hermiticity conditions (i.e., to ensure that the IR will in fact be an IUR), it must be impossible to act on $|p;0,0\rangle$ successively with $A^{1/2}$ more than p times, and so $A^{1/2}$ must annihilate some state $|p;l,0\rangle$ with $l \leqslant p/2$, i.e., this state must correspond to the zero eigenvalue of $A^{\# -1/2}A^{1/2}$. Now, from Eq. (5.11), we have

$$Y_0|p;l,s\rangle = [s(s + p - 2l + 1) - l(p - 2l)] \mid p;l,s\rangle, \tag{6.3}$$

and so, on substitution of X and Y_0 into the second of Eqs. (5.17), we get

$$A^{\# -1/2}A^{1/2} \mid p;l,s\rangle = -(2l - s - p)(2l - s + 1) \mid p;l,s\rangle. \tag{6.4}$$

Hence $A^{\# -1/2}A^{1/2} \mid p;l,0\rangle$ vanishes only when $l = p/2$. Hence we see that hermiticity is satisfied only if p is a nonnegative integer, and then $A^{1/2} \mid p;p/2,0\rangle = 0$.

From (6.4), we see also that $A^{1/2} \mid p;l,2l - p\rangle = 0$, so, when $l \geqslant p/2$, $s \geqslant (2l - p)$. On the other hand, from substitution of X and Y_0 into the first of Eqs. (5.17), we get

$$A^{-1/2}A^{\# 1/2} \mid p;l,s\rangle = -(s+1)(s+p+2) \mid p;l,s\rangle. \tag{6.5}$$

so $A^{\# 1/2} \mid p;l,s\rangle$ never vanishes. Thus we see that, whereas for $l \leqslant \frac{1}{2}p$, s lies in the range $0 \leqslant s \leqslant 2l$, for $l \geqslant \frac{1}{2}p$ the range of s is $(2l - p) \leqslant s \leqslant 2l$. This guarantees that $(p - 2l + 2s) \geqslant 0$, so that $(X + 2\mathbf{1})$ is a negative semidefinite operator, as required.

In a similar way, by the use of Eqs. (5.16), we get

$$A^{1/2}A^{\# -1/2} \mid p;l,s\rangle = -(2l - s - p - 1)(2l - s) \mid p;l,s\rangle, \tag{6.6}$$

$$A^{\# 1/2}A^{-1/2} \mid p;l,s\rangle = -s(s+p+1) \mid p;l,s\rangle. \tag{6.7}$$

One may easily check that the range of s guarantees that $A^{\# -1/2}A^{1/2}$ and $A^{1/2}A^{\# -1/2}$ are positive semidefinite, whereas $A^{\# 1/2}A^{-1/2}$ and $A^{-1/2}A^{\# 1/2}$ are negative semidefinite, as is also required by hermiticity.

The set of states obtained are depicted in Fig. 1; they are seen to be in total agreement with the expression (4.17) obtained using induced representations for the PIUR π^p of $SU(2)_\Lambda R^4$. However, Eqs. (6.4)–(6.7) also enable us to obtain the matrix elements within π^p of the operators $q^{(\#)}_{\pm 1/2}$ as follows.

First choose the relative phases of $\mid p;l,s\rangle$, $\mid p;l+\frac{1}{2},s\rangle$, and $\mid p;l+\frac{1}{2},s+1\rangle$ so that the matrix elements $\langle p;l+\frac{1}{2},s \mid A^{1/2}_l |p;l,s\rangle$, and $\langle p;l+\frac{1}{2},s+1 \mid A^{\# 1/2}_l |p;l,s\rangle$ are real and nonnegative. Then using Eqs. (6.4) and (6.5) in (5.20) and (5.21), respectively, yields

$$A^{1/2} \mid p;l,s\rangle = \left(\frac{(2l+1)(s+p-2l)(2l-s+1)}{2(l+1)}\right)^{1/2} \mid p;l+\tfrac{1}{2},s\rangle, \tag{6.8}$$

$$A^{\# 1/2} \mid p;l,s\rangle = \left(\frac{(2l+1)(s+1)(s+p+2)}{2(l+1)}\right)^{1/2} \mid p;l+\tfrac{1}{2},s+1\rangle. \tag{6.9}$$

Using these equations and (6.4) and (6.5) again, further yields

$$A^{\# -1/2} \mid p;l,s\rangle = \left(\frac{(2l+1)(s+p-2l+1)(2l-s)}{2l}\right)^{1/2} \mid p;l-\tfrac{1}{2},s\rangle, \tag{6.10}$$

$$A^{-1/2} \mid p;l,s\rangle = -\left(\frac{(2l+1)s(s+p+1)}{2l}\right)^{1/2} \mid p;l-\tfrac{1}{2},s-1\rangle. \tag{6.11}$$

We can now use Eqs. (5.15) to obtain,

$$Y_+ \mid p;l,s\rangle = -[s(s+p+1)(s+p-2l) \times (2l-s+1)]^{1/2} \mid p;l,s-1\rangle, \tag{6.12}$$

$$Y_- \mid p;l,s\rangle = [(s+1)(s+p+2) \times (s+p-2l)(2l-s)]^{1/2} \mid p;l,s+1\rangle. \tag{6.13}$$

Finally, by going back to the definitions of $A^{(\#) \pm 1/2}$ in terms of $q^{(\#)}_{\pm 1/2}$, we obtain the following results for the action of $q^{(\#)}_{\pm 1/2}$ on $|p;l,s\rangle$:

$$q_{\pm 1/2}\,|p;l,m,s\rangle = \left(\frac{(l\pm m+1)(s+p-2l)(2l-s+1)}{2(l+1)(2l+1)}\right)^{1/2} \times |p;l+\tfrac{1}{2},m\pm\tfrac{1}{2},s\rangle \mp \left(\frac{(l\mp m)s(s+p+1)}{2l(2l+1)}\right)^{1/2} \times |p;l-\tfrac{1}{2},m\pm\tfrac{1}{2},s-1\rangle, \tag{6.14}$$

$$q^{\#}_{\pm 1/2}\,|p;l,m,s\rangle = \left(\frac{(l+m+1)(s+1)(s+p+2)}{2(l+1)(2l+1)}\right)^{1/2} \times |p;l+\tfrac{1}{2},m\pm\tfrac{1}{2},s+1\rangle \pm \left(\frac{(l\mp m)(s+p-2l+1)(2l-s)}{2l(2l+1)}\right)^{1/2} \times |p;l-\tfrac{1}{2},m\pm\tfrac{1}{2},s\rangle. \tag{6.15}$$

This completes the analysis of the representation π^p. Note that, instead of X (or, equivalently, Y_0), one could use either of the Hermitian operators $(Y_+ + Y_-)$ or $i(Y_+ - Y_-)$ to label the states. This was done for the case of SU(3) in an SU(2) basis by Hughes and Yadegar[8]; the eigenvalues of these operators would, however, be rather difficult to obtain since their eigenstates would not be connected by $A^{(\#)\pm 1/2}$ in such a straightforward manner as those of X.

We end this section by futher consideration of the representation π^0 and its connection with the two dimensional isotropic harmonic oscillator. For this IUR, $p=0$, l is nondegenerate, and every state corresponds to the value $s=2l$. The matrix elements of $q^{(\#)}_{\pm 1/2}$ are easily obtained from Eqs. (6.14) and (6.15) to be

$$q_{\pm 1/2}\,|l,m\rangle = \mp[(l\mp m)]^{1/2}\,|l-\tfrac{1}{2},m\pm\tfrac{1}{2}\rangle, \tag{6.16}$$
$$q^{\#}_{\pm 1/2}\,|l,m\rangle = [(l\pm m+1)]^{1/2}\,|l+\tfrac{1}{2},m\pm\tfrac{1}{2}\rangle. \tag{6.17}$$

The relation of this IUR to the oscillator can be seen immediately from the fact that it, alone amongst the IUR of $C(\mathrm{SU}(2)_A(T_2\times T_2^{\#}))$, can be realized by the two boson creation and annihilation operators

$$a_x^\dagger = x + \frac{1}{2}\frac{\partial}{\partial x},\quad a_y^\dagger = y + \frac{1}{2}\frac{\partial}{\partial y},$$
$$a_x = x - \frac{1}{2}\frac{\partial}{\partial x},\quad a_y = y - \frac{1}{2}\frac{\partial}{\partial y},$$

where one puts

$$l_0 = \tfrac{1}{2}(a_x^\dagger a_x - a_y^\dagger a_y),\quad l_- = -a_x^\dagger a_y,\quad l_+ = -a_y^\dagger a_x,$$
$$q_{1/2} = -ia_y^\dagger,\quad q_{-1/2} = -ia_x^\dagger,\quad q^{\#}_{1/2} = ia_x,\quad q^{\#}_{-1/2} = -ia_y.$$

On substitution for the boson operators in terms of x and y, one recovers the expressions for l_0, $l_\pm$, and $q^{(\#)}_{\pm 1/2}$ obtained in Eqs. (4.2) and (4.11). Finally, note that the basis for π^0 chosen in this paper diagonalizes

$$l_0 = \frac{1}{8}\left(\frac{\partial^2}{\partial y^2} - \frac{\partial^2}{\partial x^2}\right) - \tfrac{1}{2}(y^2 - x^2),$$

and not the angular momentum operator

$$-i\left(x\frac{\partial}{\partial y} - y\frac{\partial}{\partial x}\right).$$

7. CONCLUSION

In this paper we have applied both group theoretical and Lie algebraic techniques to classify and analyze the projective unitary irreducible representations of the group $\mathrm{SU}(2)_A R^4$. The application of the Mackey induced representation method has the advantage of greater general applicability, whereas the Lie algebraic shift operator technique facilitates the solution of state labelling problems and the calculation of matrix elements. For the particular group considered here, both methods give exactly the same results, so the Lie algebraic approach has not in fact fallen short of complete generality. This is because we have been dealing with a simply connected group; in cases where the group is not simply connected, as for instance the group SO(3), Lie algebraic techniques are not sufficiently fine to distinguish between representations of the group and of locally isomorphic groups of different connectivity properties.

Another example to which the method of induced representations adopted in Sec. 4 to compute the PIUR of $\mathrm{SU}(2)_A R^4$ can be applied is that of the five-dimensional group $\mathrm{Sl}(2)_A T^2$. This has a Lie algebra $\mathrm{Sl}(2)_A R_2$ with basis $\{L_0, L_\pm, Q_{\pm 1/2}\}$, whose nonvanishing commutation relations are as given in Eq. (1.1), but with the $Q^{\#}_{\pm 1/2}$ omitted. This group also has nontrivial PIUR which yield to the Mackey machinery (see, for instance, Lipsman[18]). In particular, there is one PIUR in which $Q_{\pm 1/2}$ are represented by the operators $-ix$, $\partial/\partial x$, respectively, acting on analytic vectors of $L^2(R)$. At the Lie algebra level, for the intertwining representation of Sl(2), L_+, L_-, and L_0 are represented by

$$\frac{i}{2}x^2,\quad \frac{i}{2}\frac{\partial^2}{\partial x^2},\ \text{and}\ \left(\frac{1}{2}x\frac{\partial}{\partial x} + \frac{1}{4}\mathbf{1}\right),$$

respectively.

This is a rather famous representation, called the metaplectic representation, and was first considered by Shale.[19] On restriction to Sl(2), it reduces to a direct sum of two irreducible representations, one acting on even functions, the other on odd functions, in $L^2(R^2)$. Denoting the eigenvalues of the Sl(2) Casimir L^2 by $l(l+1)$, as we did for SU(2), it can be shown that these two representations of Sl(2) correspond to the values $l = -\frac{1}{4}, -\frac{3}{4}$ [these, of course, would violate the hermiticity conditions for SU(2), but are perfectly permissible for Sl(2)]. An unpleasant feature of the metaplectic representation is that the operator $[\frac{1}{2}x\partial/\partial x + \frac{1}{4}\mathbf{1}]$ has no eigenvectors in $L^2(R^2)$, so it is not possible to further analyze the representations with respect to the subgroup generated by this operator.

The metaplectic representation is a true projective representation, and becomes an ordinary representation only if we go to a twofold cover of Sl(2), as in the case of the spin representations of SO(3). This fact means that as far as parts (ii) and (iii) of the Mackey procedure summarized in Sec. 4 is concerned, the intertwining representation introduces an extra factor system ω' of Sl(2)–a so-called Mackey "obstruction." This determines the factor system ω'' of Sl(2) which is required in part (iii) of the Mackey procedure to restore the overall factor system to ω. To develop this further would take us too far afield. We do note, however, as an aside, that the metaplectic representation together with the operators $-ix$ and $\partial/\partial x$, provide, modulo imaginary constants, a representation of the five-dimensional graded Lie algebra Osp(2,1),[20,21] This is because the above forms of the $Q_{\pm 1/2}$ satisfy the anticommutation relations $\{Q_{1/2}, Q_{-1/2}\} = -4iL_0$, $\{Q_{1/2}, Q_{1/2}\} = 4iL_+$, $\{Q_{-1/2}, Q_{-1/2}\} = -4iL_-$. However, from the above values of l, one easily sees that this representation does not satisfy the usual SU(2)-type hermiticity relations taken for the even part of Osp(2,1).

At the Lie algebra level if, by choosing the hermiticity relations $L_0^\dagger = L_0$, $L_+^\dagger = L_-$, one replaces Sl(2) by SU(2), the shift operator techniques can be applied successfully to the analysis of irreducible representations of the central extension algebra $C(\mathrm{SU}(2)_A T_2)$. These, however, do not yield unitary representations since it is not possible to close $\{Q_{1/2}, Q_{-1/2}\}$ under hermiticity in a manner compatible with the above hermiticity relations for L_0 and $L_\pm$. Even in the case of $C(\mathrm{Sl}(2)_A T_2)$ the shift operator methods may be applied with some success–the above l values for the metaplectic representation were obtaind by shift operator techniques similar to those employed in this paper. We hope to discuss these cases in more detail in a later paper.

[1] J.W.B. Hughes and J. Yadegar, J. Phys. A: Math. Nucl. Gen. **9**, 1569 (1976).
[2] A.O. Barut and H. Kleinert, Phys. Rev. **156**, 1541 (1966).
[3] N.B. Backhouse, J. Phys. A: Math. Nucl. Gen. **10**, L1 (1977).
[4] M. Bander and C. Itzykson, Rev. Mod. Phys. **38**, 330 (1966).
[5] J.W.B. Hughes, Proc. Phys. Soc. **91**, 810 (1967).
[6] N.B. Backhouse, Physica **72**, 505 (1973).
[7] J.W.B. Hughes and J. Yadegar, J. Math. Phys. **19**, 2068 (1978).
[8] J.W.B. Hughes and J. Yadegar, J. Phys. A: Math. Nucl. Gen. **9**, 1581 (1976).
[9] M.E. Major, J. Math. Phys. **18**, 1952 (1977).
[10] G.W. Mackey, Acta Math. **99**, 269 (1958).
[11] V. Bargmann, Ann. Math. **59**, 1 (1954).
[12] K.R. Parthasarathy, Multipliers on Locally Compact Groups, Lecture Notes Math. **93**, (Springer-Verlag, Berlin, 1969).
[13] N.B. Backhouse, Q. J. Math. Oxford **21**, 277 (1970).
[14] N.B. Backhouse and C.J. Bradley, Proc. Am. Math. Soc. **36**, 260 (1972).
[15] N.B. Backhouse, Proc. Am. Math. Soc. **41**, 294 (1973).
[16] L. Baggett and A. Kleppner, J. Func. Anal. **14**, 299 (1973).
[17] J.M. Jauch and E.L. Hill, Phys. Rev. **57**, 641 (1940).
[18] R.L. Lipsman, *Group Representations*, Lecture Notes Math. **388**, (Springer-Verlag, Berlin, 1974).
[19] D. Shale, Trans. Am. Math. Soc. **103**, 149 (1962).
[20] L. Corwin, Y. Ne'eman, and S. Sternberg, Rev. Mod. Phys. **47**, 573 (1975).
[21] M. Scheunert, W. Nahm, and V. Rittenberg, J. Math. Phys. **18**, 155 (1977).

Spacelike geodesics in the Nordström geometry

Naresh Dadhich

Department of Mathematics, University of Poona, Poona-411 007, India
(Received 28 July 1978)

We study here the spacelike geodesics in the Nordström geometry. We show that (i) a radially falling tachyon can hit the singularity $r=0$ if either it is very energetic or the source is highly charged, (ii) nonradial orbits never end into $r=0$, and (iii) unstable circular orbits exist in the region which is bounded from below as well as from above. In the case of $e^2/m^2=1$, scattering cross sections are calculated.

I. INTRODUCTION

In the Schwarzschild geometry spacelike geodesics have recently been studied by many authors.[1-4] Honig *et al.*[3] pursue to completeness the investigation initiated by Hettle and Helliwell[1] and Raychaudhari.[2] They consider the extended Schwarzschild spacetime and demonstrate that spacelike geodesics with apsidal distances $<2m$ which are incident from one asymptotically flat portion of the Schwarzschild manifold must emerge into the other asymptotically flat portion. They argue that a global past-future distinction is not tenable. Narlikar and Dhurandar[4] have discussed the causal and noncausal aspects of tachyon bounce and have shown that a tachyon dropped from a radial distance $<2.56m$ always arrives before it went in whereas the one droppd from radial distance $>3.27m$ always arrives later than its starting instant. They have made out a case for relevance of tachyons in the astrophysical settings.

Davies[5] and Narlikar and Sudarshan[6] have considered the question of tachyons in the cosmological realm, and they have proved that a premordial tachyon in a big-bang universe is eventually reflected at a time barrier. Dhurandhar[7] has investigated propoagation of tachyons inside a white hole while Vaidya[8] and Dadhich[9] have found solutions which are interpreted as describing the gravitational field of a neutral and a charged tachyon respectively.

We shall here study the spacelike geodesics in the Nordström geometry reprsenting the field of a charged black hole. Banerjee and Dutta Choudhury[10] have recently considered spacelike radial and circular geodesics in this geometry. In here would occur nonlinear partial differential equations of third and fourth orders which we shall not attempt to solve exactly; instead we shall focus on the behavior of the effective potential which would tell us almost everything about the motion in a physically illuminating manner. In the particular case of $e^2/m^2=1$, the mathematical manipulations are simple so for this case we do scattering cross-section calculation.

II. THE NORDSTRÖM GEOMETRY

The gravitational field of a charged black hole is described by the line element

$$ds^2=\Delta dt^2-\Delta^{-1}dr^2-r^2(d\theta^2+\sin^2\theta d\phi^2),$$
$$\Delta=1-2m/r+e^2/r^2. \quad (1)$$

We employ the Lagrangian method and write

$$2\mathscr{L}=\Delta\dot{t}^2-\Delta^{-1}\dot{r}^2-r^2(\dot{\theta}^2+\sin^2\theta\dot{\phi}^2), \quad (2)$$

$$\mathscr{L}=\begin{cases}+1 & \text{(timelike)}\\ 0 & \text{(null)}\\ -1 & \text{(spacelike)}\end{cases}$$

Since $\mathscr{L}$ is spherically symmetric and is independent of t, we set $\theta=\pi/2$ and write the conjugate momenta

$$-p_\phi\equiv\frac{\partial\mathscr{L}}{\partial\dot{\phi}}=r^2\dot{\phi}=l \quad (3)$$

and

$$p_t\equiv\frac{\partial\mathscr{L}}{\partial\dot{t}}=\Delta\dot{t}=E. \quad (4)$$

E and l are constants of motion and they refer to energy and angular momentum of the tachyon respectively. From Eqs. (2), (3), and (4) we obtain

$$\dot{r}^2=E^2-\Delta(2\mathscr{L}+l^2/r^2). \quad (5)$$

On differentiation we get

$$\ddot{r}=-\left(\frac{m}{r^2}-\frac{e^2}{r^3}\right)\left(2\mathscr{L}+\frac{l^2}{r^2}\right)+l^2\frac{\Delta}{r^3}$$
$$=-\left(\frac{m}{r^2}-\frac{e^2}{r^3}\right)+\frac{l^2}{r^3}\left(1-\frac{3m}{r}+\frac{2e^2}{r^2}\right). \quad (6)$$

For $e^2/m^2\leqslant 1$, outer and inner horizons are defined by

$$r_\pm=m\,[1\pm(1-e^2/m^2)^{1/2}]. \quad (7)$$

III. GEODESICS

The timelike and null geodesics in the Nordström geometry have already been investigated,[11-14] so we shall study here spacelike geodesics.

A. Radial geodesics: From Eq. (5) it follows, for $l=0$ and $2\mathscr{L}=-1$, that radially incoming and outgoing tachyons will bounce back at

$$r_{1\pm}=[m/(1+E^2)]\,\{1\pm[1-(e^2/m^2)(1+E^2)]^{1/2}\} \quad (8)$$

provided

$$e^2/m^2\leqslant(1+E^2)^{-1}. \quad (9)$$

Let us write Eq. (9) in the form

$$1+E^2=\alpha(m^2/e^2), \quad 0<\alpha\leqslant 1; \quad (10)$$

then Eq. (8) would read as

 0022-2488/79/071555-04$01.00

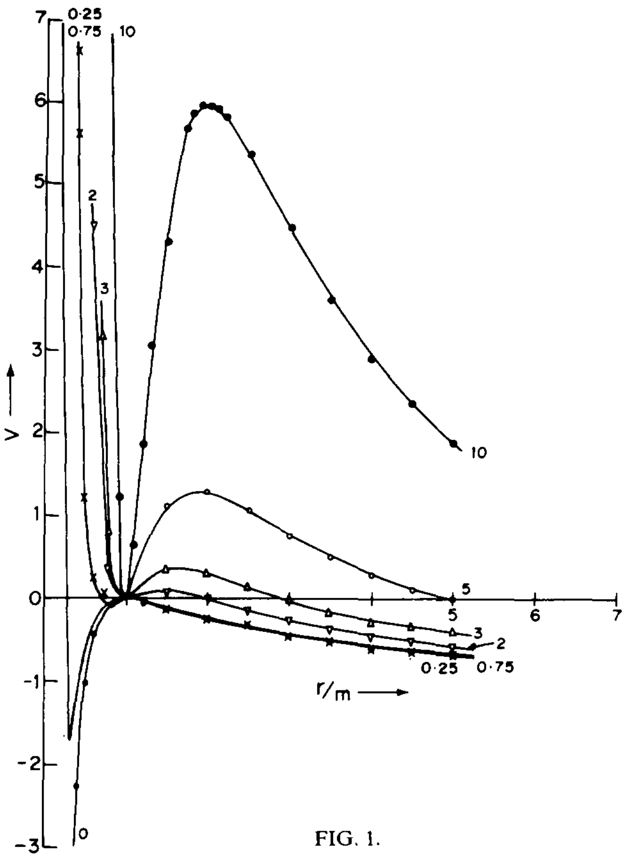

FIG. 1.

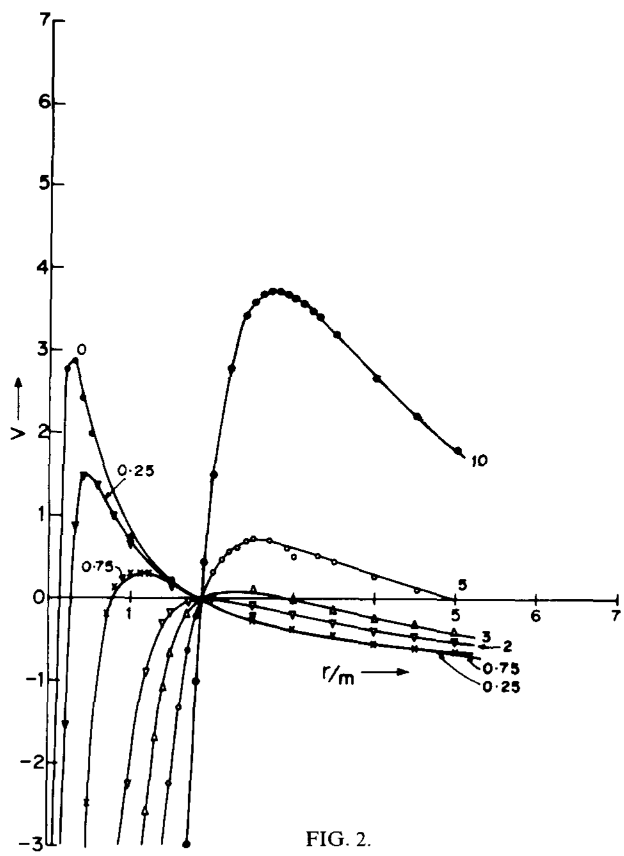

FIG. 2.

FIGS. 1. and 2. The effective potential $V(r)$ is plotted against r/m for various values of l/m (indicated at the tips of the curves). Figure 1 refers to $e^2/m^2 = 1$, and Fig. 2 refers to $e^2/m^2 = 0.1$. In the latter case $V(r)$ will become positive and will go on without limit as $r \to 0$; this has not been shown in Fig. 2.

$$r_{1\pm} = (r_0/\alpha)\,[1 \pm (1-\alpha)^{1/2}], \tag{11}$$

where

$$r_0 = e^2/m. \tag{12}$$

r_0 is the classical radius (e^2/mc^2) of the black hole.

We here observe that

$$r_+ > r_{1+} \geqslant r_0 \geqslant r_{1-} > r_-. \tag{13}$$

It is clear that highly charged ($e^2/m^2 \to 1$) black hole or highly energetic tachyons would violate the condition (9) or (10) which would mean tachyons encounter no bounce and may hit the singularity $r = 0$ which is inaccessible to tardyons. The (radial) tachyons satisfying (10) would have bounce only in the region bounded by the horizons r_+ and r_-.

Here $r_{1\pm}$ mark, respectively, outer and inner turning points for incoming and outgoing tachyons and both $r_{1\pm}$ merge into r_0 when $\alpha = 1$. That is, tachyons satisfying (10) trapped below r_{1-} can never escape and only those which are thrown out above r_{1+} will escape to infinity.

On putting $e = 0$, in Eq. (8) we get back the bounce radius in the Schwarzschild case[4]:

$$r_1 = 2m(1 + E^2)^{-1}. \tag{14}$$

In this case, the bounce always occurs, and no tachyon can hit the singularity $r = 0$.

We shall now consider nonradial trajectories.

B. The effective potential: The effective potential follows from Eq. (5) as

$$V(r) = \Delta\,(l^2/r^2 - 1), \tag{15}$$

which blows up as $r \to 0$ and it tends to -1 as $r \to \infty$ for all values of l. This means that no tachyon could drop down to $r = 0$.

The behavior of $V(r)$ is shown in the Figs. 1 and 2 for the cases $e^2/m^2 = 0.1$ and 1. Setting $V'(r) = dV/dr = 0$, we get a cubic equation

$$mr^3 + r^2(l^2 - e^2) - 3ml^2 r + 2e^2 l^2 = 0. \tag{16}$$

This equation will have maximum two positive roots. We shall not proceed to find the exact roots of this equation. However, behavior of the effective potential is demonstrated in the Figs. 1 and 2. The effective potential has a maximum in the positive values and hence there could occur unstable circular orbits whenever $E^2 = V(r_c)$, where r_c is a positive root of Eq. (16).

The case $e^2/m^2 = 1$: Now Eq. (16) could be factorized to read as

$$(r - m)(mr^2 + l^2 r - 2ml^2) = 0. \tag{17}$$

Since $r = m$ is the null surface, the only root which would give a circular orbit is

$$r_c = (l^2/2m)\,[(1 - 8m^2/l^2)^{1/2} - 1]. \tag{18}$$

This means $r_c > m$ for $l^2/m^2 > 1$ and $r_c < m$ for $l^2/m^2 < 1$.

To find the turning point (r_t) we put $\dot{r} = 0$ in Eq. (5) and write the Equation

$$r_t^4 - \frac{2m}{1+E^2} r_t^3 - \frac{l^2 - e^2}{1+E^2} r_t^2 + \frac{2ml^2}{1+E^2} r_t - \frac{e^2 l^2}{1+E^2} = 0, \quad (19)$$

which could be rewritten as

$$E^2 r_t^4 + (r_t^2 - l^2)(r_t^2 - 2mr_t + e^2) = 0.$$

This implies that if $r_t > r_+$ or $r_t < r_-$ then $r_t^2 < l^2$ and for $r_+ \geqslant r_t \geqslant r_-$ we should have $r_t^2 > l^2$.

C. Limits for circular orbits: For circular orbits both $\dot{r}$ and $\ddot{r}$ should simultaneously vanish which imply that

$$r_{c+} < (3m/2)\,[1 + (1 - 8e^2/9m^2)^{1/2}] \quad (20)$$

and

$$r_{c-} > (3m/2)\,[1 - (1 - 8e^2/9m^2)^{1/2}]. \quad (21)$$

These are the limits for the farthest and the closest circular orbits and the circular orbits could only occur in the region bounded by these. (For example, in the case of $e^2/m^2 = 1$, $r_{c+} = 2m$ and $r_{c-} = m$). Banerjee and Dutta Chaudhury[10] have obtained the limits (8), (20), and (21).

D. Trajectories in the r–ϕ plane: From Eqs. (3) and (5) with $2\mathscr{L} = -1$, we get

$$\frac{dr}{d\phi} = \pm \frac{r^2}{l}\left[E^2 - \Delta\left(\frac{l^2}{r^2} - 1\right)\right]^{1/2} \quad (22)$$

as the differential equation describing a spacelike (tachyon) trajectory in the r–ϕ plane. On differentiation, we have

$$\frac{1}{r}\frac{d^2r}{d\phi^2} = \frac{2}{r^2}\left(\frac{dr}{d\phi}\right)^2 + \frac{m}{l^2}r - \frac{3m}{r} - \frac{e^2}{l^2} + \frac{e^2}{r^2} + 1, \quad (23)$$

which implies that a trajectory is symmetric with respect to the apsidal line.

We shall now consider bending of trajectories in the r–ϕ plane. A trajectory bends towards or away from the source according as it has a maxima or minima along the apsidal line in the r–ϕ plane. We could equivalently say

$$\left(\frac{1}{r}\frac{d^2r}{d\phi^2}\right)_{\text{apse}} < 1 \quad (24)$$

for bending in and

$$\left(\frac{1}{r}\frac{d^2r}{d\phi^2}\right)_{\text{apse}} > 1 \quad (25)$$

for bending out. Using Eq. (23), we write

$$(r_t - r_0)(r_t^2 - 3l^2) < r_0 l^2 \quad \text{(bending in)} \quad (26)$$

and

$$(r_t - r_0)(r_t^2 - 3l^2) > r_0 l^2 \quad \text{(bending out)} \quad (27)$$

where $r_0 = e^2/m$. We recover the Schwarzschild conditions[3] when $e = 0$.

From Eqs. (26) and (27) it is clear that for bending out orbits we should at least have either $r_t > r_0 > \sqrt{3}l$ or $r_t < r_0 < \sqrt{3}l$ and for bending in orbits either $\sqrt{3}l \leqslant r_t \leqslant r_0$ or $\sqrt{3}l \geqslant r_t \geqslant r_0$ or even if $r_t > r_0$, there may occur a bending in orbit satisfying (26). Thus we have two (inner and outer) limits for bending in and bending out orbits separated by r_0.

IV. SCATTERING CROSS SECTIONS

It is clear from the Figs. 1 and 2 and the discussion above that no tachyon orbit can end into the singularity $r = 0$ except for radial tachyons satisfying the condition (9). In general, it is quite cumbersome to handle the cubic equation (16), and we shall therefore consider the case $e^2/m^2 = 1$ and calculate the scattering cross sections.

A. Tetrad, local observers, etc.: We first define the timelike Killing observer[3] (at fixed $r > r_+, \theta, \phi$). The orthonormal tetrad associated with the observer's world line is given by

$$t^i_{(0)} = \delta^i_0 \Delta^{-1/2}, \quad t^i_{(1)} = \delta^i_1 \Delta^{-1/2},$$
$$t^i_{(2)} = \delta^i_2 r^{-1}, \quad t^i_{(3)} = \delta^i_3 (r \sin\theta)^{-1}. \quad (28)$$

The 4-velocity u^i of a tachyon on a spacelike world line can be written as

$$u^i = (v^2 - 1)^{-1/2}\,(t^i_{(0)} + v_r t^i_{(1)} + v_\theta t^i_{(2)} + v_\phi t^i_{(3)}), \quad (29)$$

where $v = (v_r^2 + v_\theta^2 + v_\phi^2)^{1/2}$ is the local velocity as measured by the observer. We take $\pi/2$ and $v_\theta = 0$.

The local energy, local linear momentum and local tangential momentum are given as follows:

$$E_{(\text{loc})} = u_i t^i_{(0)} = (v^2 - 1)^{-1/2} = E\Delta^{-1/2}, \quad (30)$$

$$p_{r\,(\text{loc})} = -u_i t^i_{(1)} = v_r (v^2 - 1)^{-1/2} = \dot{r}\Delta^{-1/2}, \quad (31)$$

$$p_{\phi\,(\text{loc})} = -u_i t^i_{(3)} = v_\phi (v^2 - 1)^{-1/2} = r\dot{\phi}. \quad (32)$$

From Eqs. (3) and (32)

$$l = r v_\phi (v^2 - 1)^{-1/2}, \quad (33)$$

which implies that

$$v_{\phi(\infty)} = \lim_{r\to\infty} (v^2 - 1)^{1/2} l/r = 0 \quad (34)$$

and hence

$$\begin{aligned} v_{(\infty)} &= v_{r(\infty)} \\ &= \lim_{r\to\infty} (v^2 - 1)^{-1/2}\dot{r}\Delta^{-1/2} \\ &= [v^2_{(\infty)} - 1]^{1/2}\dot{r}(\infty). \end{aligned} \quad (35)$$

From (30) it follows

$$[v^2_{(\infty)} - 1]^{-1/2} = \lim_{r\to\infty} E\Delta^{-1/2} = E,$$

which means that the energy at infinity is related to the local energy by

$$E_{(\infty)} = E = \Delta^{1/2} E_{(\text{loc})}.$$

Finally let us write the impact parameters by using Eqs. (31), (33), and (35) as

$$\begin{aligned} b &= (\text{angular momentum/linear momentum})_{r\to\infty} \\ &= l\,[E v_{(\infty)}]^{-1} = l(E^2 + 1)^{-1/2}. \end{aligned} \quad (36)$$

B. Cross-section calculations: For a given energy at infinity $\{E=[v^2_{(\infty)}-1]^{1/2}\}$ of an incoming tachyon, we find the critical value of the angular momentum, l_{cr}, for which the maximum value of the effective potential $V_{max}(l_{cr})=E^2$. From Fig. 2 it is evident that there are two such values $l_{cr}>r_t$ and $l_{cr}<r_t$ and the effective potential $V(r)$, suddenly jumps, for all l, to positive value as r approaches $0.1m$ and then goes on without limit. Thus there would be no falling into $r=0$. We can then note that

$$l>l_{cr}>r_+,\quad V_{max}(l)>V_{max}(l_{cr})=E^2, \tag{37a}$$

$$l_{cr}>l>r_+,\quad V_{max}(l)<V_{max}(l_{cr})=E^2, \tag{37b}$$

$$r_-<l'_{cr}<l<r_+,\quad V_{max}(l)<V_{max}(l'_{cr})=E^2, \tag{37c}$$

$$r_-<l<l'_{cr}<r_+,\quad V_{max}(l)>V_{max}(l'_{cr})=E^2, \tag{37d}$$

$$r\to 0,\quad V(l\neq 0)\to\infty>E^2. \tag{37e}$$

The case (37a) is of the ordinary scattering above the event horizon while in (37d) the bounce occurs inside the black hole and tachyon will be scattered into the extended Nordström spacetime.[11] It falsely appears in (37b) and (37c) that tachyon would end on to $r=0$, but it does not do so owing to (37e). Hence, here again the tachyon would be scattered in to the extended Nordström spacetime from inside the black hole.

Now setting $\dot r=\ddot r=0$ in Eqs. (5) and (6), we write (i.e., $E^2=V_{max}$)

$$E^2=\left(\frac{l^2}{m^2}\right)\left(\frac{m}{r_c}\right)\left(\frac{\Delta^2}{1-r_0/r_c}\right), \tag{38}$$

where r_c is a positive root of Eq. (16) and r_0 is given by (12).

The case of $e^2/m^2=1$: Now Eq. (38) takes the form

$$l^2/m^2=E^2(m/r_c)^{-1}(1-m/r_c)^{-1}, \tag{39}$$

which by using Eq. (18) gives us the critical values of l as

$$l^2/m^2=(1+2E^2)(1+E^2)^{-1}. \tag{40}$$

We define the cross section of scattering as

$$\sigma=\pi b^2_{cr}=\pi m^2\left(\frac{1+2E^2}{1+E^2}\right)^2, \tag{41}$$

where Eqs. (36) and (40) are used. Clearly, tachyons are scattered from above the event horizon ($r>m$) as could be seen in from the Fig. 1.

In the limiting case, $E\gg 1$ (i.e.,$v_{(\infty)}\to 1$) we can write from Eq. (41)

$$\sigma=4\pi m^2(1-E^{-2}) \tag{42}$$

while in the photon limit ($v_{(\infty)}=1, E\to\infty$)

$$\sigma=4\pi m^2. \tag{43}$$

This means that the scattering cross section for tachyons is smaller than that for photons.

Let us compare these results with those for tardyons ($v<1$) where ($E\gg 1$, $v_{(\infty)}\to 1$)

$$\sigma=4\pi m^2(1+E^{-2}), \tag{44}$$

which leads to the same photon limit as above.

Finally consider the limiting case $E=0$, i.e., $v_{(\infty)}\to\infty$, we shall get [by not using Eq. (44) but an unapproximated relation analogous to Eq. (41)]

$$\sigma=\pi m^2.$$

V. DISCUSSION

The radially in-falling tachyons alone can hit the singularity $r=0$ provided [the condition (9)] they are highly energetic or the black hole has large charge ($e^2\to m^2$). They will have bounce in the region bounded by the two horizons. This is what one would expect, for radial tachyons can only have turning points where space and time have interchanged their roles.

The circular orbits can occur in the region defined by Eqs. (20) and (21) and obviously the orbits will be unstable. The scattering cross-section calculation shows for the case $e^2=m^2$ that the cross section is smaller for tachyons than that for photons.

Finally, the tachyons, which fall in from one asymptotically flat spacetime and have a bounce inside the event horizon, will come out into the other asymptotically flat spacetime. We have also considered elsewhere the trajectories of charged tachyon in the Nordström geometry.[15]

ACKNOWLEDGMENTS

It is a pleasure to thank the Raman Research Institute (Bangalore) and the Centre for Theoretical Studies (Indian Institute of Science, Bangalore) for the excellent hospitality which helped in preparing the manuscript. The author also thanks Mrs. Jayanthi Ramachandran for her help in the computations.

[1]R.O. Hettle and T.M. Helliwell, Nuovo Cimento B **13**, 82 (1973).
[2]A.K. Raychaudhuri, J. Math. Phys **15**, 856 (1974).
[3]E. Honig, K. Lake, and R.C. Roeder, Phys. Rev. D **10**, 3155 (1974).
[4]J.V. Narliker and S.V. Dhurandhar, Pramana **6**, 388 (1976).
[5]P.C.W. Davies, Nuovo Cimento B **25**, 571 (1975).
[6]J.V. Narliker and E.C.G. Sudarshan, Mon. Not. Roy. Astron. Soc. **175**, 105 (1976).
[7]S.V. Dhurandhar, Pramana **8**, 133 (1977).
[8]P.C. Vaidya, Current Sci. **40**, 65 (1971).
[9]N. Dadhich, Ind. J. Pure Appl. Math. **7**, 151 (1976).
[10]A. Banerjee and S.B. Dutta Choudhury, Aust. J. Phys. **30**, 251 (1977).
[11]J.C. Graves and D.R. Brill, Phys. Rev. **120**, 1507 (1960).
[12]E.P.T. Liang, Phys. Rev. D **9**, 3257 (1974).
[13]A. Armenti, Jr., Nuovo Cimento B **25**, 442 (1975).
[14]N. Dadhich and P.P. Kale, Pramana **9**, 71 (1977).
[15]N. Dadhich, Phys. Lett. A **68**, 291 (1978).

Prolongation structure for Langmuir solitons

A. Roy Chowdhury and T. Roy

Department of Physics, Jadavpur University, Calcutta—700032, India
(Received 14 August 1978)

A systematic analysis of nonlinear partial differential equations governing the formation and evolution of Langmuir solitons has been undertaken with the help of differential forms. The technique of Wahlquist and Estabrook has been applied in conjunction with the representation theory of Lie groups to derive the 3×3 inverse scattering formalism previously derived heuristically by Yajima.

1. INTRODUCTION

The observation of solitonlike structures in several domains of physical sciences has revealed many phenomena not explainable on the basis of the usual linearizing approach to the nonlinear equations. These special kinds of solutions of the nonlinear equations gives rise to physical phenomena forming a subject of study, under the name of synergetics.[1] While synergetics is concerned with the ideas behind physical facts of "soliton" structure and its application. The theoretical researchers have been primarily concerned with finding linear eigenvalue problem associated with the nonlinear equation under consideration. Up til now there have been no logical procedure for getting down to the inverse scattering problem, except only the approach of Wahlquist and Estabrook using the formalism of differential forms, Cartan–Ehresmann connections, and Lie groups. The initial work was done by Wahlquist and Estabrook themselves by demonstrating the utility of the method in the case of the nonlinear Schrödinger equation[2] and the KdV equation.[3] Further work along this line was done by Morris[4] in extending the method to the case of two space variables, and lastly the work of Dodd and Gibbon,[5] who have applied the formalism to the case of a higher order KdV equation. In this paper we have followed the ideas laid down in Ref. 1 to obtain the prolongation variables and the associated linear eigenvalue problem, for the case of Langmuir solitons, whose solutions were exhaustively studied by Yajima *et al.*,[6] from a 3×3 eigenvalue problem. Our main concern in this paper is not the analysis of the solutions but obtaining the inverse problem associated with the nonlinear equation in some logical steps.

2. FORMULATION

The equations under study deal strictly with the interaction of one-dimensional Langmuir waves with sound waves propagating in one direction, in particular, the phenomena being concerned with the sonic Langmuir soliton. The system of equations for the ion sound wave under the action of ponderomotive force due to a high frequency field and for the Langmuir wave was formulated by Zakharov[7]:

$$i\frac{\partial E}{\partial t}+\frac{1}{2}\frac{\partial^2 E}{\partial x^2}-nE=0,$$
$$\frac{\partial^2 n}{\partial t^2}-\frac{\partial^2 n}{\partial x^2}-2\frac{\partial^2|E|^2}{\partial x^2}=0. \tag{1}$$

It is rather interesting to note that only an approximate version of these equations were studied in Ref. 6 in which the second of Eqs. (1) was replaced by

$$\frac{\partial n}{\partial t}+\frac{\partial n}{\partial x}+\frac{\partial}{\partial x}|E|^2=0. \tag{2}$$

The differential forms equivalent to these equations are

$$\begin{aligned}
&\alpha_1=d\varphi\wedge dt-zdx\wedge dt,\\
&\alpha_2=-id\varphi\wedge dx+id\varphi\wedge dt+\tfrac{1}{2}dz\wedge dt-n\varphi dx\wedge dt,\\
&\alpha_3=d\varphi^*\wedge dt-z^*dx\wedge dt,\\
&\alpha_4=id\varphi^*\wedge dt-id\varphi^*\wedge dt+\tfrac{1}{2}dz^*\wedge dt-n\varphi^*dx\wedge dt,\\
&\alpha_5=dn\wedge dt-dn\wedge dx+(\varphi^*z+\varphi z^*)dx\wedge dt,
\end{aligned} \tag{3}$$

where instead of the original field variable E we have used $\varphi(x,t)$ connected to E by

$$\varphi(x,t)=E(x,t)e^{i(t/2-x)}. \tag{4}$$

These five differential 2-forms live in a six-dimensional space of primitive variables φ, φ^*, z, z^*, x, t. Our chief concern is the two-dimensional manifold of solutions which are really the integrals of these differential forms. Annulling this set of forms using $d\varphi=\varphi_x dx+\varphi_t dt$, we get back the original partial differential equations. But another very important point worth mentioning is that this set of forms must be "closed" under exterior differentiation, which is seen to hold for the above α_i's. That is,

$$d\alpha_i=\sum_{j=1}^{5}\eta_{ij}\wedge\alpha_j, \tag{5}$$

where η_{ij} are some set of 1-forms.

3. PROLONGATION STRUCTURE AND PSEUDOPOTENTIALS

Our next task is then a search for several different 1-forms (Pfaffians) having the structure

$$\omega_k=dy_k+F^k(x,t,\varphi,\varphi^*,z,z^*)dx+G\,k\,(x,t,\varphi,\varphi^*,z,z^*)dt, \tag{6}$$

such that their exterior derivatives are in the ring of the initial set of forms,

$$d\omega_k=\sum_{i=1}^{5}f_i^k\alpha_i. \tag{7}$$

Or, expanding Eq. (7), we get

$$dF^k\wedge dx+dG^k\wedge dt=\sum_{i=1}^{5}f_i^k\alpha_i$$

or

 0022-2488/79/071559-03$01.00

$$\sum_j\left(\frac{\partial F^k}{\partial\psi^j}d\psi^j\wedge dx+\frac{\partial G^k}{\partial\psi^j}d\psi^j\wedge dt\right)=\sum f_i^k\alpha_i, \tag{8}$$

where ψ^j denotes the set of premitive variables $\psi^j\equiv\{x, t, \varphi, \varphi^*, z, z^*\}$ occuring as argument of F^k and G^k. But for the purpose of prolongation we allow the functions F^k and G^k to depend also on the prolongation variables y_i's. Then the closure equation will read

$$d\omega_k-\sum_{i=1}^{5}f_i^k\alpha_i-\sum_{j=1}^{n}\eta_j^k\wedge\omega_j\equiv 0, \tag{9}$$

where $j=n$ is the number of prolongation variables to be used and η_j^k some set of 1-forms. In our actual calculation we write Eq. (9) in full forms as

$$\sum_j\left(\frac{\partial F^k}{\partial\psi^j}d\psi^j\wedge dx-i\frac{\partial G^k}{\partial\psi^j}d\psi^j\wedge dt\right)$$
$$\equiv\sum_{i=1}^{5}f_i^k\alpha_i+\{a_i^kdx+b_i^kdt$$
$$+c_i^kdz+d_i^kdz^*+g_i^kd\varphi+h_i^kd\varphi^*$$
$$+e_i^kdn\}\wedge\{dy^i+F^idx-iG^idt\}. \tag{10}$$

Equating the coefficients of different independent 2-forms in Eq. (10) we get the following information regarding the dependence of F^k and G^k on the primitive variables,

$$\begin{aligned}&F_z=F_{z^*}=0,\quad F_{nz}=0,\\&F_\varphi=-2G_z,\quad G_{zz}=G_{zz^*}=G_{z^*z^*}=0,\\&F_{\varphi^*}=+2G_{z^*},\quad G_{\varphi z^*}+G_{\varphi^* z}=0\\&F_\eta=iG_n,\quad G_{\varphi z}=F_{\varphi z}=0,\end{aligned} \tag{11}$$

and the identity

$$z(2G_z-G_\varphi)+z^*(G_{\varphi^*}-2G_{z^*})-F_\varphi n\varphi$$
$$+F_{\varphi^*}\eta\varphi^*-G_n(\varphi^*z+\varphi z^*)+F_iG-G_iF=0. \tag{12}$$

Further information regarding the dependence of F and G on the set ψ^j can be obtained by repeated differentiation of (12) using (11). And in this way we arrive at

$$F^k\equiv X_0+2\oint x_6+2\varphi^*x_7+nx_4$$
$$G^k\equiv X_5+zx_6+z^*x_7+(\varphi^*z-\varphi z^*)x_8+\varphi x_9$$
$$+\varphi^*x_{10}+\varphi^*\varphi x_{11}-inx_4. \tag{13}$$

For the present we restrict ourselves to the special case when $x_8=0$. Then substituting these forms of F^k and G^k again in Eq. (12), we get

$$\begin{aligned}&[x_0,x_5]=0,\quad [x_{11},x_6]=0,\\&[x_0,x_6]=-i(x_9+2x_6),\quad [x_{11},x_7]=0,\\&2[x_6,x_7]=ix_{11}-x_4,\\&[x_6,x_9]=0,\quad [x_9,x_4]=2ix_6,\\&[x_{10},x_4]=2ix_7,\quad [x_{10},x_7]=0,\end{aligned} \tag{14}$$

along with the constraints

$$\begin{aligned}&[x_0,x_9]=[x_5,x_6],\\&[x_0,x_{10}]=[x_5,x_7],\\&[x_0,x_{11}]=2[x_9,x_7]-2[x_{10},x_6],\\&[x_0,x_4]=-[x_5,x_4].\end{aligned} \tag{15}$$

The next is to obtain a closure of the algebra suggested by the above commutator. The resulting Lie algebra is then generated by the generators X_i, whose differential representation is obtained. To obtain the closure, we note that a simplification results if we set

$$x_4=\lambda x_{11}. \tag{16}$$

Also it is to be noted that all possible Jacobi indentities can never be satisfied unless we define a new generator x_{12},

$$x_{12}=[x_4,x_0], \tag{17}$$

and the complete algebra truns out to be

$$\begin{aligned}&[x_0,x_9]=\tau x_6+\lambda x_9,\\&[x_9,x_5]=2(1-\varphi^2)x_6-4i\varphi x_9,\\&[x_{10},x_6]=\alpha x_4+\beta x_0,\\&[x_9,x_{10}]=\gamma x_4+\delta x_0,\\&[x_7,x_5]=-x_{10}-2ix_7,\\&[x_5,x_{10}]=ax_{10}+bx_7,\\&2[x_9,x_7]=x_{12}+2[x_{10},x_6],\\&[x_6,x_{12}]=-2x_6,\\&[x_7,x_{12}]=-2x_7,\\&[x_9,x_{12}]=-2(x_9+x_6),\\&[x_{10},x_{12}]=6x_7-x_{10},\end{aligned} \tag{18}$$

where some arbitrary constants have crept in, whose values are to be adjusted in a suitable representation of the Lie algebra. The most obvious fact that such a algebraic structure may have many possible representations compels us to search for a differential realization of the generators on only three variables to make contact with the calculation of Yajima *et al.* One such form of the generators are given below:

$$X_{10}=\frac{(1-\zeta)}{2}\cdot y_1\frac{\partial}{\partial y_2}-\frac{(1+\zeta)}{2}\cdot y_3\frac{\partial}{\partial y_2},$$

$$X_9=-\frac{\zeta-1}{2\zeta}\cdot y_2\frac{\partial}{\partial y_1}+\frac{\zeta+1}{2\zeta}\cdot y_2\frac{\partial}{\partial y_3},$$

$$X_7=y_1\frac{\partial}{\partial y_2}-y_2\frac{\partial}{\partial y_1},$$

$$X_4=y_1\frac{\partial}{\partial y_1}-y_3\frac{\partial}{\partial y_3}+y_1\frac{\partial}{\partial y_3}-y_3\frac{\partial}{\partial y_1},$$

$$X_5=i\left(\frac{2\zeta^2}{3}-2\zeta\right)y_1\frac{\partial}{\partial y_1}-\frac{4i\zeta^2}{3}\cdot y_2\frac{\partial}{\partial y_2}$$
$$+i\left(\frac{2\zeta^2}{3}+2\zeta\right)y_3\frac{\partial}{\partial y_3},$$

$$X_6=i\left(y_2\frac{\partial}{\partial y_1}+y_2\frac{\partial}{\partial y_3}\right),$$

$$X_0=3i\zeta y_1\frac{\partial}{\partial y_1}+i\zeta y_2\frac{\partial}{\partial y_2}-i\zeta y_3\frac{\partial}{\partial y_3}. \tag{19}$$

4. INVERSE SCATTERING TRANSFORM AND PROLONGATION

In our above computation we have obtained a Casimir type differential operator realization on three prolongation variables y_1, y_2, and y_3. Other possible forms arise when the multiplicating factors to the differential operators $\partial/\partial y_1, \partial/\partial y_2$, and $\partial/\partial y_3$ are transcendental functions of y_1, y_2, y_3, but these representation may not be useful for the ultimate goal of an eigenvalue problem. Substituting the forms of generators X_i given in (19) in (6) leads to the system of matrix equations

$$Y_x = MY, \qquad Y_t = NY, \tag{20}$$

where Y is a vector with components (y_1, y_2, y_3) and matrices M and N are given by

$$M = \begin{pmatrix} 3i\zeta - ni/2\zeta & -\phi^* & -ni/2\zeta \\ \phi/2\zeta & i\zeta & \phi/2\zeta \\ -ni/2\zeta & \phi^* & -i\zeta + ni/2\zeta \end{pmatrix}, \quad n + \tfrac{1}{2}\phi^2 = \psi, \quad \tfrac{1}{2}i\phi_x - \phi = \chi,$$

$$N = \begin{pmatrix} \dfrac{i}{2\zeta}\psi + i\left(\dfrac{2\zeta^2}{2} - 2\zeta\right) & -\zeta\phi^* - \chi^* & \dfrac{i\psi}{2\zeta} \\ \dfrac{\phi}{2} + \dfrac{1}{2\zeta}\chi & -\dfrac{4i\zeta^2}{3} & -\dfrac{\phi}{2} + \dfrac{1}{2\zeta}\chi \\ -\dfrac{i}{2\zeta}\psi & -\zeta\phi^* + \chi^* & -\dfrac{i}{2\zeta}\psi + i\left(\dfrac{2\zeta^2}{3} + 2\zeta\right) \end{pmatrix}, \tag{21}$$

where we have adjusted the constants to tally with those of Ref. 6.

5. CONCLUSION

From our above discussions and the papers cited above, it is quite transparent that the only deductive approach up til now to the problem of inverse scattering transform is that of differential forms and Lie algebra. In general it is possible to have other representations of different dimension for the algebra given by equations (14), (18), which may yield more general prolongation structure for the nonlinear equation. Derivations of such representations, conservation laws, Bäcklund transformation, and the theorem of permutability will be discussed in a future communication.

[1] *Synergetics*, edited by H. Haken (workshop report) (Springer–Verlag, Berlin, 1977).
[2] F.B. Estabrook and H.D. Wahlquist, J. Math. Phys. **17**, 1293 (1976).
[3] H.D. Wahlquist and F.B. Estabrook, J. Math. Phys. **16**, 1 (1975).
[4] H.C. Morris, J. Math. Phys. **17**, 1867 (1977); **18**, 285, 530 (1977).
[5] R. Dodd and J.D. Gibbon, Proc. Roy. Soc. London Ser. A **358**, 287 (1978).
[6] N. Yajima and M. Oikawa, Prog. Theor. Phys. **56**, 1719 (1976).
[7] V.E. Zakharov, Zh. Eksp. Teor. Fiz. **62**, 1745 (1972) [Sov. Phys. JETP **35**, 908 (1972)].

Coupling coefficients: General theory

R. Dirl

Institut für Theoretische Physik, TU Wien, A-1040 Wien, Karlsplatz 13, Austria
(Received 5 June 1978)

Coupling coefficients for projective representations of finite groups are determined quite generally either by means of a general projection procedure, or as a linear combination of their corresponding Clebsch–Gordan coefficients. There we demonstrate that coupling— and Clebsch–Gordan coefficients are uniquely connected by special Clebsch–Gordan coefficients up to well-determined numerical factors.

INTRODUCTION

This paper deals with an important application of group theory to physics, namely the problem of decomposing Kronecker products of irreducible matrix representations (unirreps) into the direct sum of their irreducible constituents. The problem of computing coupling coefficients for projective unirreps (not necessarily belonging to equivalent standard factor systems) of a given group is quite similar to the problem of determining Clebsch–Gordan coefficients (CG coefficients). For the latter problem a new method has been discussed recently[1] and its utility has been demonstrated with the use of an important example.[2–4] Because of this similarity we transfer this method in such a way that it is applicable to the present problem. The method consists of considering the columns of the Kronecker product as G adapted vectors (i.e., vectors which transform according to projective unirreps belonging to the corresponding factor system) and of identifying the multiplicity index in terms of special column indices of the considered Kronecker product.

The material is organized as follows: In Sec. I we state the defining equations for coupling coefficients. There it is assumed that the projective unirreps composing the Kronecker products do not necessarily belong to equivalent factor systems. In the following section we rewrite the formulas of the general method[1] in such a way that coupling coefficients can be computed. A further possibility of determining coupling coefficients is given in Sec. III. There we show that coupling and CG coefficients are connected by a special unitary transformation. The matrix elements of these unitary transformations are uniquely determined (up to an uninteresting phase factor) and are proportional to special CG coefficients. In Sec. IV we briefly discuss physical applications which should show the importance of coupling coefficients for operator equivalences and related problems.

I. COUPLING COEFFICIENTS: STATEMENT OF THE GENERAL PROBLEM

From the outset we assume that the considered group G is finite and that two complete sets of projective unirreps for G are known which may belong to inequivalent standard factor systems R and S,

$$\mathbb{D}^{\alpha}:\{\mathbb{D}^{\alpha}(x):x\in G\},\quad \alpha\in A_{G(R)},\tag{I.1}$$

$$\mathbb{D}^{\beta}:\{\mathbb{D}^{\beta}(x):x\in G\},\quad \beta\in A_{G(S)},\tag{I.2}$$

$\mathbb{D}^{\alpha}$ ($\mathbb{D}^{\beta}$) denotes n_{α} (n_{β})-dimensional projective unirreps of G belonging to the factor system R (S). $A_{G(R)}$ ($A_{G(S)}$) denotes the set of all equivalence classes of G with respect to R (S). The matrix elements of the corresponding projective unirreps satisfy the representation property, the orthogonality, and completeness relations [see Eqs. (I.3–8) of Ref. 1].

It is known[5,6] that coupling coefficients are matrix elements of special subduction matrices, where the supergroup is the direct product group $G\times G$ and the subgroup the Kronecker product $G[x]G$. Like in the case of determining CG coefficients we have to start from unirreps of the direct product group. But in contrast to the first problem we have to consider the following projective unirreps,

$$\mathbb{D}^{\alpha\beta^*}:=\{\mathbb{D}^{\alpha\beta^*}(x,y)=\mathbb{D}^{\alpha}(x)\otimes\mathbb{D}^{\beta}(y)^*:x,y\in G\},$$
$$\alpha\in A_{G(R)},\quad \beta\in A_{G(S)},\tag{I.3}$$

which however belong to the standard factor system

$$Q((x,y),(x',y'))=R(x,x')S^*(y,y')\tag{I.4}$$

and whose matrix elements are given by

$$\mathbb{D}^{\alpha\beta^*}_{pr,qs}(x,y):=\mathbb{D}^{\alpha}_{pq}(x)\mathbb{D}^{\beta}_{rs}(y)^*,$$
$$p,q=1,2,\ldots,n_{\alpha},\quad r,s=1,2,\ldots,n_{\beta}.\tag{I.5}$$

The "subduction problem" which we want to consider is defined by

$$\mathbb{D}^{\alpha\beta^*}\downarrow G[x]G\sim\sum_{\gamma\in A_{G(RS^*)}}\oplus\, m_{\alpha\beta^*;\gamma}\mathbb{D}^{\gamma},\tag{I.6}$$

where the projective unirreps $\mathbb{D}^{\gamma}$ of G must belong to the factor system

$$T(x,y):=Q((x,x),(y,y))=R(x,y)S^*(x,y),\tag{I.7}$$

$$\mathbb{D}^{\gamma}(x)\mathbb{D}^{\gamma}(y)=T(x,y)\mathbb{D}^{\gamma}(xy),\quad \gamma\in A_{G(RS^*)}=A_{G(T)}.\tag{I.8}$$

The quantities $m_{\alpha\beta^*;\gamma}$ are called "multiplicities" which declare how many times the projective unirrep $\mathbb{D}^{\gamma}$, $\gamma\in A_{G(RS^*)}$ is contained in the reducible representation $\mathbb{D}^{\alpha\beta^*}\downarrow G[x]G$. The multiplicities have to be calculated by means of the usual character formula

$$m_{\alpha\beta^*;\gamma}=\frac{1}{|G|}\sum_{x\in G}\mathbb{X}^{\alpha}(x)\mathbb{X}^{\beta^*}(x)\mathbb{X}^{\gamma^*}(x).\tag{I.9}$$

The characters are given by the traces of the corresponding unirreps. The equation

$$n_{\alpha}n_{\beta}=\sum_{\gamma\in A_{G(RS^*)}}m_{\alpha\beta^*;\gamma}n_{\gamma}\tag{I.10}$$

is a trivial consequence of Eq. (I.5) but nevertheless important, if checking the multiplicity formula.

 0022-2488/79/071562-04$01.00

Now we are in the position to define unitary "coupling matrices" (whose matrix elements are the coupling coefficients) by means of

$$\{F^{\alpha\beta}\}^{\dagger}\mathbb{D}^{\alpha\beta *}(x)F^{\alpha\beta} = \sum_{\gamma\in A_{G(RS^*)}} \oplus\, m_{\alpha\beta *;\gamma}\mathbb{D}^{\gamma}(x),$$

$$\text{for all } x\in G, \qquad (I.11)$$

where we have introduced the notation

$$\mathbb{D}^{\alpha\beta *}(x) = \mathbb{D}^{\alpha\beta *}(x,x). \qquad (I.12)$$

Using this definition we can write Eq. (I.6) in more detail,

$$\sum_{p,q=1}^{n_\alpha}\sum_{r,s=1}^{n_\beta}\{F^{\alpha\beta}_{pr;\gamma wk}\}^*\mathbb{D}^{\alpha\beta *}_{pr,qs}(x)F^{\alpha\beta}_{qs;\gamma' w' l} = \delta_{\gamma\gamma'}\delta_{ww'}\mathbb{D}^{\gamma}_{kl}(x),$$

$$\text{for all } x\in G \text{ and } w,w' = 1,2,\dots,m_{\alpha\beta *;\gamma}, \qquad (I.13)$$

where the index w (w') is hereafter called the "multiplicity index."

II. CALCULATION OF COUPLING COEFFICIENTS

By utilizing the unitarity of the coupling matrices $F^{\alpha\beta}$, Eq. (I.13) can be rewritten in the following form,

$$\sum_{q=1}^{n_\alpha}\sum_{s=1}^{n_\beta}\mathbb{D}^{\alpha\beta *}_{pr,qs}(x)F^{\alpha\beta}_{qs;\gamma wl} = \sum_{k=1}^{n_\gamma}\mathbb{D}^{\gamma}_{kl}(x)F^{\alpha\beta}_{pr;\gamma wk}$$

$$\text{for all } x\in G \text{ and } w = 1,2,\dots,m_{\alpha\beta *;\gamma}, \qquad (II.1)$$

which is already one of the key equations for the present method. In order to verify this assertion we collect the $n_\alpha n_\beta$ matrix elements $F^{\alpha\beta}_{pr;\gamma wk}$ for fixed γ, w, and k to the column vectors $\vec{\mathbb{F}}^{\alpha\beta;\gamma w}_{k}$, whose components

$$\{\vec{\mathbb{F}}^{\alpha\beta;\gamma w}_{k}\}_{pr} = F^{\alpha\beta}_{pr;\gamma wk}$$

are just the coupling coefficients. Consequently Eq. (II.1) turns out to be

$$\mathbb{D}^{\alpha\beta *}(x)\vec{\mathbb{F}}^{\alpha\beta;\gamma w}_{l} = \sum_{k=1}^{n_\gamma}\mathbb{D}^{\gamma}_{kl}(x)\vec{\mathbb{F}}^{\alpha\beta;\gamma w}_{k},$$

$$\text{for all } x\in G,\quad \gamma\in A_{G(RS^*)},\quad w = 1,2,\dots,m_{\alpha\beta *;\gamma},$$

$$l = 1,2,\dots,n_\gamma, \qquad (II.2)$$

and shows that the vectors $\vec{\mathbb{F}}^{\alpha\beta;\gamma w}_{k}$ can be seen as G adapted vectors (which transform according to projective unirreps of G belonging to the factor system RS^*) of a $n_\alpha n_\beta$-dimensional Euclidean space $\mathscr{V}^{\alpha\beta *}$. The unitarity of the coupling matrices written down in terms of the scalar product

$$\langle\vec{\mathbb{F}}^{\alpha\beta;\gamma w}_{k},\vec{\mathbb{F}}^{\alpha\beta;\gamma' w'}_{l}\rangle = \delta_{\gamma\gamma'}\delta_{ww'}\delta_{kl} \qquad (II.3)$$

suggests together with the transformation law (II.2) how the multiplicity index w can be determined.

For this purpose we introduce by means of

$$\mathbb{A}^{\alpha\beta *}(G) := \left\{\frac{1}{|G|}\sum_{x\in G}F(x)\mathbb{D}^{\alpha\beta *}(x);F(x)\in\mathbb{C}\right\} \qquad (II.4)$$

an $n_\alpha n_\beta$-dimensional projective matrix representations of the group algebra $\mathbb{A}(G)$.[7,8] However, these representations should not be confused with those introduced in Ref. 1. The matrices

$$\mathbb{H}^{\alpha\beta;\gamma}_{kl} := \frac{n_\gamma}{|G|}\sum_{x\in G}\mathbb{D}^{\gamma *}_{kl}(x)\mathbb{D}^{\alpha\beta *}(x),$$

$$\gamma\in A_{G(RS^*)}\quad k,l = 1,2,\dots,n_\gamma \qquad (II.5)$$

form a representation of the units of $\mathbb{A}(G)$ and are represented by the zero matrix if the corresponding multiplicity is zero. These matrices satisfy the usual relations [compare Eqs. (II.7)–(II.10) of Ref. 1]. The following important formulas,

$$\mathbb{H}^{\alpha\beta;\gamma'}_{kl}\vec{\mathbb{F}}^{\alpha\beta;\gamma w}_{m} = \delta_{\gamma\gamma'}\delta_{lm}\vec{\mathbb{F}}^{\alpha\beta;\gamma w}_{k}, \quad \text{for all } w = 1,2,\dots,m_{\alpha\beta *;\gamma}, \qquad (II.6)$$

being a consequence of these relations and the transformation law (II.2), are further key equations for the present method.

As in Ref. 1 our procedure must be as follows: We construct by means of the projection operators $\mathbb{H}^{\alpha\beta;\gamma}_{aa}$ (for a given $\gamma\in A_{G(RS^*)}$ with $m_{\alpha\beta *;\gamma} > 0$ and an appropriately chosen index a) $m_{\alpha\beta *;\gamma}$-dimensional subspaces $\mathscr{V}^{\alpha\beta;\gamma}_{a}$ of $\mathscr{V}^{\alpha\beta *}$,

$$\mathscr{V}^{\alpha\beta;\gamma}_{a} = \{\mathbb{H}^{\alpha\beta;\gamma}_{aa}\vec{\mathbb{A}} : \vec{\mathbb{A}}\in\mathscr{V}^{\alpha\beta *}\}, \qquad (II.7)$$

$$\dim\mathscr{V}^{\alpha\beta;\gamma}_{a} = m_{\alpha\beta *;\gamma}. \qquad (II.8)$$

Clearly any orthonormal basis of $\mathscr{V}^{\alpha\beta;\gamma}_{a}$ represents a part of the columns of the coupling matrices, where the remaining columns must be computed by means of Eqs. (II.6), so that (II.2), (II.3), and (I.13) are satisfied. Instead of applying Schmidt's procedure in order to obtain an orthonormal basis of $\mathscr{V}^{\alpha\beta;\gamma}_{a}$, we proceed in the same way as in Ref. 1. We apply the corresponding projection operators $\mathbb{H}^{\alpha\beta;\gamma}_{aa}$ to each element of the orthonormal basis $\{\vec{\mathbb{B}}_{qs} : q = 1,2,\dots,n_\alpha;$ $s = 1,2,\dots,n_\beta\}$ (with $\{\vec{\mathbb{B}}_{qs}\}_{pr} = \delta_{pq}\delta_{rs}$) of $\mathscr{V}^{\alpha\beta *}$,

$$\{\vec{\mathbb{B}}^{\alpha\beta *;\gamma(qs)}_{a}\}_{pr} := \{\mathbb{H}^{\alpha\beta;\gamma}_{aa}\vec{\mathbb{B}}_{qs}\}_{pr} = B^{\alpha\beta *}_{pr;\gamma(qs)a}$$

$$= \frac{n_\gamma}{|G|}\sum_{x\in G}\mathbb{D}^{\alpha}_{pq}(x)\mathbb{D}^{\beta}_{rs}(x)^*\mathbb{D}^{\gamma}_{aa}(x)^*. \qquad (II.9)$$

We realize that exactly $m_{\alpha\beta *;\gamma}$ linear independent vectors $\vec{\mathbb{B}}^{\alpha\beta *;\gamma(qs)}_{a}$ must exist.

Our approach of computing the coupling matrices is now as follows: In case we can find just $m_{\alpha\beta *;\gamma}$ vectors $\vec{\mathbb{B}}^{\alpha\beta *;\gamma(q_v s_v)}_{a}$ satisfying

$$\|\vec{\mathbb{B}}^{\alpha\beta *;\gamma(q_v s_v)}_{a}\|^2 = \frac{n_\gamma}{|G|}\sum_{x\in G}\mathbb{D}^{\alpha}_{q_v q_v}(x)\mathbb{D}^{\beta}_{s_v s_v}(x)^*\mathbb{D}^{\gamma}_{aa}(x)^* > 0, \qquad (II.10)$$

$$\langle\vec{\mathbb{B}}^{\alpha\beta *;\gamma(q_v s_v)}_{a},\vec{\mathbb{B}}^{\alpha\beta *;\gamma(q_{v'} s_{v'})}_{a}\rangle = \frac{n_\gamma}{|G|}\sum_{x\in G}\mathbb{D}^{\alpha}_{q_v q_{v'}}(x)\mathbb{D}^{\beta}_{s_v s_{v'}}(x)^*\mathbb{D}^{\gamma}_{aa}(x)^*$$

$$= 0 \Longleftrightarrow (q_v s_v) \neq (q_{v'} s_{v'}), \qquad (II.11)$$

the following vectors

$$\vec{\mathbb{F}}^{\alpha\beta;\gamma(q_v s_v)}_{a} := \|\vec{\mathbb{B}}^{\alpha\beta *;\gamma(q_v s_v)}_{a}\|^{-1}\vec{\mathbb{B}}^{\alpha\beta *;\gamma(q_v s_v)}_{a} \qquad (II.12)$$

are already a part of the columns of the coupling matrices $F^{\alpha\beta}$. The remaining vectors have to be calculated by means of (II.6),

$$\{\vec{\mathbb{F}}^{\alpha\beta;\gamma(q_v s_v)}_{k}\}_{pr} = \{\mathbb{H}^{\alpha\beta;\gamma}_{ka}\vec{\mathbb{F}}^{\alpha\beta;\gamma(q_v s_v)}_{a}\}_{pr}$$

$$= \|\vec{\mathbb{B}}_a^{\alpha\beta^*;\gamma(q_v s_v)}\|^{-1} \frac{n_\gamma}{|G|} \sum_{x\in G} \mathbb{D}^\alpha_{pq_v}(x)\mathbb{D}^\beta_{rs_v}(x)^*\mathbb{D}^\gamma_{ka}(x)^*. \tag{II.13}$$

Therefore, our method allows us to determine the multiplicity index w in terms of special column indices $q_v s_v$) of the Kronecker product $\mathbb{D}^{\alpha\beta^*}$. Of course, as mentioned in Ref. 1, there remains the difficult task of again solving for whether we can always trace back the multiplicity index to special column indices of the corresponding Kronecker product. Apart from this, the present formulas are identical to those given in Ref. 1, if we replace the projective unirreps $\mathbb{D}^\beta$ by their complex conjugate values.

III. CONNECTION BETWEEN COUPLING AND CG COEFFICIENTS

In order to be able to create a connection between coupling and CG coefficients we have to realize that any projective unirrep $\mathbb{D}^{\beta^*}$ must be equivalent to a uniquely determined projective unirrep $\mathbb{D}^{\bar\beta}$ with $\bar\beta \in A_{G(S^*)}$. This equivalence implies that there must exist a n_β-dimensional unitary matrix U^β satisfying

$$\mathbb{D}^\beta(x)^* = U^{\beta\dagger}\mathbb{D}^{\bar\beta}(x)U^\beta, \quad \text{for all } x\in G \tag{III.1}$$

which is uniquely determined up to an uninteresting phase factor. Inserting (III.1) into (I.11) we obtain

$$\{(\mathbb{1}_\alpha \otimes U^\beta)F^{\alpha\beta}\}^\dagger \mathbb{D}^\alpha(x)\otimes\mathbb{D}^{\bar\beta}(x)\{(\mathbb{1}_\alpha\otimes U^\beta)F^{\alpha\beta}\}$$
$$= \sum_{\gamma\in A_{G(RS^*)}} \oplus\, m_{\alpha\beta^*;\gamma}\mathbb{D}^\gamma(x) \quad \text{for all } x\in G, \tag{III.2}$$

where $\mathbb{1}_\alpha$ denotes the n_α-dimensional unit matrix. Consequently, a comparison between (I.17) of Ref. 1 and (III.2) yields

$$(\mathbb{1}_\alpha\otimes U^\beta)F^{\alpha\beta} = C^{\alpha\bar\beta}, \tag{III.3}$$

$$F^{\alpha\bar\beta} = (\mathbb{1}_\alpha\otimes U^\beta)^\dagger C^{\alpha\bar\beta}, \quad \bar\beta\in A_{G(S^*)} \tag{III.4}$$

which describes the connection between coupling and CG matrices. Introducing the notation

$$F^{\alpha\beta}_{pr;\gamma wk} = \begin{bmatrix}\alpha & \bar\beta & \gamma w\\ p & r & k\end{bmatrix}, \tag{III.5}$$

$$C^{\alpha\bar\beta}_{qs;\gamma wl} = \begin{pmatrix}\alpha & \bar\beta & \gamma w\\ q & s & l\end{pmatrix}, \tag{III.6}$$

we obtain for (III.3) and (III.4)

$$\begin{pmatrix}\alpha & \bar\beta & \gamma w\\ p & r & k\end{pmatrix} = \sum_{s=1}^{n_\beta} U^\beta_{rs}\begin{bmatrix}\alpha & \bar\beta & \gamma w\\ p & s & k\end{bmatrix}, \tag{III.7}$$

$$\begin{bmatrix}\alpha & \bar\beta & \gamma w\\ p & r & k\end{bmatrix} = \sum_{s=1}^{n_\beta} U^{\beta*}_{sr}\begin{pmatrix}\alpha & \bar\beta & \gamma w\\ p & s & k\end{pmatrix}. \tag{III.8}$$

In order to show that the matrix elements of the unitary matrices U^β are special CG coefficients we consider the special case

$$\{F^{\alpha\beta}\}^\dagger\mathbb{D}^\alpha(x)\otimes\mathbb{D}^\beta(x)^* F^{\alpha\beta} = \sum_{\gamma\in A_{G(1)}}\oplus\, m_{\alpha\beta^*;\gamma}\mathbb{D}^\gamma(x),$$
$$\text{with } \alpha,\beta\in A_{G(S)}, \tag{II.9}$$

where both unirreps of G are assumed to belong to the same factor system, so that on the right-hand side only ordinary vector unirreps for G may occur. The corresponding multiplicity formula reduces for the special case $\gamma = 0$ (identity representation) to

$$m_{\alpha\beta^*;O} = \delta_{\alpha\beta} \tag{III.10}$$

which has the well-known consequence,[5,9] that the identity representation occurs in $D^{\beta\beta^*}$ exactly once. Equation (II.10) turns out to be

$$\|\vec{\mathbb{B}}_1^{\alpha\beta^*;O(qs)}\|^2 = n_\alpha^{-1}\delta_{\alpha\beta}\delta_{qs} \tag{III.11}$$

from which it follows that

$$\vec{\mathbb{B}}_1^{\beta\beta^*;O(qq)} = \vec{\mathbb{B}}_1^{\beta\beta^*;O(q'q')}, \quad \text{for all } q,q' = 1,2,\dots,n_\beta \tag{III.12}$$

if taking Eq. (II.9) into account. Of course (III.12) proves (III.10). Therefore, the corresponding coupling coefficients take, due to (II.13), the form

$$F^{\beta\beta}_{pr;O(11)1} = \begin{bmatrix}\beta & \bar\beta & 0\\ p & r & 1\end{bmatrix} = \frac{1}{\sqrt{n_\beta}}\delta_{pr} \tag{III.13}$$

where we have chosen $q = 1$ and where the superfluous multiplicity index is suppressed. Inserting this special case into (III.7) we obtain

$$U^\beta_{rp} = \sqrt{n_\beta}\begin{pmatrix}\beta & \bar\beta & 0\\ p & r & 1\end{pmatrix}, \tag{III.14}$$

which proves our proposition that the matrix elements of the unitary matrices U^β are (up to well-determined numerical factors) special CG coefficients. However in this connection it should be noted, that the omitted multiplicity index $w = (11)$ originates from $D^{\beta\beta^*}$ and should not be confused with special column indices of $D^{\beta\bar\beta} := \{\mathbb{D}^\beta(x)\otimes\mathbb{D}^{\bar\beta}(x) : x\in G\}$.

The special CG coefficients defining the unitary matrices U^β are readily obtained by means of Eq. (II.39) of Ref. 1, where $\mathbb{D}^{\alpha\beta}(x) = \mathbb{D}^\alpha(x)\otimes\mathbb{D}^\beta(x)$ has to be replaced by $D^{\beta\bar\beta}(x) = \mathbb{D}^\beta(x)\otimes\mathbb{D}^{\bar\beta}(x)$ and $\mathbb{D}^\gamma(x)$ by one,

$$C^{\beta\bar\beta}_{pr;O(q_0s_0)1} = \begin{pmatrix}\beta & \bar\beta & 0\\ p & r & 1\end{pmatrix}$$
$$= \frac{1}{\sqrt{|G|}}\left\{\sum_{x\in G}\mathbb{D}^\beta_{q_0q_0}(x)\mathbb{D}^{\bar\beta}_{s_0s_0}(x)\right\}^{-1/2}$$
$$\times\sum_{y\in G}\mathbb{D}^\beta_{pq_0}(y)\mathbb{D}^{\bar\beta}_{rs_0}(y). \tag{III.15}$$

Thereby the index (q_0s_0) is chosen in such a way that the norm of the corresponding vector is different from zero.

Thus we arrive at the final formulas

$$\begin{bmatrix}\alpha & \bar\beta & \gamma w\\ p & r & k\end{bmatrix} = \sqrt{n_\beta}\sum_{s=1}^{n_\beta}\begin{pmatrix}\beta & \bar\beta & 0\\ r & s & 1\end{pmatrix}^*\begin{pmatrix}\alpha & \bar\beta & \gamma w\\ p & s & k\end{pmatrix},$$
$$\alpha\in A_{G(R)},\quad \beta\in A_{G(S)};\quad \bar\beta\in A_{G(S^*)},\quad \gamma\in A_{G(RS^*)},$$
$$w = 1,2,\dots,m_{\alpha\beta^*;\gamma},\quad k = 1,2,\dots,n_\gamma,$$
$$p = 1,2,\dots,n_\alpha,\quad r = 1,2,\dots,n_\beta, \tag{III.16}$$

which connect coupling coefficients with their corresponding CG coefficients and vice versa. These formulas allow us to calculate readily coupling coefficients, if the corresponding CG coefficients are known. Concerning the multiplicity index w which occurs on both sides of Eq. (III.16) we have to make some remarks. If we assume that the CG coefficients are known, the multiplicity index w originates (if possible) from special column indices of $\mathbb{D}^{\alpha\bar{\beta}}(x) := \mathbb{D}^{\alpha}(x) \otimes \mathbb{D}^{\bar{\beta}}(x)$. This implies however that we cannot expect that the multiplicity index can also be traced back in general to special column indices of $\mathbb{D}^{\alpha\beta *}(x) = \mathbb{D}^{\alpha}(x) \otimes \mathbb{D}^{\beta}(x)^*$. The same argument holds in the inverse direction. Concluding this section we remark that the special case of ordinary vector representations is contained in a consistent way in our formulas.

IV. PHYSICAL APPLICATIONS

For many physical applications, like in crystal field theory or solid state physics, it is often necessary to construct appropriate matrix Hamiltonians which are composed of irreducible tensor operators. A typical problem of this kind is to find convenient operator bases for subspaces of the considered Hilbert space, which are irreducible with respect to a given group G. Starting from operators

$$E^{\beta}_{ij} = n_{\beta} |G|^{-1} \sum_{x\in G} D^{\beta}_{ij}(x)^* U(x) \tag{IV.1}$$

which are of the type (II.5), it is known[5,6] that the operators

$$T^{\beta;\gamma w}_{k} := \sum_{i,j} \begin{bmatrix} \beta & \bar{\beta} & \gamma w \\ i & j & k \end{bmatrix} E^{\beta}_{ij},$$
$$\gamma \in A_G, \quad w = 1,2,\ldots,m_{\beta\bar{\beta};\gamma}, \quad k = 1,2,\ldots,n_{\gamma} \tag{IV.2}$$

are irreducible tensor operators with respect to G,

$$U(x) T^{\beta;\gamma w}_{k} U(x)^{\dagger} = \sum_{k=1}^{n_{\gamma}} D^{\gamma}_{lk}(x) T^{\beta;\gamma w}_{l}. \tag{IV.3}$$

Thereby U denotes a unitary representation of G, respectively, the unirreps of G are assumed to be ordinary vector representations. Obviously Eq. (IV.2) shows the importance of coupling coefficients for the theory of equivalent operators. Hence, any component of an irreducible tensor operator acting nontrivially in a fixed irreducible subspace (which is labeled by β and if necessary by further quantum numbers), must be a unique linear combination of the operators (IV.2)

$$A^{\beta;\mu}_{a} = \sum_{w} B_w T^{\beta;\mu w}_{a}, \quad B_w \in \mathbb{C}. \tag{IV.4}$$

Of course the coefficients B_w may not depend on the index a, but reflects the special choice of the coupling coefficients, which are not uniquely determined for the nonsimple reducible case.[10] Nevertheless physical quantities must be independent from this unitary equivalence.

Another problem which we want to discuss briefly stems from solid state physics, where sometimes matrix Hamiltonians are considered, like in the effective mass $\vec{k}\cdot\vec{p}$ theory. These matrix Hamiltonians are composed of two different types of irreducible tensor operators. The problem is to find such a correct linear combination of these irreducible tensor operators, so that the Hamiltonian is invariant with respect to the Kronecker product,

$$H^{\beta} = \sum_{\substack{\gamma\mu w \\ kl}} G^{\gamma\mu w}_{kl} T^{\beta;\gamma w}_{k} \otimes Z^{\mu}_{l}, \quad G^{\gamma\mu w}_{kl} \in \mathbb{C}, \tag{IV.5}$$

$$V(y) Z^{\mu}_{l} V(y)^{\dagger} = \sum_{n} D^{\mu *}_{nl}(y) Z^{\mu}_{n}, \tag{IV.6}$$

$$U(x) \otimes V(x) H^{\beta} (U(x) \otimes V(x))^{\dagger} = H^{\beta} \quad \text{for all } x \in G. \tag{IV.7}$$

A simple inspection of Eq. (IV.7) leads, for the unknown coefficients $G^{\gamma\mu w}_{kl}$, to the result

$$G^{\gamma\mu w}_{kl} = G^{\mu w} \delta_{\gamma\mu} \delta_{kl}, \tag{IV.8}$$

which implies

$$H^{\beta} = \sum_{\mu w} G^{\mu w} \sum_{k} T^{\beta;\mu w}_{k} \otimes Z^{\mu}_{k}. \tag{IV.9}$$

Since the matrix elements of $T^{\beta;\mu w}_{k}$ are just the coupling coefficients, Eq. (IV.9) represents the generalization of Eq. (2.15) of Ref. 11 to the nonmultiplicity free case. This example shows once more the importance of coupling coefficients for physical applications.

V. CONCLUDING REMARKS

In the present paper we have demonstrated how a general method for computing CG coefficients has to be transferred in order to make this procedure applicable for determining coupling coefficients. A further possibility for calculating coupling coefficients presupposes the explicit knowledge of the corresponding CG coefficients, since these two types of coefficients are connected by simple unitary transformations. There we have shown, that even for the nonmultiplicity free case, the matrix elements of these unitary transformations are proportional to special CG coefficients, which are uniquely determined up to a common and therefore uninteresting phase factor.

[1] R. Dirl, "Clebsch–Gordan coefficients: General theory" J. Math. Phys. **20**, 659 (1979).
[2] R. Dirl, "Multiplicities for space group representations," J. Math. Phys. **20**, 664 (1979).
[3] R. Dirl, "Clebsch–Gordan coefficients for space groups," J. Math. Phys. **20**, 671 (1979).
[4] R. Dirl, "Clebsch–Gordan coefficients for *Pn3n*," J. Math. Phys. **20**, 679 (1979).
[5] P. Kasperkovitz and R. Dirl, J. Math. Phys. **15**, 1203 (1974).
[6] R. Dirl and P. Kasperkovitz, *Gruppentheorie, Anwendungen in der Atom- und Festkörperphysik* (Vieweg, Braunschweig, 1977).
[7] R. Dirl, J. Math. Phys. **18**, 2065 (1977).
[8] R. Dirl, "Induced projective representations," in *Group Theoretical Methods in Physics*, edited by R.T. Sharp and B. Kolman (Academic, New York, 1977).
[9] A.J. Coleman, "Induced and subduced representations," in *Group Theory and Its Applications*, edited by M. Loebl (Academic, New York, 1968).
[10] M. Hamermesh, *Group Theory and Its Applications to Physical Problems* (Addison-Wesley, Reading, Massachusetts, 1962).
[11] J.L. Birman, T.-K. Lee, and Rh. Berenson, Phys. Rev. B **14**, 318 (1976).

Complex conjugation of space group representations

R. Dirl

Institut für Theoretische Physik, TU Wien, A—1040 Wien, Karlsplatz 13, Austria
(Received 5 June 1978)

Useful relations are derived which allow us to determine for every unirrep of nonsymmorphic space groups, which contain the inversion as a point group operation, the equivalent complex conjugate representation.

INTRODUCTION

In this paper we start our discussion concerning the problem of determining coupling coefficients for such nonsymmorphic space groups, which contain inversion as point group operation. Since we want to prefer the second possibility to compute these coefficients, the problem reduces to calculate those unitary matrices which connect space group unirreps with their corresponding complex conjugate representations. Due to Eq. (III.1), respectively (III.4) of Ref. 1, a first step towards a general solution of our problem is done, if we can determine for every space group unirrep the equivalent complex conjugate representation. In doing so we obtain three different types of defining equations for the equivalence classes of the projective unirreps of the little cogroups $P^{\vec{q}} \simeq G^{\vec{q}}/T$ whose corresponding representations are equivalent in a generalized sense.

The present paper is organized as follows: In Sec. I the basic definitions and notations concerning space groups and their unirreps are briefly reviewed, where special choices for the sets $P{:}P^{\vec{q}}$ of left coset representatives can be introduced, since it is assumed that inversion is contained as a point group operation in the considered space groups. These special sets should extremely simplify the following considerations. In the following section we investigate the defining equations for the equivalence classes, whose corresponding representations are equivalent. This discussion leads us in Sec. III to three different types of defining equations for the equivalence classes characterizing projective unirreps of the little cogroups $P^{\vec{q}}$, whose corresponding representations are equivalent in a generalized sense, if $\vec{q}$ belongs to the surface of the fundamental domain ΔBZ of the Brillouin zone. The reason for this result is due to the fact that only for $\vec{q}$'s lying on the surface of ΔBZ are the corresponding factor systems nontrivial. Especially if $\vec{q}$ belongs to the surface of ΔBZ we are confronted with two different types of nontrivial defining equations for the equivalence classes of $P^{\vec{q}}$, depending whether the inversion I belongs to $P^{\vec{q}}$ or not. If $\vec{q}$ does not lie on the surface of ΔBZ, we obtain a further type of defining equations for the equivalence classes of $P^{\vec{q}}$, which is usual for vector representations, since the corresponding unirreps are such ones.

I. SPACE GROUP REPRESENTATIONS

In this section we briefly recall the basic definitions and notations concerning space groups and their representations, which are used throughout this and the following papers.

$$G = \{(\alpha|\vec{\tau}(\alpha) + \vec{t}): \alpha\in P, \quad t\in T\}, \tag{I.1}$$

$$(\alpha|\vec{\tau}(\alpha) + \vec{t})(\beta|\vec{\tau}(\beta) + \vec{t}') = (\alpha\beta|\vec{\tau}(\alpha\beta) + \vec{t}(\alpha,\beta) + D(\alpha)\vec{t}' + \vec{t}), \tag{I.2}$$

$$\vec{t}(\alpha,\beta) = \vec{\tau}(\alpha) + D(\alpha)\vec{\tau}(\beta) - \vec{\tau}(\alpha\beta). \tag{I.3}$$

G denotes the considered space group, $\vec{t}$ primitive lattice translations, $\vec{\tau}(\alpha)$ nonprimitive lattice translations, and $D = \{D(\alpha){:}\alpha\in P\}$ a faithful representation of the point group $P \simeq G/T$ of the crystal.

It is known[2-4] that the matrix elements of the unirreps of a nonsymmorphic space group G can be written in the following form:

$$D^{(\kappa,\vec{q})\uparrow G}_{\sigma a,\sigma' b}(\beta|\vec{\tau}(\beta) + \vec{t}) = \Delta^{\vec{q}}(\sigma,\beta\sigma')\, e^{-i\vec{q}(\sigma)\cdot\vec{t}}\, B^{\vec{q}}_{\sigma,\sigma'}(\beta)\mathbb{R}^{\kappa}_{ab}(\sigma^{-1}\beta\sigma'),$$

$$\vec{q}\in\Delta BZ, \quad \kappa\in A_{P^{\vec{q}}(S^{\vec{q}})}, \quad a,b = 1,2,\cdots n_\kappa, \quad \sigma,\sigma'\in P{:}P^{\vec{q}} \tag{I.4}$$

$$P^{\vec{q}} := \{\alpha{:}D(\alpha)\vec{q} = \vec{q} + \vec{Q}\{\vec{q}(\alpha)\};\alpha\in P\}, \tag{I.5}$$

$$\Delta^{\vec{q}}(\gamma,\gamma') := \delta_{\gamma P^{\vec{q}},\gamma' P^{\vec{q}}}, \quad \text{for all } \gamma,\gamma'\in P, \tag{I.6}$$

$$\vec{q}(\gamma) := D(\gamma)\vec{q}, \quad \text{for all } \gamma\in P, \tag{I.7}$$

$$B^{\vec{q}}_{\sigma,\sigma'}(\beta) := \exp[-i\vec{q}(\sigma)\cdot\{\vec{\tau}(\beta) + D(\beta)\vec{\tau}(\sigma') - \vec{\tau}(\sigma)\}], \quad \text{for all } \beta\in P. \tag{I.8}$$

Thereby ΔBZ denotes the fundamental domain of the Brillouin zone, $G^{\vec{q}}$ the group of the $\vec{q}$ vector, $\vec{Q}\{\vec{q}(\alpha)\}$ reciprocal lattice vectors, $\sigma,\sigma'\in P{:}P^{\vec{q}}$ left coset representatives of $P^{\vec{q}} \simeq G^{\vec{q}}/T$ with respect to P, and $\mathbb{R}^{\kappa} := \{\mathbb{R}^{\kappa}(\alpha){:}\alpha\in P^{\vec{q}}\}$ n_K-dimensional projective unirreps of $P^{\vec{q}}$ which belong to the factor system

$$S^{\vec{q}}(\alpha,\beta) = \exp[-i\vec{q}\cdot(D(\alpha) - \mathbb{1})\vec{\tau}(\beta))], \quad \text{for all } \alpha,\beta\in P^{\vec{q}}. \tag{I.9}$$

(In this connection we have to note that the projective unirreps $\mathbb{R}^{\kappa}$ reduce to ordinary vector representations of $P^{\vec{q}}$ if $\vec{q}$ does not lie on the surface of ΔBZ.) The characters of the space group unirreps are immediately obtained from Eq. (I.4),

$$X^{(\kappa,\vec{q})\uparrow G}(\beta|\vec{\tau}(\beta) + \vec{t}) = \sum_{\sigma\in P:P^{\vec{q}}} \Delta^{\vec{q}}(\sigma,\beta\sigma)\, B^{\vec{q}}_{\sigma,\sigma}(\beta)\mathbb{X}^{\kappa}(\sigma^{-1}\beta\sigma)\, e^{-i\vec{q}(\sigma)\cdot\vec{t}}, \tag{I.10}$$

$$\mathbb{X}^{\kappa}(\alpha) := \operatorname{trace}\mathbb{R}^{\kappa}(\alpha), \quad \text{for all } \alpha\in P^{\vec{q}}. \tag{I.11}$$

This character formula will be used together with the orthogonality relations for characters in the following section in order to determine those vector unirreps of G which are

 0022-2488/79/071566-04$01.00

equivalent to their complex conjugate representations, i.e.,

$$\{D^{(\kappa,\vec{q})\uparrow G}(\alpha|\vec{\tau}(\alpha)+\vec{t})\}^* \sim D^{(\bar{\kappa},\vec{q}')\uparrow G}(\alpha|\vec{\tau}(\alpha)+\vec{t}) \quad \text{for all } (\alpha|\vec{\tau}(\alpha)+\vec{t})\in G \qquad (I.12)$$

or briefly

$$\{(\kappa,\vec{q})\uparrow G\}^* = (\bar{\kappa},\vec{q}')\uparrow G. \qquad (I.13)$$

Concerning the characters of (I.10) we have to note that they may not depend on any special choice for $P{:}P^{\vec{q}}$, although the corresponding unirreps depend in any way on such a choice for $P{:}P^{\vec{q}}$.

Since it is assumed from the outset that we are restricting our consideration to space groups which contain inversion as a point group operation, it is possible to specify the sets $P{:}P^{\vec{q}}$ as discussed in the following. (Thereby it should be noted that $I\in P\simeq G/T$ does not imply that ΔBZ must be identical to the basic domain of the Brillouin zone, i.e., ΔBZ will be in general an integer multiple of this basic domain.) Hence, if $\vec{q}$ does not belong to the surface of ΔBZ, we can choose in any way

$$I\in P{:}P^{\vec{q}}. \qquad (I.14)$$

This has a consequence that the order of $P{:}P^{\vec{q}}$ must be an even integer, i.e.,

$$|P{:}P^{\vec{q}}| = 2n, \quad n\in\mathbb{N}, \qquad (I.15)$$

since due to the definition of $P^{\vec{q}}$ and Eq. (I.14), i.e.,

$$\vec{q}(I)\neq\vec{q}+\vec{Q}, \qquad (I.16)$$

it follows from

$$\vec{q}(\sigma)\neq\vec{q}(\sigma I)+\vec{Q}, \quad \text{for all } \sigma\in P{:}P^{\vec{q}}, \qquad (I.17)$$

that we can choose $I\sigma$ as further left coset representative, if $\sigma\in P{:}P^{\vec{q}}$. This assertion can be proven by means of relations of the type

$$\vec{q}(\sigma_i I) = \vec{q}(\sigma_j) + \vec{Q}: \quad \sigma_i,\sigma_j\in P{:}P^{\vec{q}} \qquad (I.18)$$

which admit as solutions [use the definition (I.5)]

$$\sigma_i^{-1}\sigma_j I\in P^{\vec{q}} \Longleftrightarrow \sigma_j = I\sigma_i\alpha, \quad \text{with } \alpha\in P^{\vec{q}}. \qquad (I.19)$$

The special choice $\alpha = e$ completes the proof of our proposition. Hence we establish our sets $P{:}P^{\vec{q}}$ as follows

$$P{:}P^{\vec{q}} = \{\sigma_1,\dots,\sigma_k\}\cup\{I\sigma_1,\dots,I\sigma_k\},$$
$$k = \tfrac{1}{2}|P{:}P^{\vec{q}}|, \quad \sigma_1 = e \qquad (I.20)$$

with

$$\sigma_j\neq I\sigma_i, \quad \text{for all } i,j = 1,2,\dots,k. \qquad (I.21)$$

If $\vec{q}$ belongs to the surface of ΔBZ, there may occur two cases

$$I\in P{:}P^{\vec{q}}, \qquad (I.22)$$

$$I\in P^{\vec{q}}. \qquad (I.23)$$

For the first case the same argument holds as before, whereas for the second case we cannot expect further simplifications concerning the corresponding sets $P{:}P^{\vec{q}}$. Finally it should be noted that just (I.24), respectively (I.22), is realized for nearly all cases, while the case (I.23) holds true only for few points of the surface of ΔBZ.

II. COMPLEX CONJUGATION OF SPACE GROUP UNIRREPS

Equation (I.12) written down in more detail requires that there must exist for each $(\kappa,\vec{q})\uparrow G\in A_G$ a unitary matrix $U^{(\kappa,\vec{q})}$ satisfying

$$\{D^{(\kappa,\vec{q})\uparrow G}(\beta|\vec{\tau}(\beta)+\vec{t})\}^*$$
$$= U^{(\kappa,\vec{q})+} D^{(\bar{\kappa},\vec{q}')\uparrow G}(\beta|\vec{\tau}(\beta)+\vec{t})U^{(\bar{\kappa},\vec{q})}, \quad \text{for all } (\beta|\vec{\tau}(\beta)+\vec{t})\in G, \qquad (II.1)$$

where $(\kappa,\vec{q})\uparrow G$, $(\bar{\kappa},\vec{q}')\uparrow G\in A_G$. Hence

$$X^{(\kappa,\vec{q})\uparrow G}(\beta|\vec{\tau}(\beta)+\vec{t})^* = X^{(\bar{\kappa},\vec{q}')\uparrow G}(\beta|\vec{\tau}(\beta)+\vec{t}), \quad \text{for all} (\beta|\vec{\tau}(\beta)+\vec{t})\in G. \qquad (II.2)$$

Inspecting Eq. (II.7) of Ref. 5 for the special case

$$m_{\kappa,\vec{q})(\bar{\kappa},\vec{q}');(0,\vec{0})}$$
$$= \delta_{\vec{q},\vec{q}'}\frac{1}{|P|}\sum_{\substack{\beta\in P\\ \sigma,\tau\in P:P^{\vec{q}}}} \delta_{\vec{q}(\sigma)+\vec{q}(\tau),\vec{Q}\{\vec{q}(\sigma)+\vec{q}(\tau)\}}\Delta^{\vec{q}}(\sigma,\beta\sigma)$$
$$\times\Delta^{\vec{q}}(\tau,\beta T)B^{\vec{q}}_{\sigma,\sigma}(\beta)B^{\vec{q}}_{\tau,\tau}(\beta)\mathbb{X}^{\kappa}(\sigma^{-1}\beta\sigma)\mathbb{X}^{\bar{\kappa}}(\tau^{-1}\beta\tau), \qquad (II.3)$$

which can be seen as the defining equation for $(\bar{\kappa},\vec{q}')\uparrow G\in A_G$, if

$$m_{(\kappa,\vec{q})(\bar{\kappa},\vec{q}');(0,\vec{0})} = 1 \qquad (II.4)$$

is required. We obtain as first result $\vec{q}' = \vec{q}$, i.e.,

$$\{(\kappa,\vec{q})\uparrow G\}^* = (\bar{\kappa},\vec{q})\uparrow G \quad \text{for all } \vec{q}\in\Delta BZ. \qquad (II.5)$$

Therefore,

$$\bar{\kappa}\in A_{P^{\vec{q}}(S^{\vec{q}})} \qquad (II.6)$$

which is an important consequence of (II.5) and leads to the proposition that the factor systems $S^{\vec{q}}$ and $S^{\vec{q}*}$ must be equivalent for all $\vec{q}\in\Delta BZ$.

$$S^{\vec{q}}(\alpha,\beta)^* = \omega^{\vec{q}}(\alpha)\omega^{\vec{q}}(\beta)\omega^{\vec{q}}(\alpha\beta)^*S^{\vec{q}}(\alpha,\beta),$$
$$\text{for all } \alpha,\beta\in P^{\vec{q}} \text{ and } |\omega^{\vec{q}}(\alpha)| = 1. \qquad (II.7)$$

Later on we shall prove this proposition for the various cases which may occur due to (I.14), (I.22), or (I.23). In order to be able to determine $\bar{\kappa}$ we consider Eq. (II.2) by utilizing (II.5) and (I.10),

$$\sum_{\sigma\in P:P^{\vec{q}}}\Delta^{\vec{q}}(\sigma,\beta\sigma)B^{\vec{q}*}_{\sigma,\sigma}(\beta)\mathbb{X}^{\kappa}(\sigma^{-1}\beta\sigma)^*e^{i\vec{q}(\sigma)\cdot\vec{t}}$$
$$= \sum_{\tau\in P:P^{\vec{q}}}\Delta^{\vec{q}}(\tau,\beta\tau)B^{\vec{q}}_{\tau,\tau}(\beta)\mathbb{X}^{\bar{\kappa}}(\tau^{-1}\beta\tau)e^{-i\vec{q}(\tau)\cdot\vec{t}}. \qquad (II.8)$$

Irrespective to the different cases (I.14), (I.22), and (I.23), which shall be discussed separately, we simplify Eq. (II.8) in the following section by utilizing the orthogonality relations for the unirreps of the translation group T. These simplifications lead in accordance to the above-mentioned cases to three different types of defining equations for $\bar{\kappa}$.

III. COMPLEX CONJUGATION OF PROJECTIVE POINT GROUP UNIRREPS

According to (I.14), (I.22), or (I.23) we have to distinguish several cases. If $\vec{q}$ does not belong to the surface of ΔBZ we take the special choice (I.20) for the left coset representatives, which yields for Eq. (II.8)

$$\sum_{\sigma\in P:P^{\vec{q}}} \Delta^{\vec{q}}(\sigma,\beta\sigma)e^{-i\vec{q}(\sigma)\cdot\vec{t}}\{B^{\vec{q}*}_{I\sigma,I\sigma}(\beta)X^{\kappa}(\sigma^{-1}\beta\sigma)^{*} - B^{\vec{q}}_{\sigma,\sigma}(\beta)X^{\bar{\kappa}}(\sigma^{-1}\beta\sigma)\} = 0. \qquad \text{(III.1)}$$

Thereby we have already taken into account that the unirreps of $P^{\vec{q}}$ are ordinary vector representations. Because of

$$B^{\vec{q}}_{I\sigma,I\sigma}(\beta)B^{\vec{q}}_{\sigma,\sigma}(\beta) = 1, \quad \text{for all } \beta\in\sigma P^{\vec{q}}\sigma^{-1} \qquad \text{(III.2)}$$

and the orthogonality relations for the unirreps of the translation group T, Eq. (III.1) yields the nonsurprising result

$$X^{\kappa}(\alpha)^{*} = X^{\bar{\kappa}}(\alpha), \quad \text{for all } \alpha\in P^{\vec{q}} \text{ and } \kappa,\bar{\kappa}\in A_{P^{\vec{q}}(1)}. \qquad \text{(III.3)}$$

This result implies for the corresponding vector unirreps of $P^{\vec{q}}$ that there must exist unitary matrices U^{κ} which are defined by

$$R^{\kappa}(\alpha)^{*} = U^{\kappa +}R^{\bar{\kappa}}(\alpha)U^{\kappa}, \quad \text{for all } \alpha\in P^{\vec{q}} \text{ and } \kappa\in A_{P^{\vec{q}}(1)}. \qquad \text{(III.4)}$$

However we are confronted with a quite different situation, if $\vec{q}$ is an element of the surface of ΔBZ. According to Eqs. (I.22) and (I.23) we have to distinguish two cases. Presupposing (I.22) is realized we obtain for Eq. (II.8)

$$\sum_{\sigma\in P:P^{\vec{q}}} \Delta^{\vec{q}}(\sigma,\beta\sigma)e^{-i\vec{q}(\sigma)\cdot\vec{t}}\{B^{\vec{q}*}_{I\sigma,I\sigma}(\beta)\mathrm{X}^{\kappa}(\sigma^{-1}\beta\sigma)^{*} - B^{\vec{q}}_{\sigma,\sigma}(\beta)\mathrm{X}^{\bar{\kappa}}(\sigma^{-1}\beta\sigma)\} = 0. \qquad \text{(III.5)}$$

As in the previous case, if we use the orthogonality relations for the unirreps of the translation group T, we arrive at a less trivial defining equation for $\bar{\kappa}\in A_{P^{\vec{q}}(S^{\vec{q}})}$, namely,

$$\mathrm{X}^{\bar{\kappa}}(\sigma^{-1}\beta\sigma) = B^{\vec{q}*}_{\sigma,\sigma}(\beta)B^{\vec{q}*}_{I\sigma,I\sigma}(\beta)\mathrm{X}^{\kappa}(\sigma^{-1}\beta\sigma)^{*}, \quad \text{for all } \beta\in\sigma P^{\vec{q}}\sigma^{-1}. \qquad \text{(III.6)}$$

A straightforward calculation yields for

$$B^{\vec{q}}_{\sigma,\sigma}(\beta)B^{\vec{q}}_{I\sigma,I\sigma}(\beta) = \exp[i\vec{\mathrm{Q}}\{\vec{q}(\sigma^{-1}\beta^{-1}\sigma)\}\cdot\vec{\tau}(I)] \quad \text{for all } \beta\in\sigma^{-1}P^{\vec{q}}\sigma \qquad \text{(III.7)}$$

which implies

$$\mathrm{X}^{\bar{\kappa}}(\alpha) = \exp[-i\vec{\mathrm{Q}}\{\vec{q}(\alpha^{-1})\}\cdot\vec{\tau}(I)]\mathrm{X}^{\kappa}(\alpha)^{*}, \quad \text{for all } \alpha\in P^{\vec{q}}. \qquad \text{(III.8)}$$

These equations uniquely define $\bar{\kappa}\in A_{P^{\vec{q}}(S^{\vec{q}})}$. In order to complete our proof we have to show due to Eq. (II.7) the correctness of

$$S^{\vec{q}}(\alpha,\beta)^{*} = \exp\{i[\vec{\mathrm{Q}}\{\vec{q}(\alpha^{-1})\} + \vec{\mathrm{Q}}\{\vec{q}(\beta^{-1})\} - \vec{\mathrm{Q}}\{\vec{q}(\beta^{-1}\alpha^{-1})\}]\cdot\vec{\tau}(I)\}S^{\vec{q}}(\alpha,\beta), \quad \text{for all } \alpha,\beta\in P^{\vec{q}} \qquad \text{(III.9)}$$

which would prove our proposition that $S^{\vec{q}}$ and $S^{\vec{q}*}$ must be equivalent factor systems. Together with (I.9) and

$$\vec{\mathrm{Q}}\{\vec{q}(\alpha\beta)\} = \vec{\mathrm{Q}}\{\vec{q}(\alpha)\} + D(\alpha)\vec{\mathrm{Q}}\{\vec{q}(\beta)\}, \quad \text{for all } \alpha,\beta\in P^{\vec{q}} \qquad \text{(III.10)}$$

which is a trivial consequence of the definition of the little cogroups $P^{\vec{q}}$, it is easy to verify Eq. (III.9). Inserting (III.7) and (III.8) into (II.3) we immediately obtain (II.4).

Concluding this case we remark that Eq. (III.8) requires the existence of unitary matrices U^{κ} which are defined by

$$\mathbb{R}^{\kappa}(\alpha)^{*} = \exp[i\vec{\mathrm{Q}}\{\vec{q}(\alpha^{-1})\}\cdot\vec{\tau}(I)]U^{\kappa +}\mathbb{R}^{\bar{\kappa}}(\alpha)U^{\kappa}, \quad \text{for all } \alpha\in P^{\vec{q}}. \qquad \text{(III.11)}$$

Contrary to the previous case [compare Eq. (III.4)], there enter into these defining equations unimodular factors which have to be taken into account.

For the last case where $I\in P^{\vec{q}}$ is realized, Eq. (II.8) turns out to be

$$\sum_{\sigma\in P:P^{\vec{q}}} \Delta^{\vec{q}}(\sigma,\beta\sigma)e^{-i\vec{q}(\sigma)\cdot\vec{t}}\{B^{\vec{q}*}_{\sigma,\sigma}(\beta)\mathrm{X}^{\kappa}(\sigma^{-1}\beta\sigma)^{*} - B^{\vec{q}}_{\sigma,\sigma}(\beta)\mathrm{X}^{\bar{\kappa}}(\sigma^{-1}\beta\sigma)\} = 0. \qquad \text{(III.12)}$$

since

$$I\in P^{\vec{q}} \Longleftrightarrow \vec{q}(I) = \vec{q} + \vec{\mathrm{Q}}\{\vec{q}(I)\}. \qquad \text{(III.13)}$$

Hence

$$\mathrm{X}^{\bar{\kappa}}(\sigma^{-1}\beta\sigma) = \{B^{\vec{q}*}_{\sigma,\sigma}(\beta)\}^{2}\mathrm{X}^{\kappa}(\sigma^{-1}\beta\sigma)^{*}, \text{ for all } \beta\in\sigma P^{\vec{q}}\sigma^{-1}. \qquad \text{(III.14)}$$

A straightforward calculation yields for

$$\{B^{\vec{q}}_{\sigma,\sigma}(\beta)\}^{2} = \exp[i\vec{\mathrm{Q}}\{\vec{q}(I)\}\cdot\vec{\tau}(\sigma^{-1}\beta\sigma)] \qquad \text{(III.15)}$$

so that the defining equation for $\bar{\kappa}\in A_{P^{\vec{q}}(S^{\vec{q}})}$ is given by

$$\mathrm{X}^{\bar{\kappa}}(\alpha) = \exp[-i\vec{\mathrm{Q}}\{\vec{q}(I)\}\cdot\vec{\tau}(\alpha)]\mathrm{X}^{\kappa}(\alpha)^{*}, \text{ for all } \alpha\in P^{\vec{q}}. \qquad \text{(III.16)}$$

As in the previous case a simple calculation shows the correctness of

$$S^{\vec{q}}(\alpha,\beta)^{*} = \exp\{i\vec{\mathrm{Q}}\{(I)\}\cdot[\vec{\tau}(\alpha) + \vec{\tau}(\beta) - \vec{\tau}(\alpha\beta)]\} \times S^{\vec{q}}(\alpha,\beta), \text{ for all } \alpha,\beta\in P^{\vec{q}}, \qquad \text{(III.17)}$$

where we have to use the definition (I.9) of the factor system $S^{\vec{q}}$.

Due to Eq. (III.16) we conclude that the equivalence between the projective unirreps which are connected by Eq. (III.16) must be expressible in the following way,

$$\mathbb{R}^{\kappa}(\alpha)^{*} = \exp[i\vec{\mathrm{Q}}\{\vec{q}(I)\}\cdot\vec{\tau}(\alpha)]U^{\kappa +}\mathbb{R}^{\bar{\kappa}}(\alpha)U^{\kappa}, \quad \text{for all } \alpha\in P^{\vec{q}}. \qquad \text{(III.18)}$$

These defining equations are similar to Eqs. (III.11), but are in general quite different from Eqs. (III.4) because of the occurrence of the nontrivial unimodular factors.

IV. CONCLUDING REMARKS

The aim of this paper was to derive equations which allow us to identify those equivalence classes whose corresponding space group unirreps are linked by complex conjugation, where we have considered space groups which contain inversion as a point group operation. We have shown that this problem can be reduced to the corresponding one for the equivalence classes of the little cogroups $P^{\vec{q}}$. Thereby we have obtained three different types of defining equations. Two types of these equations are nontrivial and are realized only if $\vec{q}$ belongs to the surface of ΔBZ, where the corresponding factor systems $S^{\vec{q}}$ are nontrivial. The reason for these nontrivial equations is due to the fact that the factor systems $S^{\vec{q}}$ and $S^{\vec{q}^*}$ are equivalent. Finally we remark that this feature of the factor systems $S^{\vec{q}}$ is inherent to the representation theory for all nonsymmorphic space groups, which contain inversion as a point group operation. A discussion concerning the same problem for space groups not containing inversion as a point group operation will be carried out in a forthcoming paper.

[1]R. Dirl, "Coupling coefficients: General theory," J. Math. Phys. **20**, 1562 (1979).

[2]J.L. Birman, "Theory of Crystal Space Groups and Infra-red Raman Processes of Insulated Crystals," in *Handbuch der Physik*, edited by S. Flügge (Springer, Berlin, 1974).

[3]Rh. Berenson and J.L. Birman, J. Math. Phys. **16**, 227 (1975).

[4]R. Dirl, J. Math. Phys. **18**, 2065 (1977).

[5]R. Dirl, "Multiplicities for space group representations," J. Math. Phys. (in press).

Coupling coefficients for space groups

R. Dirl

Institut für Theoretische Physik, TU Wien, A—1040 Wien, Karlsplatz 13, Austria
(Received 5 June 1978)

Unitary matrices connecting coupling with their corresponding Clebsch–Gordan coefficients are determined quite generally for nonsymmorphic space groups which contain the inversion as point group operation. The matrix elements of these unitary matrices factorize into two independent parts where a special choice for the sets of the left coset representatives simplifies the calculations.

INTRODUCTION

Preferring the second possibility of computing coupling coefficients for nonsymmorphic space groups, we shall report on in this article how the unitary matrices can be determined, which connect space group unirreps with their equivalent complex conjugate representations. The present method consists of calculating the matrix elements of these unitary matrices in terms of special CG coefficients, where the latter ones are determined by means of a general procedure, which has been discussed quite generally in Ref. 1. As in the previous paper we remark that only such space groups are considered which contain the inversion as point group operation.

We organize our paper as follows: In Sec. I we start by specializing the general formulas of Ref. 1 in order to be able to compute the above-mentioned special CG coefficients. The reason for calculating these special CG coefficients is that they define up to well-determined numerical factors the desired unitary matrices which connect space group unirreps to their complex conjugate counterpart. In Sec. III we determine for our three different cases (which may only occur for nonsymmorphic space groups) the corresponding unitary matrices quite generally. Thereby we show that the matrix elements of these unitary matrices have a special structure. Namely, they factorize into a part which concerns only the elements of the left cosets $P{:}P^{\vec{q}}$ and a second part which concerns the projective unirreps of the little cogroups $P^{\vec{q}}$.

I. COUPLING COEFFICIENTS FOR NONSYMMORPHIC SPACE GROUPS

According to the general considerations of Ref. 2 there exist two possibilities to compute coupling coefficients for space group representations. For the first possibility we have to specialize the general formulas (II.9)–(II.13) of Ref. 2 to ordinary space group representations by taking Eq. (I.4–I.8) of Ref. 3 into account. Thereby we would obtain immediately formulas which are identical to (I.3), (I.5), (I.6), (I.8), and (I.9) of Ref. 1, if $\vec{q}'(\underline{\tau}')$ is replaced by $-\vec{q}'(\underline{\tau}')$ and $B^{\vec{q}'}_{\underline{\tau}',\underline{\sigma}'}(\beta)\mathbb{R}^{\kappa'}_{d'c'}(\underline{\tau}'^{-1}\beta\underline{\sigma}')$ by its complex conjugate value.

For the second possibility which is in accordance to Eq. (III.1) of Ref. 2, respectively Eq. (II.1) of Ref. 3, we have to compute solely unitary matrices satisfying

$$\{D^{(\kappa,\vec{q})\uparrow G}(\beta\mid\vec{\tau}(\beta)+\vec{t})\}^{*} = U^{(\kappa,\vec{q})+}D^{(\bar{\kappa},\vec{q})\uparrow G}(\beta\mid\vec{\tau}(\beta)+\vec{t})U^{(\kappa,\vec{q})}, \quad \text{for all } (\beta\mid\vec{\tau}(\beta)+\vec{t})\in G. \tag{I.1}$$

Thereby it is assumed that the corresponding CG coefficients of the considered space groups are already known. With the aid of the general formula (III.8) of Ref. 2 the desired coupling coefficients are readily obtained,

$$\left[\begin{matrix}(\kappa,\vec{q}) & \overline{(\kappa',\vec{q}')} \\ \underline{\tau},d & \underline{\tau}',d'\end{matrix}\middle|\begin{matrix}(\kappa_0,\vec{q}_0) & w \\ \underline{\sigma},j & \end{matrix}\right] = \sum_{\underline{\tau}'',d''} U^{(\kappa',\vec{q}')*}_{\underline{\tau}'',d'';\underline{\tau}',d'}\left(\begin{matrix}(\kappa,\vec{q}) & \overline{(\kappa',\vec{q}')} \\ \underline{\tau},d & \underline{\tau}''d''\end{matrix}\middle|\begin{matrix}(\kappa_0,\vec{q}_0) & w=(\underline{\sigma}_v,c_v;\underline{\sigma}'_v,c'_v) \\ \underline{\sigma},j & \end{matrix}\right). \tag{I.2}$$

As already pointed out in Ref. 2 it should be noted that the multiplicity index w occuring on both sides of Eq. (I.2) is explained in terms of special column indices of the following Kronecker product,

$$\begin{aligned} D^{(\kappa,\vec{q})\uparrow G;\overline{(\kappa',\vec{q})\uparrow G}}(\beta\mid\vec{\tau}(\beta)+\vec{t}) &= D^{(\kappa,\vec{q})\uparrow G}(\beta\mid\vec{\tau}(\beta)+\vec{t})\otimes D^{\overline{(\kappa',\vec{q}')\uparrow G}}(\beta\mid\vec{\tau}(\beta)+\vec{t}) \\ &= D^{(\kappa,\vec{q})\uparrow G}(\beta\mid\vec{\tau}(\beta)+\vec{t})\otimes D^{(\bar{\kappa}',\vec{q}')\uparrow G}(\beta\mid\vec{\tau}(\beta)+\vec{t}). \end{aligned} \tag{I.3}$$

Consequently our problem reduces to the determination of unitary matrices satisfying (I.1) for the various cases which were considered in Ref. 3.

Due to Eq. (III.14) of Ref. 2 we have to compute

$$U^{(\kappa,\vec{q})}_{\underline{\tau}',d';\underline{\tau},d} = \sqrt{n_\kappa |P:P^{\vec{q}}|}\left(\begin{matrix}(\kappa,\vec{q}) & (\bar{\kappa},\vec{q}) \\ \underline{\tau},d & \underline{\tau}',d'\end{matrix}\middle|\begin{matrix}(0,\vec{0}) \\ \underline{e}1\end{matrix}\right), \tag{I.4}$$

 0022-2488/79/071570-05$01.00

where $(0,\vec{0})$ denotes the identity representation of the space group and the superfluous multiplicity index usually appearing (in our notation) on the right-hand side of (I.4) has been omitted. These special CG coefficients are because of (III.10) of Ref. 2 uniquely determined up to a common and therefore uninteresting phase factor. Using the notations of Ref. 1 we have to consider the following equations in order to obtain the desired unitary matrices (I.4):

$$\|\vec{\mathbf{B}}_{e,1}^{(\kappa,\vec{q})(\bar{\kappa},\vec{q}),(0,\vec{0})(\sigma,c;\sigma',c')}\|^2 = \left\{\begin{matrix}(\kappa,\vec{q}) & (\bar{\kappa},\vec{q}) \\ \sigma,c & \sigma',c'\end{matrix}\middle|\begin{matrix}(0,\vec{0}) & (\sigma,c;\sigma',c') \\ e,1 & \end{matrix}\right\}$$

$$= \delta_{\vec{q}(\sigma)+\vec{q}(\sigma');\vec{Q}[\vec{q}(\sigma)+\vec{q}(\sigma')]}\frac{1}{|P|}\sum_{\beta\in P}\Delta^{\vec{q}}(\sigma,\beta\sigma)\Delta^{\vec{q}}(\sigma',\beta\sigma')B^{\vec{q}}_{\sigma,\sigma}(\beta)B^{\vec{q}}_{\sigma',\sigma'}(\beta)\mathbb{R}^{\kappa}_{cc}(\sigma^{-1}\beta\sigma)\mathbb{R}^{\bar{\kappa}}_{c'c'}(\sigma'^{-1}\beta\sigma')>0. \tag{I.5}$$

Equation (I.5) is a special case of (I.5) of Ref. 1. This equation will be used as first step of our procedure in order to determine those values for $(\sigma,c;\sigma',c')$ that (I.5) is valid. Presupposing that this task is solved we obtain the special CG coefficients (I.4) by considering the corresponding special case of Eq. (I.8) of Ref. 1,

$$\left(\begin{matrix}(\kappa,\vec{q}) & (\bar{\kappa},\vec{q}) \\ \tau,d & \tau',d'\end{matrix}\middle|\begin{matrix}(0,\vec{0}) & (\sigma,c;\sigma',c') \\ e,1 & \end{matrix}\right)$$

$$= \|\vec{\mathbf{B}}_{e,1}^{(\kappa,\vec{q})(\bar{\kappa},\vec{q}),(0,\vec{0})(\sigma,c;\sigma',c')}\|^{-1}\delta_{\vec{q}(\tau)+\vec{q}(\tau');\vec{Q}[\vec{q}(\tau)+\vec{q}(\tau')]}\frac{1}{|P|}\sum_{\beta\in P}\Delta^{\vec{q}}(\tau,\beta\sigma)\Delta^{\vec{q}}(\tau',\beta\sigma')B^{\vec{q}}_{\tau,\sigma}(\beta)B^{\vec{q}}_{\tau',\sigma'}(\beta)\mathbb{R}^{\kappa}_{dc}(\tau^{-1}\beta\sigma)\mathbb{R}^{\bar{\kappa}}_{d'c'}(\tau'^{-1}\beta\sigma'). \tag{I.6}$$

In the following section we shall compute these CG coefficients quite generally for the various cases which have been discussed in Ref. 3.

II. STRUCTURE OF THE UNITARY MATRICES $U^{(\kappa,\vec{q})}$

As already pointed out we shall consider in this section different cases depending on $\vec{q}$ and its corresponding little cogroup $P^{\vec{q}}$. If $\vec{q}$ does not belong to the surface of ΔBZ we can choose in any way

$$I\in P:P^{\vec{q}}, \tag{II.1}$$

Since the inversion I cannot belong to $P^{\vec{q}}$. For this case we have to distinguish two further possibilities: Either $\vec{q}$ is an element of a general star (i.e., $P^{\vec{q}}=\{e\}$), or belongs to a star of higher symmetry (i.e., $P\supseteq P^{\vec{q}}\supset\{e\}$). If however $\vec{q}$ is an element of the surface of ΔBZ we have to discuss the following two situations:

$$I\in P:P^{\vec{q}} \tag{II.2}$$

$$I\in P^{\vec{q}}, \tag{II.3}$$

where the corresponding little cogroup $P^{\vec{q}}$ must be in any way nontrivial. Clearly (II.1) coincides with (II.2), whereas (II.3) is quite different to the foregoing situations. Nevertheless (II.3) is only realized for few points of the surface of ΔBZ.

A. $\vec{q}\notin$ surface of ΔBZ

From Eq. (I.5) we obtain for

$$\|\vec{\mathbf{B}}_{e,1}^{(\kappa,\vec{q})(\bar{\kappa},\vec{q});(0,\vec{0})(\sigma,c;\sigma',c')}\|^2$$

$$=\delta_{\sigma',I\sigma}\frac{1}{|P|}\sum_{\beta\in P}\Delta^{\vec{q}}(\sigma,\beta\sigma)B^{\vec{q}}_{\sigma,\sigma}(\beta)B^{\vec{q}}_{I\sigma,I\sigma}(\beta)$$

$$\times R^{\kappa}_{cc}(\sigma^{-1}\beta\sigma)R^{\bar{\kappa}}_{c'c'}(\sigma^{-1}\beta\sigma), \tag{II.4}$$

where (II.1) is used and we have to note that the unirreps of $P^{\vec{q}}$ are ordinary vector representations.

If $\vec{q}$ belongs to a general star, Eq. (II.4) reduces to

$$\|\vec{\mathbf{B}}_{e,1}^{(0,\vec{q})(0,\vec{q});(0,\vec{0})(\sigma,\sigma')}\|^2=\delta_{\sigma',I\sigma}\frac{1}{|P|},\quad \text{for all } \sigma\in P, \tag{II.5}$$

where the identical representation of $P^{\vec{q}}$ is denoted by 0 and the superfluous column indices $c=c'=1$ are omitted. Choosing $\sigma=e$, (I.6) yields

$$\left(\begin{matrix}(0,\vec{q}) & (0,\vec{q}) \\ \tau & \tau'\end{matrix}\middle|\begin{matrix}(0,\vec{0}) & (e;I) \\ e,1 & \end{matrix}\right)$$

$$=\delta_{\tau',I\tau}\frac{1}{\sqrt{|P|}}B^{\vec{q}}_{I\tau,I}(\tau);\quad \tau,\tau'\in P. \tag{II.6}$$

Hence we arrive due to Eq. (I.4) for the simplest case to the result

$$U^{(0,\vec{q})}_{\tau';\tau}=\left(\begin{matrix}(0,\vec{q}) & (0,\vec{q}) \\ \tau & \tau'\end{matrix}\middle|\begin{matrix}(0,\vec{0}) & (e;I) \\ e,1 & \end{matrix}\right)=\delta_{\tau',I\tau}B^{\vec{q}}_{I\tau,I}(\tau). \tag{II.7}$$

Concluding this simple case we can convince ourselves readily that $U^{(0,\vec{q})}$ is a unitary matrix satisfying (I.1).

If $\vec{q}$ is an element of a star of higher symmetry, Eq. (II.4) becomes

$$\|\vec{\mathbf{B}}_{e,1}^{(\kappa,\vec{q})(\bar{\kappa},\vec{q});(0,\vec{0})(\sigma,c;\sigma',c')}\|^2$$

$$=\delta_{\sigma',I\sigma}\frac{1}{|P|}\sum_{\alpha\in P^{\vec{q}}}R^{\kappa}_{cc}(\alpha)R^{\bar{\kappa}}_{c'c'}(\alpha) \tag{II.8}$$

since

$$B^{\vec{q}}_{\sigma,\sigma}(\beta)B^{\vec{q}}_{I\sigma,I\sigma}(\beta)=1,\quad \text{for all } \beta\in\sigma P^{\vec{q}}\sigma^{-1} \tag{II.9}$$

[see Eq. (III.2) of Ref. 3]. In order to gain more insight into the structure of (II.8) we define unitary matrices U^{κ} satisfying

$$R^{\kappa}(\alpha)^{*} = U^{\kappa+} R^{\bar{\kappa}}(\alpha) U^{\kappa}, \quad \text{for all } \alpha\in P^{\mathfrak{q}} \tag{II.10}$$

since $R^{\kappa *}$ and $R^{\bar{\kappa}}$ must be equivalent vector unirreps of $P^{\mathfrak{q}}$. Inserting (II.10) into (II.8), this equation turns out to be

$$\|\vec{\mathbf{B}}_{e,1}^{(\kappa,\mathfrak{q})(\bar{\kappa},\mathfrak{q});(0,\vec{0})(\sigma,c;\sigma',c')}\|^2 = \delta_{\sigma',I\sigma}\frac{1}{n_\kappa|P:P^{\mathfrak{q}}|}\,|U^{\kappa}_{c'c}|^2 \tag{II.11}$$

if the orthogonality relations for the matrix elements of the unirreps of $P^{\mathfrak{q}}$ are taken into account. The special choice

$$\sigma = e, \quad c = 1 \tag{II.12}$$

yields

$$\|\vec{\mathbf{B}}_{e,1}^{(\kappa,\mathfrak{q})(\bar{\kappa},\mathfrak{q});(0,\vec{0})(e,1;I,c')}\|^2 = \frac{|U^{\kappa}_{c'1}|^2}{n_\kappa|P:P^{\mathfrak{q}}|} \tag{II.13}$$

which is greater than zero, if c' is appropriately chosen, since at least one matrix element of the first column of U^{κ} must be different from zero. Inserting the special column indices

$$\sigma = e, \quad c = 1; \quad \sigma' = I, \quad c' = c_0' \tag{II.14}$$

of the Kronecker product $D^{(\kappa,\mathfrak{q})\uparrow G(\bar{\kappa},\mathfrak{q})\uparrow G}$ into the general formula (I.6), we obtain

$$\begin{pmatrix} (\kappa,\mathfrak{q}) & (\bar{\kappa},\mathfrak{q}) & (0,\vec{0}) & (e,1;I,c_0') \\ \tau,d & \tau',d' & e,1 & \end{pmatrix}$$

$$= \frac{\sqrt{n_\kappa|P:P^{\mathfrak{q}}|}}{|U^{\kappa}_{c_0'1}|}\,\delta_{\tau',I\tau}e^{i\mathfrak{q}(\tau)\cdot\vec{t}(\tau,I)}\frac{1}{|P|}$$

$$\times\sum_{\alpha\in P^{\mathfrak{q}}} R^{\kappa}_{d1}(\alpha)R^{\bar{\kappa}}_{d'c_0'}(\alpha). \tag{II.15}$$

Thereby we have used the relation

$$B^{\mathfrak{q}}_{\tau,e}(\beta)B^{\mathfrak{q}}_{I\tau,I}(\beta) = e^{i\mathfrak{q}(\tau)\cdot\vec{t}(\tau,I)}, \quad \text{for all } \beta\in P^{\mathfrak{q}} \tag{II.16}$$

which is independent of $\beta\in\tau P^{\mathfrak{q}}$, since $\mathfrak{q}(\tau^{-1}\beta) = \mathfrak{q}$ for all $\beta\in\tau P^{\mathfrak{q}}$. Utilizing Eq. (II.10) in order to simplify (II.15), we arrive at the result

$$\begin{pmatrix} (\kappa,\mathfrak{q}) & (\bar{\kappa},\mathfrak{q}) & (0,\vec{0}) & (e,1;I,c_0') \\ \tau,d & \tau'd' & e,1 & \end{pmatrix}$$

$$= \delta_{\tau',I\tau}\frac{e^{i\mathfrak{q}(\tau)\cdot\vec{t}(\tau,I)}}{\sqrt{n_\kappa|P:P^{\mathfrak{q}}|}}\,U^{\kappa}_{d'd}, \tag{II.17}$$

where the phase factor $U^{\kappa *}_{c_0'1}|U^{\kappa}_{c_0'1}|^{-1}$ has been fixed by one. Hence, due to Eq. (I.4) we obtain

$$U^{(\kappa,\mathfrak{q})}_{\tau'd';\tau,d} = \delta_{\tau',I\tau}\,e^{i\mathfrak{q}(\tau)\cdot\vec{t}(\tau,I)}U^{\kappa}_{d'd}. \tag{II.18}$$

With the aid of

$$B^{\mathfrak{q}}_{I\sigma,I\sigma'}(\beta)\exp[-i\mathfrak{q}(\sigma)\cdot\vec{t}(\sigma,I) + i\mathfrak{q}(\sigma')\cdot\vec{t}(\sigma',I)]$$

$$= B^{\mathfrak{q}*}_{\sigma,\sigma'}(\beta), \qquad \text{for all } \beta\in\sigma P^{\mathfrak{q}}\sigma'^{-1} \tag{II.19}$$

it can be shown by a straightforward calculation that Eq. (I.1) is valid, presupposing that (II.10) holds. Equations (II.18) shows obviously that the matrix elements of $U^{(\kappa,\mathfrak{q})}$ factorizes into a part which depends only on left coset representatives and a second part which concerns only the complex conjugation of the vector unirreps of $P^{\mathfrak{q}}$. The computation of the unitary matrices U^{κ}, which are defined by Eq. (II.10) can be carried out in two different ways.

At the first possibility we can inspect Eq. (II.10) for an appropriated chosen set of generating elements of $P^{\mathfrak{q}}$. The second possibility consists of computing directly Eq. (II.15), i.e.,

$$U^{\kappa}_{d'd} = \sqrt{\frac{n_\kappa}{|P^{\mathfrak{q}}|}}\left\{\sum_{\alpha\in P^{\mathfrak{q}}} R^{\kappa}_{11}(\alpha)R^{\bar{\kappa}}_{c_0'c_0'}(\alpha)\right\}^{-1/2}\sum_{\beta\in P^{\mathfrak{q}}}$$

$$\times R^{\kappa}_{d1}(\beta)R^{\bar{\kappa}}_{d'c_0'}(\beta). \tag{II.20}$$

B. $\mathfrak{q}\in$surface of ΔBZ

If $\mathfrak{q}$ belongs to the surface of ΔBZ we have to distinguish between the cases (II.2) and (II.3). Considering the case $I\in P:P^{\mathfrak{q}}$, Eq. (II.4) achieves the form

$$\|\vec{\mathbf{B}}_{e,1}^{(\kappa,\mathfrak{q})(\bar{\kappa},\mathfrak{q}),(0,\vec{0})(\sigma,c;\sigma',c')}\|^2$$

$$= \delta_{\sigma',I\sigma}\frac{1}{|P|}\sum_{\alpha\in P^{\mathfrak{q}}}\exp[i\vec{Q}\{\mathfrak{q}(\alpha^{-1})\}\cdot\vec{t}(I)]\mathbb{R}^{\kappa}_{cc}(\alpha)\mathbb{R}^{\bar{\kappa}}_{c'c'}(\alpha) \tag{II.21}$$

at which we have already used

$$B^{\mathfrak{q}}_{\sigma,\sigma}(\beta)B^{\mathfrak{q}}_{I\sigma,I\sigma}(\beta)$$

$$= \exp[iQ\{\mathfrak{q}(\sigma^{-1}\beta\sigma)\}\cdot\vec{t}(I)], \quad \text{for all } \beta\in\sigma P^{\mathfrak{q}}\sigma^{-1} \tag{II.22}$$

[see Eq. (III.7) of Ref. 3]. Similar to the previous case we define by

$$\mathbb{R}^{\kappa}(\alpha)^{*} = \exp[i\vec{Q}\{\mathfrak{q}(\alpha^{-1})\}\cdot\vec{t}(I)]\,U^{\kappa+}\mathbb{R}^{\bar{\kappa}}(\alpha)U^{\kappa},$$

$$\text{for all } \alpha\in P^{\mathfrak{q}} \tag{II.23}$$

unitary matrices U^{κ}; $\kappa\in A_{P^{\mathfrak{q}}(S^{\mathfrak{q}})}$, which must exist in accordance to Eq. (III.11) of Ref. 3. However it should be noted that in contrary to Eq. (II.10) there enter into Eq. (II.23) projective unirreps which give rise to the unimodular factors $\exp[i\vec{Q}\{\mathfrak{q}(\alpha^{-1})\}\cdot\vec{t}(I)]$. The reason for the occurence of these unimodular factors originates from the fact that the factor systems $S^{\mathfrak{q}}$ and $S^{\mathfrak{q}*}$ are equivalent. Introducing (II.23) into (II.20) we obtain

$$\|\vec{\mathbf{B}}_{e,1}^{(\kappa,\mathfrak{q})(\bar{\kappa},\mathfrak{q}),(0,\vec{0})(\sigma,c;\sigma',c')}\|^2 = \delta_{\sigma',I\sigma}\frac{|U^{\kappa}_{c'c}|^2}{n_\kappa|P:P^{\mathfrak{q}}|} \tag{II.24}$$

if taking the orthogonality relations for the matrix elements of the projective unirreps of $P^{\mathfrak{q}}$ into account. The special choice

$$\sigma = e, \quad c = 1 \tag{II.25}$$

yields

$$\sigma' = I, \quad c' = c_0' \tag{II.26}$$

where c_0' is appropriately chosen so that

$$\|\vec{\mathbf{B}}_{e,1}^{(\kappa,\mathfrak{q})(\bar{\kappa},\mathfrak{q}),(0,\vec{0})(e,1;Ic_0')}\|^2 = \frac{|U^{\kappa}_{c_0'1}|^2}{n_\kappa|P:P^{\mathfrak{q}}|} > 0 \tag{II.27}$$

is satisfied. Inserting the special column indices (II.25) and (II.26) into (I.6) it follows for

$$\begin{pmatrix} (\kappa,\vec{q}) & (\bar{\kappa},\vec{q}) & (0,\vec{0}) & (e,1;I,c_0') \\ \underline{\tau},d & \underline{\tau}',d & e,1 & \end{pmatrix}$$

$$= \frac{\sqrt{n_\kappa |P:P^{\vec{q}}|}}{|U^\kappa_{c_0'1}|} \delta_{\underline{\tau}',I\underline{\tau}} e^{i\vec{q}(\underline{\tau})\cdot\vec{t}(\underline{\tau},I)}$$

$$\times \frac{1}{|P|} \sum_{\alpha\in P^{\vec{q}}} \exp[i\vec{Q}\{\vec{q}(\alpha^{-1})\}\cdot\vec{\tau}(I)] \mathbb{R}^\kappa_{d1}(\alpha) \mathbb{R}^{\bar{\kappa}}_{d'c_0'}(\alpha).$$

(II.28)

This result is achieved with the aid of

$$B^{\vec{q}}_{\underline{\tau},e}(\beta) B^{\vec{q}}_{I\underline{\tau},I}(\beta) = \exp[i\vec{q}(\underline{\tau})\cdot\vec{t}(\underline{\tau},I)]$$

$$\times \exp[i\vec{Q}\{\vec{q}(\beta^{-1}\underline{\tau})\}\cdot\vec{\tau}(I)],$$

$$\text{for all } \beta\in\underline{\tau}P^{\vec{q}}, \qquad \text{(II.29)}$$

where we have to note that the first factor in (II.29) is independent of $\beta\in\underline{\tau}P^{\vec{q}}$. Utilizing once again (II.23), Eq. (II.28) becomes

$$\begin{pmatrix} (\kappa,\vec{q}) & (\bar{\kappa},\vec{q}) & (0,\vec{0}) & (e,1;I,c_0') \\ \underline{\tau},d & \underline{\tau}',d' & e,1 & \end{pmatrix}$$

$$= \delta_{\underline{\tau}',I\underline{\tau}} \frac{e^{i\vec{q}(\underline{\tau})\cdot\vec{t}(\underline{\tau},I)}}{\sqrt{n_\kappa |P:P^{\vec{q}}|}} U^\kappa_{d'd}, \qquad \text{(II.30)}$$

where the phase factor $U^{\kappa *}_{c_0'1} |U^\kappa_{c_0'1}|^{-1}$ has been chosen arbitrarily as one, like for (II.17). Thus (I.4) achieves for this case the final form

$$U^{(\kappa,\vec{q})}_{\underline{\tau}'d';\underline{\tau},d} = \delta_{\underline{\tau}',I\underline{\tau}} e^{i\vec{q}(\underline{\tau})\cdot\vec{t}(\underline{\tau},I)} U^\kappa_{d'd}. \qquad \text{(II.31)}$$

As in the previous case there remains to carry out a simple calculation by using relations of the type

$$B^{\vec{q}}_{\underline{\tau},\underline{\sigma}}(\beta) B^{\vec{q}}_{I\underline{\tau},I\underline{\sigma}}(\beta)$$

$$= \exp[i\vec{q}(\underline{\tau})\cdot\vec{t}(\underline{\tau},I) - i\vec{q}(\underline{\sigma})\cdot\vec{t}(\underline{\sigma},I)]$$

$$\times \exp[i\vec{Q}\{\vec{q}((\underline{\tau}^{-1}\beta\underline{\sigma})^{-1})\}\cdot\vec{\tau}(I)],$$

$$\text{for all } \beta\in\underline{\tau}P^{\vec{q}}\underline{\sigma}^{-1} \qquad \text{(II.32)}$$

in order to show the correctness of the defining equation (I.1) for the unitary matrices $U^{(\kappa,\vec{q})}$. Comparing (II.18) with (II.31) we realize that we are confronted with a similar situation, since (II.31) factorizes in the same way as (II.18), but the unitary matrices U^κ occuring in (II.31) are defined by more complicated equations [see Eq. (II.23)].

As in the previous case there are two possibilities to compute the unitary matrices U^κ. Either we inspect Eq. (II.23) for a set of generating elements of $P^{\vec{q}}$, or we compute

$$U^\kappa_{d'd} = \sqrt{\frac{n_\kappa}{|P^{\vec{q}}|}} \left\{ \sum_{\alpha\in P^{\vec{q}}} e^{i\vec{Q}\{\vec{q}(\alpha^{-1})\}\cdot\vec{\tau}(I)} \mathbb{R}^\kappa_{11}(\alpha) \mathbb{R}^{\bar{\kappa}}_{c_0'c_0'}(\alpha) \right\}^{1/2}$$

$$\times \sum_{\beta\in P^{\vec{q}}} e^{i\vec{Q}\{\vec{q}(\beta^{-1})\}\cdot\vec{\tau}(I)} \mathbb{R}^\kappa_{d1}(\beta) \mathbb{R}^{\bar{\kappa}}_{d'c_0'}(\beta). \qquad \text{(II.33)}$$

Presupposing that the second possibility $I\in P^{\vec{q}}$ is realized, Eq. (II.4) takes the form

$$\|\vec{\mathbf{B}}^{(\kappa,\vec{q})(\bar{\kappa},\vec{q});(0,\vec{0})(\sigma,c;\sigma',c')}_{e,1}\|^2$$

$$= \delta_{\sigma,\sigma'} \frac{1}{|P|} \sum_{\alpha\in P^{\vec{q}}} e^{i\vec{Q}\{\vec{q}(I)\}\cdot\vec{\tau}(\alpha)} \mathbb{R}^\kappa_{cc}(\alpha) \mathbb{R}^{\bar{\kappa}}_{c'c'}(\alpha). \qquad \text{(II.34)}$$

Thereby we have already used the relations

$$\{B^{\vec{q}}_{\underline{\sigma},\underline{\sigma}}(\beta)\}^2 = e^{i\vec{Q}\{\vec{q}(I)\}\cdot\vec{\tau}(\underline{\sigma}^{-1}\beta\underline{\sigma})}, \quad \text{for all } \beta\in\underline{\sigma}P^{\vec{q}}\underline{\sigma}^{-1} \qquad \text{(II.35)}$$

[see Eq. (III.15) of Ref. 3]. As in the foregoing case we define by means of

$$\mathbb{R}^\kappa(\alpha)^* = e^{i\vec{Q}\{\vec{q}(I)\}\cdot\vec{\tau}(\alpha)} U^{\kappa +} \mathbb{R}^{\bar{\kappa}}(\alpha) U^\kappa, \quad \text{for all } \alpha\in P^{\vec{q}}, \qquad \text{(II.36)}$$

unitary matrices U^κ; $\kappa\in A_{P^{\vec{q}}(S^{\vec{q}})}$, whose existence is guaranteed by Eq. (III.16) of Ref. 3. Taking (II.36) and the orthogonality relations for the matrix elements of the projective unirreps of $P^{\vec{q}}$ into account, Eq. (II.34) turns out to be

$$\|\vec{\mathbf{B}}^{(\kappa,\vec{q})(\bar{\kappa},\vec{q}),(0,\vec{0})(\sigma,c;\sigma',c')}_{e,1}\|^2 = \delta_{\sigma,\sigma'} \frac{|U^\kappa_{c'c}|^2}{n_\kappa |P:P^{\vec{q}}|}. \qquad \text{(II.37)}$$

Consequently,

$$\|\vec{\mathbf{B}}^{(\kappa,\vec{q})(\bar{\kappa},\vec{q}),(0,\vec{0})(e,1;e,c')}_{e,1}\|^2 = \frac{|U^\kappa_{c'1}|^2}{n_\kappa |P:P^{\vec{q}}|} > 0 \qquad \text{(II.38)}$$

implies that

$$\underline{\sigma} = e, \quad c = 1; \quad \underline{\sigma}' = e, \quad c' = c_0' \qquad \text{(II.39)}$$

is an appropriated choice in order to determine the corresponding CG coefficients. Thus Eq. (I.6) achieves the form

$$\begin{pmatrix} (\kappa,\vec{q}) & (\bar{\kappa},\vec{q}) & (0,\vec{0}) & (e,1;e,c_0') \\ \underline{\tau},d & \underline{\tau}',d' & e,1 & \end{pmatrix}$$

$$= \frac{\sqrt{n_\kappa |P:P^{\vec{q}}|}}{|U^\kappa_{c_0'1}|} \delta_{\underline{\tau}',\underline{\tau}} \frac{1}{|P|}$$

$$\times \sum_{\alpha\in P^{\vec{q}}} e^{i\vec{Q}\{\vec{q}(I)\}\cdot\vec{\tau}(\alpha)} \mathbb{R}^\kappa_{d1}(\alpha) \mathbb{R}^{\bar{\kappa}}_{d'c_0'}(\alpha), \qquad \text{(II.40)}$$

where we have inserted (II.39) into (I.6) and

$$\{B^{\vec{q}}_{\underline{\tau},e}(\beta)\}^2 = e^{i\vec{Q}\{\vec{q}(I)\}\cdot\vec{\tau}(\underline{\tau}^{-1}\beta)}, \quad \text{for all } \beta\in\underline{\tau}P^{\vec{q}} \qquad \text{(II.41)}$$

was taken into account. Formula (II.40) turns out to be

$$\begin{pmatrix} (\kappa,\vec{q}) & (\bar{\kappa},\vec{q}) & (0,\vec{0}) & (e,1;e,c_0') \\ \underline{\tau},d & \underline{\tau}',d' & e,1 & \end{pmatrix}$$

$$= \delta_{\underline{\tau}',\underline{\tau}} \frac{1}{\sqrt{n_\kappa |P:P^{\vec{q}}|}} U^\kappa_{d'd} \qquad \text{(II.42)}$$

if Eq. (II.36) is utilized once again. Due to (I.4) we arrive at the final result

$$U^{(\kappa,\vec{q})}_{\underline{\tau}'d';\underline{\tau},d} = \delta_{\underline{\tau}',\underline{\tau}} U^\kappa_{d'd}, \qquad \text{(II.43)}$$

which has the same structure as (II.18) and (II.31), but requires the inspection of (II.36), if the corresponding U^κ shall be determined. With the aid of the following equations

$$\{B^{\vec{q}}_{\sigma,\sigma'}(\beta)\}^2 = e^{i\vec{Q}\{\vec{q}(I)\}\cdot\vec{\tau}(\sigma^{-1}\beta\sigma')}, \quad \text{for all } \beta\in\sigma P^{\vec{q}}\sigma'^{-1} \tag{II.44}$$

the defining equation (I.1) for the corresponding unitary matrices $U^{(\kappa,\vec{q})}$ is readily verified.

As in the previous cases there exist two possibilities for determining the unitary matrices U^κ. Either to inspect Eq. (II.36) for a set of generating elements of $P^{\vec{q}}$, or by computing (II.40), i.e.,

$$U^\kappa_{d'd} = \sqrt{\frac{n_\kappa}{|P^{\vec{q}}|}} \left\{ \sum_{\alpha\in P^{\vec{q}}} e^{i\vec{Q}\{\vec{q}(I)\}\cdot\vec{\tau}(\alpha)} \mathbb{R}^\kappa_{11}(\alpha)\mathbb{R}^{\bar{\kappa}}_{c_0'c_0'}(\alpha)\right\}^{-1/2}$$

$$\times \sum_{\beta\in P^{\vec{q}}} e^{i\vec{Q}\{\vec{q}(I)\}\cdot\vec{\tau}(\beta)} \mathbb{R}^\kappa_{d1}(\beta)\mathbb{R}^{\bar{\kappa}}_{d'c_0'}(\beta). \tag{II.45}$$

Concluding this section we remark that the special case of symmorphic space groups is contained in a consistent way in our formulas (II.18), (II.31), and (II.43) (if $I\in P\simeq G/T$ is satisfied). Thereby the unimodular factors reduce to one and the defining equations (II.10), (II.23), and (II.36) for the matrices U^κ become of the type (II.10).

III. CONCLUDING REMARKS

The aim of this paper was to demonstrate the utility of the present method for computing those unitary matrices, which connect space group unirreps with their equivalent complex conjugate representations. Concerning the considered space groups it was assumed that they must contain the inversion as point group operation. Thereby we have shown that the matrix elements of these unitary matrices factorize into two characteristic parts due to the special choice for the sets $P : P^{\vec{q}}$ of the left coset representatives. Presupposing these unitary matrices have been determined, coupling coefficients are readily obtained in terms of the corresponding CG coefficients, if the latter ones are known and vice versa. The same problem for space groups not containing the inversion as point group operation should be discussed in a forthcoming paper.

[1]R. Dirl, "Clebsch–Gordan coefficients for space groups," J. Math. Phys. **20**, 671 (1979).
[2]R. Dirl, "Coupling coefficients: General theory," J. Math. Phys. **20**, 1562 (1979).
[3]R. Dirl, "Complex conjugation of space group representations," J. Math. Phys. **20**, 1566 (1979).

Coupling coefficients for *Pn*3*n*

R. Dirl

Institut für Theoretische Physik, TU Wien, A—1040 Wien, Karlspatz 13, Austria
(Received 5 June 1978)

A general method for calculating Clebsch–Gordan coefficients is applied in order to compute for typical examples of $Pn3n$ those unitary matrices which connect coupling—with their corresponding Clebsch–Gordan coefficients.

INTRODUCTION

For the present paper we illustrate with the aid of some typical examples concerning the nonsymmorphic space group $Pn3n$ how the proposed method works. We compute for these examples those unitary matrices which connect uniquely coupling coefficients with their corresponding CG coefficients.

The organization of the material of this paper is as follows: In Sec. I we recall briefly some definitions and notations concerning the space group $Pn3n$. Furthermore we list for our examples the little cogroups $P^{\vec{q}}$, suitable chosen sets $P{:}P^{\vec{q}}$ of left coset representatives, and complete sets of projective unirreps for $P^{\vec{q}}$. In Sec. II we determine for our examples the equivalence classes of the projective unirreps of $P^{\vec{q}}$, which are linked by complex conjugation. Thereby we have to distinguish three different cases due to our general procedure. The unitary matrices connecting projective unirreps of $P^{\vec{q}}$ with their equivalent complex conjugate representations are computed in the following section. The corresponding unitary matrices which connect vector unirreps of $Pn3n$ with their equivalent complex conjugate representations are readily obtained by means of general formulas. Presupposed CG coefficients for $Pn3n$ are known the corresponding coupling coefficients follow immediately from our procedure.

I. DISCUSSION OF VARIOUS CASES: DEFINITIONS AND NOTATIONS

As in Ref. 1 we choose the nonsymmorphic space group $Pn3n$ as an example in order to demonstrate the utility of the present method when computing the unitary matrices which connect space group unirreps with their equivalent complex conjugate representations. Throughout this paper we use the same definitions and notations as introduced in Ref. 1. For the sake of simplicity some of them are recalled. The nontrivial lattice translations are defined by

$$\vec{\tau}(n)=\vec{0},\quad \text{for all } n\in O, \tag{I.1}$$

$$\vec{\tau}(In)=(1/2,1/2,1/2),\quad \text{for all } n\in O, \tag{I.2}$$

where the group element I of the point group $O_h=O\times\{E,I\}$ denotes the inversion. The fundamental domain ΔBZ of the corresponding Brillouin zone is given by Eq. (I.3) of Ref. 1. The orthogonal matrices $D(\alpha)$; $\alpha\in P\simeq G/T$ are readily obtainable from Table 1.4 of Ref. 2. Complete sets of projective unirreps with their corresponding factor systems of the little cogroups $P^{\vec{q}}$ are listed in full detail in Ref. 3 for all points of the surface of ΔBZ and for some $\vec{q}$'s lying inside of ΔBZ.

The present method can be applied for computing coupling coefficients, if the unirreps of the considered space group are known, where the special choice for the sets $P:P^{\vec{q}}$ [due to Eq. (I.20) of Ref. 4] has to be taken into account in any way.

Due to Ref. 4, three different cases may happen when determining the unitary matrices $U^{(\kappa,\vec{q})}$. In the following we shall consider for these cases typical examples and list complete sets of projective unirreps of the corresponding little cogroups $P^{\vec{q}}$, where for the equivalence classes $\kappa\in A_{P^{\vec{q}}(S^{\vec{q}})}$ the same symbols are used as in Refs. 1 and 3. The matrices are written down for appropriated chosen sets of generating elements of $P^{\vec{q}}$.

Case (i):

$$\vec{q}=\pi(0,y,0)\in\Delta BZ\Leftarrow 0<y<1, \tag{I.3}$$

$$P^{\vec{q}}=P^{\Delta}=\{E,\sigma_x\}\,\textcircled{s}\,\{E,C_{2y},C_{4y}^{\pm}\}=C_{4v}, \tag{I.4}$$

$$O_h:P^{\Delta}=\{E,C_{4z}^{-},C_{4x}^{+};I,IC_{4z}^{-},IC_{4x}^{+}\}, \tag{I.5}$$

$$R^{\mu}:\mu=0,1,\quad C_{4y}^{+}\to 1;\quad \sigma_x\to(-1)^{\mu}, \tag{I.6}$$

$$\mu=2,3,\quad C_{4y}^{+}\to -1;\quad \sigma_x\to(-1)^{\mu}, \tag{I.7}$$

$$\mu=5,\quad C_{4y}^{+}\to\begin{vmatrix}-i&0\\0&i\end{vmatrix};\quad \sigma_x\to\begin{vmatrix}0&1\\1&0\end{vmatrix}. \tag{I.8}$$

Case (ii):

$$\vec{q}=\pi(x,1,x)\in\Delta BZ\Leftarrow 0<x<1, \tag{I.9}$$

$$P^{\vec{q}}=P^{S}=\{E,\sigma_y\}\times\{E,\sigma_{de}\}, \tag{I.10}$$

$$P:P^{S}=\{E,C_{4y}^{+},C_{4z}^{\pm},C_{4x}^{\pm};I,IC_{4y}^{+},IC_{4z}^{\pm},IC_{4x}^{\pm}\}, \tag{I.11}$$

$$\mathbb{R}^{(0)\uparrow P^{S}}:\sigma_{de}\to\begin{vmatrix}1&0\\0&-1\end{vmatrix};\sigma_y\to\begin{vmatrix}0&-1\\1&0\end{vmatrix}. \tag{I.12}$$

Case (iii,a):

$$\vec{q}=\pi(0,1,0)\in\Delta BZ, \tag{I.13}$$

$$P^{\vec{q}}=P^{X}=\{E,I\}\times P^{\Delta}=\{E,I\}\times C_{4v}, \tag{I.14}$$

$$P:P^{X}=O:P^{\Delta}=\{E,C_{31}^{\pm}\}, \tag{I.15}$$

$$\mathbb{R}^{(\kappa,\mu)\uparrow P^{X}}:\kappa=0,1;\quad \mu=5,$$

$$C_{4y}^{+}\to\begin{vmatrix}-i&0\\0&i\end{vmatrix};\quad \sigma_x\to\begin{vmatrix}0&1\\1&0\end{vmatrix};\quad I\to(-1)^{\kappa}\begin{vmatrix}i&0\\0&-i\end{vmatrix}, \tag{I.16}$$

$$\mathbb{R}^{(\mu=0)\uparrow P^{X}}:$$

$$C_{4y}^{+}\to\begin{vmatrix}1&0\\0&1\end{vmatrix};\quad \sigma_x\to\begin{vmatrix}1&0\\0&-1\end{vmatrix};\quad I\to\begin{vmatrix}0&-1\\1&0\end{vmatrix}, \tag{I.17}$$

 0022-2488/79/071575-04$01.00

$\mathbb{R}^{(\mu=2)\uparrow P^X}$:

$$C^+_{4y}\rightarrow\begin{vmatrix}-1 & 0\\ 0 & -1\end{vmatrix};\quad \sigma_x\rightarrow\begin{vmatrix}1 & 0\\ 0 & -1\end{vmatrix};\quad I\rightarrow\begin{vmatrix}0 & -1\\ 1 & 0\end{vmatrix}. \tag{I.18}$$

Case (iii,b):

$$\vec{q}=\pi(1,1,1)\in\Delta BZ, \tag{I.19}$$

$$P^{\vec{q}}=P^R=O_h, \tag{I.20}$$

$$P:P^R=\{E\}, \tag{I.21}$$

$\mathbb{R}^{(\kappa,\mu=2)\uparrow O}h:\kappa=0,1;\omega=e^{-i2\pi/3}$,

$$C_{2x}\rightarrow\begin{vmatrix}1 & 0\\ 0 & 1\end{vmatrix};\quad C_{2z}\rightarrow\begin{vmatrix}1 & 0\\ 0 & 1\end{vmatrix},$$

$$C^+_{31}\rightarrow\begin{vmatrix}\omega & 0\\ 0 & \omega^2\end{vmatrix};\quad C_{2d}\rightarrow\begin{vmatrix}0 & 1\\ 1 & 0\end{vmatrix};\quad I\rightarrow(-1)^\kappa\begin{vmatrix}i & 0\\ 0 & -i\end{vmatrix}, \tag{I.22}$$

$\mathbb{R}^{(\mu=0)\uparrow O}h$:

$$C_{2x}\rightarrow\begin{vmatrix}1 & 0\\ 0 & 1\end{vmatrix};\quad C_{2z}\rightarrow\begin{vmatrix}1 & 0\\ 0 & 1\end{vmatrix},$$

$$C^+_{31}\rightarrow\begin{vmatrix}1 & 0\\ 0 & 1\end{vmatrix};\quad C_{2d}\rightarrow\begin{vmatrix}1 & 0\\ 0 & -1\end{vmatrix};\quad I\rightarrow\begin{vmatrix}0 & -1\\ 1 & 0\end{vmatrix}, \tag{I.23}$$

$\mathbb{R}^{(\mu=4)\uparrow O_h}$:

$$C_{2x}\rightarrow\begin{vmatrix}-1 & 0 & 0 & 0 & 0 & 0\\ 0 & 1 & 0 & 0 & 0 & 0\\ 0 & 0 & -1 & 0 & 0 & 0\\ 0 & 0 & 0 & -1 & 0 & 0\\ 0 & 0 & 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 0 & 0 & -1\end{vmatrix};$$

$$C_{2z}\rightarrow\begin{vmatrix}1 & 0 & 0 & 0 & 0 & 0\\ 0 & -1 & 0 & 0 & 0 & 0\\ 0 & 0 & -1 & 0 & 0 & 0\\ 0 & 0 & 0 & 1 & 0 & 0\\ 0 & 0 & 0 & 0 & -1 & 0\\ 0 & 0 & 0 & 0 & 0 & -1\end{vmatrix},$$

$$C^+_{31}\rightarrow\begin{vmatrix}0 & 0 & 1 & 0 & 0 & 0\\ 1 & 0 & 0 & 0 & 0 & 0\\ 0 & 1 & 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0 & 0 & 1\\ 0 & 0 & 0 & 1 & 0 & 0\\ 0 & 0 & 0 & 0 & 1 & 0\end{vmatrix};$$

$$C_{2d}\rightarrow\begin{vmatrix}0 & 0 & 1 & 0 & 0 & 0\\ 0 & -1 & 0 & 0 & 0 & 0\\ 1 & 0 & 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0 & 0 & -1\\ 0 & 0 & 0 & 0 & 1 & 0\\ 0 & 0 & 0 & -1 & 0 & 0\end{vmatrix},$$

$$I\rightarrow\begin{vmatrix}0 & 0 & 0 & -1 & 0 & 0\\ 0 & 0 & 0 & 0 & -1 & 0\\ 0 & 0 & 0 & 0 & 0 & -1\\ 1 & 0 & 0 & 0 & 0 & 0\\ 0 & 1 & 0 & 0 & 0 & 0\\ 0 & 0 & 1 & 0 & 0 & 0\end{vmatrix}. \tag{I.24}$$

II. COMPLEX CONJUGATION OF SPACE GROUP UNIRREPS

As already pointed out in Refs. 5 and 6, coupling coefficients for a given group can be determined either by the general method given in Ref. 5, or, as proposed in Ref. 6 [see Eq. (I.2) of Ref. 6] by computing the unitary matrices $U^{(\kappa,\vec{q})}$ which are defined by Eq. (I.1) of Ref. 6. The second approach presupposes the explicite knowledge of the corresponding CG coefficients of the considered space group. According to our method the first step consists of determining the vector unirreps of $Pn3n$ which are linked by complex conjugation.

$$\{(\kappa,\vec{q})\uparrow G\}^*=(\bar{\kappa},\vec{q})\uparrow G;\quad \vec{q}\in\Delta BZ,\quad \kappa,\bar{\kappa}\in A_{P^{\vec{q}}(S^{\vec{q}})}. \tag{II.1}$$

Thereby it should be noted that for the sake of simplicity the prime of κ' and $\vec{q}'$ [occuring in Eq. (I.1) of Ref. 6] have been omitted, respectively; Eq. (II.5) of Ref. 4 has been taken into account. Hence in order to determine $\bar{\kappa}$ for a given κ we have to inspect one of the equations of (III.3), (III.8), and (III.16) of Ref. 4 depending on the case which shall be considered.

Case (i): Since $\vec{q}$ does not belong to the surface of ΔBZ, Eq. (III.3) of Ref. 4 must be inspected in order to be able to determine the equivalence classes of $A_{P^{\vec{q}}(S^{\vec{q}})}$ which are linked by complex conjugation,

$$X^\kappa(\alpha)^*=X^{\bar{\kappa}}(\alpha),\quad \text{for all }\alpha\in P^\Delta. \tag{II.2}$$

Identifying in accordance to our notation the symbol μ by κ we obtain immediately from Eqs. (I.6–8)

$$X^\kappa(\alpha)^*=X^\kappa(\alpha),\quad \text{for all }\alpha\in P^\Delta, \tag{II.3}$$

i.e., the vector unirreps of P^Δ are all equivalent to their complex conjugates.

$$\kappa^*=\bar{\kappa}=\kappa. \tag{II.4}$$

Case (ii): Since $\vec{q}$ belongs to the surface of ΔBZ and the inversion I does not belong to P^S we have to investigate Eq. (III.8) of Ref. 4.

$$\mathbb{X}^\kappa(\alpha)^*\exp[-i\vec{Q}\{\vec{q}(\alpha^{-1})\}\cdot\vec{\tau}(I)]=\mathbb{X}^{\bar{\kappa}}(\alpha),\quad \text{for all }\alpha\in P^S. \tag{II.5}$$

Due to (I.10) there exists only one projective unirrep for P^S which has as trivial consequence

$$\kappa=(0)\uparrow P^S:\kappa^*=\kappa. \tag{II.6}$$

A simple calculation yields for the unimodular factors appearing in Eq. (II.5)

$$\exp[-i\vec{Q}\{\vec{q}(\sigma_{de})\}\cdot\vec{\tau}(I)]=1, \tag{II.7}$$

$$\exp[-i\vec{Q}\{\vec{q}(\sigma_y)\}\cdot\vec{\tau}(I)]=-1, \tag{II.8}$$

which imply that nontrivial defining equations have to be considered when the corresponding unitary matrix U^{κ} should be computed.

Case (iii,a): Like in the previous case, $\vec{q}$ belongs to the surface of ΔBZ, but the inversion I is an element of P^X. Thus we have to consider Eq. (III.16) of Ref. 4,

$$\mathbb{X}^{\kappa}(\alpha)^* \exp[-i\vec{Q}\{\vec{q}(I)\}\cdot\vec{\tau}(\alpha)] = \mathbb{X}^{\bar{\kappa}}(\alpha), \quad \text{for all } \alpha\in P^X. \tag{II.9}$$

Because of

$$\vec{Q}\{\vec{q}(I)\} = 2\pi(0,-1,0), \tag{II.10}$$

$$\vec{\tau}(\alpha) = (0,0,0), \quad \text{for all } \alpha\in C_{4v}, \tag{II.11}$$

$$\vec{\tau}(I\alpha) = \vec{\tau}(\alpha I) = (1/2,1/2,1/2), \quad \text{for all } \alpha\in C_{4v}, \tag{II.12}$$

the following equations are readily verified:

$$\exp[-i\vec{Q}\{\vec{q}(I)\}\cdot\vec{\tau}(\alpha)] = 1, \quad \text{for all } \alpha\in C_{4v}, \tag{II.13}$$

$$\exp[-i\vec{Q}\{\vec{q}(I)\}\cdot\vec{\tau}(I\alpha)] = -1, \quad \text{for all } \alpha\in C_{4v}. \tag{II.14}$$

A simple inspection of (II.9) yields for the unirreps (I.16)

$$\{(\kappa=0,\mu=5)\uparrow P^X\}^* = (\kappa=1,\mu=5)\uparrow P^X. \tag{II.15}$$

This result can be verified, e.g., by

$$\begin{aligned}&\{\mathbb{X}^{(\kappa,\mu=5)\uparrow P^X}(C_{4y}^{+}I)\}^*\\ &= 2(-1)^{\kappa}\\ &= \exp[i\vec{Q}\{\vec{q}(I)\}\cdot\vec{\tau}(C_{4y}^{+}I)]\overline{\mathbb{X}^{(\kappa,\mu=5)\uparrow P^X}}(C_{4y}^{+}I)\\ &= 2(-1)^{\bar{\kappa}+1}, \quad \text{for } \kappa=0,1\end{aligned} \tag{II.16}$$

or by means of other group elements whose characters of the corresponding unirrep are different from zero. By similar arguments it can be shown that

$$\{(\mu)\uparrow P^X\}^* = (\mu)\uparrow P^X, \quad \text{for } \mu=0,2, \tag{II.17}$$

is valid, i.e., the unirreps (I.17) and (I.18) are equivalent to their complex conjugate representations.

Case (iii,b): This example is equal to the previous case, apart from the fact that the little cogroup P^R is the whole point group O_h. Considering the first two sets (I.22) of inequivalent unirreps for P^R we obtain

$$\{(0,\mu=2)\uparrow O_h\}^* = (1,\mu=2)\uparrow O_h \tag{II.18}$$

which can be proven, e.g., by means of

$$\begin{aligned}&\{\mathbb{X}^{(\kappa,\mu=2)\uparrow O_h}(C_{31}^{+}I)\}^*\\ &= (-1)^{\kappa}(\omega+\omega^2)\\ &= \exp[i\vec{Q}\{\vec{q}(I)\}\cdot\vec{\tau}(C_{31}^{+}I)]\overline{\mathbb{X}^{(\kappa,\mu=2)\uparrow O_h}}(C_{31}^{+}I)\\ &= (-1)^{\bar{\kappa}+1}(\omega+\omega^2), \quad \text{for } \kappa=0,1\end{aligned} \tag{II.19}$$

or by other group elements of O_h whose characters of the corresponding unirrep are different from zero. Similarly we can show

$$\{(\mu)\uparrow O_h\}^* = (\mu)\uparrow O_h, \quad \text{for } \mu=0,4. \tag{II.20}$$

III. COMPUTATION OF THE UNITARY MATRICES U^{κ}

The next step of our procedure is to determine by means of Eqs. (II.10), (II.23), or (II.36) of Ref. 6 the unitary matrices U^{κ}.

Case (i): due to Eq. (II.4) and (I.6)–(I.8) we obtain the results,

$$U^{\kappa} = 1 \quad \text{for} \quad \kappa=\mu=0,1,2,3$$

and

$$U^{\kappa} = \begin{vmatrix}0 & 1\\ 1 & 0\end{vmatrix} \quad \text{for} \quad \kappa=\mu=5. \tag{III.1}$$

Case (ii): Because of Eq. (II.6) we have to determine a unitary matrix U satisfying

$$\mathbb{R}^*(\alpha)\exp[-i\vec{Q}\{\vec{q}(\alpha^{-1})\}\cdot\vec{\tau}(I)] = U^{+}\mathbb{R}^{\bar{\kappa}}(\alpha)U, \quad \text{for all } \alpha\in P^S, \tag{III.2}$$

where we have omitted for the sake of simplicity the index $\kappa=(0)\uparrow P^S$. Taking (II.7,8) into account we arrive at nontrivial defining equations for this unitary matrix U which have as solution

$$U = \begin{vmatrix}1 & 0\\ 0 & -1\end{vmatrix}. \tag{III.3}$$

Thereby we have to note, that, although the projective unirrep (I.12) is real-valued, Eq. (II.8) prevents the trivial solution for U, i.e., to be the two-dimensional unit matrix. This is an example which is characteric for nonsymmorphic space groups.

Case (iii,a): Due to Eqs. (II.15,17) we have to consider three different cases when calculating

$$\mathbb{R}^{\kappa}(\alpha)^*\exp[-i\vec{Q}\{\vec{q}(I)\}\cdot\vec{\tau}(\alpha)] = U^{\kappa +}\mathbb{R}^{\bar{\kappa}}(\alpha)U^{\kappa}, \quad \text{for all } \alpha\in P^X. \tag{III.4}$$

Thereby the symbols of the equivalence classes $\{(\kappa,5)\uparrow P^X : \kappa=0,1;(\mu)\uparrow g\, P^X: \mu=0,2\} = A_{P^X(S^X)}$ are replaced by κ. The unimodular factors appearing in (III.4) turn out to be for the considered set of generating elements as

$$\exp[-i\vec{Q}\{\vec{q}(I)\}\cdot\vec{\tau}(C_{4y}^{+})] = 1, \tag{III.5}$$

$$\exp[-i\vec{Q}\{\vec{q}(I)\}\cdot\vec{\tau}(\sigma_x)] = \exp[-i\vec{Q}\{\vec{q}(I)\}\cdot\vec{\tau}(I)] = -1. \tag{III.6}$$

Equations (III.4) allow us to compute the matrices U^{κ}, if restricting them to the set of generating elements and taking Eqs. (III.5,6) into account. We obtain as results

$$U^{\kappa} = \begin{vmatrix}0 & -1\\ 1 & 0\end{vmatrix}, \quad \text{for } \kappa=(0,\mu=5)\uparrow P^X, \tag{III.7}$$

$$U^{\kappa} = \begin{vmatrix}0 & 1\\ 1 & 0\end{vmatrix}, \quad \text{for } \kappa=(\mu)\uparrow P^X;\ \ \mu=0,2. \tag{III.8}$$

As in the previous case we have to note that for both cases (III.8) a nontrivial unitary matrix U^{κ} is obtained, although the corresponding projective unirreps (I.17,18) are real-valued.

Case (iii,b): Because of Eqs. (II.18) and (II.20) we have to compute unitary matrices U^{κ} for three different cases by means of

$$\mathbb{R}^{\kappa}(\alpha)^*\exp[-i\vec{Q}\{\vec{q}(I)\}\cdot\vec{\tau}(\alpha)] = U^{\kappa +}\mathbb{R}^{\bar{\kappa}}(\alpha)U^{\kappa}, \quad \text{for all } \alpha\in O_h, \tag{III.9}$$

where an abbreviated notation is used for the equivalence classes of $A_{O_h(S^{\kappa})} = \{(\kappa,\mu=2)\uparrow O_h : \kappa=0,1;\ (\mu)\uparrow O_h$

$:\mu = 0,4\}$. The unimodular factors appearing in Eqs. (III.9) take for the generating elements of O_h the following values.

$$\exp[-i\vec{Q}\{\vec{q}(I)\}\cdot\vec{\tau}(\alpha)] = 1, \quad \text{for } \alpha = C_{2x}, C_{2z}, C_{31}^{+}, C_{2d}, \tag{III.10}$$

$$\exp[-i\vec{Q}\{\vec{q}(I)\}\cdot\vec{\tau}(I)] = -1. \tag{III.11}$$

Thus we arrive to the results

$$U^{\kappa} = \begin{vmatrix} 0 & 1 \\ 1 & 0 \end{vmatrix}, \quad \text{for } \kappa = (0,\mu = 2)\uparrow O_h, \tag{III.12}$$

$$U^{\kappa} = \begin{vmatrix} 1 & 0 \\ 0 & -1 \end{vmatrix}, \quad \text{for } \kappa = (\mu = 0)\uparrow O_h, \tag{III.13}$$

$$U^{\kappa} = \begin{vmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 \end{vmatrix}, \quad \text{for } \kappa = (\mu = 4)\uparrow O_h \tag{III.14}$$

which are readily obtained from Eqs. (III.9). Although the projective unirreps (I.23,24) are real-valued, we obtain because of (III.11) for the corresponding U^{κ} nontrivial unitary matrices.

IV. COUPLING COEFFICIENTS FOR *Pn3n*

The last step of our procedure is to insert the matrix elements of U^{κ} into the corresponding formula (II.18), or (II.31), respectively (II.43) of Ref. 6. The knowledge of the unitary matrices $U^{(\kappa,\vec{q})}$ allows us to write down explicitly the coupling coefficients in terms of CG coefficients by means of Eq. (I.2) of Ref. 6, presupposed the corresponding CG coefficients are already determined (see Ref. 1). Concerning the final formula for $U^{(\kappa,\vec{q})}$ we shall use in the following the matrix notation for U^{κ}. The formulas read:

$$U^{(\kappa,\vec{q})}_{\tau',\tau} = \delta_{\tau',I\tau}\, e^{i\vec{q}(\tau)\cdot\vec{t}(\tau,I)} U^{\kappa}; \quad \tau,\tau' \in P{:}P^{\vec{q}}, \tag{IV.1}$$

$$U^{(\kappa,\vec{q})}_{\tau',\tau} = \delta_{\tau',\tau} U^{\kappa}; \quad \tau,\tau' \in P{:}P^{\vec{q}}, \tag{IV.2}$$

where Eq. (IV.1) corresponds to (II.18), respectively (II.31) of Ref. 6 depending on whether $\vec{q}$ is not an element of the surface of ΔBZ, respectively belongs to the surface of ΔBZ (with $I \notin P^{\vec{q}}$). Equation (IV.2) is identical with Eq. (II.43) of Ref. 6 and is concerned with the case where $\vec{q}$ belongs to the surface of ΔBZ and $I \in P^{\vec{q}}$.

Before considering our examples we recall the definition of the special vectors

$$\vec{t}(\tau,I) = \vec{\tau}(\tau) + D(\tau)\vec{\tau}(I) - \vec{\tau}(\tau I), \quad \text{for all } \tau \in P{:}P^{\vec{q}}. \tag{IV.3}$$

The identities

$$\vec{q}(\tau)\cdot\vec{t}(\tau,I) = \vec{q}(\tau I)\cdot\vec{t}(I\tau,I), \quad \text{for all } \tau \in P{:}P^{\vec{q}} \tag{IV.4}$$

are readily verified and leads to a further simplification when determining the unimodular factors occuring in Eq. (IV.1). These identities being valid for any nonsymmorphic space group, are a consequence of the special structure of the sets $P{:}P^{\vec{q}}$ of the left coset representatives [see Eq. (I.20) of Ref. 4].

Case (i): The unimodular factors occuring in Eq. (IV.1) turn out to be

$$e^{i\vec{q}(\tau)\cdot\vec{t}(\tau,I)} = 1, \quad \text{for all } \tau \in O_h{:}P^{\Delta}. \tag{IV.5}$$

In order to prove this result we have to take Eq. (IV.3), Table 1.4 of Ref. 2 and the symmetry relations (IV.4) into account. Hence we arrive to the final result

$$U^{(\kappa,\vec{q})}_{\tau',\tau} = \delta_{\tau',I\tau} \quad \text{for } \kappa = \mu = 0,1,2,3$$

and

$$U^{(5,\vec{q})}_{\tau',\tau} = \delta_{\tau',I\tau} \begin{vmatrix} 0 & 1 \\ 1 & 0 \end{vmatrix}. \tag{IV.6}$$

Case (ii): The unimodular factors appearing in Eq. (IV.1) are given by

$$e^{i\vec{q}(\tau)\cdot\vec{t}(\tau,I)} = 1, \quad \text{for } \tau = E, \tag{IV.7}$$

$$e^{i\vec{q}(\tau)\cdot\vec{t}(\tau,I)} = -1, \quad \text{for } \tau = C_{4z}^{+}, C_{4x}^{-}, \tag{IV.8}$$

$$e^{i\vec{q}(\tau)\cdot\vec{t}(\tau,I)} = e^{i\pi x}, \quad \text{for } \tau = C_{4y}^{+}, C_{4z}^{-}, C_{4x}^{+}. \tag{IV.9}$$

The remaining factors follow from the symmetry relations (IV.4). Inserting (IV.7)–(IV.9) and (III.3) into (IV.1) we obtain immediately the corresponding unitary matrix $U^{(\kappa,\vec{q})}$.

Case (iii,a,b): The unitary matrices $U^{(\kappa,\vec{q})}$ follow from (IV.2) by inserting either (III.7), (III.8) or (III.12)–(III.14) depending on the considered case.

Concluding this section we remark that the coupling coefficients are readily obtainable by means of Eq. (I.2) of Ref. 6.

V. CONCLUDING REMARKS

With the aid of some typical examples concerning the nonsymmorphic space group $Pn3n$ we have illustrated the usefulness of the present method. Thereby we have shown that a suitable choice of the sets $P{:}P^{\vec{q}}$ of the left coset representatives simplify the considerations extremely.

[1]R. Dirl, "Clebsch–Gordan coefficients for *Pn3n*," J. Math. Phys. **20**, 664 (1979).

[2]C.J. Bradley and A.P. Cracknell, *The Mathematical Theory of Symmetry in solids* (Clarendon, Oxford, 1972).

[3]R. Dirl, J. Math. Phys. **18**, 2065 (1977).

[4]R. Dirl, "Complex conjugation of space group representations," J. Math. Phys. **20**, 1566 (1979).

[5]R. Dirl, "Coupling coefficients: General theory," J. Math. Phys. **20**, 1562 (1979).

[6]R. Dirl, "Coupling coefficients for space groups," J. Math. Phys. **20**, 1570 (1979).

Group theoretical analysis of second order phase transitions in magnetic structures

P. Rudra and M. K. Sikdar

Department of Physics, University of Kalyani, Kalyani, West Bengal, 741235, India
(Received 21 March 1977; revised manscript received 14 July 1978)

Landau's analysis of second order phase transitions in nonmagnetic crystals, using group theoretical results, have been extended to magnetic transitions. Corepresentation theory of magnetic space groups, having both linear and antilinear operations, has been used to obtain the criteria for second order phase transitions in magnetic structures. The possible magnetic structures after a second order phase transition from the paramagnetic phase with the space group D_{2h}^{7}, which is the case for $CuCl_2 \cdot 2H_2O$, have been obtained as an example.

1. INTRODUCTION

Phase transitions in physical systems are of immense interest and are widely studied. A suitable criterion of classifying[1] the phase transitions is according to the continuity or otherwise of the Gibb's potentials G and its derivatives $G^{(n)}$ of order n with respect to the temperature where $n = 1,2,\cdots$. Ehrenfest classified an nth order phase transitions as one in which $G^{(0)} = G, G^{(1)}, \ldots, G^{(n-1)}$ are all continuous whereas, $G^{(n)}$ is discontinuous in the two phase across the transition temperature. In the case of a *second* order phase transition the structure characterized by an order-parameter changes. The order parameters can be scalars, vectors, tensors, or spinors. In the case of ferroelectric transitions the order parameter is a 3-component vector, the electric polarization vector. In the case of transition from one magnetic phase to another the order parameters are[2,3] the thermodynamic average $\langle S_\alpha \rangle$ which is a 3-component axial vector. In the case of a second-order phase transition the change in structure is manifested in a change in group symmetry of the corresponding parameter.

Landau[1,4] worked out the group theoretical analysis of the second order phase transition and found criteria to determine the possible structure after such a transition when the group of the higher symmetry phase is given. His argument depends on two essentials facts.[4] In the expansion of the Gibb's potential G, the term cubic in the order parameters η_i must vanish and terms of the form

$$\eta_i \frac{\partial \eta_j}{\partial x_\alpha} - \eta_j \frac{\partial \eta_i}{\partial x_\alpha}$$

must not occur in the integrand of the integral giving Gibb's potential of the total crystal. These physical conditions can be stated in group theoretical formulations. If G is the group of the higher symmetry phase and η_i belongs to the irreducible representation D^{μ} of G, then

(i) the symmetrized triple product representation $D^{[\mu \otimes \mu \otimes \mu]}$ must not contain the identity representation, and

(ii) the antisymmetrized Kronecker inner direct product representation $D^{\{\mu \otimes \mu\}}$ must not contain the representation V of G formed by the polar vectors of the coordinate.

It is clear that the Landau analysis does not cover the case of transition to helical structures, which is left out by condition (ii). Lyubarskii showed[4] that the group G' of the system in the lower symmetry phase after a second order phase transition is the largest subgroup of G such that $D^{\mu} \downarrow G'$ contains the identity representation of G'. The order parameters η_i which are responsible for a second order phase transition between the structures of symmetries G and G' are those which belong to the identity representation of G'. Goldrich and Birman[5] appended an extra condition on the possible symmetry group G' in the case of transition to the ferroelectric phase.

Landau considered those phase transitions in which the order parameters are either scalars or polar vectors. In the case of phase transition involving magnetic structures the order parameters, as has been mentioned earlier, are components of an axial vector. Kovalev[6] and Dzialoshinskii[7] extended Landau–Lyubarskii's group theoretical analysis to magnetic transition involving axial vector order parameters which moreover are not invariant under the time-reversal operator θ. If G is the chemical group in the higher symmetry paramagnetic phase, the group of the magnetic structure in this phase is $G \cup \theta G$. But they did not take into account the antilinear nature of θ and used the usual representation theory of linear groups instead of using the corepresentation theory[8,9] of Wigner which is the proper theory of these magnetic groups.

Cracknell[10] and Cracknell and Sedaghat[11] have previously applied corepresentation theory of magnetic groups to the second-order phase transitions in magnetic structures. Their results involve calculations with the full group G, involving the infinite number of lattice translations of all the point group elements. Kopsky,[12] and Backhouse and Gard[13,14] have considered a method of reduction of symmetrized and antisymmetrized powers of corepresentations of magnetic groups.

In a previous publication[15] we have given suitable forms of the corepresentation matrices of the elements of magnetic space groups $M(G)$. In this paper we have utilized these forms to give workable expressions to the two Landau conditions for transition involving magnetic structures. The expressions reduce to summations over elements of the little cogroup $\bar{K}(\mathbf{k})$ of the reciprocal vector $\mathbf{k}$ and some other sub-

 0022-2488/79/071579-15$01.00

sets defined in the text. The order of these different subsets including $\bar{K}(\mathbf{k})$ can at most be equal to the order of the point group portion of the magnetic space group. In the applications of group theory to physical problems such simplications are essential for the efficacy of the method in tackling the problem. Since the irreducible corepresentation $D^{\mathbf{k}\mu}$ characterized by the reciprocal vector $\mathbf{k}$ in the first Brillouin zone of $M(G)$ and another index μ can be of three types of Wigner's classification,[8,16,17] we find that nine different cases arise. Criteria to determine whether a particular $D^{\mathbf{k}\mu}$ is an active corepresentation[4] satisfying the Landau conditions have been obtained for all these nine cases. The analysis has been done for projective corepresentations $D^{\lambda\mathbf{k}\mu}$ pertaining to a factor system λ. This has been done since spinor corepresentations can be considered as projective corepresentations. The case of the transition from the paramagnetic phase with the chemical space group D_{2h}^{7} to a lower magnetic symmetry has been worked out as an example. The groups in the possible lower symmetry phase after a second order phase transition as also the order-parameters in terms of the thermodynamic average of the spin operators $\langle S_\alpha \rangle$ that mediate the transformations have also been obtained. It should be pointed out that the analysis presented above can have any magnetic space group as that of the higher symmetry phase. The paramagnetic phase is just a special case.

We shall mention here that Dimmock[18] pointed out that the Landau condition (ii) is not valid for transitions involving helical magnetic structure. Brazovskii and Dzialoshinkii[19,20] showed that in a real phase transition fluctuations near the critical temperature prevent the transition being of a pure second order. Cracknell, Loreno, and Przystawa have recently shown[21] that the thermodynamic method and the group theoretical methods are not identical. The former gives additional results that the latter does not give.

We conclude this section by explaining the notations used in this paper. For ready comparison the notations of Ref. 15 have been retained here. A reciprocal lattice vector $\mathbf{K}$ is defined as $\mathbf{K}\cdot\mathbf{R} = 2\pi$ integer, where R is a direct lattice vector. The vector space of all $\mathbf{K}$ is denoted by $\Sigma(\mathbf{K})$. The vector space obtained by $\gamma\mathbf{K}$, $\mathbf{K}\in\Sigma(\mathbf{K})$ is denoted by $\Sigma(\gamma\mathbf{K})$. A bar over any space group means the collection of all proper & improper rotations (affixed with the time reversal operator θ where appropriate) appearing in the space group elements.

2. MAGNETIC SPACE GROUPS AND LANDAU CONDITIONS

We consider the case when the system has in the higher symmetry phase the group characterized by the magnetic space group

$$M = GUa_0G, \quad a_0^2\in G. \tag{1}$$

Here G is a linear space group with elements $(\mathbf{n} + \mathbf{t}(u)|u)$, where $\mathbf{n}$ are the lattice translations, u are the proper or improper rotations, and $\mathbf{t}(u)$ are fixed nonprimitive translations associated with u. The antilinear operator a_0 is given by

$$a_0 = \theta(\mathbf{c}|\gamma). \tag{2}$$

Here θ is the time reversal operator which is antilinear and commutes with all the space transformations. The projective corepresentations $D^{\lambda\mathbf{k}\mu}$ of M belonging to the projective factor system $\lambda(\alpha,\beta)$, $\alpha,\beta\in M$, are given in terms of[9,15]:

(i) the $\mathbf{k}$ vectors in the fundamental region of the first Brillouin zone,

(ii) the little cogroup $\bar{K}(\mathbf{k})$, defined as

$$\bar{K}(\mathbf{k}) = \{u\in\bar{G}\,|u\mathbf{k} - \mathbf{k}\in\Sigma(\mathbf{K})\}, \tag{3a}$$

where $\Sigma(\mathbf{K})$ is the vector space of the reciprocal lattice vectors $\mathbf{K}$, and

$$\bar{G} = \{u|(\mathbf{n} + \mathbf{t}(u)|u)\in G\}, \tag{3b}$$

(iii) the left coset representatives α_i, $i = 1,2,\ldots,r$, of $\bar{K}(\mathbf{k})$ in the factor group $\bar{G} = G/\mathscr{T}$ (where $\mathscr{T}$ is the primitive translational subgroup of M and hence of G), (iv) the nonprimitive translations $\mathbf{t}(\alpha_i) = \mathbf{a}_i$, and (v) the irreducible projective representations $\Gamma^{\lambda\mathbf{k}\mu}(u)$ of $\bar{K}(\mathbf{k})$. It should be noted[15] that (henceforth we omit the superscript λ if no confusion can arise)

$$\Gamma^{\mathbf{k}\mu}(u_1)\Gamma^{\mathbf{k}\mu}(u_2) = \omega(u_1,u_2)\Gamma^{\mathbf{k}\mu}(u_1u_2),$$

$$\omega(u_1,u_2) = \lambda(u_1,u_2)\exp i(u_1^{-1}\mathbf{k} - \mathbf{k})\cdot\mathbf{t}(u_2),$$

for all $u_1,u_2\in\bar{K}(\mathbf{k})$, and

$$\Gamma^{\mathbf{k}\mu}(u) = 0, \text{ for } u\notin\bar{K}(\mathbf{k}). \tag{4}$$

The characters of the matrices $\Gamma^{\mathbf{k}\mu}(u)$ are denoted by $\psi^{\mathbf{k}\mu}(u)$. The matrix elements of the corepresentations $D^{\lambda\mathbf{k}\mu}$ are given in Ref. 15. It should be noted that in that reference we have inadvertently dropped a factor involving $\lambda(\alpha,\beta)$ in the expressions for $\Phi_m^{\lambda}(i,j,\mathbf{k};\mathbf{t}(u),u)$, $m = 1,2,3,4$. The correct expressions given here involve the factor system $\lambda(\alpha,\beta)$:

$$\Phi_1^{\lambda}(i,j,\mathbf{k};\mathbf{t}(u),u) = f^{\lambda}(i,j,u)\exp i\mathbf{k}_i\cdot[\mathbf{t}(u) + u\mathbf{a}_j - \mathbf{a}_i],$$

$$\Phi_2^{\lambda}(i,j,\mathbf{k};\mathbf{t}(u),u) = \Phi_3^{\lambda}(i,j,\mathbf{k};\mathbf{t}(u),u)$$
$$= f^{\lambda}(i,j,\gamma^{-1}u\gamma)^*\exp(-i\gamma\mathbf{k}_i)[\mathbf{t}(u) + u\mathbf{c} - \mathbf{c} + u\gamma\mathbf{a}_j - \gamma\mathbf{a}_i],$$

$$\Phi_4^{\lambda}(i,j,\mathbf{k};\mathbf{t}(u),u)$$
$$= f^{\lambda}(i,j,u\gamma^2)\exp i\mathbf{k}_i[\mathbf{t}(u) + u\mathbf{c} + u\gamma\mathbf{c} + u\gamma^2\mathbf{a}_j - \mathbf{a}_i],$$

where

$$f^{\lambda}(i,j,u) = \lambda(\alpha_i^{-1},u\alpha_j)\lambda(u,\alpha_j)$$
$$\times[\lambda(\alpha_i,\alpha_i^{-1})^*\lambda(\alpha_j,\alpha_j^{-1})^*]^{1/2}. \tag{5}$$

The phase of the square root of the complex number in $f^{\lambda}(i,j,u)$ is in the following manner:

$$[\lambda(\alpha_i,\alpha_i^{-1})^*\lambda(\alpha_j,\alpha_j^{-1})^*]^{1/2}$$
$$= \exp(-i/2)[\sigma(\alpha_i,\alpha_i^{-1}) + \sigma(\alpha_j,\alpha_j^{-1})],$$

where $\lambda(\alpha,\beta) = \exp i\sigma(\alpha,\beta)$, $0\leqslant\sigma(\alpha,\beta) < 2\pi$, for all α,β.

We now define the corepresentation

$$\bar{D}^{\mathbf{k}\mu}(\mathbf{n} + \mathbf{t}(u)|u)$$
$$= \lambda(u,\theta\gamma)^*\lambda(\theta\gamma,\gamma^{-1}u\gamma)D^{\lambda\mathbf{k}\mu}(\gamma^{-1}[-\mathbf{c}$$

$+ uc + n + t(u)] | \gamma^{-1} u \gamma)^*$.

The Wigner typology of a corepresentation $D^{\lambda \mathbf{k}\mu}$ is now defined as follows

(i) $D^{\lambda \mathbf{k}\mu}$ is of type (a) if $\bar{D}^{\lambda \mathbf{k}\mu} = P^{-1} D^{\lambda \mathbf{k}\mu} P$, with $PP^* = \lambda(\theta\gamma,\theta\gamma) D^{\lambda \mathbf{k}\mu}(\mathbf{c} + \gamma\mathbf{c}|\gamma^2)$.

(ii) $D^{\lambda \mathbf{k}\mu}$ is of type (b) if $\bar{D}^{\lambda \mathbf{k}\mu} = P^{-1} D^{\lambda \mathbf{k}\mu} P$, with $PP^* = -\lambda(\theta\gamma,\theta\gamma) D^{\lambda \mathbf{k}\mu}(\mathbf{c} + \gamma\mathbf{c}|\gamma^2)$.

(iii) $D^{\lambda \mathbf{k}\mu}$ is of type (c) if $\bar{D}^{\lambda \mathbf{k}\mu}$ is not equivalent to $D^{\lambda \mathbf{k}\mu}$. This matrix P, which appears in the matrix elements of the corepresentation matrices,[15] satisfies the relations:

$$\sum_{j'n'} \Gamma^{\lambda \mathbf{k}\mu}_{mn'}(\alpha_i^{-1} u \alpha_{j'}) \Phi^{\lambda}_1(i,j',\mathbf{k};\mathbf{t}(u),u) P_{j'n',jn} \exp i\mathbf{k}_i\cdot\mathbf{n}$$

$$= \lambda(u,\theta\gamma)^* \lambda(\theta\gamma,\gamma^{-1}u\gamma) \sum_{j'n'} \Gamma^{\lambda \mathbf{k}\mu}_{n'n}(\alpha_{j'}^{-1}\gamma^{-1}u\gamma\alpha_j)^*$$

$$\times \Phi^{\lambda}_2(j',j,\mathbf{k};\mathbf{t}(u),u) P_{im,j'n'} \exp(-i\gamma\mathbf{k}_{j'})\cdot\mathbf{n}, \qquad (6)$$

$$(PP^*)_{im,jn} = \pm\lambda(\theta\gamma,\theta\gamma)\Gamma^{\lambda \mathbf{k}\mu}_{mn}(\alpha_i^{-1}\gamma^2\alpha_j) \exp i\mathbf{k}_i$$

$$\cdot(\mathbf{c} + \gamma\mathbf{c} + \gamma^2\mathbf{a}_j - \mathbf{a}_i).$$

To make our analysis general we assume that the order parameters η_is defining the phase transition may transform according to a corepresentation $D^{\lambda \mathbf{k}\mu}$ of M belonging to a factor system $\lambda(\alpha,\beta)$. The Landau conditions for such a transition to be of second order are[1,6,10]:

(1) The antisymmetric Kronecker direct product corepresentation $D^{\{\lambda \mathbf{k}\mu \otimes \lambda \mathbf{k}\mu\}}$ must not have any irreducible component common with V, the corepresentation formed by polar vectors,

$$\sum_{R\in G} \chi^{\{\lambda \mathbf{k}\mu \otimes \lambda \mathbf{k}\mu\}}(R)\chi(R)^* = 0.$$

(2) The symmetric triple product corepresentation $D^{[\lambda \mathbf{k}\mu \otimes \lambda \mathbf{k}\mu \otimes \lambda \mathbf{k}\mu]}$ must not contain the identity corepresentation,

$$\sum_{R\in G} \chi^{[\lambda \mathbf{k}\mu \otimes \lambda \mathbf{k}\mu \otimes \lambda \mathbf{k}\mu]}(R) = 0.$$

Here we denote by $\chi^{\{\lambda \mathbf{k}\mu \otimes \lambda \mathbf{k}\mu\}}(R)$ the character of $D^{\{\lambda \mathbf{k}\mu \otimes \lambda \mathbf{k}\mu\}}(R)$, by $\chi^{[\lambda \mathbf{k}\mu \otimes \lambda \mathbf{k}\mu \otimes \lambda \mathbf{k}\mu]}(R)$ the character of $D^{[\lambda \mathbf{k}\mu \otimes \lambda \mathbf{k}\mu \otimes \lambda \mathbf{k}\mu]}(R)$, and by $\chi(R)$ the character of $V(R)$. It can be verified with a little manipulation with the defining relations that if $D^{\lambda \mathbf{k}\mu}$ belongs to the factor system $\lambda(\alpha,\beta)$, $D^{\{\lambda \mathbf{k}\mu \otimes \lambda \mathbf{k}\mu\}}$ belongs to the factor system $\lambda(\alpha,\beta)^2$, and $D^{[\lambda \mathbf{k}\mu \otimes \lambda \mathbf{k}\mu \otimes \lambda \mathbf{k}\mu]}$ belongs to the factor system $\lambda(\alpha,\beta)^3$. Also, the corepresentation V belongs to the $\mathbf{k}$ vector $\mathbf{k} = \mathbf{0}$. $\chi^{\{\lambda \mathbf{k}\mu \otimes \lambda \mathbf{k}\mu\}}$ is first expressed as a sum of irreducible characters $\chi^{\lambda^2 \mathbf{0}\nu}$ and then the summation over $R\in G$ on the left-hand side of the first Landau condition is evaluated. The two Landau conditions will then reduce to[9,22,23]:

$$\mathscr{A} \equiv \sum_{\nu} b^{\{\lambda \mathbf{k}\mu \otimes \lambda \mathbf{k}\mu\}}_{\lambda^2 \mathbf{0}\nu} \sum_{R\in G} \chi^{\lambda^2 \mathbf{0}\nu}(R)\chi(R)^* = 0,$$

where

$$b^{\{\lambda \mathbf{k}\mu \otimes \lambda \mathbf{k}\mu\}}_{\lambda^2 \mathbf{0}\nu} = \sum_{R\in G} \chi^{\{\lambda \mathbf{k}\mu \otimes \lambda \mathbf{k}\mu\}}(R)\chi^{\lambda^2 \mathbf{0}\nu}(R)^*$$

and

$$\chi^{\lambda^2 \mathbf{0}\nu}(R) = \text{character of } D^{\lambda^2 \mathbf{0}\nu}(R), \qquad (7)$$

and

$$\mathscr{S} \equiv \sum_{R\in G} \chi^{[\lambda \mathbf{k}\mu \otimes \lambda \mathbf{k}\mu \otimes \lambda \mathbf{k}\mu]}(R) = 0. \qquad (8)$$

Here $\mathscr{A}$ is simply the numerator of $\Sigma_{R\in G}\chi^{\{\lambda \mathbf{k}\mu \otimes \lambda \mathbf{k}\mu\}}(R) \times \chi(R)^*$. In order to evaluate b in Eq. (7) and $\mathscr{S}$ in Eq. (8) we will use[10,13,14]

$$\chi^{\{\lambda \mathbf{k}\mu \otimes \lambda \mathbf{k}\mu\}}(R) = \tfrac{1}{2}[(\chi^{\lambda \mathbf{k}\mu}(R))^2 - \lambda(R,R)\chi^{\lambda \mathbf{k}\mu}(R^2)]$$

and

$$\chi^{[\lambda \mathbf{k}\mu \otimes \lambda \mathbf{k}\mu \otimes \lambda \mathbf{k}\mu]}(R)$$

$$= \tfrac{1}{6}[(\chi^{\lambda \mathbf{k}\mu}(R))^3 - 3\lambda(R,R)\chi^{\lambda \mathbf{k}\mu}(R)\chi^{\lambda \mathbf{k}\mu}(R^2)$$

$$+ 2\lambda(R,R)\lambda(R^2,R)\chi^{\lambda \mathbf{k}\mu}(R^3)]. \qquad (9)$$

Using Eqs. (7), (8), (9), and the forms of the corepresentation matrices given in Ref. 15 we get the group theoretical forms of the Landau conditions for a second order phase transition from one magnetic phase to another magnetic phase.

The summation over $R\in G$ in all these equations actually involves two summations one over the lattice translations $\mathbf{n}$, and the second over the proper and improper rotations u. The first summation will give factors like $\Sigma_{\mathbf{n}} \exp i\mathbf{k}\cdot\mathbf{n} = N\Delta(\mathbf{k})$, where N is the number of elements in the translation subgroup $\mathscr{T}$ and

$$\Delta(\mathbf{k}) = \begin{cases} 1, & \text{if } \mathbf{k}\in\Sigma(\mathbf{K}) \\ 0, & \text{if } \mathbf{k}\notin\Sigma(\mathbf{K}) \end{cases}. \qquad (10)$$

Because of the factors $\Delta(\mathbf{k})$, the summation over u will ultimately be restricted (cf. Appendix) over different subsets of $\bar{G}$ defined below:

$\bar{M} - \bar{G} = \{\theta u\gamma | u\in\bar{G}\}$,

$\bar{K}(\mathbf{k}) = \{u\in\bar{G} | u\mathbf{k} - \mathbf{k}\in\Sigma(\mathbf{K})\}$,

$\bar{K}_-(\mathbf{k}) = \{u\in\bar{G} | u\mathbf{k} + \mathbf{k}\in\Sigma(\mathbf{K})\}$,

$\bar{K}_{-\gamma}(\mathbf{k}) = \{u\in\bar{G} | u\mathbf{k} + \mathbf{k}\in\Sigma(\gamma\mathbf{K})\}$,

$\bar{Q}(\mathbf{k}) - \bar{K}(\mathbf{k}) = \{\theta u\gamma\in\bar{M} - \bar{G} | u\gamma\mathbf{k} + \mathbf{k}\in\Sigma(\gamma\mathbf{K})\}$,

$$
\begin{aligned}
&\bar{Q}_{-\gamma}(\mathbf{k}) - \bar{K}_{-\gamma}(\mathbf{k}) = \{\theta u\gamma \in \bar{M} - \bar{G}\,|\,u\gamma\mathbf{k} - \mathbf{k} \in \Sigma(\mathbf{K})\},\\
&\bar{K}^{(1,2)}_{-}(\mathbf{k}) = \{(u_1,u_2) \in (\bar{G},\bar{G})\,|\,u_1\mathbf{k} + u_2\mathbf{k} + \mathbf{k} \in \Sigma(\mathbf{K})\},\\
&\bar{K}^{(1,2)}_{-\gamma}(\mathbf{k}) = \{(u_1,u_2) \in (\bar{G},\bar{G})\,|\,u_1\mathbf{k} + u_2\mathbf{k} + \mathbf{k} \in \Sigma(\gamma\mathbf{K})\},\\
&\bar{K}^{(v)}_{-}(\mathbf{k}) = \{u \in \bar{G}\,|\,u\mathbf{k} + v^{-1}u\mathbf{k} + \mathbf{k} \in \Sigma(\mathbf{K})\},\\
&\bar{K}^{(v)}_{-\gamma}(\mathbf{k}) = \{u \in \bar{G}\,|\,u\mathbf{k} + v^{-1}u\mathbf{k} + \mathbf{k} \in \Sigma(\gamma\mathbf{K})\},\\
&\bar{K}^{(2)}_{-}(\mathbf{k}) = \{u \in \bar{G}\,|\,u\mathbf{k} + u^2\mathbf{k} + \mathbf{k} \in \Sigma(\mathbf{K})\},\\
&\bar{K}^{(2)}_{-\gamma}(\mathbf{k}) = \{u \in \bar{G}\,|\,u\mathbf{k} + u^2\mathbf{k} + \mathbf{k} \in \Sigma(\gamma\mathbf{K})\},\\
&\bar{Q}^{(1,2)}_{-}(\mathbf{k}) - \bar{K}^{(1,2)}_{-}(\mathbf{k}) = \{(\theta u_1\gamma,\theta u_2\gamma) \in (\bar{M} - \bar{G},\bar{M} - \bar{G})\,|\,u_1\gamma\mathbf{k} + u_2\gamma\mathbf{k} - \mathbf{k} \in \Sigma(\gamma\mathbf{K})\},\\
&\bar{Q}^{(1,2)}_{-\gamma}(\mathbf{k}) - \bar{K}^{(1,2)}_{-\gamma}(\mathbf{k}) = \{(\theta u_1\gamma,\theta u_1\gamma) \in (\bar{M} - \bar{G},\bar{M} - \bar{G})\,|\,u_2\gamma\mathbf{k} + u_2\gamma\mathbf{k} - \mathbf{k} \in \Sigma(\mathbf{K})\},\\
&\bar{Q}^{(v)}_{-}(\mathbf{k}) - \bar{K}^{(v)}_{-}(\mathbf{k}) = \{\theta u\gamma \in \bar{M} - \bar{G}\,|\,u\gamma\mathbf{k} + v^{-1}u\gamma\mathbf{k} - \mathbf{k} \in \Sigma(\gamma\mathbf{K})\},\\
&\bar{Q}^{(v)}_{-\gamma}(\mathbf{k}) - \bar{K}^{(v)}_{-\gamma}(\mathbf{k}) = \{\theta u\gamma \in \bar{M} - \bar{G}\,|\,u\gamma\mathbf{k} + v^{-1}u\gamma\mathbf{k} - \mathbf{k} \in \Sigma(\mathbf{K})\}.
\end{aligned}
\tag{11}
$$

In order to find out whether a corepresentation $D^{\lambda\mathbf{k}\mu}$ will induce a second order phase transition we have first to check to which Wigner typology $D^{\lambda\mathbf{k}\mu}$ belongs. This is done by using the Wigner–Bradley criterion[8,9,16]

$$
\mathscr{C} \equiv \sum_{\substack{\theta u\gamma \in \bar{Q}(\mathbf{k}) - \bar{K}(\mathbf{k})\\ (u\gamma)^2 \in \bar{K}(\mathbf{k})}} \lambda(\theta u\gamma,\theta u\gamma)\exp i\mathbf{k}[\mathbf{t}(u) + u\gamma\mathbf{t}(u) + u\mathbf{c} + u\gamma u\mathbf{c}]\psi^{\lambda\mathbf{k}\mu}(u\gamma u\gamma)
$$

$$
= \begin{cases} +|\bar{K}(\mathbf{k})| & \text{for type (a) corepresentation,}\\ -|\bar{K}(\mathbf{k})| & \text{for type (b) corepresentation,}\\ \quad 0 & \text{for type (c) corepresentation,}\end{cases}
\tag{12}
$$

where $|\bar{K}(\mathbf{k})|$ denotes the order of the little cogroup $\bar{K}(\mathbf{k})$. In order to calculate $\mathscr{S}$ and $b^{\{\lambda\mathbf{k}\mu \otimes \lambda\mathbf{k}\mu\}}_{\lambda'\mathbf{0}\nu}$ we have to consider separately the three types of corepresentations for $D^{\lambda\mathbf{k}\mu}$ and $D^{\lambda'\mathbf{0}\nu}$. Because of the particular block structures of the corepresentation matrices for the three types, each of these two summations breaks up in different parts involving twenty separate subsummations S_1 to S_{20}. Use of Eq. (10) then reduces all these subsummations to those over the subsets defined in Eq. (11). In Table I we give the expressions for $\mathscr{S}$ and $b^{\{\lambda\mathbf{k}\mu \otimes \lambda\mathbf{k}\mu\}}_{\lambda'\mathbf{0}\nu}$ in terms of S_i, $i = 1,2,\ldots,20$, defined below:

$$
\begin{aligned}
S_1 = \sum_{u \in \bar{K}(k)} &\psi^{\lambda\mathbf{k}\mu}(u)\sum_{i_1}\sum_{i_2}\sum_{i_3} \exp i(\mathbf{k} + \alpha_{i_1}^{-1}\alpha_{i_2}^{-1}\mathbf{k} + \alpha_{i_1}^{-1}\alpha_{i_3}\mathbf{k})\cdot\mathbf{t}(u)\exp i\alpha_{i_2}(\alpha_{i_2}^{-1}\alpha_{i_1}u^{-1}\alpha_{i_1}^{-1}\alpha_{i_2}\mathbf{k} - \mathbf{k})\cdot(\mathbf{a}_{i_2} - \mathbf{a}_{i_1})\\
&\times \exp i\alpha_{i_3}(\alpha_{i_3}^{-1}\alpha_{i_1}u^{-1}\alpha_{i_1}^{-1}\alpha_{i_3}\mathbf{k} - \mathbf{k})\cdot(\mathbf{a}_{i_3} - \mathbf{a}_{i_1})f^{\lambda}(i_1,i_1,\alpha_{i_1}u\alpha_{i_1}^{-1})f^{\lambda}(i_2,i_2,\alpha_{i_1}u\alpha_{i_1}^{-1})f^{\lambda}(i_3,i_3,\alpha_{i_1}u\alpha_{i_1}^{-1})\\
&\times \psi^{\lambda\mathbf{k}\mu}(\alpha_{i_2}^{-1}\alpha_{i_1}u\alpha_{i_1}^{-1}\alpha_{i_2})\psi^{\lambda\mathbf{k}\mu}(\alpha_{i_3}^{-1}\alpha_{i_1}u\alpha_{i_1}^{-1}\alpha_{i_3}),
\end{aligned}
$$

where

$$
(\alpha_{i_1}^{-1}\alpha_{i_2},\alpha_{i_1}^{-1}\alpha_{i_3}) \in \bar{K}^{(1,2)}_{-}(\mathbf{k}),
$$

$$
\begin{aligned}
S_2 = \sum_{u \in \bar{K}(\mathbf{k})} &\psi^{\lambda\mathbf{k}\mu}(u)\sum_{i_1}\sum_{i_2}\lambda(\alpha_{i_1}u\alpha_{i_1}^{-1},\alpha_{i_1}u\alpha_{i_1}^{-1})\exp i(\mathbf{k} + \alpha_{i_1}^{-1}\alpha_{i_2}\mathbf{k} + u^{-1}\alpha_{i_1}^{-1}\alpha_{i_2}\mathbf{k})\cdot\mathbf{t}(u)\exp i\alpha_{i_2}(\alpha_{i_2}^{-1}\alpha_{i_1}u^{-2}\alpha_{i_1}^{-1}\alpha_{i_2}\mathbf{k} - \mathbf{k})\\
&\times(\mathbf{a}_{i_2} - \mathbf{a}_{i_1})f^{\lambda}(i_1,i_1,\alpha_{i_1}u\alpha_{i_1}^{-1})f^{\lambda}(i_2,i_2,\alpha_{i_1}u^2\alpha_{i_1}^{-1})\psi^{\lambda\mathbf{k}\mu}(\alpha_{i_2}^{-1}\alpha_{i_1}u^2\alpha_{i_1}^{-1}\alpha_{i_2}),
\end{aligned}
$$

where

$$
\alpha_{i_1}^{-1}\alpha_{i_2} \in \bar{K}^{(u)}_{-}(\mathbf{k}),
$$

$$
S_3 = \sum_{u \in \bar{K}^{(2)}(\mathbf{k}),\, u^3 \in \bar{K}(\mathbf{k})} \psi^{\lambda\mathbf{k}\mu}(u^3)\exp i\mathbf{k}\cdot(\mathbf{t}(u) + u\mathbf{t}(u) + u^2\mathbf{t}(u))\sum_{i_1}\lambda(\alpha_{i_1}u\alpha_{i_1}^{-1},\alpha_{i_1}u\alpha_{i_1}^{-1})\lambda(\alpha_{i_1}u^2\alpha_{i_1}^{-1},\alpha_{i_1}u\alpha_{i_1}^{-1})f^{\lambda}(i_1,i_1,\alpha_{i_1}u^3\alpha_{i_1}^{-1}),
$$

$$
\begin{aligned}
S_4 = \sum_{u \in \bar{K}(\mathbf{k})} &\psi^{\lambda\mathbf{k}\mu}(u)\sum_{i_1}\sum_{i_2}\lambda(\alpha_{i_1}^{-1}u\alpha_{i_1}^{-1},\alpha_{i_1}u\alpha_{i_1}^{-1})\lambda(\alpha_{i_1}u^2\alpha_{i_1}^{-1},\theta\gamma)^{*}\lambda(\theta\gamma,\gamma^{-1}\alpha_{i_1}u^2\alpha_{i_1}^{-1}\gamma)\exp i(\mathbf{k} - \alpha_{i_1}^{-1}\gamma\alpha_{i_2}\mathbf{k}\\
&- u^{-1}\alpha_{i_1}^{-1}\gamma\alpha_{i_2}\mathbf{k}\cdot\mathbf{t}(u)\exp(-i\gamma\alpha_{i_2})\,(\alpha_{i_2}^{-1}\gamma^{-1}\alpha_{i_1}u^{-2}\alpha_{i_1}^{-1}\gamma\alpha_{i_2}\mathbf{k} - \mathbf{k})\cdot(\mathbf{c} + \gamma\mathbf{a}_{i_2} - \mathbf{a}_{i_1})
\end{aligned}
$$

$$\times f^{\lambda}(i_1,i_1,\alpha_{i_1}u\alpha_{i_1}^{-1})f^{\lambda}(i_2,i_2,\gamma^{-1}\alpha_{i_1}u^2\alpha_{i_2}^{-1}\gamma)^*\psi^{\lambda\mathbf{k}\mu}(\alpha_{i_2}^{-1}\gamma^{-1}\alpha_{i_1}u^2\alpha_{i_1}^{-1}\gamma\alpha_{i_2})^*,$$

where

$$\theta\alpha_{i_1}^{-1}\gamma\alpha_{i_2}\in\bar{Q}_{-\gamma}^{(u)}(\mathbf{k})-\bar{K}_{-\gamma}^{(u)}(\mathbf{k}),$$

$$S_5=\sum_{u\in\bar{K}(\mathbf{k})}\psi^{\lambda\mathbf{k}\mu}(u)^*\sum_{i_1}\sum_{i_2}\lambda(\gamma\alpha_{i_2}u\alpha_{i_2}^{-1}\gamma^{-1},\gamma\alpha_{i_2}u\alpha_{i_2}^{-1}\gamma^{-1})\lambda(\gamma\alpha_{i_2}u\alpha_{i_2}^{-1}\gamma^{-1},\theta\gamma,)^*\lambda(\theta\gamma,\alpha_{i_2}u\alpha_{i_2}^{-1})\exp-i(\mathbf{k}-\alpha_{i_2}^{-1}\gamma^{-1}\alpha_{i_1}\mathbf{k}$$

$$-u^{-1}\alpha_{i_2}^{-1}\gamma^{-1}\alpha_{i_1}\mathbf{k})\cdot\mathbf{t}(u)\exp(-i\alpha_{i_1})(\alpha_{i_1}^{-1}\gamma\alpha_{i_2}u^{-2}\alpha_{i_2}^{-1}\gamma^{-1}\alpha_{i_1}\mathbf{k}-\mathbf{k})\cdot(\mathbf{c}+\gamma\mathbf{a}_{i_2}-\mathbf{a}_{i_1})f^{\lambda}(i_1,i_1,\gamma\alpha_{i_2}u^2\alpha_{i_2}^{-1}\gamma^{-1})f^{\lambda}(i_2,i_2,\alpha_{i_2}u\alpha_{i_2}^{-1})^*$$

$$\times\psi^{\lambda\mathbf{k}\mu}(\alpha_{i_1}^{-1}\gamma\alpha_{i_2}u^2\alpha_{i_2}^{-1}\gamma^{-1}\alpha_{i_1}),$$

where

$$\theta\alpha_{i_2}^{-1}\gamma^{-1}\alpha_{i_1}\in\bar{Q}_{-}^{(u)}(\mathbf{k})\bar{K}_{-}^{(u)}(\mathbf{k}),$$

$$S_6=\sum_{u\in\bar{K}(\mathbf{k})}\psi^{\lambda\mathbf{k}\mu}(u)^*\sum_{i_1}\sum_{i_2}\lambda(\gamma\alpha_{i_1}u\alpha_{i_1}^{-1}\gamma^{-1},\gamma\alpha_{i_1}u\alpha_{i_1}^{-1}\gamma^{-1})\lambda(\gamma\alpha_{i_1}u\alpha_{i_1}^{-1}\gamma^{-1},\theta\gamma)^*\lambda(\theta\gamma,\alpha_{i_1}u\alpha_{i_1}^{-1})\lambda(\gamma\alpha_{i_1}u^2\alpha_{i_1}^{-1}\gamma^{-1},\theta\gamma)^*$$

$$\times\lambda(\theta\gamma,\alpha_{i_1}u^2\alpha_{i_1}^{-1})\exp-i(\mathbf{k}+\alpha_{i_1}^{-1}\alpha_{i_2}\mathbf{k}+u^{-1}\alpha_{i_1}^{-1}\alpha_{i_2}\mathbf{k})\cdot\mathbf{t}(u)\exp(-i\alpha_{i_2})(\alpha_{i_2}^{-1}\alpha_{i_1}u^{-2}\alpha_{i_1}^{-1}\alpha_{i_2}\mathbf{k}-\mathbf{k})\cdot(\mathbf{a}_{i_2}-\mathbf{a}_{i_1})$$

$$\times f^{\lambda}(i_1,i_1,\alpha_{i_1}u\alpha_{i_1}^{-1})^*f^{\lambda}(i_2,i_2,\alpha_{i_1}u^2\alpha_{i_1}^{-1})^*\psi^{\lambda\mathbf{k}\mu}(\alpha_{i_2}^{-1}\alpha_{i_1}u^2\alpha_{i_1}^{-1}\alpha_{i_2})^*,$$

where

$$\alpha_{i_1}^{-1}\alpha_{i_2}\in\bar{K}_{-\gamma}^{(u)}(\mathbf{k})$$

$$S_7=\sum_{u\in\bar{K}_{-\gamma}^{(2)}(\mathbf{k}),u^3\in\bar{K}(\mathbf{k})}\psi^{\lambda\mathbf{k}\mu}(u^3)^*\exp(-i\mathbf{k})\cdot[\mathbf{t}(u)+u\mathbf{t}(u)+u^2\mathbf{t}(u)]$$

$$\times\sum_{i_1}\lambda(\gamma\alpha_{i_1}u\alpha_{i_1}^{-1}\gamma^{-1},\gamma\alpha_{i_1}u\alpha_{i_1}^{-1}\gamma^{-1})\lambda(\gamma\alpha_{i_1}u^2\alpha_{i_1}^{-1}\gamma^{-1},\gamma\alpha_{i_1}u\alpha_{i_1}^{-1}\gamma^{-1})$$

$$\times\lambda(\gamma\alpha_{i_1}u^3\alpha_{i_1}^{-1}\gamma^{-1},\theta\gamma)^*\lambda(\theta\gamma,\alpha_{i_1}u\alpha_{i_1}^{-1})f^{\lambda}(i_1,i_1,\alpha_{i_1}u^3\alpha_{i_1}^{-1})^*,$$

$$S_8=\sum_{u\in\bar{K}(\mathbf{k})}\psi^{\lambda\mathbf{k}\mu}(u)\sum_{i_1}\sum_{i_2}\sum_{i_3}\lambda(\alpha_{i_1}u\alpha_{i_1}^{-1},\theta\gamma)^*\lambda(\theta\gamma,\gamma^{-1}\alpha_{i_1}u\alpha_{i_1}^{-1}\gamma)\exp i(\mathbf{k}+\alpha_{i_1}^{-1}\alpha_{i_2}\mathbf{k}-\alpha_{i_1}^{-1}\gamma\alpha_{i_3}\mathbf{k})\cdot\mathbf{t}(u)\exp i\alpha_{i_2}(\alpha_{i_2}^{-1}\alpha_{i_1}u\alpha_{i_1}^{-1}$$

$$\alpha_{i_2}\mathbf{k}-\mathbf{k})\cdot(\mathbf{a}_{i_2}-\mathbf{a}_{i_1})\exp(-i\gamma\alpha_{i_3})(\alpha_{i_3}^{-1}\gamma^{-1}\alpha_{i_1}u^{-1}\alpha_{i_1}^{-1}\gamma\alpha_{i_3}\mathbf{k}-\mathbf{k})\cdot(\mathbf{c}+\gamma\mathbf{a}_{i_3}-\mathbf{a}_{i_1})f^{\lambda}(i_1,i_1,\alpha_{i_1}u\alpha_{i_1}^{-1})f^{\lambda}(i_2,i_2,\alpha_{i_1}u\alpha_{i_1}^{-1})$$

$$\times f^{\lambda}(i_3,i_3,\gamma^{-1}\alpha_{i_1}u\alpha_{i_1}^{-1}\gamma)^*\psi^{\lambda\mathbf{k}\mu}(\alpha_{i_2}^{-1}\alpha_{i_1}u\alpha_{i_1}^{-1}\alpha_{i_2})\psi^{\lambda\mathbf{k}\mu}(\alpha_{i_3}^{-1}\gamma^{-1}\alpha_{i_1}u\alpha_{i_1}^{-1}\gamma\alpha_{i_3})^*,$$

where

$$(\theta\alpha_{i_1}^{-1}\gamma^{-1}\alpha_{i_1},\theta\alpha_{i_1}^{-1}\gamma^{-1}\alpha_{i_2})\in\bar{Q}_{-}^{(1,2)}(\mathbf{k})-\bar{K}_{-}^{(1,2)}(\mathbf{k})$$

$$S_9=\sum_{u\in\bar{K}(\mathbf{k})}\psi^{\lambda\mathbf{k}\mu}(u)\sum_{i_1}\sum_{i_2}\sum_{i_3}[\lambda(\alpha_{i_1}u\alpha_{i_1}^{-1},\theta\gamma)^*\lambda(\theta\gamma,\gamma^{-1}\alpha_{i_1}u\alpha_{i_1}^{-1}\gamma)]^2$$

$$\times\exp i(\mathbf{k}-\alpha_{i_1}^{-1}\gamma\alpha_{i_2}\mathbf{k}-\alpha_{i_1}^{-1}\gamma\alpha_{i_3}\mathbf{k})\cdot\mathbf{t}(u)\exp(-i\gamma\alpha_{i_2})(\alpha_{i_2}^{-1}\gamma^{-1}\alpha_{i_1}u^{-1}\alpha_{i_1}^{-1}\gamma\alpha_{i_2}\mathbf{k}-\mathbf{k})$$

$$\cdot(\mathbf{c}+\gamma\mathbf{a}_{i_2}-\mathbf{a}_{i_1})\exp(-i\gamma\alpha_{i_3})(\alpha_{i_3}^{-1}\gamma^{-1}\alpha_{i_1}u^{-1}\alpha_{i_1}^{-1}\gamma\alpha_{i_3}\mathbf{k}-\mathbf{k})$$

$$\cdot(\mathbf{c}+\gamma\mathbf{a}_{i_3}-\mathbf{a}_{i_1})f^{\lambda}(i_1,i_1,\alpha_{i_1}u\alpha_{i_1}^{-1})f^{\lambda}(i_2,i_2,\gamma^{-1}\alpha_{i_1}u\alpha_{i_1}^{-1}\gamma)^*$$

$$\times f^{\lambda}(i_3,i_3,\gamma^{-1}\alpha_{i_1}u\alpha_{i_1}^{-1}\gamma)^*\psi^{\lambda\mathbf{k}\mu}(\alpha_{i_2}^{-1}\gamma^{-1}\alpha_{i_1}u\alpha_{i_1}^{-1}\gamma\alpha_{i_2})^*\psi^{\lambda\mathbf{k}\mu}(\alpha_{i_3}^{-1}\gamma^{-1}\alpha_{i_1}u\alpha_{i_1}^{-1}\gamma\alpha_{i_3})^*,$$

where

$$(\theta\alpha_{i_1}^{-1}\gamma\alpha_{i_2},\theta\alpha_{i_1}^{-1}\gamma\alpha_{i_3})\in\overline{Q}_{-\gamma}^{(1,2)}(\mathbf{k})-\overline{K}_{-\gamma}^{(1,2)}(\mathbf{k}),$$

$$S_{10}=\sum_{u\in K(\mathbf{k})}\sum\psi^{\lambda\mathbf{k}\mu}(u)^*\sum_{i_1}\sum_{i_2}\sum_{i_3}[\lambda(\gamma\alpha_{i_1}u\alpha_{i_1}^{-1}\gamma^{-1},\theta\gamma)^*\lambda(\theta\gamma,\alpha_{i_1}u\alpha_{i_1}^{-1})]^3$$

$$\times\exp(-i)(\mathbf{k}+\alpha_{i_1}^{-1}\alpha_{i_2}\mathbf{k}+\alpha_{i_1}^{-1}\alpha_{i_3}\mathbf{k})\cdot\mathbf{t}(u)\exp(-i\alpha_{i_2})$$

$$\times(\alpha_{i_2}^{-1}\alpha_{i_1}u^{-1}\alpha_{i_1}^{-1}\alpha_{i_2}\mathbf{k}-\mathbf{k})\cdot(\mathbf{a}_{i_2}-\mathbf{a}_{i_1})\exp-i\alpha_{i_3}(\alpha_{i_3}^{-1}\alpha_{i_1}u^{-1}\alpha_{i_1}^{-1}\alpha_{i_3}\mathbf{k}-\mathbf{k})\cdot(\mathbf{a}_{i_3}-\mathbf{a}_{i_1})f^{\lambda}(i_1,i_1,\alpha_{i_1}u\alpha_{i_1}^{-1})^*$$

$$\times f^{\lambda}(i_2,i_2,\alpha_{i_1}u\alpha_{i_1}^{-1})^*f^{\lambda}(i_3,i_3,\alpha_{i_1}u\alpha_{i_1}^{-1})^*\psi^{\lambda\mathbf{k}\mu}(\alpha_{i_2}^{-1}\alpha_{i_1}u\alpha_{i_1}^{-1}\alpha_{i_2})^*\psi^{\lambda\mathbf{k}\mu}(\alpha_{i_3}^{-1}\alpha_{i_1}u\alpha_{i_1}^{-1}\alpha_{i_3})^*$$

where

$$(\alpha_{i_1}^{-1}\alpha_{i_2},\alpha_{i_1}^{-1}\alpha_{i_3})\in\overline{K}_{-\gamma}^{(1,2)}(\mathbf{k}),$$

$$S_{11}=\sum_{u\in\overline{K}(\mathbf{k})}\psi^{\lambda\mathbf{k}\mu}(u)\sum_{i_1}\sum_{i_2}\exp i(\alpha_{i_1}^{-1}\alpha_{i_2}\mathbf{k}+\mathbf{k})\cdot\mathbf{t}(u)\exp i\alpha_{i_2}(\alpha_{i_2}^{-1}\alpha_{i_1}u^{-1}\alpha_{i_1}^{-1}\alpha_{i_2}\mathbf{k}-\mathbf{k})\cdot(\mathbf{a}_{i_2}-\mathbf{a}_{i_1})f^{\lambda}(i_1,i_1,\alpha_{i_1}u\alpha_{i_1}^{-1})$$

$$\times f^{\lambda}(i_2,i_2,\alpha_{i_1}u\alpha_{i_1}^{-1})\psi^{\lambda\mathbf{k}\mu}(\alpha_{i_2}^{-1}\alpha_{i_1}u\alpha_{i_1}^{-1}\alpha_{i_2})\psi^{\lambda\,2\mathbf{k}=0\nu}(\alpha_{i_1}u\alpha_{i_1}^{-1})^*,$$

where

$$\alpha_{i_1}^{-1}\alpha_{i_2}\in\overline{K}_{-}(\mathbf{k}),$$

$$S_{12}=\sum_{u\in\overline{K}(\mathbf{k}),u^2\in\overline{K}(\mathbf{k})}\lambda(u,u)\psi^{\lambda\mathbf{k}\mu}(u^2)\exp i\mathbf{k}\cdot[\mathbf{t}(u)+u\mathbf{t}(u)]\sum_{i_1}[\lambda(\alpha_{i_1}^{-1},\alpha_{i_1}u\alpha_{i_1}^{-1})\lambda(u,\alpha_{i_1}^{-1})^*]^2\psi^{\lambda\,2\mathbf{k}=0\nu}(\alpha_{i_1}u\alpha_{i_1}^{-1})^*,$$

$$S_{13}=\sum_{u\in\overline{K}(\mathbf{k})}\psi^{\lambda\mathbf{k}\mu}(u)\sum_{i_1}\sum_{i_2}[\lambda(\alpha_{i_1}u\alpha_{i_1}^{-1},\theta\gamma)\lambda(\theta\gamma,\gamma^{-1}\alpha_{i_1}u\alpha_{i_1}^{-1}\gamma)^*]^2\exp i(\alpha_{i_1}^{-1}\alpha_{i_2}\mathbf{k}+\mathbf{k})\cdot\mathbf{t}(u)\exp i\alpha_{i_2}(\alpha_{i_2}^{-1}\alpha_{i_1}u^{-1}\alpha_{i_1}^{-1}\alpha_{i_2}\mathbf{k}-\mathbf{k})$$

$$\cdot(\mathbf{a}_{i_2}-\mathbf{a}_{i_1})f^{\lambda}(i_1,i_1,\alpha_{i_1}u\alpha_{i_1}^{-1})f^{\lambda}(i_2,i_2,\alpha_{i_1}u\alpha_{i_1}^{-1})\psi^{\lambda\mathbf{k}\mu}(\alpha_{i_2}^{-1}\alpha_{i_1}u\alpha_{i_1}^{-1}\alpha_{i_2})\psi^{\lambda\,2\mathbf{k}=0\nu}(\gamma^{-1}\alpha_{i_1}u\alpha_{i_1}^{-1}\gamma),$$

where

$$\alpha_{i_1}^{-1}\alpha_{i_2}\in\overline{K}_{-}(\mathbf{k}),$$

$$S_{14}=\sum_{u\in\overline{K}(\mathbf{k}),u^2\in\overline{K}(\mathbf{k})}\lambda(u,u)\psi^{\lambda\mathbf{k}\mu}(u^2)\exp i\mathbf{k}\cdot[\mathbf{t}(u)+u\mathbf{t}(u)]\sum_{i_1}[\lambda(\alpha_{i_1}^{-1},\alpha_{i_1}u\alpha_{i_1}^{-1})\lambda(u,\alpha_{i_1}^{-1})^*\lambda(\alpha_{i_1}u\alpha_{i_1}^{-1},\theta\gamma)$$

$$\lambda(\theta\gamma,\gamma^{-1}\alpha_{i_1}u\alpha_{i_1}^{-1}\gamma)^*]^2\psi^{\lambda\,2\mathbf{k}=0\nu}(\gamma^{-1}\alpha_{i_1}u\alpha_{i_1}^{-1}\gamma),$$

$$S_{15}=\sum_{u\in\overline{K}(\mathbf{k})}\psi^{\lambda\mathbf{k}\mu}(u)^*\sum_{i_1}\sum_{i_2}[\lambda(\gamma\alpha_{i_1}u\alpha_{i_1}^{-1}\gamma^{-1},\theta\gamma)^*\lambda(\theta\gamma,\alpha_{i_1}u\alpha_{i_1}^{-1})]^2\exp(-i)(\mathbf{k}+\alpha_{i_1}^{-1}\alpha_{i_2}\mathbf{k})\cdot\mathbf{t}(u)\exp(-i\alpha_{i_2})(\alpha_{i_2}^{-1}\alpha_{i_1}u^{-1}\alpha_{i_1}^{-1}$$

$$\times\alpha_{i_2}\mathbf{k}-\mathbf{k})\cdot(\mathbf{a}_{i_2}-\mathbf{a}_{i_1})f^{\lambda}(i_1,i_1,\alpha_{i_1}u\alpha_{i_1}^{-1})^*f^{\lambda}(i_2,i_2,\alpha_{i_1}u\alpha_{i_1}^{-1})^*\psi^{\lambda\mathbf{k}\mu}(\alpha_{i_2}^{-1}\alpha_{i_1}u\alpha_{i_1}^{-1}\alpha_{i_2})^*\psi^{\lambda\,2\mathbf{k}=0\nu}(\gamma\alpha_{i_1}u\alpha_{i_1}^{-1}\gamma^{-1})^*,$$

where

$$\alpha_{i_1}^{-1}\alpha_{i_2}\in\overline{K}_{-\gamma}(\mathbf{k}),$$

$$S_{16}=\sum_{u\in\overline{K}(\mathbf{k})}\psi^{\lambda\mathbf{k}\mu}(u)\sum_{i_1}\sum_{i_2}\lambda(\alpha_{i_1}u\alpha_{i_1}^{-1},\theta\gamma)^*\lambda(\theta\gamma,\gamma^{-1}\alpha_{i_1}u\alpha_{i_1}^{-1}\gamma)\exp-i(\alpha_{i_1}^{-1}\gamma\alpha_{i_2}\mathbf{k}-\mathbf{k})\cdot\mathbf{t}(u)\exp(-i\gamma\alpha_{i_2})(\alpha_{i_2}^{-1}\gamma^{-1}\alpha_{i_1}u^{-1}\alpha_{i_1}^{-1}\gamma\alpha_{i_2}$$

$$\times\mathbf{k}-\mathbf{k})\cdot(\mathbf{c}+\gamma\mathbf{a}_{i_2}-\mathbf{a}_{i_1})f^{\lambda}(i_1,i_1,\alpha_{i_1}u\alpha_{i_1}^{-1})f^{\lambda}(i_2,i_2,\gamma^{-1}\alpha_{i_1}u\alpha_{i_1}^{-1}\gamma)^*\psi^{\lambda\mathbf{k}\mu}(\alpha_{i_2}^{-1}\gamma^{-1}\alpha_{i_1}u\alpha_{i_1}^{-1}\gamma\alpha_{i_2})^*\psi^{\lambda\,2\mathbf{k}=0\nu}(\alpha_{i_1}u\alpha_{i_1}^{-1})^*,$$

where

$$\theta\alpha_{i_1}^{-1}\gamma\alpha_{i_2}\in\overline{Q}_{-\gamma}(\mathbf{k})-\overline{K}_{-\gamma}(\mathbf{k}),$$

$$S_{17}=\sum_{\substack{u\in\bar{K}_{-\gamma}(\mathbf{k})\\ u^2\in\bar{K}(\mathbf{k})}}\psi^{\lambda\,\mathbf{k}\mu}(u^2)^{*}\exp(-i\mathbf{k})\cdot[\mathbf{t}(u)+u\mathbf{t}(u)]\sum_{i_1}\lambda\,(\gamma\alpha_{i_1}u\alpha_{i_1}^{-1}\gamma^{-1},\gamma\alpha_{i_1}u\alpha_{i_1}^{-1}\gamma^{-1})\lambda\,(\gamma\alpha_{i_1}u^2\alpha_{i_1}^{-1}\gamma^{-1},\theta\gamma)^{*}\lambda\,(\theta\gamma,\alpha_{i_1}u^2\alpha_{i_1}^{-1})$$

$$\times f^{\lambda}(i_1,i_1,\alpha_{i_1}u^2\alpha_{i_1}^{-1})^{*}\psi^{\lambda\,{}^2\mathbf{k}=0\nu}(\gamma\alpha_{i_1}u\alpha_{i_1}^{-1}\gamma^{-1})^{*},$$

$$S_{18}=\sum_{u\in\bar{K}(\mathbf{k})}\psi^{\lambda\,\mathbf{k}\mu}(u)^{*}\sum_{i_1}\sum_{i_2}\exp(-i)(\alpha_{i_1}^{-1}\alpha_{i_2}\mathbf{k}+\mathbf{k})\cdot\mathbf{t}(u)\exp(-i\alpha_{i_2})(\alpha_{i_2}^{-1}\alpha_{i_1}u^{-1}\alpha_{i_1}^{-1}\alpha_{i_2}\mathbf{k}-\mathbf{k})\cdot(\mathbf{a}_{i_2}-\mathbf{a}_{i_1})f^{\lambda}(i_1,i_1,\alpha_{i_1}u\alpha_{i_1}^{-1})^{*}$$

$$\times f^{\lambda}(i_2,i_2,\alpha_{i_1}u\alpha_{i_1}^{-1})^{*}\psi^{\lambda\,\mathbf{k}\mu}(\alpha_{i_2}^{-1}\alpha_{i_1}u\alpha_{i_1}^{-1}\alpha_{i_2})^{*}\psi^{\lambda\,{}^2\mathbf{k}=0\nu}(\alpha_{i_1}u\alpha_{i_1}^{-1}),$$

where

$$\alpha_{i_1}^{-1}\alpha_{i_2}\in\bar{K}_{-\gamma}(\mathbf{k}),$$

$$S_{19}=\sum_{u\in\bar{K}(\mathbf{k})}\psi^{\lambda\,\mathbf{k}\mu}(u)\sum_{i_1}\sum_{i_2}\lambda\,(\alpha_{i_1}u\alpha_{i_1}^{-1},\theta\gamma)\lambda\,(\theta\gamma,\gamma^{-1}\alpha_{i_1}u\alpha_{i_1}^{-1}\gamma)^{*}\exp-i(\alpha_{i_1}^{-1}\gamma\alpha_{i_1}\mathbf{k}-\mathbf{k})\cdot\mathbf{t}(u)\exp(-i\gamma\alpha_{i_2})(\alpha_{i_2}^{-1}\gamma^{-1}\alpha_{i_1}u^{-1}\alpha_{i_1}^{-1}\gamma\alpha_{i_2}$$

$$\times\mathbf{k}-\mathbf{k})\cdot(\mathbf{c}+\gamma\mathbf{a}_{i_2}-\mathbf{a}_{i_1})f^{\lambda}(i_1,i_1,\alpha_{i_1}u\alpha_{i_1}^{-1})f^{\lambda}(i_2,i_2,\gamma^{-1}\alpha_{i_1}u\alpha_{i_1}^{-1}\gamma)^{*}\psi^{\lambda\,\mathbf{k}\mu}(\alpha_{i_2}^{-1}\gamma^{-1}\alpha_{i_1}u\alpha_{i_1}^{-1}\gamma\alpha_{i_2})^{*}\psi^{\lambda\,{}^2 0\nu}(\gamma^{-1}\alpha_{i_1}u\alpha_{i_1}^{-1}\gamma),$$

where

$$\theta\alpha_{i_1}^{-1}\gamma\alpha_{i_2}\in\bar{Q}_{-\gamma}(\mathbf{k})-\bar{K}_{-\gamma}(\mathbf{k}),$$

$$S_{20}=\sum_{u\in\bar{K}_{-\gamma}(\mathbf{k}),u^2\in\bar{K}(\mathbf{k})}\psi^{\lambda\,\mathbf{k}\mu}(u^2)^{*}\exp(-i\mathbf{k})\cdot[\mathbf{t}(u)+u\mathbf{t}(u)]\sum_{i_1}\lambda\,(\gamma\alpha_{i_1}u\alpha_{i_1}^{-1}\gamma^{-1},\gamma\alpha_{i_1}u\alpha_{i_1}^{-1}\gamma^{-1})$$

$$\times\lambda\,(\gamma\alpha_{i_1}u^2\alpha_{i_1}^{-1}\gamma^{-1},\theta\gamma)^{*}\lambda\,(\theta\gamma,\alpha_{i_1}u^2\alpha_{i_1}^{-1})\,[\lambda\,(\gamma\alpha_{i_1}u\alpha_{i_1}^{-1}\gamma^{-1},\theta\gamma)\lambda\,(\theta\gamma,\alpha_{i_1}u\alpha_{i_1}^{-1})^{*}]^2f^{\lambda}(i_1,i_1,\alpha_{i_1}u^2\alpha_{i_1}^{-1})^{*}\psi^{\lambda\,{}^2 0\nu}(\alpha_{i_1}u\alpha_{i_1}^{-1}).\quad(13)$$

In all these expressions the α_i's denote the coset representatives of the little cogroup $\bar{K}(\mathbf{k})$ of $D^{\lambda\,\mathbf{k}\mu}$ (*not* of $D^{\lambda\,{}^2 0\nu}$) in the factor group $\bar{G}$. In deducing these expressions we have used the fact that the little cogroup $\bar{K}(\mathbf{k}=0)$ for the reciprocal vector $\mathbf{k}=\mathbf{0}$ is the full point group $\bar{G}$.

To obtain the magnetic space group structure M_1 which the system can acquire after a second order phase transition, we find out the largest subgroup M_1 of M such that $D^{\lambda\,\mathbf{k}\mu}\downarrow M_1$ contains the identity corepresentation of M_1. This is done as follows. We first find out the irreducible corepresentations $D^{\lambda\,\mathbf{k}\mu}$ that satisfy the two Landau conditions. For each of these corepresentations we obtain the bases $\Phi_i^{\lambda\,\mathbf{k}\mu}$ [these are not to be confused with the Φ's in Eq. (5)], constructed from the order parameters. The linear combinations of these bases

$$\Psi^{\lambda\,\mathbf{k}\mu}=\sum_i C_i^{\lambda\,\mathbf{k}\mu}\Phi_i^{\lambda\,\mathbf{k}\mu}\quad(14)$$

with arbitrary $C_i^{\lambda\,\mathbf{k}\mu}$'s are then operated on by the Wigner operators,[9] O_R, $R\in M$. The demand that

$$O_R\Psi^{\lambda\,\mathbf{k}\mu}=\Psi^{\lambda\,\mathbf{k}\mu}\quad(15)$$

cannot be satisfied for all $R\in M$, with nontrivial $C_i^{\lambda\,\mathbf{k}\mu}$'s. The largest subgroups $M_1^{\lambda\,\mathbf{k}\mu}$'s (in the sense that there exists no subgroup M_1' such that $M_1^{\lambda\,\mathbf{k}\mu}\subset M_1'\subset M$) that can be formed consistent with Eq. (15) are the possible structures at the lower symmetry phase. Equation (15) defines relations between the $C_i^{\lambda\,\mathbf{k}\mu}$'s and thus gives the $\Psi^{\lambda\,\mathbf{k}\mu}$'s that mediate a second order phase transition.

In a transition between magnetic phases the Gibb's potential is written[2,3,7] in powers of the thermodynamic averages of the three spin operators:

$$S_{\pm}(i)=\sum_{m=-J_i}^{J_i}\sqrt{J_i(J_i+1)-m(m\pm1)}$$

$$\times|J_i,m\pm1;i\rangle\langle J_i,m;i|,$$

$$S_z(i)=\sum_{m=-J_i}^{J_i}m|J_i,m;i\rangle\langle J_i,m;i|,\quad(16)$$

where $|J_i,m;i\rangle$ is the spin state at the ith site with the total angular momentum J_i and the z component m. These spin operators are the order parameters appearing in the Gibb's potential.[2,3] Since $|J_i,m;i\rangle$ transforms as the irreducible representation D^{J_i} of the full rotation group, the spin operators $S_{\pm}(i),S_z(i)$ will transform as $D^{J_i}\otimes D^{J_i*}$. Thus in both the cases of integral and half-integral spins the order parameters transform according to irreducible corepresentations belonging to the factor system $\lambda\,(\alpha,\beta)$ which is always unity for all elements $\alpha,\beta\in M$.

3. AN EXAMPLE: $CuCl_2\cdot 2H_2O$

Extensive experiments[24] have been performed on $CuCl_2\cdot 2H_2O$, which is a dilute antiferromagnet with the Neel temperature[25] $T_N=4.3$ K. We shall apply the theory given in Sec. 2 to this crystal and shall find out the magnetic symmetries to which it may go as a result of a second order phase transition from the paramagnetic phase.

The chemical symmetry of $CuCl_2{\cdot}2H_2O$ is $G = Pbmn \equiv D^{7}_{2h}$ with the lattice constants $a = 7.38$ a.u., $b = 8.04$ a.u., $c = 3.72$ a.u. and the atomic positions are[25]

Cu in 2(a):$(0,0,0),(\frac{1}{2},\frac{1}{2},0)$,

O in 4(e): $\pm(0,v_1,0)$, $\pm(\frac{1}{2},\frac{1}{2}+v_1,0)$, with $v_1 = 0.2390$,

Cl in 4(h): $\pm(u_1,0,w_1)$, $\pm(\frac{1}{2}-u_1,\frac{1}{2},w_1)$,
with $u_1 = 0.2420$ and $w_1 = 0.3804$,

H in 8(i): $\pm(u_2,v_2,w_2)$, $\pm(\frac{1}{2}-u_2,\frac{1}{2}-v_2,w_2)$,
$\pm(u_2,-v_2,w_2)$, $\pm(\frac{1}{2}-u_2,\frac{1}{2}+v_2,w_2)$,
with $u_2 = 0.0822$, $v_2 = 0.3065$, $w_2 = 0.1295$.

The primitive lattice is generated by the vectors

$$\mathbf{a}_1 = (a,0,0),\quad \mathbf{a}_2 = (0,b,0),\quad \mathbf{a}_3 = (0,0,c).$$

The chemical symmetry group G has the elements[25]

$$G = \{(\mathbf{n}|E)\} \circledS \{(0|E),(0|U^y),(0|I),(0|IU^y),(\mathbf{t}_0|U^x),(\mathbf{t}_0|U^z),(\mathbf{t}_0|IU^x),(\mathbf{t}_0|IU^z)\}$$

where $\mathbf{n} = \Sigma_{i=1}^{3} n_i\mathbf{a}_i$, and $\mathbf{t}_0 = (\mathbf{a}_1 + \mathbf{a}_2)/2$. The reciprocal lattice vectors are

$$\mathbf{b}_1 = (2\pi/a,0,0),\quad \mathbf{b}_2 = (0,2\pi/b,0),\quad \mathbf{b}_3 = (0,0,2\pi/c).$$

The first Brillouin zone for this reciprocal lattice has been given by Zak *et al.*[26]

The magnetic symmetry of this crystal in the paramagnetic phase is $M = G \cup \theta(0|E)G$.

In this case the lattice $\Sigma(\gamma\mathbf{K})$ is the same as the reciprocal lattice $\Sigma(\mathbf{K})$. In Table II we have given the name of the **k**-points within the fundamental region of the first Brillouin zone, the coordinates of these **k**-points, the little cogroup $\bar{K}(\mathbf{k})$, the coset representatives α_i's of $\bar{K}(\mathbf{k})$ in $G/\mathscr{T}$, the projective factor systems $\omega(\alpha,\beta)$, the irreducible representations $\Gamma^{\mathbf{k}\mu}$ of $\bar{K}(\mathbf{k})$, and the different sets given in Eq. (11) that are required in determining the possibility of a second order phase transition. The table contains only those **k**-points for which a second order magnetic transition is possible. Table VI gives the possible magnetic symmetries $M^{\lambda\mathbf{k}\mu}$ which may result after such a transition caused by the irreducible corepresentations $D^{\lambda\mathbf{k}\mu}$ of M that satisfy the Landau conditions.

ACKNOWLEDGMENT

One of us (M.K.S.) is grateful to the University Grants Commission (India) for the award of a research fellowship during the time this work was done.

APPENDIX

In this Appendix we indicate how the quantities $b^{\{\lambda\mathbf{k}\mu \otimes \lambda\mathbf{k}\mu\}}_{\lambda',\mathbf{k}=0,\nu}$ and $\mathscr{S}$ are expressed in Table I in terms of the S_i's of Eq. (13). Since the mode of obtaining the expressions in that Table from Eqs. (7), (8), (9) are very similar we here work out a typical case: the expression for $\mathscr{S}$ when $D^{\lambda\mathbf{k}\mu}$ is of type (b).

Inserting the expression for $D^{\lambda\mathbf{k}\mu}$ in Eq. (9) and performing the summation over the lattice translations we very easily obtain

$$\begin{aligned}\mathscr{S} = (2N/3)\Big\{&2\sum_{i_1i_2i_3}\Delta(\mathbf{k}_{i_1}+\mathbf{k}_{i_2}+\mathbf{k}_{i_3})\sum_{u\in\bar{G}}\prod_{a=1}^{3}[\Phi^{\lambda}_{1}(i_a,i_a,\mathbf{k};\\ &\mathbf{t}(u),u)\psi^{\lambda\mathbf{k}\mu}(\alpha_{i_a}^{-1}u\alpha_{i_a})] + 3\sum_{i_1i_2}\sum_{u\in G}\lambda(u,u)\Delta(\mathbf{k}_{i_1}+\mathbf{k}_{i_2}\\ &+u^{-1}\mathbf{k}_{i_2})\exp i\mathbf{k}_{i_2}\cdot[\mathbf{t}(u)+u\mathbf{t}(u)-\mathbf{t}(u^2)]\Phi^{\lambda}_{1}(i_1,i_1,\\ &\mathbf{k};\mathbf{t}(u),u)\times\Phi^{\lambda}_{1}(i_2,i_2,\mathbf{k};\mathbf{t}(u^2),u^2)\psi^{\lambda\mathbf{k}\mu}(\alpha_{i_1}^{-1}u\alpha_{i_1})\\ &\psi^{\lambda\mathbf{k}\mu}(\alpha_{i_2}^{-1}u^2\alpha_{i_2})\\ &+\sum_{i}\sum_{u\in G}\lambda(u,u)\lambda(u^2,u)\Delta(\mathbf{k}_i+u^{-1}\mathbf{k}_i\\ &+u^{-2}\mathbf{k}_i)\exp(\mathbf{k}_i)\cdot[\mathbf{t}(u)+u\mathbf{t}(u)+u^2\mathbf{t}(u)-\mathbf{t}(u^3)]\\ &\times\Phi^{\lambda}_{1}(i,i,\mathbf{k};\mathbf{t}(u^3),u^3)\psi^{\lambda\mathbf{k}\mu}(\alpha_i^{-1}u^3\alpha_i)\Big\}.\end{aligned}\tag{A1}$$

If we ignore the numerical factors 2 and 3, then the three summations within the braces are just S_1,S_2, and S_3, respectively of Table I.

We now show how the first summation can be expressed as S_1. The factor $\Delta(\mathbf{k}_{i_1}+\mathbf{k}_{i_2}+\mathbf{k}_{i_3})$ ensures that the summations over i_1,i_2, and i_3 are zero unless $\alpha_{i_1}\mathbf{k}+\alpha_{i_2}\mathbf{k}+\alpha_{i_3}\mathbf{k} = \mathbf{K}$, a reciprocal lattice vector. In this case the value of the factor is one. It is now evident that

$$\alpha_{i_1}^{-1}\alpha_{i_2}\mathbf{k}+\alpha_{i_1}^{-1}\alpha_{i_3}\mathbf{k} = -\mathbf{k}+\alpha_{i_1}^{-1}\mathbf{K} = -\mathbf{k}+\mathbf{K}'.\tag{A2}$$

$\mathbf{K}'$ is another reciprocal lattice vector.

Thus the summation over i_1,i_2, and i_3 will have nonvanishing contribution only when

$$(\alpha_{i_1}^{-1}\alpha_{i_2},\alpha_{i_1}^{-1}\alpha_{i_3})\in\bar{K}^{(1,2)}_{-}(\mathbf{k}),\quad\text{defined in Eq. (11).}\tag{A3}$$

The factor $\psi^{\lambda\mathbf{k}\mu}(\alpha_{i_1}^{-1}u\alpha_{i_1})$ ensures that unless

$$v = \alpha_{i_1}^{-1}u\alpha_{i_1}\in\bar{K}(\mathbf{k}),\tag{A4}$$

the contribution to the summation is zero. So the summation over $u\in\bar{G}$ can be restricted to $v\in\bar{K}(\mathbf{k})$. The two factors

$$\psi^{\lambda\mathbf{k}\mu}(\alpha_{i_2}^{-1}u\alpha_{i_2})\psi^{\lambda\mathbf{k}\mu}(\alpha_{i_3}^{-1}u\alpha_{i_3})$$

will be

$$\psi^{\lambda\mathbf{k}\mu}(\alpha_{i_2}^{-1}\alpha_{i_1}v\alpha_{i_1}^{-1}\alpha_{i_2})\psi^{\lambda\mathbf{k}\mu}(\alpha_{i_3}^{-1}\alpha_{i_1}v\alpha_{i_1}^{-1}\alpha_{i_3}).$$

$$\prod_{a=1}^{3}\Phi^{\lambda}_{1}(i_a,i_a,\mathbf{k};\mathbf{t}(u),u) = B\prod_{a=1}^{3}f^{\lambda}(i_a,i_a,\alpha_{i_1}v\alpha_{i_1}^{-1}),$$

where

$$B = \prod_{j=1}^{3}\exp i\mathbf{k}_j(\mathbf{t}(u)+u\mathbf{a}_{i_j}-\mathbf{a}_j)$$

$$= \exp i\mathbf{K}[\mathbf{t}(u)+u\mathbf{a}_{i_1}-\mathbf{a}_{i_1}]\exp i(u^{-1}\mathbf{k}_{i_2}-\mathbf{k}_{i_2})\cdot(\mathbf{a}_{i_2}-\mathbf{a}_{i_1})$$

$$\times \exp i(u^{-1}\mathbf{k}_{i_3} - \mathbf{k}_{i_3})\cdot(\mathbf{a}_{i_3} - \mathbf{a}_{i_1}). \tag{A5}$$

The last two factors in B are equal to

$$\prod_{j=2,3} \exp i\alpha_{i_j}(\alpha_{i_j}^{-1}\alpha_{i_1} v \alpha_{i_1}^{-1}\alpha_{i_j}\mathbf{k} - \mathbf{k})\cdot(\mathbf{a}_{i_j} - \mathbf{a}_{i_1}). \tag{A6}$$

To simplify $\exp i\mathbf{K}\cdot[\mathbf{t}(u) + u\mathbf{a}_{i_1} - \mathbf{a}_{i_1}]$ we proceed like this:

$$\begin{aligned}(\mathbf{a}_{i_1}|\alpha_{i_1})^{-1}(\mathbf{t}(u)|u)(\mathbf{a}_{i_1}|\alpha_{i_1}) &= (-\alpha_{i_1}^{-1}\mathbf{a}_{i_1} + \alpha_{i_1}^{-1}\mathbf{t}(u) + \alpha_{i_1}^{-1}u\mathbf{a}_{i_1}|\alpha_{i_1}^{-1}u\alpha_{i_1})\\ &= (\mathbf{m} + \mathbf{t}(\alpha_{i_1}^{-1}u\alpha_{i_1})|\alpha_{i_1}^{-1}u\alpha_{i_1})\in G.\end{aligned}$$

Here

$$\mathbf{m} = -\alpha_{i_1}^{-1}\mathbf{a}_{i_1} + \alpha_{i_1}^{-1}\mathbf{t}(u) + \alpha_{i_1}^{-1}u\mathbf{a}_{i_1} - \mathbf{t}(\alpha_{i_1}^{-1}u\alpha_{i_1})$$

$$= \alpha_{i_1}^{-1}(\mathbf{t}(u) + u\mathbf{a}_{i_1} - \mathbf{a}_{i_1}) - \mathbf{t}(v)$$

is a reciprocal lattice vector. Hence

$$\begin{aligned}\exp i\mathbf{K}\cdot(\mathbf{t}(u) + u\mathbf{a}_i - \mathbf{a}_{i_1}) &= \exp i\mathbf{K}\cdot\alpha_{i_1}(\mathbf{m} + \mathbf{t}(v))\\ &= \exp i\mathbf{K}\cdot\alpha_{i_1}\mathbf{t}(v) = \exp i\alpha_{i_1}^{-1}\mathbf{K}\cdot\mathbf{t}(v)\\ &= \exp i(\mathbf{k} + \alpha_{i_1}^{-1}\alpha_{i_2}\mathbf{k} + \alpha_{i_1}^{-1}\alpha_{i_3}\mathbf{k})\cdot\mathbf{t}(v)\end{aligned} \tag{A7}$$

Changing the symbol of summation from v to u we find that the first term within the braces of $\mathscr{S}$ in Eq. (A1) is $2S_1$ where S_1 is given in Eq. (13). All the other expressions in Table I are obtained in similar procedures. In actual practice only a few number of terms contribute to the S_i's.

TABLE I. Expressions for $\mathscr{S}$ and $b_{\lambda'\mathbf{k}=0\nu}^{\{\lambda\mathbf{k}\mu\otimes\lambda\mathbf{k}\mu\}}$ in terms of the summations S_i, $i = 1,\ldots,20$.

Type of the Corepresentation $D^{\lambda\mathbf{k}\mu}$	Expression for $\mathscr{S}$	Type of Corepresentation $D^{\lambda'\mathbf{k}=0\nu}$	Expression for $b_{\lambda'\mathbf{k}=0\nu}^{\{\lambda\mathbf{k}\mu\otimes\lambda\mathbf{k}\mu\}}$
Type (a)	$N/6(S_1 + 3S_2 + 2S_3)$	Type (a)	$N/2(S_{11} - S_{12})$
		Type (b)	$N(S_{11} - S_{12})$
		Type (c)	$N/2(S_{11} + S_{13} - S_{12} - S_{14})$
Type (b)	$2N/3(2S_1 + 3S_2 + S_3)$	Type (a)	$N(2S_{11} - S_{12})$
		Type (b)	$2N(2S_{11} - S_{12})$
		Type (c)	$N(2S_{11} + 2S_{13} - S_{12} - S_{14})$
Type (c)	$N/6(S_1 + 3S_2 + 2S_3 + 3S_4 + 3S_5 + 3S_6 + 2S_7 + 3S_8 + 3S_9 + S_{10})$	Type (a)	$N/2(S_{11} - S_{12} + S_{15} + 2S_{16} - S_{17})$
		Type (b)	$N(S_{11} - S_{12} + S_{15} + 2S_{16} - S_{17})$
		Type (c)	$N/2(S_{11} - S_{12} + S_{13} - S_{14} + S_{15} + 2S_{16} - S_{17} + S_{18} + 2S_{19} - S_{20})$

TABLE II. The **k**-points in the fundamental region of the Brillouin zone of the $CuCl_2\cdot 2H_2O$ lattice (D_{2h}^7), the little cogroup $\bar{K}(\mathbf{k})$, and other parameters required to determine whether second order magnetic phase transition occurs. We list only those **k**-points which are compatible with the Landau conditions.

Name of the **k**-point	Γ	Z	R	S
Coordinates of **k** in terms of $\mathbf{b}_1,\mathbf{b}_2,\mathbf{b}_3$	(0,0,0)	$(0,0,\pi/c)$	$(\pi/a,\pi/b,\pi/c)$	$(\pi/a,\pi/b,0)$
Elements of the little cogroup $\bar{K}(\mathbf{k})$	D_{2h}: $E,U^x,U^y,U^z,I,IU^x,IU^y,IU^z$	D_{2h}: $E,U^xU^y,U^z,I,IU^x,IU^y,IU^z$	D_{2h}: $E,U^x,U^y,U^z,I,IU^x,IU^y,IU^z$	D_{2h}: $E,U^x,U^y,U^z,I,IU^x,IU^y,IU^z$
Coset representatives α_i's of $\bar{K}(\mathbf{k})$ in $G/\mathscr{T}$	$\alpha_1 = E$	$\alpha_1 = E$	$\alpha_1 = E$	$\alpha_1 = E$
Projective factor system	Vector representation $\omega(u_1,u_2) = 1$, for all $u_1,u_2\in\bar{K}(\mathbf{k})$	Vector representation	As given in Table III	As given in Table III
Irreducible projective representation $\Gamma^{\lambda\mathbf{k}\mu}$	As given in Table IV	As given in Table IV	As given in Table V	As given in Table V
$\bar{K}(\mathbf{k}) = \bar{K}_{-\gamma}(\mathbf{k})$	$\bar{K}(\mathbf{k})$	$\bar{K}(\mathbf{k})$	$\bar{K}(\mathbf{k})$	$\bar{K}(\mathbf{k})$
$\bar{Q}(\mathbf{k}) - \bar{K}(\mathbf{k}) = \bar{Q}_{-\gamma}(\mathbf{k}) - \bar{K}_{-\gamma}(\mathbf{k})$	$\theta\bar{K}(\mathbf{k})$	$\theta\bar{K}(\mathbf{k})$	$\theta\bar{K}(\mathbf{k})$	$\theta\bar{K}(\mathbf{k})$

TABLE II. Continued.

Name of the **k**-point	Γ	Z	R	S
Types of corepresentation	Type (a)	Type (a)	Type (a)	Type (a)
$b_{\lambda'\mathbf{k}=0\nu}^{\{\lambda\mathbf{k}\mu\otimes\lambda\mathbf{k}\mu\}}$ for different $\Gamma^{\lambda\mathbf{k}\mu}$ and $\Gamma^{\lambda'\mathbf{k}=0\nu}$	0 for all μ,ν	0 for all μ,ν	$8N\delta_{\nu,2}$ for all μ	$8N\delta_{\nu,2}$ for all μ
$\mathscr{A}$ for different $\Gamma^{\lambda\mathbf{k}\mu}$	Zero	Zero	Zero	Zero
The different subsets required to calculate $\mathscr{S}$ when $\mathscr{A}=0$	$\bar{K}_{-}^{(1,2)}(\mathbf{k})=\{(E,E)\}$ $\bar{K}_{-}^{(v)}(\mathbf{k})=\{E\}$, $\forall\ v\in\bar{K}(\mathbf{k})$ $\{v\in\bar{K}_{-}^{(2)}(\mathbf{k})\|v^3\in\bar{K}(\mathbf{k})\}$ $=\bar{K}(\mathbf{k})$	$\bar{K}_{-}^{(1,2)}(\mathbf{k})=\{\phi\}$ $\equiv$vacuous $\bar{K}_{-}^{(v)}(\mathbf{k})=\{\phi\},\forall v\in\bar{K}(\mathbf{k})$ $\{v\in\bar{K}_{-}^{(2)}(\mathbf{k})\|v^3\in\bar{K}(\mathbf{k})\}=\{\phi\}$	As for the point Z	As for the point Z
Values of $\mathscr{S}$ for different $\Gamma^{\lambda\mathbf{k}\mu}$	$8N\delta_{\mu,1}$	0 for all μ	0 for all μ	0 for all μ

TABLE III. Projective factor system $\omega(u_1,u_2)$ or the **k**-points R and S.

u_1 \ u_2	E	U^y	I	IU^y	U^x	U^z	IU^x	IU^z
E	1	1	1	1	1	1	1	1
U^y	1	1	1	1	-1	-1	-1	-1
I	1	1	1	1	1	1	1	1
IU^y	1	1	1	1	-1	-1	-1	-1
U^x	1	1	1	1	-1	-1	-1	-1
U^z	1	1	1	1	1	1	1	1
IU^x	1	1	1	1	-1	-1	-1	-1
IU^z	1	1	1	1	1	1	1	1

TABLE IV. Irreducible projective representations $\Gamma^{\lambda\mathbf{k}\mu}$ for the **k**-points Γ and Z.

$\Gamma^{\lambda\mathbf{k}\mu}(u)$ \ u	E	U^y	U^z	U^x	I	IU^y	IU^z	IU^x
$\Gamma^{\mathbf{k}1}$	1	1	1	1	1	1	1	1
$\Gamma^{\mathbf{k}2}$	1	-1	-1	1	1	-1	-1	1
$\Gamma^{\mathbf{k}3}$	1	1	-1	-1	1	1	-1	-1
$\Gamma^{\mathbf{k}4}$	1	-1	1	-1	1	-1	-1	-1
$\Gamma^{\mathbf{k}5}$	1	1	1	1	-1	-1	-1	-1
$\Gamma^{\mathbf{k}6}$	1	-1	-1	1	-1	1	1	-1
$\Gamma^{\mathbf{k}7}$	1	1	-1	-1	-1	-1	1	1
$\Gamma^{\mathbf{k}8}$	1	-1	1	-1	-1	1	-1	1

TABLE V. Irreducible projective representations $\Gamma^{\lambda\mathbf{k}\mu}$ for the **k**-points R and S. Here α,β are arbitrary constants.

u	E	U^y	I	IU^y	U^x	U^z	IU^x	IU^z
$\Gamma^{\lambda\mathbf{k}1}$	$\begin{pmatrix}1&0\\0&1\end{pmatrix}$	$\begin{pmatrix}0&e^{i\alpha}\\e^{-i\alpha}&0\end{pmatrix}$	$\begin{pmatrix}1&0\\0&1\end{pmatrix}$	$\begin{pmatrix}0&e^{i\alpha}\\e^{-i\alpha}&0\end{pmatrix}$	$\begin{pmatrix}i&0\\0&-i\end{pmatrix}$	$\begin{pmatrix}0&ie^{i\alpha}\\-ie^{-i\alpha}&0\end{pmatrix}$	$\begin{pmatrix}i&0\\0&-i\end{pmatrix}$	$\begin{pmatrix}0&ie^{i\alpha}\\-ie^{i\alpha}&0\end{pmatrix}$
$\psi^{\lambda\mathbf{k}1}$	2	0	2	0	0	0	0	0
$\Gamma^{\lambda\mathbf{k}2}$	$\begin{pmatrix}1&0\\0&1\end{pmatrix}$	$\begin{pmatrix}0&e^{i\beta}\\e^{-i\beta}&0\end{pmatrix}$	$\begin{pmatrix}-1&0\\0&-1\end{pmatrix}$	$\begin{pmatrix}0&-e^{i\beta}\\-e^{-i\beta}&0\end{pmatrix}$	$\begin{pmatrix}i&0\\0&-i\end{pmatrix}$	$\begin{pmatrix}0&ie^{i\beta}\\-ie^{i\beta}&0\end{pmatrix}$	$\begin{pmatrix}-i&0\\0&i\end{pmatrix}$	$\begin{pmatrix}0&-ie^{i\beta}\\ie^{i\beta}&0\end{pmatrix}$
$\psi^{\lambda\mathbf{k}2}$	2	0	-2	0	0	0	0	0

TABLE VI. Magnetic space groups $M^{\lambda \mathbf{k}\mu}$ compatible with the irreducible corepresentation $D^{\lambda \mathbf{k}\mu}$ of M satisfying the Landau conditions. Since all these corepresentations here are of type (a), the bases of M invariant under $M^{\lambda \mathbf{k}\mu}$ depend on whether we take the + sign or the − sign in the matrices for the antilinear elements (cf. Ref. 12). This is reflected in the + or − sign within the square bracket in the last column. The Φ^{μ}_{m} is the mth base of the μth irreducible corepresentation $D^{\lambda \mathbf{k}\mu}$ formed with the spin operators S_x, S_y, S_z. $\rho = \exp i\pi/4$. In the 6th column (p), (f), and (a) means paramagnetic, ferromagnetic, and antiferromagnetic arrangements, respectively.

k-point within the 1st Brillouin	Lattice translation of the magnetic lattice $\mathscr{T}^M$	Nonsymmorphic translation	Irred. Corep. $\Gamma^{\lambda \mathbf{k}\mu}$	Magnetic space group $M^{\lambda \mathbf{k}\mu} = G^{\lambda \mathbf{k}\mu} \cup a_0 G^{\lambda \mathbf{k}\mu}$: $\bar{G}^{\lambda \mathbf{k}\mu} = G^{\lambda \mathbf{k}\mu}/\mathscr{T}^M$	a_0	Bases of M invariant under $M^{\lambda \mathbf{k}\mu}$
Γ: $\mathbf{k}=0$	$\mathbf{a}_1'=\mathbf{a}_1$ $\mathbf{a}_2'=\mathbf{a}_2$ $\mathbf{a}_3'=\mathbf{a}_3$	$\mathbf{t}_0=(\mathbf{a}_1'+\mathbf{a}_2')/2$	Γ^2	$(0\|E),(0\|I),(\mathbf{t}_0\|U^x),$ $(\mathbf{t}_0\|IU^x)$	(p)$\theta(0\|E)$ (f)$\theta(0\|U^y)$	$[+]\psi=\Phi^2_1$; $[-]i\psi$; $[+]i\psi$; $[-]\psi$.
	primitive $V'=V$		Γ^3	$(0\|E),(0\|I),(0\|U^y),$ $(0\|IU^y)$	(p)$\theta(0\|E)$ (f)$\theta(\mathbf{t}_0\|U^x)$	$[+]\psi=\Phi^3_1$; $[-]i\psi$; $[+]i\psi=$; $[-]\psi$.
			Γ^4	$(0\|E),(0\|I),(\mathbf{t}_0\|U^z),$ $(\mathbf{t}_0\|IU^z)$	(p)$\theta(0\|E)$ (f)$\theta(0\|U^y)$	$[+]\psi=\Phi^4_1$; $[-]i\psi$; $[+]i\psi$; $[-]\psi$.
			Γ^5	$(0\|E),(0\|U^y),(\mathbf{t}_0\|U^x),$ $(\mathbf{t}_0\|U^z)$	(p)$\theta(0\|E)$ (f)$\theta(0\|I)$	$[+]\psi=\Phi^5_1$; $[-]i\psi$; $[+]i\psi$; $[-]\psi$.
			Γ^6	$(0\|E),(0\|IU^y),(\mathbf{t}_0\|U^x)$ $(\mathbf{t}_0\|IU^z)$	(p)$\theta(0\|E)$ (f)$\theta(0\|I)$	$[+]\psi=\Phi^6_1$; $[-]i\psi$; $[+]i\psi$; $[-]\psi$.
			Γ^7	$(0\|E),(0\|U^y),(\mathbf{t}_0\|IU^x),$ $(\mathbf{t}_0\|IU^z)$	(p)$\theta(0\|E)$ (f)$\theta(0\|I)$	$[+]\psi=\Phi^7_1$; $[-]i\psi$; $[+]i\psi$; $[-]\psi$.
			Γ^8	$(0\|E),(0\|IU^y),$ $(\mathbf{t}_0\|U^z),(\mathbf{t}_0\|IU^x)$	(p)$\theta(0\|E)$ (f)$\theta(0\|I)$	$[+]\psi=\Phi^8_1$; $[-]i\psi$; $[+]i\psi$; $[-]\psi$.
Z: $\mathbf{k}=\mathbf{b}_3/2$	$\mathbf{a}_1'=\mathbf{a}_1$ $\mathbf{a}_2'=\mathbf{a}_2$ $\mathbf{a}_3'=2\mathbf{a}_3$	$\mathbf{t}_0=(\mathbf{a}_1'/2)+$ $\mathbf{a}_2'/2$	Z^1	$(0\|E),(0\|I),(0\|U^y),(0\|IU^y)$ $(\mathbf{t}_0\|U^x),(\mathbf{t}_0\|U^z),(\mathbf{t}_0\|U^x),$ $(\mathbf{t}_0\|IU^z)$	(p)$\theta(0\|E)$ (a)$\theta(\mathbf{t}\|E)$	$[+]\psi=\Phi^1_1$; $[-]i\psi$; $[+]i\psi$; $[-]\psi$.
	primitive	$\mathbf{t}=\mathbf{a}_3'/2$ $\mathbf{t}_1=\mathbf{t}_0+\mathbf{t}$	Z^2	$(0\|E),(0\|I),(\mathbf{t}_0\|U^x)$ $(\mathbf{t}\|U^y),(\mathbf{t}\|IU^y),(\mathbf{t}_1\|U^z),$ $(\mathbf{t}_0\|IU^x),(\mathbf{t}_1\|IU^z)$	(p)$\theta(0\|E)$ (a)$\theta(t\|E)$	$[+]\psi=\Phi^2_1$; $[-]i\psi$; $[+]i\psi$; $[-]\psi$.
	$V'=2V$		Z^3	$(0\|E),(0\|I),(0\|U^y),$ $(0\|IU^y),(\mathbf{t}_1\|U^z),(\mathbf{t}_1\|IU^z)$ $(\mathbf{t}_1\|U^x),(\mathbf{t}_1\|IU^x)$	(p)$\theta(0\|E)$ (a)$\theta(\mathbf{t}\|E)$	$[+]\psi=\Phi^3_1$; $[-]i\psi$; $[+]i\psi$; $[-]\psi$.
			Z^4	$(0\|E),(0\|I),(\mathbf{t}\|U^y),$ $(\mathbf{t}\|IU^y),(\mathbf{t}_0\|U^z),(\mathbf{t}_0\|IU^z)$ $(\mathbf{t}_1\|U^x),(\mathbf{t}_1\|IU^x)$	(p)$\theta(0\|E)$ (a)$\theta(\mathbf{t}\|E)$	$[+]\psi=\Phi^4_1$; $[-]i\psi$; $[+]i\psi$; $[-]\psi$.
			Z^5	$(0\|E),(\mathbf{t}\|I),(0_0\|U^y),$ $(\mathbf{t}\|IU^y),(\mathbf{t}_0\|U^z),(\mathbf{t}_1\|IU^z),$ $(\mathbf{t}_0\|U^x),(\mathbf{t}_1\|IU^x)$	(p)$\theta(0\|E)$ (a)$\theta(\mathbf{t}\|E)$	$[+]\psi=\Phi^5_1$; $[-]i\psi$; $[+]i\psi$; $[-]\psi$.

TABLE VI. Continued.

k-point within the 1st Brillouin	Lattice translation of the magnetic lattice $\mathscr{T}^M$	Nonsymmorphic translation	Irred. Corep. $\Gamma^{\lambda k\mu}$	Magnetic space group $M^{\lambda k\mu} = G^{\lambda k\mu} \cup a_0 G^{\lambda k\mu}$ $\bar{G}^{\lambda k\mu} = G^{\lambda k\mu}/\mathscr{T}^M$	a_0	Bases of M invariant under $M^{\lambda k\mu}$
			Z^6	$(\mathbf{0}\|E),(\mathbf{t}\|I),(\mathbf{t}\|U^y)$ $(\mathbf{0}\|IU^y),(\mathbf{t}_0\|U^x),(\mathbf{t}_1\|IU^x)$, $(\mathbf{t}_1\|U^z),(\mathbf{t}_0\|IU^z)$	(p)$\theta(\mathbf{0}\|E)$ (a)$\theta(\mathbf{t}\|E)$	$[+]\psi = \Phi^6_1$; $[-]i\psi$; $[+]i\psi$; $[-]\psi$.
			Z^7	$(\mathbf{0}\|E),(\mathbf{t}\|I),(\mathbf{0}\|U^y)$, $(\mathbf{t}\|IU^y),(\mathbf{t}_1\|U^z),(\mathbf{t}_0\|IU^z)$, $(\mathbf{t}_1\|U^x),(\mathbf{t}_0\|IU^x)$	(p)$\theta(\mathbf{0}\|E)$ (a)$\theta(\mathbf{t}\|E)$	$[+]\psi = \Phi^7_1$; $[-]i\psi$; $[+]i\psi$; $[-]\psi$.
			Z^8	$(\mathbf{0}\|E),(\mathbf{t}\|I),(\mathbf{t}\|U^y)$ $(\mathbf{0}\|IU^y),(\mathbf{t}_0\|U^z),(\mathbf{t}_1\|IU^z)$, $(\mathbf{t}_1\|U^x),(\mathbf{t}_0\|IU^x)$	(p)$\theta(\mathbf{0}\|E)$ (a)$\theta(\mathbf{t}\|E)$	$[+]\psi = \Phi^8_1$; $[-]i\psi$; $[+]i\psi$; $[-]\psi$.
R:	$\mathbf{a}'_1 = \mathbf{a}_1$ $+ \mathbf{a}_2$	$\mathbf{t}_0 = \mathbf{a}'_1/2$	R^1	$(\mathbf{0}\|E),(\mathbf{0}\|I),(\mathbf{0}\|U^y),(\mathbf{0}\|IU^y)$	(p)$\theta(\mathbf{0}\|E)$ (a)$\theta(\mathbf{t}\|E)$	$[+]\psi = e^{i\alpha/2}\Phi^1_1 + e^{-i\alpha/2}\Phi^1_2$;$[-]i\psi$; $[+]i\psi$; $[-]\psi$.
$\mathbf{k} = \mathbf{b}_1/2$ $+ \mathbf{b}_2/2$ $+ \mathbf{b}_3/2$	$\mathbf{a}'_2 = \mathbf{a}_1$ $+ \mathbf{a}_3$	$\mathbf{t} = \mathbf{a}'_1/2$ $+ \mathbf{a}'_2/2$ $+ \mathbf{a}'_3/2$		$(\mathbf{0}\|E),(\mathbf{0}\|I),(\mathbf{t}\|U^y),(\mathbf{t}\|IU^y)$	(p)$\theta(\mathbf{0}\|E)$ (a)$\theta(\mathbf{t}\|E)$	$[+]\psi = e^{i\alpha/2}\Phi^1_1 - \mathrm{i}e^{-i\alpha/2}\Phi^1_2$; $[-]-i\psi$; $[+]-i\psi$; $[-]\psi$.
	$\mathbf{a}'_3 = \mathbf{a}_2$ $+ \mathbf{a}_3$	$\mathbf{t}_1 = \mathbf{t} - \mathbf{t}_0$		$(\mathbf{0}\|E),(\mathbf{0}\|I),(\mathbf{t}_0\|U^z)$ $(\mathbf{t}_0\|IU^z)$	(p)$\theta(\mathbf{0}\|E)$ (a)$\theta(\mathbf{t}\|E)$	$[+]\psi = i\rho e^{i\alpha/2}\Phi^1_1 - i\rho^* e^{-i\alpha/2}\Phi^1_2$; $[-]-i\psi$; $[+]-i\psi$; $[-]\psi$.
	all f.c.c.			$(\mathbf{0}\|E),(\mathbf{0}\|I),(\mathbf{t}_1\|U^z)$, $(\mathbf{t}_1\|IU^z)$	(p)$\theta(\mathbf{0}\|E)$ (a)$\theta(\mathbf{t}\|E)$	$[+]\psi = \rho e^{i\alpha/2}\Phi^1_1 + \rho^* e^{-i\alpha/2}\Phi^1_2$; $[-]i\psi$; $[+]i\psi$; $[-]\psi$.
	$V' = 2V$		R^2	$(\mathbf{0}\|E),(\mathbf{t}\|I),(\mathbf{0}\|U^y)$, $(\mathbf{t}\|IU^y)$	(p)$\theta(\mathbf{0}\|E)$ (a)$\theta(\mathbf{t}\|E)$	$[+]\psi = e^{i\beta/2}\Phi^2_1 + e^{-i\beta/2}\Phi^2_2$;$[-]i\psi$; $[+]i\psi$; $[-]\psi$.
				$(\mathbf{0}\|E),(\mathbf{t}\|I),(\mathbf{t}\|U^y)$, $(\mathbf{t}\|IU^y)$	(p)$\theta(\mathbf{0}\|E)$ (a)$\theta(\mathbf{t}\|E)$	$[+]\psi = ie^{i\beta/2}\Phi^2_1 - ie^{-i\beta/2}\Phi^2_2$; $[-]-i\psi$; $[+]-i\psi$; $[-]\psi$.
				$(\mathbf{0}\|E),(\mathbf{t}\|I),(\mathbf{t}_0\|U^z)$, $(\mathbf{t}_1\|IU^z)$	(p)$\theta(\mathbf{0}\|E)$ (a)$\theta(\mathbf{t}\|E)$	$[+]\psi = i\rho e^{i\beta/2}\Phi^2_1 - i\rho^* e^{-i\beta/2}\Phi^2_2$; $[-]-i\psi$; $[+]-i\psi$; $[-]\psi$.
				$(\mathbf{0}\|E),(\mathbf{t}\|I),(\mathbf{t}_1\|U^z)$, $(\mathbf{t}_0\|IU^z)$	(p)$\theta(\mathbf{0}\|E)$ (a)$\theta(\mathbf{t}\|E)$	$[+]\psi = \rho e^{i\beta/2}\Phi^2_1 + \rho^* e^{-i\beta/2}\Phi^2_2$; $[-]i\psi$; $[+]i\psi$; $[-]\psi$.

TABLE VI. Continued.

k-point within the 1st Brillouin	Lattice translation of the magnetic lattice $\mathscr{T}^M$	Nonsymmorphic translation	Irred. Corep. $\Gamma^{\lambda \mathbf{k}\mu}$	$\overline{G}^{\lambda \mathbf{k}\mu} = G^{\lambda \mathbf{k}\mu}/\mathscr{T}^M$	Magnetic space group $M^{\lambda \mathbf{k}\mu} = G^{\lambda \mathbf{k}\mu} \cup a_0 G^{\lambda \mathbf{k}\mu}$ a_0	Bases of M invariant under $M^{\lambda \mathbf{k}\mu}$
S: $\mathbf{k} = \mathbf{b}_1/2$	$\mathbf{a}_1' = \mathbf{a}_1 + \mathbf{a}_2$ $\mathbf{a}_2' = \mathbf{a}_2 - a_1$	$\mathbf{t}_0 = \mathbf{a}_1'/2$ $\mathbf{t} = \mathbf{a}_1'/2$	S^1	$(\mathbf{0}\|E),(\mathbf{0}\|I),(\mathbf{0}\|U^y),$ $(\mathbf{0}\|IU^y)$	(p)$\theta(\mathbf{0}\|E)$ (a)$\theta(\mathbf{t}\|E)$	$[+]\psi = e^{i\alpha/2}\Phi_1^1 + e^{-i\alpha/2}\Phi_2^1$; $[-]i\psi$; $[+]i\psi$; $[-]\psi$.
$+\mathbf{b}_2/2$	$\mathbf{a}_3' = \mathbf{a}_3$	$+\mathbf{a}_2'/2$		$(\mathbf{0}\|E),(\mathbf{0}\|I),(\mathbf{t}\|U^y),$ $(\mathbf{t}\|IU^y)$	(p)$\theta(\mathbf{0}\|E)$ (a)$\theta(\mathbf{t}\|E)$	$[+]\psi = ie^{i\alpha/2}\Phi_1^1 - ie^{-i\alpha/2}\Phi_2^1$; $[-]-i\psi$; $[+]-i\psi$; $[-]\psi$.
	one f.c.c. $V' = 2V$	$\mathbf{t}_1 = \mathbf{t} - \mathbf{t}_0$		$(\mathbf{0}\|E),(\mathbf{0}\|I),(\mathbf{t}_0\|U^z),$ $(\mathbf{t}_0\|IU^z)$	(p)$\theta(\mathbf{0}\|E)$ (a)$\theta(\mathbf{t}\|E)$	$[+]\psi = i\rho e^{i\alpha/2}\Phi_1^1 - i\rho^* e^{-i\alpha/2}\Phi_2^1$; $[-]-i\psi$; $[+]-i\psi$; $[-]\psi$.
				$(\mathbf{0}\|E),(\mathbf{0}\|I),(\mathbf{t}_1\|U^z),$ $(\mathbf{t}_1\|IU^z)$	(p)$\theta(\mathbf{0}\|E)$ (a)$\theta(\mathbf{t}\|E)$	$[+]\psi = \rho e^{i\alpha/2}\Phi_1^1 + \rho^* e^{-i\alpha/2}\Phi_2^1$; $[-]i\psi$; $[+]i\psi$; $[-]\psi$.
			S^2	$(\mathbf{0}\|E),(\mathbf{t}\|I),(\mathbf{0}\|U^y)$ $(\mathbf{t}\|IU^y)$	(p)$\theta(\mathbf{0}\|E)$ (a)$\theta(\mathbf{t}\|E)$	$[+]\psi = e^{i\beta/2}\Phi_1^2 + e^{-i\beta/2}\Phi_2^2$; $[-]i\psi$; $[+]i\psi$; $[-]\psi$.
				$(\mathbf{0}\|E),(\mathbf{t}\|I),(\mathbf{t}\|U^y),$ $(\mathbf{0}\|IU^y)$	(p)$\theta(\mathbf{0}\|E)$ (a)$\theta(\mathbf{t}\|E)$	$[+]\psi = ie^{i\beta/2}\Phi_1^2 - ie^{-i\beta/2}\Phi_2^2$; $[-]-i\psi$; $[+]-i\psi$; $[-]\psi$.
				$(\mathbf{0}\|E),(\mathbf{t}\|I),(\mathbf{t}_0\|U^z)$ $(\mathbf{t}_1\|IU^z)$	(p)$\theta(\mathbf{0}\|E)$ (a)$\theta(\mathbf{t}\|E)$	$[+]\psi = i\rho e^{i\beta/2}\Phi_1^2 - i\rho^* e^{-i\beta/2}\Phi_2^2$;$[-]-i\psi$; $[+]-i\psi$; $[-]\psi$.
				$(\mathbf{0}\|E),(\mathbf{t}\|I),(\mathbf{t}_1\|U^z),$ $(\mathbf{t}_0\|IU^z)$	(p)$\theta(\mathbf{0}\|E)$ (a)$\theta(\mathbf{t}\|E)$	$[+]\psi = \rho e^{i\beta/2}\Phi_1^2 + \rho^* e^{-i\beta/2}\Phi_2^2$;$[-]i\psi$; $[+]i\psi$; $[-]\psi$.

[1]L.D. Landau and E.M. Lifshits, *Statistical Physics* (Pergamon, London, 1959).
[2]E. Stanley, *Introduction to Phase Transitions and Critical Phenomena* (Clarendon, Oxford, 1971).
[3]K.P. Belov, *Magnetic Transitions* (Consultants Bureau, New York, 1961).
[4]G.Y. Lyubarskii, *Group Theory and its Applications in Physics* (Pergamon, London, 1960).
[5]F.E. Goldrich and J.L. Birman, Phys. Rev. **167**, 528 (1968).
[6]O.V. Kovalev, Sov. Phys.-Solid State **5**, 2309, 2315 (1964).
[7]I.E. Dzialoshinskii, Sov. Phys.-JETP **5**, 1259 (1957).
[8]E.P. Wigner, *Group Theory* (Academic, New York, 1959).
[9]C.J. Bradley and B.L. Davies, Rev. Mod. Phys. **40**, 359 (1968).
[10]A.P. Cracknell, J. Phys. C **4**, 2488 (1971).
[11]A.P. Cracknell and A.K. Sedaghat, J. Phys. C **5**, 977 (1972).
[12]V. Kopsky, J. Phys. C **9**, 3391, 3405 (1976).
[13]N.B. Backhouse, J. Math. Phys. **15**, 119 (1974).
[14]P. Gard and N.B. Backhouse, J. Phys. A **8**, 450 (1976).
[15]P. Rudra and M.K. Sikdar, J. Math. Phys. **17**, 463 (1976).
[16]J.O. Dimmock and R.G. Wheeler, J. Phys. Chem. Sol. **23**, 729 (1962).
[17]J.O. Dimmock, J. Math. Phys. **4**, 1307 (1963).
[18]J.O. Dimmock, Phys. Rev. **130**, 1337 (1963).
[19]I.E. Dzialoshinskii, Sov. Phys.-JETP **19**, 960 (1964).
[20]S.A. Brazovskii and I.E. Dzialoshinskii, Sov. Phys.-JETP Lett. **21**, 164 (1975).
[21]A.P. Cracknell, J. Lorenc, and J.A. Przystawa, J. Phys. C **9**, 1731 (1976).
[22]P. Rudra, J. Math. Phys. **15**, 2031 (1974).
[23]G.F. Karavaev, Sov. Phys.-Solid State **6**, 2943 (1965).
[24]H. Umebayashi, G. Shirane, B.C. Frazer, and D.E. Cox, J. Appl. Phys. **38**, 1461 (1967); Phys. Rev. **167**, 519 (1968).
[25]S.J. Joshua, Phys. Stat. Sol. **38**, 643 (1970).
[26]J. Zak, A. Casher, M. Glueck, and Y. Gur, *Irreducible Representations of Space Groups* (Benjamin, New York, 1969).

Successive observations in relativistic quantum theory

M. D. Srinivas

Department of Theoretical Physics, University of Madras, Guindy Campus, Madras-600025, India
(Received 5 April 1978)

Relativistic quantum theory is formulated as a theory of successive local observations, consistent with all the fundamental requirements of relativistic invariance. Based on the causal structure of the space–time of special relativity, it is argued that the nonrelativistic notion of a "time-ordered sequence of instantaneous observations" should be replaced in a relativistic theory should be replaced in a relativistic theory by the notion of a "causally ordered sequence of local observations." The fundamental statistical law of relativistic quantum theory is then formulated such that it provides an unambigous prescription for the statistical correlations between the outcomes of any causally ordered sequence of local observations. It is also shown that the fundamental statistical law of the theory is consistent with the so-called "principle of local causes" which essentially states that the probability connections between a set of local events do not depend on local observations carried out in space–time regions which lie outside of their causal past. Finally, an equivalent formulation of the theory in terms of a relativistically covariant generalization of the collapse postulate is presented.

INTRODUCTION

A completely satisfactory reconciliation of the theories of relativity and quantum mechanics continues to be one of the foremost objectives of present day theoretical physics. Apart from the difficulties inherent in the relativistic formulation of the quantum theory of interacting systems, there appear to be several conceptual problems in achieving a unification of these two theories[1] (see also Refs. 2–18). Many of these conceptual problems are essentially due to the fundamentally nonrelativistic character of several of the basic concepts of quantum theory in its conventional formulation. In particular, we may mention the following features [see (D1), Appendix C] of the conventional formulation of (nonrelativistic) quantum theory, which clearly seem to be in conflict with the basic tenets of the theory of relativity:

(a) The Schrödinger picture notion of the state *of* a system at *various instants of time.*

(b) The collapse postulate which posits an *instantaneous transition* of the above state of the system (or equivalently a transition of the "Heisenberg picture state" across a *constant time hypersurface*).

(c) The restriction that only those experiments which are *performed instantaneously* can be represented as observables in the theory.

(d) The fact that the observable predictions of the theory which are in the form of statistical correlations between the outcomes of a sequence of observations depend on the *time ordering* of these observations.

The present investigation is addressed to a consideration of some of the conceptual difficulties mentioned above. In particular an attempt is made at developing a general framework of relativistic quantum theory which is [not only free from the above features (a)–(d), but also] in complete conformity with all the fundamental requirements of relativistic invariance and, at the same time, provides a consistent description of successive (local) observations. For this purpose, we shall base our considerations on a particular (generalized version of a) conceptual reformulation of quantum theory (as a theory of successive observations) outlined by Wigner.[1,3–5] The most important feature of this formulation (which also reproduces the entire observational content of conventional quantum theory) is that the fundamental statistical law of the theory is postulated in such a way that it directly provides all the statistical correlations between the results of successive measurements, without taking recourse to the collapse postulate. Our discussion in Secs. 1 (see also Refs. 19–24) shows that such an exclusion of the collapse postulate (from the basic assumptions of the theory) can be achieved in complete generality, provided (i) we include in the characterization of an observable, a specification of the associated "measurement transformations," [see (D2)] and (ii) we formulate the theory in terms of the notion of the "state preparation procedure (SPP)" according to which the system was prepared, instead of the conventional notions of the "state *of* the system" in the Schrödinger or Heisenberg picture.

Once we adopt the above formulation of quantum theory, the only modification that seems to be necessary in connection with (a) and (b) above, is that the notion of an "(instantaneous) SPP carried out at time t" should be extended to that of a "(local) SPP carried out in a space–time region O." In order to overcome the difficulties mentioned in (c) and (d), it will be argued in Sec. 2 (and Appendices A and B) that it is necessary to (iii) extend the notion of an "(instantaneous) observation procedure carried out at time t'" to that of a "(local) observation procedure (OP) carried out in a space–time region O," and (iv) replace the conventional notion of a "time-ordered sequence of (instantaneous) OP." by the (relativistic) notion of a "causally ordered sequence (COS) of local OP."

The basic postulates of a general framework of relativistic (local) quantum theory [which incorporates the above modifications (i)–(iv)] are given in Sec. 2 and it is shown that the theory satisfies all the fundamental requirements of relativistic invariance (whch are discussed in Appendix A). In Sec. 3 it is shown that the fundamental statistical law of the

 0022-2488/79/071593-19$01.00

theory is consistent with the (relativistic) "*principle of local causes*" [(D3)]—which essentially states that the joint probabilities for observing various outcomes of some OP $\{(X_{i_k},O_{i_k})|k = 1,2,...,n\}$ (in the notation of Sec. 2), when an ensemble of systems [prepared according to some SPP (μ,O_0)] is subjected to a "causally orderable sequence" $S = \{(X_1,O_1), (X_2,O_2),...,X_{i_1},O_{i_1}),...,(X_{i_n},O_{i_n}),...,(X_{i_r},O_{i_r})\}$ of OP do not depend upon on those OP $(X_2,O_2)\in S$, where O_2 has a null intersection with the causal past of all the regions $\{O_{i_k}|k = 1,2,...,n\}$. Finally in Sec. 4 it is demonstrated that the fundamental statistical law of relativistic quantum theory also implies a relativistic version of the collapse principle (which is now formulated in terms of local SPP).

It may be noted that our formulation of the fundamental statistical law of relativistic quantum theory constitutes a significant generalization of some of the earlier investigations[7-13] of successive observations in relativistic quantum theory. Although many of these investigations employ the notion of local OP, their treatment of successive observations is based on some sort of a relativistically covariant generalization of the conventional Heisenberg picture formulation of the collapse postulate. In this context, it should be pointed out that the "Heisenberg picture of state *of* the system in a given space–time region" is not a particularly useful concept—because such a "local state *of* the system in a given space–time region [(D4)]" depends on the sequence of observations one has chosen to perform on the system "earlier." Also, if it is assumed for example that a transition or collapse of such a state occurs across the backward light cone of the region of observation (as is done for example in Refs. 9–13), then the theory will be restricted to a consideration of *only* those sequences of local OP where the regions of observation do not intersect each others' light cones. Our considerations (cf. Sec. 2 and Appendix B) in fact show that fundamental statistical law of relativistic quantum theory provides the statistical correlations between the outcomes of *any* "causally orderable sequence" of OP.

1. A CONCEPTUAL REFORMULATION OF NONRELATIVISTIC QUANTUM THEORY

In this section we shall outine a formulation of nonrelativistic quantum theory which clearly reveals the fact that all the observable predictions of the theory are in the form of statistical correlations between the results of successive measurements performed on a system. We shall pay particular attention to those features of the theory which will have to be modified when relativistic considerations are brought in.

In the conventional formulation of quantum theory[25] states are represented by density operators and observables are represented by self-adjoint operators on a (separable) Hilbert space $\mathscr{H}$. The fundamental statistical law of the theory is then usually formulated in terms of two prescriptions which may be referred to as the "*Born statistical formula*" and the "*collapse postulate*." The Born statistical formula prescribes that in an experiment which is designed to measure the observable A and is performed on an ensemble of systems in state ρ, the outcome will be found to lie in a Borel set $E\in B(R)$ with probability

$$\mathrm{Tr}(\rho P(E)), \tag{1.1}$$

where $E\to P(E)$ is the spectral measure associated with the self-adjoint operator A. The collapse postulate fixes the state of a system after an experiment has been performed, and plays a crucial role in the conventional formulation of quantum theory in the prediction of joint probabilities for a sequence of observations. In order to state this postulate clearly we shall adopt the (conventional) Heisenberg picture description according to which the state of a system does not change in between observations. Also, we shall consider an observable $A(t)$ (where "t" refers to the instant of observation) which has a completely discrete spectrum and is given by

$$A(t) = \sum_i \lambda_i P_i(t). \tag{1.2}$$

The collapse postulate may now be stated as follows: If the state of the system, prior to the observation of $A(t)$ at time "t" is given by the density operator ρ, then the state of the system immediately after an observation of $A(t)$ in which the outcome was found to lie in $E\in B(R)$, will be given by the density operator $\rho'/\mathrm{Tr}\rho'$, where

$$\rho' = \sum_{\lambda_i\in E} P_i(t)\rho P_i(t). \tag{1.3}$$

From the above two prescriptions, we can obtain the joint probabilities for a sequence of observations. Consider a set of observables $\{A_i(t_i)|i = 1,2,...,r\}$ which have the spectral resolution

$$A_i(t_i) = \sum_j \lambda_j^{(i)} P_j^{(i)}(t_i). \tag{1.4}$$

Now, we can conclude from Eqs. (1.1) and (1.3) that whenever $t_1 < t_2 < \cdots < t_r$, the joint probability for observing the value of $A_i(t_i)$ to be $\lambda_{j_i}^{(i)}$ when an ensemble of systems in state ρ (before t_1) are subjected to a sequence of experiments to measure $A_1(t_1),A_2(t_2),...,A_r(t_r)$, is given by the formula

$$\mathrm{Tr}[P_{j_r}^{(r)}(t_r)\cdots P_{j_1}^{(1)}(t_1)\rho P_{j_1}^{(1)}(t_1)\cdots P_{j_r}^{(r)}(t_r)]. \tag{1.5}$$

The above formula can be extended so as to include also situations where $t_1 \leqslant t_2 \leqslant \cdots \leqslant t_r$ provided that we also stipulate that whenever $t_i = t_{i+1}$ then only such observables $[A_i(t_i)$ and $A_{i+1}(t_{i+1})]$ are to be considered, for which the corresponding projection operators satisfy

$$P_{j_i}^{(i)}(t_i)P_{j_{i+1}}^{(i+1)}(t_{i+1}) = P_{j_{i+1}}^{(i+1)}(t_{i+1})P_{j_i}^{(i)}(t_i), \tag{1.6}$$

for all j_i, j_{i+1}.

We shall refer to Eq. (1.5) as the Wigner formula as it was first obtained by Wigner.[3-5] A very important feature of this formula is that it *incorporates all the observable consequences* of *the collapse postulate*. In fact, Wigner has outlined a "*conceptual reformulation of quantum theory*"[1,3-5] *in which Eq.* (1.5) (*which provides all the probability connections between outcomes* of *successive observations*) *is* (*itself*) *adopted as the fundamental statistical prescription of the theory instead of Eqs.* (1.1) *and* (1.3).

Recent investigations show that the formalism of quantum theory outlined so far [and, therefore, the Wigner formula (1.5)] needs to be generalized mainly for the following reasons:

(a) If the observables are merely characterized as self-adjoint operators, then the collapse postulate (1.3) (which was stated only for observables with a completely discrete spectrum), cannot be generalized in any meaningful way[19,21,22,25] for observables with a continuous spectrum. In particular, therefore, the conventional framework of quantum theory does not provide any prescription [akin to the Wigner formula (1.5)] for the joint probabilities of observing various outcomes in a sequence of experiments which have nondiscrete value spaces.

(b) A succession of two (or more) experiments designed to measure some observables, cannot be characterized as the measurement of a "composite observable" [(D5)] as long as we stick to the characterization of an observable as a self-adjoint operator. It should perhaps be emphasized that this is not just a technically unsatisfactory feature of the theory; in fact it is of considerable significance in connection with the problem of constructing operationally meaningful "local observables" from the field operators in quantum field theory (see, for example, Ref. 1).

(c) Recent investigations show that the class of measurement transformations (1.3) ("ideal measurements") has to be extended to include more general "operations" in any realistic description of situations where continuous observations are performed.[22,26,27] Such a generalization is in particular essential for a derivation of the well-known photon counting formula in quantum optics.[28]

In order to overcome the above shortcomings of the formalism of (conventional) quantum theory it has been found necessary to (i) generalize the class of measurement transformations (1.3) so as to include more general "operations" (cf. below), and (ii) characterize each observation procedure (or experiment) in terms of an "operation-valued measure." [See (D6).] In the conventional description, a self-adjoint operator, or equivalently, a projection-valued measure was associated with an "observable." This however, does not completely characterize the corresponding experiment or observation procedure (OP), for which the associated "measurement transformations" will also have to be specified.

Recent investigations show that the above modifications (i) and (ii) lead to (an extension of the conventional formalism of quantum theory into) the following framework of *quantum probability theory*[19,21] which formalizes the statistics of successive observations in (nonrelativistic) quantum theory. Let $V=\mathscr{T}_S(\mathscr{H})$ be the Banach space of the set of all self-adjoint trace class operators and V^+ be its positive cone. Then the *set of events* $\mathscr{O}$ is (an appropriate subset of) the set $L_1^+(V)$ of all "operations" on V, i.e., each $A\in\mathscr{O}$ is a linear positive norm-nonincreasing operator on V. In particular, $\mathscr{O}$ includes the class of operations (measurement transformations) of the form given by Eq. (1.3). $\mathscr{O}$ is partially ordered by the relation

$$A\leqslant B\Leftrightarrow Av\leqslant Bv, \tag{1.7}$$

for all $v\in V^+$. $\mathscr{O}$ contains a unique null element which we denote by θ. $\mathscr{O}$ has a subset Σ (consisting, in general, of more than one element) of maximal elements, which have the property

$$\xi\in\Sigma\Leftrightarrow \mathrm{Tr}[\xi(v)]=\mathrm{Tr}[v], \tag{1.8}$$

for all $v\in V^+$. In particular the identity operation $I\in\Sigma$.

The conjunction $A\wedge B$ of two events, $A,B\in\mathscr{O}$ is defined by the equation

$$(A\wedge B)(v)=B(A(v)), \tag{1.9}$$

for all $v\in V$. $A\wedge B$ corresponds to the composite event "the event A followed by the event B." We say that two events $A,B\in\mathscr{O}$ are disjoint ($A\perp B$) if $A+B\in\mathscr{O}$ also; if A and B are disjoint then we may define their disjunction $A\vee B$ by the relation

$$A\vee B=A+B. \tag{1.10}$$

Given an event $A\in\mathscr{O}$ there are in general several "complementary events" $A'\in\mathscr{O}$ such that $A+A'\in\Sigma$. Hence, for each $A\in\mathscr{O}$ there exist several $\xi\in\Sigma$ such that $A\leqslant\xi$. *In fact, apart from the noncommutativity of the conjunction* (1.9), *it is the fact that the set Σ of maximal elements* (*of the event space*) *is nontrivial* (i.e., $\Sigma\neq I$) *which constitutes the most important nonclassical feature of quantum theory—and gives rise to the so called "quantum interference of probabilities"* (cf. Refs. 21 and 23 and references cited therein).

The basic assumption of nonrelativistic quantum theory can now be stated as follows: To each observation procedure (OP) or experiment, there is associated an ordered pair (X,t), where t is the instant of observation and X is an "operation-valued measure"—i.e., if $(S,B(S))$ is the value space of the experiment (usually assumed to be a standard Borel space) then X is a map

$$X{:}B(S)\rightarrow\mathscr{O},$$

which satisfies the following:

(i) $X(S)\in\Sigma$. $\quad$ (1.11)

(ii) If $\{E_i\}$ is any sequence of mutually disjoint elements of $B(S)$, then

$$X(\cup_i E_i)=\sum_i X(E_i), \tag{1.12}$$

where the right-hand side is assumed to converge in the strong operator topology on $\mathscr{O}$. It may be noted that the above definition of an OP allows for all those "observables" of the conventional formulation for which the collapse postulate makes sense, to be reformulated as OP; for example, the conventional "observable" $A(t)$ given by Eq. (1.2) can now be viewed as the OP (X_A,t), where

$$X_A(E)(v)=\sum_{\lambda_i\in E}P_i(t)vP_i(t), \tag{1.13}$$

for all $v\in V$.

We shall now characterize a state preparation procedure (SPP) as an ordered pair (μ,t), where t is the instant of preparation and μ is a linear functional

$$\mu{:}\mathscr{O}\rightarrow[0,1],$$

which is continuous under the strong operator topology and satisfies the following:

(i) $\mu(\xi)=1$, $\quad$ (1.14)

for all $\xi \in \Sigma$.

$$\text{(ii)}\ \mu(A \wedge \xi) = \mu(A), \tag{1.15}$$

for all $A \in \mathscr{O}$ and $\xi \in \Sigma$.

We may again note that to each density operator ρ (which characterizes a "state" in the conventional formulation) there is associated an SPP (μ_ρ, t), where

$$\mu_\rho(A) = \mathrm{Tr}[A(\rho)], \tag{1.16}$$

for each $A \in \mathscr{O}$.

The fundamental statistical law of nonrelativistic quantum theory may now be stated as follows: The probability $\Pr^{(\mu,t_0)}_{(X_1,t_1),\dots,(X_r,t_r)}\{X_{i_1}(E_{i_1}),\dots,X_{i_n}(E_{i_n})\}$ for observing the outcomes of experiments $\{(X_{i_k},t_{i_k})\}$ to lie in $\{E_{i_k}\}$ $(k = 1,2,\dots,n)$ when an ensemble of systems prepared according to the SPP (μ,t_0) is subjected ot the sequence of experiments $\{(X_1,t_1),\dots,(X_{i_1},t_{i_1}),\dots,(X_{i_n},t_{i_n}),\dots,(X_r,t_r)\}$ is given by the following relation whenever $t_0 \leqslant t_1 < t_2 < \cdots < t_{i_1} < \cdots < t_{i_n} < \cdots < t_r$:

$$\begin{aligned}\Pr&{}^{(\mu,t_0)}_{(X_1,t_1),\dots,(X_r,t_r)}\{X_{i_1}(E_{i_1}),\dots,(X_{i_n},t_{i_n})\}\\ &= \mu(\xi_1 \wedge \xi_2 \wedge \cdots \wedge \xi_{i_1-1} \wedge X_{i_1}(E_{i_1}) \wedge \xi_{i_1+1} \wedge \cdots\\ &\quad \wedge X_{i_n}(E_{i_n}) \wedge \cdots \wedge \xi_r),\end{aligned} \tag{1.17}$$

where $\xi_\alpha = X_\alpha(S_\alpha)$, where S_α is the value space of (X_α,t_α) $(\alpha = 1,\dots,r)$. It is clear that in Eq. (1.17), the time-ordering of the sequence of observations plays a crucial role. As in the case of the Wigner formula (1.5), we can extend Eq. (1.17) to include situations where $t_0 \leqslant t_1 \leqslant t_2 \leqslant \cdots \leqslant t_{i_1} \leqslant \cdots \leqslant t_{i_n} \leqslant \cdots \leqslant t_r$ provided we stipulate that whenever $t_\alpha = t_{\alpha+1}$ only such OP $[(X_\alpha,t_\alpha)$ and $(X_{\alpha+1},t_{\alpha+1})]$ are considered, which also satisfy

$$X_\alpha(E_\alpha) \wedge X_{\alpha+1}(E_{\alpha+1}) = X_{\alpha+1}(E_{\alpha+1}) \wedge X_\alpha(E_\alpha), \tag{1.18}$$

for all $E_\alpha \in B(S_\alpha)$, $E_{\alpha+1} \in B(S_{\alpha+1})$. Whenever Eq. (1.18) is satisfied, we shall say that the operation-valued measures X_α and $X_{\alpha+1}$ are compatible $(X_\alpha \leftrightarrow X_{\alpha+1})$.

Equation (1.17) constitutes the appropriate generalization of the Wigner formula (1.5) in the above formulation. In fact (1.17) reduces to (1.5) when we consider OP of the form of (1.13) and SPP of the form (1.16). In this context we may note some of the important features of the above formulation of nonrelativistic quantum theory. First of all, the fundamental statistical law of the theory is completely embodied in the above generalized Wigner formula (GWF), Eq. (1.17), and there is no reference to a collapse or reduction of the state of a system due to measurements. Another important feature of the above formulation (and of nonrelativistic quantum theory in general) is the assumption that every SPP and OP is an instantaneous procedure. It should also be noted that no restrictions have been placed as to what OP $\{(X,t)\}$ (and SPP $\{(\mu,t)\}$) can be (separately) carried out at any given instant "t." This, of course, does not mean that all the various OP $\{(X,t)\}$ can be carried out "simultaneously" on the same system at the instant "t." It is only the case that at any given instant "t" every OP (X,t) can be carried out, where X is an arbitrary operation-valued measure which satisfies Eqs. (1.11) and (1.12), and the same is true also of every SPP (μ,t), where μ is an arbitrary strongly continuous linear functional on $\mathscr{O}$ taking values in [0,1], which satisfies Eqs. (1.14) and (1.15).

Finally, it may be noted that our formulation of nonrelativistic quantum theory clearly reveals the following very important nonclassical feature of the quantum theoretic probabilities—which has been sometimes referred to as the "quantum interference of probabilities." This is the feature [which follows from Eq. (1.17)] that the joint probability

$$\Pr{}^{(\mu,t_0)}_{(X_1,t_1),\dots,(X_r,t_r)}\{X_{i_1}(E_{i_1}),\dots,X_{i_n}(E_{i_n})\}$$

for a set of events $\{X_{i_k}(E_{i_k}) \mid k = 1,2,\dots,n\}$ seems to *depend also* on the sequence of OP $\{(X_1,t_1),\dots,(X_r,t_r)\}$ performed on the system. However by making use of Eqs. (1.14), (1.15), (1.17), and (1.18), it can be shown that the above joint probability does not depend on all those OP $\{(X_\beta,t_\beta)\}$, where $\beta \geqslant i_n$ (i.e., $t_\beta \geqslant t_{i_n}$). In other words we have the following "(statistical) *principle of local causes*" in nonrelativistic quantum theory:

"If $t_0 \leqslant t_1 \leqslant \cdots \leqslant t_{i_1} \leqslant \cdots \leqslant t_{i_n} \leqslant \cdots \leqslant t_r$, then the joint probability

$$\Pr{}^{(\mu,t_0)}_{(X_1,t_1),\dots,(X_r,t_r)}\{X_{i_1}(E_{i_1}),\dots,X_{i_n}(E_{i_n})\}$$

depends in general only on those OP $\{(X_\alpha,t_\alpha) \mid t_\alpha < t_{i_n}\}$ performed earlier than t_{i_n}; it does not depend on what OP $\{(X_\beta,t_\beta) \mid t_\beta \geqslant t_{i_n}\}$ are performed on the system at times later than t_{i_n} (in the future)."

2. FUNDAMENTAL STATISTICAL LAW OF RELATIVISTIC QUANTUM THEORY

In this section we shall investigate how the basic statistical prescription of nonrelativistic quantum theory as given by the GWF (1.17) can be generalized in such a way that the theory will conform with all the fundamental requirements of relativistic invariance. For a profound and thoroughly systematic discussion of the conceptual basis and use of invariance principles (both in classical and quantum physics) with special reference to relativisitc invariance, the reader is referred to the pioneering studies of Wigner.[1,2,4,29,30,31] We shall mainly touch upon those details which are essential to the formulation of the basic statistical prescription of relativistic quantum theory in a form analogous to Eq. (1.17).

Since the law of nature embodied in Eq. (1.17) is a statistical law, any empirical verification of the law (which clearly involves a repetition of experiments) is possible only when the probabilities (1.17) are invariant under space–time displacements. In other words, displacement invariance is essential even for an empirically meaningful statement of the laws of nature as per quantum theory. We are, in any case, concerned with a situation where the theory is assumed to be invariant under the (larger) group $P^\uparrow_+$ of all restricted inhomogeneous Lorentz transformations. A careful formulation of this principle of relativistic invariance in terms of what may be called the "fundamental requirements of relativistic invariance" (cf. requirements I–IV in Appendix A) has been undertaken in Appendix A. This appendix also includes a discussion of those additional assumptions which need to be made in order to arrive at the framework of relativistic (local) quantum theory that we shall outline in this section.

The most important requirement is of course the requirement of *translatability* (cf. requirement I of Appendix A) of the description of physical phenomena from the language of one observer into that of any other equivalent observer. As has been emphasized by Wigner,[1,2,4] this requirement of translatability, which is often taken for granted in classical physics, has important consequences in connection with the notion of OP and SPP in relativistic quantum theory. For example, in nonrelativistic quantum theory every OP was characterized as an ordered pair (X,t), where X is an operation-valued measure and t the instant of observation. The spatial extent of the region of observation (which could be some volume V) was not specified at all. Thus, strictly speaking each OP in nonrelativistic quantum theory should be defined as an ordered triple (X,t,V), where V is the spatial region of observation. Now we can clearly see why this nonrelativisitic notion of instantaneous OP is at variance with the theory of relativity. First of all it should be noted that the restriction that all OP should be instantaneous is very unrealistic from a physical point of view. Also (theoretically) such a restriction does not even permit the possibility of considering a succession of two OP as a composite OP. Most important of course is the fact that, in a relativistic theory, we cannot restrict the class of all OP to be instantaneous actions carried over a finite region of space, because under relativistic transformations a $t=$ const hypersurface of one observer will not in general be transformed into another such surface for the equivalent observer—thus clearly violating the requirement of translatability. The preceding remarks apply equally well to the characterization of SPP in nonrelativistic quantum theory, as they were also viewed as instantaneous actions carried over a region of space.

Thus, in a relativistic theory, the requirement of translatability immediately leads to the notion of *(local)* OP (X,O) and *(local)* SPP (μ,O) whch are carried out in a space–time region O, which is not restricted to be a subset of the $t=$ const hypersurfaces of a single observer. It also follows from Eqs. (A2) and (A3) that the set $\mathscr{U}_O(M)$ ($\mathscr{U}_S(M)$) of all space–time regions in the Minkowski space M in which the various OP (SPP) of the theory are assumed to be carried out, should be a "*translatable set* of *space–time regions*" in that

$$O\in\mathscr{U}_O(M)\Rightarrow gO\in\mathscr{U}_O(M), \tag{2.1a}$$

$$O\in\mathscr{U}_S(M)\Rightarrow gO\in\mathscr{U}_S(M), \tag{2.1b}$$

for all $g\in P^{\uparrow}_{+}$, where

$$gO=\{x|g^{-1}x\in O\}.$$

The above restriction to a translatable set of space–time regions ensures that every OP and SPP in the description of an observer shall have a meaning for every equivalent observer as given by Eq. (A2). As an example of a translatable set, we can consider the set of all space–time points $\{\{x\}|x\in M\}$ in the Minkowski space, as is often done in relativistic field theories where the OP are the procedures for "measuring" the fields at various space–time points. In our investigations we shall leave the choice of the translatable set $\mathscr{U}_O(M)$ of space–time regions to be quite arbitrary [cf. however Eq. (2.3) below]. We shall only assume (for the sake of convenience) that the sets $\mathscr{U}_O(M)$ and $\mathscr{U}_S(M)$ coincide.

Once we replace the (nontranslatable) notions of instantaneous OP and instantaneous SPP of nonrelativistic quantum theory, by the notions of local OP and local SPP, we are confronted with the following questions:

(a) What are the local OP which may be considered to *causally succeed* a given local SPP?

(b) When is it that a sequence of local OP can be thought of to be a *causally ordered sequence* of successive OP performed on a system?

In Appendix B we shall discuss these and related questions on the basis of the causal properties of the space–time of special relativity. The essential point of course is to ensure that the relativistic transformation of space–time does not lead to a conflict between the notions of successive observations as employed by equivalent observers in different states of motion. For example, if (X_1,O_1) and (X_2,O_2) are two OP, neither of them can be considered to causally succeed the other if the space–time region O_2 has a nonnull intersection with both the "causal past" and "causal future" of O_1. However, the OP (X_2,O_2) can be considered to causally succeed the OP (X_1,O_1) if O_2 has a null intersection with the causal past of O_1 (i.e., $O_2>O_1$ in the notation of Appendix B). In the same way an OP (X_2,O_2) can be considered to causally succeed a SPP (μ,O_1) iff $O_2>O_1$. It may be noted that this notion of causal succession allows for the OP (X_2,O_2) to be carried out in a region O_2 which may be completely spacelike to, or partly spacelike and partly in the causal future of the space–time region O_1 in which the original OP (X_1,O_1) or SPP (μ,O_1) was carried out. Finally, any notion of a "causally ordered sequence (COS)" of procedures should be such that if the OP (X_2,O_2) causally succeeds the OP (X_1,O_1) then it should be possible to consider the sequence of these OP as a single composite OP (which may be denoted by $(X_1\circ X_2,O_1\cup O_2)$ carried out in the space–time region $O_1\cup O_2$.

The arguments of the preceding paragraph may be summarized in terms of a "*relativistic succession principle*" which is stated in Appendix B. It is also shown there that this principle leads to a consistent scheme of causal ordering of the various local OP and SPP. According to this scheme, a sequence of OP $\{(X_1,O_1),...,(X_r,O_r)\}$ can be considered as a sequence of *successive* OP which *follow* a SPP (μ,O_0) if and only if $\{O_0,O_1,...,O_r\}$ is a "*causally ordered sequence* (COS)" of space–time regions—i.e., it should satisfy the condition [Eq. (B11)].

$$O_i>O_j \text{ whenever } j<i \ (i,j=0,1,2,...,r). \tag{2.2}$$

The basic properties of COS are stated in Theorem B1 and Theorem B2 of Appendix B. Also in Theorem B3 is stated a necessary and sufficient condition for a sequence of space–time regions to be "*causally orderable*"—i.e., the sequence may be permuted into a COS.

We should emphasize here that the notion of a COS of OP allows the theory to deal with a much larger class of sequences of OP when compared with some of the earlier discussions of successive observations in relativistic quantum theory. For example, Hellwig and Krauss[11] impose the condition that the space–time regions of successive observations should not intersect each others' light cones—a very restrictive condition which seems to arise mainly because their discussion is not based on the appropriate relativistic generalization of the collapse postulate (cf. Sec. 4 below). Several other investigations of successive observations in relativistic quantum theory (for example, Ref. 9) do not consider the problem of causal ordering of a sequence of OP, as they restrict themselves to a discussion of procedures carried out at space–time points only.

The notions of local OP, local SPP, and COS of OP are the basic ingredients of a relativistic theory of successive observations. Based on these we shall now outline a framework of relativistic (local) quantum thoery. Such a framework should be in some sense a generalization of the framework of nonrelativistic quantum theory outlined in Sec. 1. In particular, the local OP and local SPP should be defined in terms of a basic space $\mathscr{O}$ of events of the theory. Moreover, since the events of our theory are nothing but the realization of particular outcomes when the local OP are carried out we are led to consider for each ("allowed") space–time region O, a corresponding set $\mathscr{O}(O)$ of "local events in O. [see (D7)] This correspondence $O \rightarrow \mathscr{O}(O)$ between space–time regions and the "local events" essentially fixes the structure of the theory.

We now state the basic postulates of relativistic (local) quantum theory. [(D8)]. Let $\mathscr{U}(M)$ be a translatable set of regions in the Minkowski space, which satisfies the condition

$$O_1, O_2 \in \mathscr{U}(M) \quad \text{and} \quad O_1 < O_2 \Rightarrow O_1 \cup O_2 \in \mathscr{U}(M). \tag{2.3}$$

With each $O \in \mathscr{U}(M)$, there is associated a set $\mathscr{O}(O)$ of "*local events (operations) in O*", where $\mathscr{O}(O)$ is a convex subset of $L^1_+(V)$, the set of all positive norm-nonincreasing linear operators on the Banach space $V = \mathscr{T}_S(\mathscr{H})$ the set of all self-adjoint trace class operators on a (separable) Hilbert space $\mathscr{H}$. Let

$$\mathscr{U} = \bigcup_{O \in \mathscr{U}(M)} \mathscr{O}(O) \tag{2.4}$$

denote the set of all local events and let Σ be the set of all (maximal) elements in $\mathscr{O}$ which satisfy Eq. (1.8). For each $O \in \mathscr{U}(M)$ the set $\mathscr{O}(O)$ satisfies the following conditions (E1)–(E5):

(E1)

(a) $I \in \mathscr{O}(O)$.

(b) $\theta \in \mathscr{O}(O)$.

(c) For each $A \in \mathscr{O}(O)$ there exists $A' \in \mathscr{O}(O)$ such that $A + A' \in \Sigma$.

(d) If $\{A_i\}$ is a sequence of elements of $\mathscr{O}(O)$ such that $\Sigma_i A_i$ converges in the strong operator topology to an element of $L^1_+(V)$, then $\Sigma_i A_i \in \mathscr{O}(O)$.

(E2) $O_1 \subset O_2 \Rightarrow \mathscr{O}(O_1) \subset \mathscr{O}(O_2)$.

(E3) If $A_1 \in \mathscr{O}(O_1)$, $A_2 \in \mathscr{O}(O_2)$ and $O_1 < O_2$, then $A_1 \wedge A_2 \in \mathscr{O}(O_1 \cup O_2)$.

(E4) If $A_1 \in \mathscr{O}(O_1)$, $A_2 \in \mathscr{O}(O_2)$ and O_1 is completely spacelike to O_2(O_1[as in Eq. (B8)]O_2), then $A_1 \wedge A_2 = A_2 \wedge A_1$.

(E5) To each $g \in P^{\uparrow}_+$ (the group of all restricted inhomogeneous Lorentz transformations) is associated a map α_g: $\mathscr{O} \rightarrow \mathscr{O}$, which satisfies the following conditions (a)–(c):

(a) $g \rightarrow \alpha_g$ is a representation of $P^{\uparrow}_+$

$\alpha_{g_1} \alpha_{g_2} = \alpha_{g_1 g_2}$,

and α_e is the identity transformation where e is the identity element of $P^{\uparrow}_+$.

(b) For each $g \in P^{\uparrow}_+$, α_g is a automorphism of $\mathscr{O}$ in the following sense: α_g is a bijection which is continuous in the strong operator topology and satisfies the following conditions (i)–(v):

(i) $\alpha_g I = I$.

(ii) $\alpha_g \theta = \theta$.

(iii) $\xi \in \Sigma \Rightarrow \alpha_g \xi \in \Sigma$.

(iv) $\alpha_g(A_1 + A_2) = \alpha_g A_1 + \alpha_g A_2$,

whenever $A_1, A_2, A_1 + A_2 \in \mathscr{O}(O)$ for some O.

(v) $\alpha_g(A_1 \wedge A_2) = (\alpha_g A_1) \wedge (\alpha_g A_2)$,

whenever $A_1 \in \mathscr{O}(O_1)$, $A_2 \in \mathscr{O}(O_2)$ for some $O_1 < O_2$.

(c) $\alpha_g \mathscr{O}(O) = \mathscr{O}(gO)$

for all $O \in \mathscr{U}(M)$ and $g \in P^{\uparrow}_+$.

Based on the above structure of the event space, we now define a local OP as follows: *A local* OP which is performed in a space–time region O, and which has a value space $[S, B(S)]$ (assumed to be a standard Borel space) is characterized by an ordered pair (X, O) where X is a map

$X: B(S) \rightarrow \mathscr{O}(O)$,

which satisfies the following conditions (OP1), (OP2):

(OP1) $X(S) \in \Sigma$.

(OP2) If $\{E_i\}$ is a sequence of mutually disjoint elements of $B(S)$, then

$$X(\bigcup_i E_i) = \sum_i X(E_i)$$

where the right-hand side is assumed to converge in the strong operator topology on $\mathscr{O}(O)$.

Thus each local OP (X, O) is characterized by an "operation-valued measure" whose range is contained in $\mathscr{O}(O)$. This is of course clearly in accordance with the meaning we attached earlier to the set $\mathscr{O}(O)$ of local events in O. As regards the relativistic transformation properties of the local OP, we shall assume the following:

(OP3) Under each relativistic transformation $g \in P^{\uparrow}_+$ the OP(X, O) will be transformed into the OP $\beta_g(X, O)$ given by the relation

$\beta_g(X, O) = (\beta_g X, gO)$,

where the operation-valued measure $\beta_g X$ is defined by the equation

$(\beta_g X)(E) = \alpha_g(X(E))$,

for all $E \in B(S)$.

It can easily be shown [from (E1)–(E5)] that $\beta_g(X, O)$ as defined above is also a local OP (carried out in the region

gO). Also $g \to \beta_g$ is a representation of $P^{\uparrow}_{+}$. The operational meaning of the OP $\beta_g(X,O)$ [in terms of that of the OP(X,O)] is discussed in Appendix A. If we denote by $\mathscr{E}(O)$ the set of all local OP which are carried out in the region O, then we have

$$\beta_g \mathscr{E}(O) = \mathscr{E}(gO). \tag{2.5}$$

We now consider how a succession $\{(X_1,O_1),(X_2,O_2)\}$ of two local OP can be viewed as a *composite* OP performed in $O_1 \cup O_2$ whenever $O_1 < O_2$ (as is required by the "relativistic succession princple" of Appendix B). In this context we may recall the well-known result of Davies and Lewis[19,27] that if the value spaces $(S_1 B(S_1))$, $(S_2, B(S_2))$ of two operation-valued measures X_1, X_2 are standard Borel spaces, then there exists a unique operation-valued measure $X_1 \circ X_2$ with value space $[S_1 \times S_2, B(S_1 \times S_2)]$ such that

$$(X_1 \circ X_2)(E_1 \times E_2) = X_1(E_1) \wedge X_2(E_2), \tag{2.6}$$

for every $E_1 \in B(S_1)$ and $E_2 \in B(S_2)$. If we now consider two OP (X_1,O_1), (X_2,O_2) such that $O_1 < O_2$, than it follows from and (2.6) that for each $E_1 \in B(S_1)$ and $E_2 \in B(S_2)$.

$$(X_1 \circ X_2)(E_1 \times E_2) \in \mathscr{O}(O_1 \cup O_2). \tag{2.7}$$

By making suitable topological assumptions [(D9)] on the local event spaces $\mathscr{O}(O)$, it may be ensured that the entire range of the operation-valued measure $X_1 \circ X_2$ is contained in $\mathscr{O}(O_1 \cup O_2)$. The $(X_1 \circ X_2, O_1 \cup O_2)$ will be a local OP which is carried out in $O_1 \cup O_2$. We shall write

$$(X_1,O_1) \circ (X_2,O_2) = (X_1 \circ X_2, O_1 \cup O_2), \tag{2.8}$$

and refer to the operation "$\circ$" (which is defined only when $O_1 < O_2$), as the operation of compounding two local OP. From the physical interpretation of the conjunction of two events in quantum theory (cf. Sec. 1), it follows that $(X_1 \circ X_2, O_1 \cup O_2)$ is the composite OP which corresponds to the succession of the two OP (X_1,O_1), (X_2,O_2) whenever $O_1 < O_2$.

The compounding of OP is associative in general. However from (E4) it follows that for any two OP (X_1,O_1), (X_2,O_2), if $O_1 \times O_2$, then

$$X_1(E_1) \wedge X_2(E_2) = X_2(E_2) \wedge X_1(E_1), \tag{2.9}$$

for all $E_1 \in B(S_1)$, $E_2 \in B(S_2)$. Whenever Eq. (2.9) is satisfied, we shall say that the OP (X_1,O_1) and (X_2,O_2) are *mutually compatible*. It is obvious that the compatibility (2.9) of two local OP (X_1,O_1), (X_2,O_2) performed in mutually spacelike regions O_1, O_2 depends crucially on the "compatibility of spacelike events" as expressed in our postulate (E4).

If $\{(X_1,O_1),\ldots,(X_r,O_r)\}$ is a COS of OP, then the composite OP which corresponds to the above sequence, is given by

$$(X_1,O_1) \circ (X_2,O_2) \circ \cdots \circ (X_r,O_r) = (X_1 \circ X_2 \circ \cdots \circ X_r, \bigcup_{i=1}^{r} O_i). \tag{2.10}$$

From Theorem B2 and the above relation (2.9), it follows tht the composite OP given by (2.10) does not depend on which causally ordered permutation of the COS $\{(X_1,O_1),\ldots,(X_r,O_r)\}$ is employed in Eq. (2.10). Finally, it may be remarked that compounding of local OP is a relativistically covariant operation—for, if $O_1 < O_2$ then for each $g \in P^{\uparrow}_{+}$ we have $gO_1 < gO_2$ and

$$\beta_g((X_1,O_1) \circ (X_2,O_2)) = (\beta_g(X_1,O_1)) \circ (\beta_g(X_2,O_2)). \tag{2.11}$$

We shall now consider the characterization of local SPP in our formalism. From the relativistic succession principle it follows that a local SPP which is carried out in a space–time region O_0, can be followed by only those OP (X_1,O_1) such that $O_1 > O_0$—i.e., only those OP which are carried out in regions O_1 such that $O_1 \subset S^{+}(O_0)$ [cf. Eq. (B2)]. Hence we are led to the following definition: A *local* SPP which is carried out in a space–time region O_0 is characterized by an ordered pair (μ, O_0) where μ is a linear functional

$$\mu: \bigcup_{O \subset S^{+}(O_0)} \mathscr{O}(O) \to [0,1],$$

which is continuous in the strong operator topology and satisfies the following conditions (SPP1), (SPP2):

(SPP1): $\mu(\xi) = 1$,

for all $\xi \in \Sigma \cap [\bigcup_{O \subset S^{+}(O_0)} \mathscr{O}(O)]$.

(SPP2) If $\{O_0,O_1,O_2\}$ is a COS and $A \in \mathscr{O}(O_1)$, $\xi \in \Sigma \cap \mathscr{O}(O_2)$, then

$$\mu(A \wedge \xi) = \mu(A).$$

The physical significance of the above linear functional "μ" is of course that it assigns probabilities to every succession of local events in $S^{+}(O_0)$. Therefore, (as in the formulation of nonrelativistic quantum theory outlined in Sec. 1), in our mathematical characterization, *an* SPP (μ,O_0) *refers to the entire history of the system "after" its preparation in the space–time region* O_0. It may be useful to compare our notion of a SPP with the closely related notion of "*state sub spacie aeternitatis*" or briefly "*state suspae*." The notion of *state suspae* is usually identified with that of the conventional "Heisenberg picture state of the system" in which case it is quite different from the notion of SPP. Sometimes the notion of *state suspae* is defined as an idealized SPP which is carried out in the infinite past. In our formalism such a SPP will be represented by a strongly continuous linear functional μ: $\mathscr{O} \to [0,1]$ (note that it is defined on *all* local events) which satisfies conditions analogous to (SPP1) and (SPP2).

As regards the relativistic transformation properties of the local SPP, we shall assume the following:

(SPP3) Under each relativistic transformation $g \in P^{\uparrow}_{+}$ the SPP(μ,O_0) will be transformed into the SPP $\tilde{\beta}_g(\mu,O_0)$ given by the relation

$$\tilde{\beta}_g(\mu,O_0) = (\tilde{\beta}_g \mu, gO_0),$$

where the linear functional $\tilde{\beta}_g \mu$ is defined by the equation

$$(\tilde{\beta}_g \mu)(A) = \mu(\alpha_{g^{-1}} A),$$

for all

$$A \in \bigcup_{O \subset S^{+}(gO_0)} \mathscr{O}(O).$$

We can easily show from (E1)–(E5) that $\tilde{\beta}_g(\mu,O_0)$ as defined above is also a local SPP (carried out in the space–time region gO_0) and that $g \to \tilde{\beta}_g$ is a representation of $P^{\uparrow}_{+}$. Also if we denote by $\mathscr{T}(O_0)$ the set of all local SPP which are carried out in O_0, then we have

$$\widetilde{\beta}_g \mathscr{T}(O_0) = \mathscr{T}(gO_0). \tag{2.12}$$

We now state the fundamental statistical law of relativistic local quantum theory as follows:

When an ensemble of systems prepared according to the SPP (μ,O_0) *is subjected to the sequence of* OP $\{(X_1,O_1),...,(X_{i_1},O_{i_1}),...,(X_{i_n},O_{i_n}),...,(X_r,O_r)\}$, *where* $\{O_0,O_1,O_2,...,O_r\}$ *is a* COS, *the joint probability for observing the outcomes of the* OP $\{(X_{i_k},O_{i_k})\}$ to lie in Borel sets $\{E_{i_k}\}$ $(k = 1,2,...,n)$ *is given by the equation*

$$\begin{aligned}\Pr{}^{(\mu,O_0)}_{(X_1,O_1),...,(X_r,O_r)}\{X_{i_1}(E_{i_1}),...,X_{i_n}(E_{i_n})\} &= \mu(\xi_1 \wedge \xi_2 \wedge \cdots \wedge \xi_{i_1-1}) \wedge X_{i_1}(E_{i_1}) \wedge \xi_{i_1+1} \\ &\quad \wedge \cdots \wedge X_{i_n}(E_{i_n}) \wedge \cdots \wedge \xi_r),\end{aligned} \tag{2.13}$$

where $\xi_\alpha = X_\alpha(S_\alpha)$ and S_α *is the value space of the* OP (X_α,O_α) $(\alpha = 1,2,...,r)$.

Equation (2.13) may be referred to as the *relativistic Wigner formula* (RWF) as it is the appropriate relativistic generalization of Eqs. (1.5), (1.17). As per the fundamental statistical law stated above, RWF incorporates the entire observational content of the theory. As we have already remarked, the above formulation of the fundamental law of relativistic quantum theory is completely free from any reference to the notion of collapse or reduction of the state of a system due to observations. However, we shall show later in Sec. 4 that the above fundamental law can also be equivalently formulated in terms of an appropriate relativistic generalization of the collapse postulate.

We shall now show that the relativistic Wigner formula (2.13) gives an unambiguous prescription for the statistical correlations between the outcomes of successive (local) observations. For this let us consider an ensemble of systems prepared according to the SPP (μ,O_0). Then a set of local OP $\{(X_1,O_1),...,(X_r,O_r)\}$ can be considered as successive experiments performed on this ensemble of systems iff $\{O_1,O_2,...,O_r\} = S$, is a causally orderable sequence and $O_\alpha \subset S^+(O_0)$ $(\alpha = 1,...,r)$. Let $\{O_1,O_2,...,O_r\}$ itself be a causally ordered permutation of S. Then Eq. (2.13) gives the joint probabilities for observing the various possible outcomes in the above set of experiments. Now if $S' = \{O_{\pi_1},O_{\pi_2},...,O_{\pi_r}\}$ is any other causally ordered permutation of S, then it follows from theorem B2 that S can be obtained from S' by carrying out successive interchanges between neighboring elements which are spacelike to each other. Therefore if we now employ the formula (2.13) for the COS of OP

$$\{(X_{\pi_1},O_{\pi_1}),(X_{\pi_2},O_{\pi_2}),...,(X_{\pi_r},O_{\pi_r})\}$$

we obtain the same result as above because [from Eq. (2.9)] the right-hand side of Eq. (2.13) remains unaltered under every interchange of two successive OP which are carried out in mutually spacelike regions. In other words, we have the following result:

Theorem 2.1: Let (μ,O_0) be an SPP and $\{(X_1,O_1),...,(X_{i_1},O_{i_1}),...,(X_{i_n},O_{i_n}),...,(X_r,O_r)\}$ be a sequence of OP such that $\{O_0,O_1,...,O_r\}$ is a COS. Then for every permutation π of the indices $\{1,2,...,r\}$ which is such that $\{O_{\pi_1},O_{\pi_2},...,O_{\pi_r}\}$ is also a COS, we have

$$\begin{aligned}&\Pr{}^{(\mu,O_0)}_{(X_1,O_1),...,(X_r,O_r)}\{X_{i_1}(E_{i_1}),...,X_{i_n}(E_{i_n})\} \\ &\quad = \Pr{}^{(\mu,O_0)}_{(X_{\pi_1},O_{\pi_1}),...,(X_{\pi_r},O_{\pi_r})}\{X_{i_1}(E_{i_1}),...,X_{i_n}(E_{i_n})\},\end{aligned} \tag{2.14}$$

for all $E_{i_k} \in B(S_{i_k})$, $(k = 1,2,...,n)$.

The above theorem clearly shows that given a causally orderable sequence of OP $\{(X_1,O_1),...,(X_r,O_r)\}$ the fundamental statistical law (2.13) prescribes unique statistical correlations between their various possible outcomes [irrespective of what causally ordered permutation of these OP is adopted while applying Eq. (2.13)]. Hence we have now established the fact that the framework of relativistic quantum theory we have described so far, provides a consistent description of successive (local) observations (in accordance with the "relativistic succession principle"). We shall now show that this framework is also in complete conformity with the fundamental requirements of relativistic invariance stated in Appendix A. First of all it may be recalled that the relativistic transformation properties of local OP (and SPP) were defined in such a way that under each relativistic transformation $g \in P^{\uparrow}_{+}$ an OP (X,O) [SPP (μ,O)] is transformed into the OP $(\beta_g X,gO)$ [SPP $(\widetilde{\beta}_g \mu,gO)$] in accordance with Eqs. (A11a), (A11b); also $g \to \beta_g$ and $g \to \widetilde{\beta}_g$ constitute representations of the relativity group $P^{\uparrow}_{+}$ in agreement with Theorem A1. Therefore, we only have to verify that the fundamental statistical law of the theory as stated above, satisfies the requirement of relativistic symmetry [cf. Eq. (A18)] in agreement with Theorem A2. For this, we may note that from (SPP3), (E5), and Eq. (2.13) it follows that

$$\begin{aligned}\Pr{}^{(\mu,O_0)}_{(X_1,O_1),...,(X_r,O_r)}\{X_{i_1}(E_{i_1}),...,X_{i_n}(E_{i_n})\} &= (\widetilde{\beta}_g \mu)[\alpha_g\{X_1(S_1) \wedge \cdots \wedge X_{i_1}(E_{i_1}) \wedge \cdots \\ &\quad \times \wedge X_{i_n}(E_{i_n}) \wedge \cdots \wedge X_r(S_r)\}] \\ &= (\widetilde{\beta}_g \mu)\{(\beta_g X_1)(S_1) \wedge \cdots (\beta_g X_{i_1})(E_{i_1}) \wedge \\ &\quad \cdots \wedge (\beta_g X_{i_n})(E_{i_n}) \wedge \cdots \wedge (\beta_g X_r)(S_r)\},\end{aligned} \tag{2.15}$$

where the second step follows from (E5bv) and (OP3). Now if we make use of the fact (cf. Theorem B1) that whenever $\{O_0,O_1,..,O_r\}$ is a COS $\{gO_0,...,gO_r\}$ is also a COS, then we recognize that the right-hand side of Eq. (2.15) is nothing but

$$\Pr{}^{(\widetilde{\beta}_g \mu,gO_0)}_{(\beta_g X_1,gO_1),...,(\beta_g X_r,gO_r)}\{(\beta_g X_{i_1})(E_{i_1}),...,(\beta_g X_{i_n})(E_{i_n})\}.$$

This establishes the fact that the relativistic Wigner formula (2.13) also satisfies the requirement of relativistic symmetry as expressed by Eq. (A18).

3. PRINCIPLE OF LOCAL CAUSES

In this section we shall derive a relativistic generalization of the (statistical) principle of local causes stated in Sec. 1. The fundamental statistical law of relativistic quantum theory [Eq. (2.13)] stated in the last section clearly reveals that the "quantum interference of probabilities" is also operative in relativistic quantum theory—in that the joint probability $\Pr^{(\mu,O_0)}_{(X_1,O_1),\ldots,(X_r,O_r)}\{X_{i_1}(E_{i_1}),\ldots,X_{i_n}(E_{i_n})\}$ seems to depend not only on the set of events $\{X_{i_k}(E_{i_k})|k=1,2,\ldots,n\}$ but also on the sequence $\{(X_1,O_1),\ldots,(X_r,O_r)\}$ of all OP performed on the system. However, our heuristic understanding of relativistic causality (sometimes referred to as Einstein causality) definitely suggests that such a dependence (if any) should be confined to only those OP $(X_\alpha,O_\alpha)\in\{(X_1,O_1),\ldots,(X_r,O_r)\}$ where O_α has a nonnull intersection with the causal past of at least one of the regions $\{O_{i_k}|k=1,2,\ldots,n\}$. In other words, we would like to show that the joint probability $\Pr^{(\mu,O_0)}_{(X_1,O_1),\ldots,(X_r,O_r)}\{X_{i_1}(E_{i_1}),\ldots,X_{i_n}(E_{i_n})\}$ does not depend on all those OP $(X_\beta,O_\beta)\in\{(X_1,O_1),\ldots,(X_r,O_r)\}$, where $O_\beta\subset S^+(\cup_{k=1}^n O_{i_k})$.

Since an analysis of the general case is somewhat complicated, we shall first consider the (much simpler) situation where a SPP (μ,O_0) is followed by two OP (X_1,O_1), (X_2,O_2), where $O_1,O_2\subset S^+(O_0)$. If $O_1<O_2$, then $\{O_0,O_1,O_2\}$ is a COS and the probabilities $\Pr^{(\mu,O_0)}_{(X_1,O_1),(X_2,O_2)}\{X_1(E_1)\}$ can be obtained from (2.13). Now, we can immediately conclude from (2.13) and (SPP2) the following:

Theorem 3.1: If (μ,O_0) is an arbitrary SPP and (X_1,O_1), (X_2,O_2) are arbitrary OP such that $\{O_0,O_1,O_2\}$ is a COS, then

$$\Pr^{(\mu,O_0)}_{(X_1,O_1),(X_2,O_2)}\{X_1(E_1)\}=\Pr^{(\mu,O_0)}_{(X_1,O_1)}\{X_1(E_1)\}, \tag{3.1}$$

for all $E_1\in B(S_1)$.

We shall paraphrase the above result as stating that the probability of observing a particular outcome in the OP (X_1,O_1) *does not depend* on what OP (X_2,O_2) is being carried out in the region O_2 provided $O_1<O_2$. In fact if (X_2',O_2) is some other procedure which can be carried out in the space–time region O_2, then it follows from (3.1) that

$$\Pr^{(\mu,O_0)}_{(X_1,O_1),(X_2,O_2)}\{X_1(E_1)\}=\Pr^{(\mu,O_0)}_{(X_1,O_1),(X_2',O_2)}\{X_1(E_1)\}. \tag{3.2}$$

In fact Eq. (3.2) is equivalent to (3.1), because we can consider (3.1) as the particular case of (3.2) when the observation procedure (X_2',O_2) corresponds to "carrying out no observations to all." Also note that in Eqs. (3.1) or (3.2), (X_2,O_2) could be a composite observation procedure corresponding to a sequence of observation procedures all carried out in subregions of $S^+(O_0\cup O_1)$.

If in addition to the fact that $\{O_0,O_1,O_2\}$ is a COS of space–time regions, we also assume that $O_1\times O_2$, then we have the relations

$$\Pr^{(\mu,O)}_{(X_1,O_1),(X_2,O_2)}\{X_1(E_1)\}=\Pr^{(\mu,O_0)}_{(X_1,O_1)}\{X_1(E_1)\}, \tag{3.3a}$$

$$\Pr^{(\mu,O_0)}_{(X_1,O_1),(X_2,O_2)}\{X_2(E_2)\}=\Pr^{(\mu,O_0)}_{(X_2,O_2)}\{X_2(E_2)\} \tag{3.3b}$$

for all $E_1\in B(S_1)$ and $E_2\in B(S_2)$. *Thus if a local* SPP (μ,O_0) *is followed by two local* OP (X_1,O_1) *and* (X_2,O_2) *which are carried out in mutually spacelike regions, then the probabilities for observing the varius outcomes in any one experiment does not depend on which particular experiment (if any) is being carried out in the other region.* This however does *not* mean that the OP (X_1,O_1) and (X_2,O_2) are *statistically independent* (cf. Ref. 23) for it could very well happen that

$$\Pr^{(\mu,O_0)}_{(X_1,O_1),(X_2,O_2)}\{X_1(E_1),X_2(E_2)\}$$
$$\neq\Pr^{(\mu,O_0)}_{(X_1,O_1),(X_2,O_2)}\{X_1(E_1)\}\Pr^{(\mu,O_0)}_{(X_1,O_1),(X_2,O_2)}\{X_2(E_2)\}. \tag{3.4}$$

In fact, the existence of statistical correlations [of the form given by Eq. (3.4)] between two events occurring in mutually spacelike regions, is typical of all those situations which are usually studied in connection with the so-called "Einstein–Podolsky–Rosen Paradox"[32–34] (see also Refs. 9, 11, 15, 17).

Theorem 3.1 is of course a particular case of the principle of local causes we are trying to establish and has a simple form because it was restricted to the case of only two observation procedures. The appropriate generalization of this result to the general case involving a sequence of OP is not particularly straightforward. We shall first obtain the most general such result which can be derived solely on the basis of the postulates stated in the last section. For this purpose we introduce the following motion of "causal precedence in a sequence." If $\{O_1,O_2,\ldots,O_r\}=S$, is a causally orderable sequence of space–time regions, we shall say that $O_\alpha\in S$ has a *causal precedence* over another element $O_\beta\in S$ *in the sequence* S if either $O_\alpha\not> O_\beta$ or, if there exists an integer $n<r$ and elements $O_{i_1},O_{i_2},\ldots,O_{i_n}\in S$ such that $O_{i_1}\not> O_{i_2},\ldots,O_{i_{n-1}}\not> O_{i_n}$, $O_{i_n}\not> O_\beta$ and $O_\alpha\not> O_{i_1}$. We can easily show that if $O_\alpha\in S$ has causal precedence over $O_\beta\in S$ in a causally orderable sequence S, then in every causally ordered permutation of S, O_α precedes O_β and vice versa. The following theorem can be easily established on the basis of Eq. (2.9), (SPP3) and the basic properties of causally ordered sequences stated in Appendix B:

Theorem 3.2: If (μ,O_0) is an SPP and $\{(X_1,O_1),\ldots,(X_{i_1},O_{i_1}),\ldots,(X_{i_n},O_{i_n}),\ldots,(X_r,O_r)\}$ is a sequence of OP such that $\{O_0,O_1,\ldots,O_r\}$ is a COS, then

$$\Pr^{(\mu,O_0)}_{(X_1,O_1),\ldots,(X_r,O_r)}\{X_{i_1}(E_{i_1}),\ldots,X_{i_n}(E_{i_n})\}$$
$$=\Pr^{\mu,O_0}_{[(X_1,O_1),\ldots,(X_r,O_r)]_{i_1,i_2,\ldots,i_n}}\{X_{i_1}(E_{i_1}),\ldots,X_{i_n}(E_{i_n})\}, \tag{3.5}$$

for all $E_{i_k}\in B(S_{i_k})$ $(k=1,2,\ldots,n)$, where $[(X_1,O_1),\ldots,(X_r,O_r)]_{i_1,i_2,\ldots,i_n}$ is a causally ordered subsequence of $\{(X_1,O_1),\ldots,(X_r,O_r)\}$ which includes apart from the OP $\{(X_{i_k},O_{i_k})|k=1,2,\ldots,n\}$ only those OP $(X_\alpha,O_\alpha)\in\{(X_1,O_1),\ldots,(X_r,O_r)\}$, where O_α has a causal precedence over at least one of the regions $\{O_{i_k}|k=1,2,\ldots,n\}$.

Let us note some of the implications of the above result. First of all, it implies that the probability (3.5) does not depend on all those OP $(X_\alpha,O_\alpha)\in\{(X_1,O_1),\ldots,(X_r,O_r)\}$ which are such that

(i) O_α has a nonnull intersection with the causal future of $\cup_{k=1}^n O_{i_k}$, or

(ii) O_α is spacelike to each of the regions $\{O_{i_k}|k=1,2,\ldots,n\}$ *and* there is a causally ordered permutation of $\{O_1,O_2,\ldots,O_r\}$ in which all the regions $\{O_{i_k}|k=1,2,\ldots,n\}$ precede O_α. However, the above Theorem 3.2 does not *entail* that the probability (3.5) is independent of all those OP $(X_\alpha,O_\alpha)\in\{(X_1,O_1),\ldots,(X_r,O_r)\}$ where (even though O_α is possibly spacelike to all the regions $\{O_{i_k}|k=1,\ldots,n\}$), O_α has a causal precedence over (at least) one of the regions $\{O_{i_k}|k = 1,2,...,n\}$ in the sequence $\{O_1,O_2,...,O_r\}$.

In order to illustrate the above remarks, let us consider for example a COS $\{O_0,O_1,O_2,O_3\}$ such that $O_1 \times O_3$, but $O_1 \not> O_2$ and $O_2 \not> O_3$. In such a case, O_1 has a causal precedence over O_3 in $\{O_1,O_2,O_3\}$. Then, we will *not* be able to conclude from Theorem 3.2 that $\Pr^{(\mu,O_0)}_{(X_1,O_1),(X_2,O_2),(X_3,O_3)}\{X_3(E_3)\}$ does not depend on the OP (X_1,O_1) (as suggested by our heuristic ideas of causality). However, we should also emphasize *that the Theorem* 3.2 *in no way implies that the above probability does in fact depend on the* OP (X_1,O_1) *carried out in* O_1. It simply has nothing to offer in such cases.

We shall now indicate how the result of Theorem 3.2 can be strengthened into the result that the probability (3.5) does not depend on all those OP $(X_\alpha,O_\alpha)\in\{(X_1,O_1),...,(X_r,O_r)\}$ whenever $O_\alpha \subset S^+(\cup_{k=1}^n O_{i_k})$ or equivalently, whenever O_α has a null intersection with the causal pasts of all the regions $\{O_{i_k}|k = 1,2,...,n\}$. It appears that such a result cannot be derived merely on the basis of the framework sketched in Sec. 2. The simplest way of establishing such a general princple of local causes seems to involve the following postulate (E6), [(D10)] which asserts a certain "decomposability of the local OP."

(E6) If the space–time region O is such that $O = O_1 \cup O_2$, where $O_1 < O_2$, then given any local OP (X,O), there exist two local OP $(X_1,O_1),(X_2,O_2)$ such that

$$(X,O) = (X_1,O_1)\circ(X_2,O_2).$$

Based on the above assumption (in addition to the postulates of Sec. 2), we can now show the following (statistical) "principle of local causes" in relativistic quantum theory:

Theorem 3.3 (*principle of local causes*): If (μ,O_0) is a SPP and $\{(X_1,O_1),...,(X_{i_1},O_{i_1}),...,(X_{i_n},O_{i_n}),...,(X_r,O_r)\}$ is a sequence of OP such that $\{O_0,O_1,...,O_r\}$ is a COS, then

$$\Pr^{(\mu,O_0)}_{(X_1,O_1),...,(X_r,O_r)}\{X_{i_1}(E_{i_1}),...,X_{i_n}(E_{i_n})\} = \Pr^{(\mu,O_0)}_{[(X_1,O_1),...,(X_r,O_r)]_{\cup_k^n O_{i_k}}}\{X_{i_1}(E_{i_1}),...,X_{i_n}(E_{i_n})\}, \tag{3.6}$$

for all $E_{i_k}\in B(S_{i_k})$ $(k = 1,2,...,n)$, where $[(X_1,O_1),...,(X_r,O_r)]_{\cup_k^n O_{i_k}}$ is the causally ordered subsequence of $\{(X_1,O_1),...,(X_r,O_r)\}$ which is obtained by excluding from the above sequence all those elements $\{(X_\alpha,O_\alpha)\}$ where

$$O_\alpha \subset S^+\left(\bigcup_{k=1}^n O_{i_k}\right).$$

We shall just sketch the proof of the above theorem. Let $\{O_{j_1},O_{j_2},...,O_{j_m}\}$ be the set of all those elements of $\{O_1,...,O_r\}\setminus\{O_{i_1},...,O_{i_n}\}$ which have a nonnull intersection. With the causal past

$$J^-\left(\bigcup_{k=1}^n O_{i_k}\right) \text{ of } \bigcup_{k=1}^n O_{i_k}.$$

If there are no such elements, then Theorem 3.3 follows directly from Theorem 3.2. Now, we define the regions $\bar{O}_{j_\gamma},\bar{\bar{O}}_{j_\gamma}$ by the equations

$$\bar{O}_{j_\gamma} = O_{j_\gamma} \cap J^-\left(\bigcup_{k=1}^n O_{i_k}\right), \tag{3.7a}$$

$$\bar{\bar{O}}_{j_\gamma} = O_{j_\gamma} \setminus \bar{O}_{j_\gamma}, \tag{3.7b}$$

for all $\gamma = 1,...,m$. It can easily be shown that $\bar{O}_{j_\gamma} < \bar{\bar{O}}_{j_\gamma}$ and as we have $O_{j_\gamma} = \bar{O}_{j_\gamma} \cup \bar{\bar{O}}_{j_\gamma}$, we can conclude from (E6) that each of the OP $(X_{j_\gamma},O_{j_\gamma})$ can be considered as the succession of two OP $(\bar{X}_{j_\gamma},\bar{O}_{j_\gamma})$, $(\bar{\bar{X}}_{j_\gamma},\bar{\bar{O}}_{j_\gamma})$ for each $\gamma = 1,2,...,m$. We now replace each of the OP $(X_{j_\gamma},O_{j_\gamma})$ by the corresponding pair $(\bar{X}_{j_\gamma},\bar{O}_{j_\gamma}),(\bar{\bar{X}}_{j_\gamma},\bar{\bar{O}}_{j_\gamma})$ in the sequence $S = \{(X_1,O_1),...,(X_r,O_r)\}$ and obtain a new COS of OP S'. Then the joint probability (3.6) for the sequence of OP S can now be expressed in terms of the joint probabilities for the sequence of OP S'. If $O_\alpha \in \{O_1,O_2,...,O_r\}$ is such that $O_\alpha \subset S^+(\cup_{k=1}^n O_{i_k})$, then O_α is distinct from all the $\{O_{j_\gamma}\}$ $(\gamma = 1,2,...,m)$ and it can be easily seen that *in the sequence* S', O_α does not have a causal precedence over *any* of the regions $\{O_{i_k}|k = 1,2,...,n\}$. Therefore, we can now apply Theorem 3.2 and conclude that the probability (3.6) does not depend on any of the OP $(X_\alpha,O_\alpha)\in S$, for which $O_\alpha \subset S^+(\cup_{k=1}^n O_{i_k})$. Hence the theorem is proven.

It may be of interest to note that if we restrict the class of all OP considered to those that can be performed at space–time points only, then the Theorem 3.3 will be a simple corollary of our Theorem 3.2. This is of course quite understandable as the additional postulate (E6), that was needed in the proof of Theorem 3.3, does not yield anything new as regards OP that are carried out at space–time points. For the class of OP considered in the conventional formulation of quantum theory [cf. Eq. (1.13)], Schlieder[9] has in fact demonstrated a result equivalent to our principle of local causes, by assuming in addition that all OP are being carried out at space–time points only.

Finally, we may mention that Stapp[35–38] has recently proved a "generalized Bell's theorem" according to which the statistical predictions of quantum theory are in conflict with the princple of local causes (in its simplest form as stated in Theorem 3.1) provided one also assumes that a certain "principle of contrafactual definiteness" is satisfied. We shall not go into a discussion of this result here, as we propose to show elsewhere that it is the "principle of contrafactual definiteness" which is in direct conflict with the statistical predictions of quantum theory. In any case, we have clearly demonstrated above that relativistic quantum theory can be formulated in such a way that even the general (statistical) principle of local causes (as stated in the Theorem 3.3) is satisfied.

4. THE COLLAPSE POSTULATE IN RELATIVISTIC QUANTUM THEORY

In this section we shall demonstrate that the fundamental statistical law of relativistic quantum theory as stated in Sec. 2 admits also an equivalent formulation in terms of a relativistic version of the collapse postulate.

We may recall that in Sec. 1 it was observed that in the conventional formulation of nonrelativistic quantum theory the fundamental statistical prescription of the theory can either be given in terms of the Wigner formula (1.5) or *equivalently* in terms of the Born statistical formula and the collapse postulate. However, it should be noted that the above assertion of equivalence is based on several implicit assumptions on the nature (and measurement) of quantum mechanical probabilities. Moreover, for a clear understanding of the collapse postulate, it is very essential that the notion of a state preparation procedure be made more precise. It is for these reasons that we shall preface our considerations on the collapse postulate in relativistic quantum theory by some remarks on the empirical content of the fundamental statistical law of relativistic quantum theory as expressed by the relativistic Wigner formula (2.13).

We shall adopt the point of view that the joint probability

$$\Pr^{(\mu,O_0)}_{(X_1,O_1),\ldots,(X_r,O_r)}\{X_{i_1}(E_{i_1}),\ldots,X_{i_n}(E_{i_n})\}$$

given by Eq. (2.13), is to be considered as the "*limit of the statistical frequency*" corresponding to the number of systems for which the outcomes of the OP $\{X_{i_k},O_{i_k}\}$ are found to lie in $\{E_{i_k}\}$ $(k = 1,2,\ldots,n)$ when an "*ensemble*" of systems "*prepared according to the* SPP (μ,O_0)" is subjected to the sequence of OP $\{(X_1,O_1),\ldots,(X_{i_1},O_{i_1}),\ldots,(X_{i_n},O_{i_n}),\ldots,(X_r,O_r)\}$. In other words, the joint probabilities (2.13) (which are the only empirical predictions of the theory) are to be treated as relative frequencies in an ensemble. Since the association of probabilities to various events was achieved in our theory by means of the notion of an SPP (cf. Sec. 2), *it follows that each* SPP *is essentially a mathematical characterization of a certain procedure (or an equivalence class of procedures) for preparing ensembles.*

An "*ensemble*" is to be understood as a (conceptually infinite) collection of "noninteracting" systems which can be well ordered into a sequence (see in this connection[39]). An "*ensemble preparation procedure* (EPP)" is a procedure (somewhat like a set of "instructions") for preparing ensembles—each of which is said to be "prepared according to the given EPP." If Ω is an ensemble prepared according to some "local EPP carried out in a space–time region O_0" then we shall denote by

$$n^{\Omega,O_0;N}_{(X_1,O_1),\ldots,(X_r,O_r)}\{X_{i_1}(E_{i_1}),\ldots,X_{i_n}(E_{i_n})\} \tag{4.1}$$

the number of systems for which the outcomes of the OP $\{(X_{i_k},O_{i_k})\}$ were found to lie in $\{E_{i_k}\}$ $(k = 1,2,\ldots,n)$, when the first N systems of the ensemble Ω were subjected to the sequence of OP $\{(X_1,O_1),\ldots,(X_{i_1},O_{i_1}),\ldots,(X_{i_n},O_{i_n}),\ldots,(X_r,O_r)\}$, where of course $\{O_0,O_1,\ldots,O_r\}$ is a COS.

The above description is of course based on the "fiction" that the ensemble Ω is a stored collection of noninteracting systems all prepared in the space–time region O_0. Strictly speaking a "physical ensemble" which is prepared according to some EPP "carried out in O_0," actually consists of systems prepared over various space–time regions—which may be for example space–time translates of or, more generally, relativistic transforms of the region O_0. Therefore in order to provide satisfactory answers to the questions

(i) what is actually meant by an "EPP carried out in a space–time region O_0?" and

(ii) what is the empirical meaning of Eq. (4.1)?
we will have to go into an analysis of the "empirical content" of the various invariance principles satisfied by the theory. This will not be undertaken here as it is quite beyond the purview of the present investigation.

It was already noted that the EPP are to some extent the operational counterparts of the mathematical objects of our theory the SPP. This connection can be made more explicit (in the context of relativistic quantum theory) as follows: An EPP carried out in O_0 is said to be a "*valid* EPP" if there exists a SPP (μ,O_0) such that for any ensemble Ω prepared according to the given EPP, we have

$$\lim_{N\to\infty} n^{\Omega,O_0;N}_{(X_1,O_1)}\{X_1(E_1)\}/N = \mu(X_1(E_1)), \tag{4.2}$$

for all $E_1 \in B(S_1)$, for all OP (X,O_1) such that $O_1 > O_0$. Such an SPP (μ,O_0) (if it exists) is unique, and we shall say that the given EPP is "*associated with the* SPP (μ,O_0)." Note that it could very well be the case that several different EPP are associated with the same SPP. However, as far as the observational predictions of quantum theory are concerned, all such EPP are equivalent—and even indistinguishable except by the specific description of the corresponding mode of preparation.

We shall hereafter say that "Ω_μ *is an ensemble of systems prepared according to the* SPP (μ,O_0)," if Ω_μ is an ensemble prepared according to some valid EPP associated with the SPP (μ,O_0). We shall now state the fundamental statistical law of relativistic quantum theory (2.13) in the following form which employs the notion of "relative frequencies in an ensemble," instead of "probabilities."

If Ω_μ is any ensemble of systems prepared according to the SPP (μ,O_0) *and if* $\{(X_1,O_1),\ldots,(X_{i_1},O_{i_1}),\ldots,(X_{i_n},O_{i_n}),\ldots,(X_r,O_r)\}$ *is a sequence of* OP *such that* $\{O_0,O_1,\ldots,O_r\}$ *is a* COS *of space–time regions, then*

$$\lim_{N\to\infty} n^{\Omega_\mu,O_0;N}_{(X_1,O_1),\ldots,(X_r,O_r)}\{X_{i_1}(E_{i_1}),\ldots,X_{i_n}(E_{i_n})\}/N$$
$$= \mu(\xi_1\wedge\xi_2\wedge\cdots\wedge X_{i_1}(E_{i_1})\wedge\cdots\wedge X_{i_n}(E_{i_n})\wedge\cdots\wedge\xi_r), \tag{4.3}$$

for all $E_{i_k}\in B(S_{i_k})$ $(k = 1,2,\ldots,n)$ where we have adopted the notation of Eqs. (2.13) *and* (4.1).

For our later considerations it will be useful to introduce the following notion of a "*composite* SPP" whch consists in carrying out a given SPP followed by a selection corresponding to the occurrence of some event. Let (μ,O_0) be a SPP and $A\in\mathscr{O}(O_1)$ such that $O_0 < O_1$ and $\mu(A)\neq O$. Then we shall denote by

$$(\mu,O_0)^*(A,O_1) = (A^*\mu,O_0\cup O_1) \tag{4.4}$$

the SPP carried out in $O_0\cup O_1$, where $A^*\mu$ is the linear functional defined by

$$(A^*\mu)(B) = \mu(A\wedge B)/\mu(A), \tag{4.5}$$

for all $B\in\mathscr{O}(O_2)$ for all $O_2\subset S^+(O_0\cup O_1)$. It can be easily verified that $(A^*\mu,O_0\cup O_1)$ as defined above satisfies all our conditions (SPP1)–(SPP3), and therefore defines a SPP carried out in the space–time region $O_0\cup O_1$. We can also easily verify that if $\{O_0,O_1,\ldots,O_r\}$ is a COS, then

$$(\cdots((\mu,O_0)^*(A_1,O_1))^*\cdots^*(A_r,O_r))$$

$$= (A_r{}^*\cdots{}^*A_1{}^*\mu, O_0\cup O_1\cup\cdots\cup O_r), \tag{4.6}$$

where the left-hand side is well-defined iff the right-hand side is well defined. If (μ,O_0) is a SPP, (X_1,O_1) is a OP such that $O_0 < O_1$ and $E_1\in B(S_1)$ such that $\mu(X_1(E_1))\neq 0$, then we shall refer to $(\mu,O_0)^*(X_1(E_1),O_1)$ as the "composite SPP (μ,O_0) followed by a selection corresponding to the ocurrence of the event $X_1(E_1)$ in O_1."

We shall now state the main result of this section:

Theorem 4.1: Under the asumptions of Sec. 2 [excepting Eq. (2.13)], the fundamental statistical law of relativistic quantum theory as given by Eq. (4.3) is equivalent to the following two "postulates" BSF and RCP, taken together:

(1) *Born statistical formula* (BSF): If Ω_μ is any ensemble prepared according to a SPP (μ,O_0) and if $\{(X_1,O_1),\ldots,(X_r,O_r)\}$ is an arbitrary sequence of OP such that $\{O_0,O_1,\ldots,O_r\}$ is a COS, then

$$\lim_{N\to\infty} n^{\Omega_\mu,O_0;N}_{(X_1,O_1),\ldots,(X_r,O_r)}\{X_1(E_1)\}/N = \mu(X_1(E_1)), \tag{4.7}$$

for all $E_1\in B(S_1)$.

(2) *Relativistic collapse postulate* (RCP): Let (μ,O_0) be an SPP and (X_1,O_1) an OP such that $O_0 < O_1$ and let $E_1\in B(S_1)$. Now, if $\mu(X_1(E_1))\neq 0$, then the procedure "prepare an ensemble according to the SPP (μ,O_0), subject this ensemble to the OP (X_1,O_1) and select all those systems for which the outcome is observed to lie in E_1" is a valid EPP carried out in the region $O_0\cup O_1$ and is associated with the SPP $(\mu,O_0)^*(X_1(E_1),O_1)$. If $\mu(X_1(E_1)) = 0$, then the above procedure does not prepare any ensemble at all.

Proof: Let us first show that Eq. (4.3) implies both BSF and RCP. The proof of BSF is immediate as Eq. (4.7) turns out to be a particular case of Eq. (4.3) once (SPP2) is taken into account. In order to show that RCP also follows from Eq. (4.3), let us consider an ensemble Ω' prepared according to the EPP stated in RCP in the case when $\mu(X_1(E_1))\neq 0$. Then there exists an ensemble Ω_μ prepared according to the SPP (μ,O_0) such that Ω' is the ensemble of all those systems which are obtained when Ω_μ is subjected to the OP(X_1,O_1) and those systems are selected for which the outcome was observed to lie in E_1. Let (X_2,O_2) be an arbitrary OP and where $O_2\subset S^+(O_0\cup O_1)$ and let us also write

$$N_1 = n^{\Omega_\mu,O_0;N}_{(X_1,O_1),(X_2,O_2)}\{X_1(E_1)\}, \tag{4.8a}$$

$$N_2 = n^{\Omega_\mu,O_0;N}_{(X_1,O_1),(X_2,O_2)}\{X_1(E_1), X_2(E_2)\}. \tag{4.8b}$$

Now from the definition of the ensemble Ω' it follows that

$$N_2 = n^{\Omega',O_0\cup O_1;N_1}_{(X_2,O_2)}\{X_2(E_2)\}. \tag{4.9}$$

Hence, we obtain

$$n^{\Omega',O_0\cup O_1;N_1}_{(X_2,O_2)}\{X_2(E_2)\}/N_1 = \frac{(N_2/N)}{(N_1/N)}. \tag{4.10}$$

Now if $N\to\infty$, then $N_1\to\infty$ also and vice versa as is evident from Eq. (4.7) and the fact that $\mu(X_1(E_1))\neq 0$. Hence we obtain from Eqs. (4.3) and (4.8)–(4.10) the following relation:

$$\lim_{N\to\infty} n^{\Omega',O_0\cup O_1;N_1}_{(X_2,O_2)}\{X_2(E_2)\}/N_1$$

$$= \mu(X_1(E_1)\wedge X_2(E_2))/\mu(X_1(E_1)), \tag{4.11}$$

for all $E_2\in B(S_2)$. Since Eq. (4.11) holds for all OP (X_2,O_2) where $O_2\subset S^+(O_0\cup O_1)$, it follows from Eq. (4.2) that Ω' is an ensemble prepared according to a valid EPP associated with the SPP $(\mu,O_0)^*(X_1(E_1),O_1)$ as stated in RCP. Finally when $\mu(X_1(E_1)) = 0$, it follows from Eq. (4.7) that $\lim_{N\to\infty}(N_1/N)\to 0$, and therefore it is evident that the procedure stated in RCP does not prepare any ensemble whatsoever in such a situation. Thus we have shown that RCP follows from the fundamental statistical law of the theory as given by Eq. (4.3). The proof of the reverse implication can be constructed without any difficulty by a repeated application of RCP, once we take $N\to\infty$ limit in the following equation:

$$n^{\Omega,O_0;N}_{(X_1,O_1),\ldots,(X_r,O_r)}\{X_{i_1}(E_{i_1}),\ldots,X_{i_k}(E_{i_k})\} = \frac{N_1}{N}\prod_{p=1}^{r-1}\frac{N_{p+1}}{N_p}, \tag{4.12}$$

where we have introduced the notation

$$N_p = n^{\Omega,O_0;N}_{(X_1,O_1),\ldots,(X_r,O_r)}\{X_1(E_1),\ldots,X_p(E_p)\}$$

and $E_\alpha = S_\alpha$ [the value space of (X_α,O_α)] for α different from any of the indices $\{i_1,\ldots,i_k\}$. This completes the proof of Theorem 4.1.

The most important feature of the relativistic version of the collapse postulate (RCP) established above, is of course that it has been formulated in terms of the relativistically covariant notion of local SPP. Also the above theorem clearly brings out the dual role played by the collapse principle in quantum theory. First (from a pragmatic point of view) it can be included as an assumption (or algorithm) of the thoery along with the Born statistical formula, in order to derive the fundamental statistical prescription of the theory for the probability connections between the results of successive observations. Secondly, RCP is also an important proposition of the theory which establishes a certain correspondence of certain composite EPP with the SPP of the theory—for, according to RCP if it is given that a certain EPP is associated with an SPP (μ,O_0), then RCP asserts that another EPP, which involves the carrying out of the former EPP followed by a process of selection of all those systems for which a particular outcome is observed in a certain OP, is to be associated with an SPP related to (μ,O_0). Therefore, it would indeed be much more appropriate to refer to the "collapse postulate" (RCP) as a *"principle of ensemble preparation by selection."*

In conclusion we may also note that the above formulation of the collapse postulate as a "principle of ensemble preparation by selection," removes in a certain sense some of the mystery that has surrounded the notion of collapse. In fact such a principle can be shown to be valid also in classical probability theory, where again an analog of our Theorem 4.1 can be proved (provided, of course that similar assumptions are made concerning the nature and measurement of probabilities). The crucial difference in this respect between classical theory and quantum theory is that *in quantum theory the* EPP *"prepare an ensemble according to the* SPP (μ,O_0), *subject this ensemble to the* OP (X_1,O_1) *(where*

$O_1 > O_0$) *and select all the systems irrespective of the outcome," will be associated with the* SPP $(\mu,O_0)^*(X_1(S_1),O_1)$ *which is in general different from* $(\mu,O_0 \cup O_1)$. This nonclassical feature of the quantum collapse principle is just another manifestation of the basic nonclassical feature of the quantum theoretical probabilities which was referred to as the quantum interference of probabilities.

APPENDIX A: FUNDAMENTAL REQUIREMENTS OF RELATIVISTIC INVARIANCE

In this appendix we shall discuss some of the fundamental requirements for a theory to be relativistically invariant and indicate how they naturally lead to some of the salient features of the framework of relativistic quantum theory outlined in Sec. 2.

The fundamental requirements of relativistic invariance postulate certain relations between the descriptions of physical phenomena and the formulations of the laws of nature as carried out in the languages adopted by different equivalent observers. In the present investigation, two observers $\mathscr{R}$ and $\overline{\mathscr{R}}$ are considered to be equivalent if the state of motion of $\overline{\mathscr{R}}$ is related to that of $\mathscr{R}$ via a relativistic transformation $g \in P^{\uparrow}_{+}$ the group of all restricted inhomogeneous Lorentz transformations; we shall then write

$$\overline{\mathscr{R}} = g\mathscr{R}. \tag{A1}$$

In quantum theory, the description of physical phenomena in the language of an observer $\mathscr{R}$ involves (i) descriptions $\{(X,O)\}$ of observation procedures (OP) carried out in space–time regions $\{O\}$, and (ii) descriptions $\{(\mu,O)\}$ of state preparation procedures (SPP) carried out in space–time regions $\{O\}$ [see (D11)]. The laws of nature are then formulated so as to provide statistical correlations between the outcomes of a sequence of OP performed on an ensemble of systems prepared according to some SPP.

In order to state the fundamental requirements of relativistic invariance, we need the following notions of "bodily identity" and "subjective identity".[40] We shall say that the two descriptions (X,O) and $(\overline{X},\overline{O})$ of OP in the languages of $\mathscr{R}$ and $\overline{\mathscr{R}}$ are *bodily identical*, if $(\overline{X},\overline{O})$ is the description in the language of $\overline{\mathscr{R}}$ of objectively the same OP which is described as (X,O) in the language of $\mathscr{R}$. We shall say that the two descriptions $(\overline{X},\overline{O})$, and (X,O) of OP are *subjectively identical* if the OP described as $(\overline{X},\overline{O})$ by $\overline{\mathscr{R}}$ bears the same relation to $\overline{\mathscr{R}}$ as the OP described as (X,O) (in the language of $\mathscr{R}$) does to $\mathscr{R}$. We can similarly define the notions of bodily and subjective identity of the two descriptions (μ,O) and $\overline{\mu},\overline{O})$ of SPP in the languages of $\mathscr{R}$ and $\overline{\mathscr{R}}$.

The fundamental requirements of relativistic invariance may now be stated as follows [(D12)]:

I. It is possible to translate the description of physical phenomena from the language of any given observer into that of any other equivalent observer. In such a translation, the space–time regions in which the various SPP and OP are carried out, are also transformed as per the associated relativistic transformation.

II. For every SPP and every sequence of OP which can be carried out by one observer, there exist subjectively identical SPP and sequence of OP whch are possible for an equivalent observer; equivalent observers "see the same world."

III. For any four equivalent observers $\mathscr{R}_1$, $\mathscr{R}_2$, $\mathscr{R}_3$, and $\mathscr{R}_4$, if $\mathscr{R}_3$ is related to $\mathscr{R}_1$ by the same relativistic transformation as $\mathscr{R}_4$ is to $\mathscr{R}_2$ then the OP (SPP) which are subjectively identical for $\mathscr{R}_1$ and $\mathscr{R}_2$ are also subjectively identical for $\mathscr{R}_3$ and $\mathscr{R}_4$—or, more specifically

(a) if the two descriptions $(X_1,O_1),(X_2,O_2)$ of OP in the languages of $\mathscr{R}_1$ and $\mathscr{R}_2$ (respectively) are subjectively identical, then the OP (X_3,O_3) in the language of $\mathscr{R}_3$ which is bodily identical to (X_1,O_1), is subjectively identical to the OP (X_4,O_4) in the language of $\mathscr{R}_4$ which is bodily identical to (X_2,O_2), and

(b) a similar condition on SPP.

IV. The laws of nature can be expressed in the language of every equivalent observer in such a way that

(a) any two equivalent observers predict the same statistical correlations between the outcomes of bodily the same sequence of OP which follow the performance of bodily the same SPP, and

(b) any two equivalent observers predict the same statistical correlations between the outcomes of two sequences of OP which are subjectively identical, when they are performed after subjectively identical SPP are carried out.

We shall first consider some of the general conclusions that can be drawn from these requirements. From requirement I it follows that given any two equivalent observers $\mathscr{R}$ and $\overline{\mathscr{R}}$ there exist bijections $B_{\overline{\mathscr{R}}\leftarrow\mathscr{R}}$ and $\widetilde{B}_{\overline{\mathscr{R}}\leftarrow\mathscr{R}}$ which, respectively, map descriptions of OP and SPP from the language of observer $\mathscr{R}$ into the descriptions of bodily identical OP and SPP in the language of $\overline{\mathscr{R}}$; i.e., if $\overline{\mathscr{R}} = g\mathscr{R}$ and (X,O) is an OP and (μ,O) is an SPP in the language of $\mathscr{R}$, then

$$(\overline{X}^B,\overline{O}) = B_{\overline{\mathscr{R}}\leftarrow\mathscr{R}}(X,O) \tag{A2a}$$

$$(\overline{\mu}^B,\overline{O}) = \widetilde{B}_{\overline{\mathscr{R}}\leftarrow\mathscr{R}}(\mu,O) \tag{A2b}$$

are the bodily identical OP and SPP in the language of $\overline{\mathscr{R}}$, where

$$\overline{O} = gO = \{x|g^{-1}x\in O\} \tag{A3}$$

is the corresponding relativistically transformed space–time region. In the same way, it follows from II that there exist bijections $S_{\overline{\mathscr{R}}\leftarrow\mathscr{R}}$ and $\widetilde{S}_{\overline{\mathscr{R}}\leftarrow\mathscr{R}}$ such that the equations

$$(\overline{X}^S,O) = S_{\overline{\mathscr{R}}\leftarrow\mathscr{R}}(X,O) \tag{A4a}$$

$$(\overline{\mu}^S,0) = \widetilde{S}_{\overline{\mathscr{R}}\leftarrow\mathscr{R}}(\mu,O) \tag{A4b}$$

relate subjectively identical procedures. From the definition of subjectively identical procedures, it follows that all equivalent observers employ identical descriptions for the space–time regions in which subjectively identical procedures are carried out.

The mappings $B_{\overline{\mathscr{R}}\leftarrow\mathscr{R}}$ and $\widetilde{B}_{\overline{\mathscr{R}}\leftarrow\mathscr{R}}$ defined above satisfy the following relations:

$$B_{\mathscr{R}_3\leftarrow\mathscr{R}_1} = B_{\mathscr{R}_3\leftarrow\mathscr{R}_2}B_{\mathscr{R}_2\leftarrow\mathscr{R}_1}, \tag{A5a}$$

$$\widetilde{B}_{\mathscr{R}_3\leftarrow\mathscr{R}_1} = \widetilde{B}_{\mathscr{R}_3\leftarrow\mathscr{R}_2}\widetilde{B}_{\mathscr{R}_2\leftarrow\mathscr{R}_1} \tag{A5b}$$

for any three equivalent observers $\mathscr{R}_1$, $\mathscr{R}_2$, and $\mathscr{R}_3$. Also $B_{\mathscr{R}\leftarrow\mathscr{R}}$ are identity transformations on the appropriate spaces. Similar relations are also satisfied by the mappings $S_{\overline{\mathscr{R}}\leftarrow\mathscr{R}}$ and $\widetilde{S}_{\overline{\mathscr{R}}\leftarrow\mathscr{R}}$.

The requirement III can now be expressed in terms of the above mappings as follows:

$$S_{g\mathscr{R}_2\leftarrow g\mathscr{R}_1} = B_{g\mathscr{R}_2\leftarrow\mathscr{R}_2} S_{\mathscr{R}_2\leftarrow\mathscr{R}_1} B_{\mathscr{R}_1\leftarrow g\mathscr{R}_1}, \tag{A6a}$$

$$\widetilde{S}_{g\mathscr{R}_2\leftarrow g\mathscr{R}_1} = \widetilde{B}_{g\mathscr{R}_2\leftarrow\mathscr{R}_2} \widetilde{S}_{\mathscr{R}_2\leftarrow\mathscr{R}_1} \widetilde{B}_{\mathscr{R}_1\leftarrow g\mathscr{R}_1} \tag{A6b}$$

for all $g\in P^{\uparrow}_{+}$ and any two equivalent observers $\mathscr{R}_1$, $\mathscr{R}_2$. Now, from the properties of the mapping B, $\widetilde{B}$, S, and $\widetilde{S}$, it follows that Eqs. (A6a) and (A6b) are equivalent to the equations

$$B_{g\mathscr{R}_2\leftarrow\mathscr{R}_2} = S_{g\mathscr{R}_2\leftarrow g\mathscr{R}_1} B_{g\mathscr{R}_1\leftarrow\mathscr{R}_1} S_{\mathscr{R}_1\leftarrow\mathscr{R}_2}, \tag{A7a}$$

$$\widetilde{B}_{g\mathscr{R}_2\leftarrow\mathscr{R}_2} = \widetilde{S}_{g\mathscr{R}_2\leftarrow g\mathscr{R}_1} \widetilde{B}_{g\mathscr{R}_1\leftarrow\mathscr{R}_1} \widetilde{S}_{\mathscr{R}_1\leftarrow\mathscr{R}_2}. \tag{A7b}$$

Therefore the requirement III can be equivalently expressed as follows:

III′. For any four equivalent observers $\mathscr{R}_1$, $\mathscr{R}_2$, $\mathscr{R}_3$, and $\mathscr{R}_4$, if $\mathscr{R}_3$ is related to $\mathscr{R}_1$ by the same relativistic transformation as $\mathscr{R}_4$ is to $\mathscr{R}_2$, then

(a) if the OP (X_1,O_1) (in the language of $\mathscr{R}_1$) is bodily identical to the OP (X_3,O_3) (in the language of $\mathscr{R}_3$), then the OP (X_2,O_2) (in the language of $\mathscr{R}_2$) which is subjectively identical to (X_1,O_1) will be bodily identical to the OP (X_4,O_4) (in the language of $\mathscr{R}_4$) which is subjectively identical to (X_2,O_2), and

(b) a similar condition on SPP.

We now define the mappings $\beta_{g,\mathscr{R}}$ and $\widetilde{\beta}_{g,\mathscr{R}}$ by the equation

$$\beta_{g,\mathscr{R}} = S_{\mathscr{R}\leftarrow g\mathscr{R}} B_{g\mathscr{R}\leftarrow\mathscr{R}}, \tag{A8a}$$

$$\widetilde{\beta}_{g,\mathscr{R}} = \widetilde{S}_{\mathscr{R}\leftarrow g\mathscr{R}} \widetilde{B}_{g\mathscr{R}\leftarrow\mathscr{R}}. \tag{A8b}$$

These mappings transform descriptions (of OP and SPP) in the language of $\mathscr{R}$ into descriptions (of some other OP and SPP) in the *same* language. From Eqs. (A5)–(A8) it follows that

$$\beta_{g_1,\mathscr{R}}\beta_{g_2,\mathscr{R}} = \beta_{g_1g_2,\mathscr{R}}, \tag{A9a}$$

$$\widetilde{\beta}_{g_1,\mathscr{R}}\widetilde{\beta}_{g_2,\mathscr{R}} = \widetilde{\beta}_{g_1g_2,\mathscr{R}} \tag{A9b}$$

for all $g_1,g_2\in P^{\uparrow}_{+}$. Also $\beta_{e,\mathscr{R}}$ and $\widetilde{\beta}_{e,\mathscr{R}}$ reduce to identity maps where e is the identity element in $P^{\uparrow}_{+}$. Therefore we have the following important result:

Theorem A 1 (*Wightman*[40]): $g\rightarrow\beta_{g,\mathscr{R}}$ and $g\rightarrow\widetilde{\beta}_{g,\mathscr{R}}$ are representations of the relativity group $P^{\uparrow}_{+}$.

Now, since Theorem A1 implies that $(\beta_{g,\mathscr{R}})^{-1} = \beta_{g^{-1},\mathscr{R}}$ [and similarly for $(\widetilde{\beta}_{g,\mathscr{R}})^{-1}$], we are led to the relations

$$\beta_{g,\mathscr{R}} = B_{\mathscr{R}\leftarrow g^{-1}\mathscr{R}} S_{g^{-1}\mathscr{R}\leftarrow\mathscr{R}}, \tag{A10a}$$

$$\widetilde{\beta}_{g,\mathscr{R}} = \widetilde{\beta}_{\mathscr{R}\leftarrow g^{-1}\mathscr{R}} \widetilde{S}_{g^{-1}\mathscr{R}\leftarrow\mathscr{R}}. \tag{A10b}$$

for each OP (X,O) and SPP (μ,O) in the language of $\mathscr{R}$, let us write

$$\beta_{g,\mathscr{R}}(X,O) = (\beta_g X, gO), \tag{A11a}$$

$$\widetilde{\beta}_{g,\mathscr{R}}(\mu,O) = (\widetilde{\beta}_g \mu, gO). \tag{A11b}$$

From the Eqs. (A8) and (A10) it follows that the transformation $\beta_{g,\mathscr{R}}(\widetilde{\beta}_{g,\mathscr{R}})$ has the following two interpretations: $\beta_{g,\mathscr{R}}(\widetilde{\beta}_{g,\mathscr{R}})$ *transforms each description* (X,O) $((\mu,O))$ *of an* OP (SPP) *in the language of* $\mathscr{R}$ *to the description* $(\beta_g X, gO)$ $((\beta_g \mu, gO))$, *in the same language, of that* OP (SPP) *which*

(i) *bears the same relation to the observer* $\mathscr{R}$ *as the* OP (SPP), *which has the description* (X,O) $((\mu,O))$ *in the language of* $\mathscr{R}$, *does to the observer* $g\mathscr{R}$.

(ii) *bears the same relation to the observer* $g^{-1}\mathscr{R}$ *as the* OP (SPP), *which has the description* (X,O) $((\mu,O))$ *in the language of* $\mathscr{R}$, *does to the observer* $\mathscr{R}$.

There is also the mapping $\gamma_{g,\mathscr{R}}$ (and similarly $\widetilde{\gamma}_{g,\mathscr{R}}$) defined by the equation

$$\gamma_{g,\mathscr{R}} = B_{\mathscr{R}\leftarrow g\mathscr{R}} S_{g\mathscr{R}\leftarrow\mathscr{R}}, \tag{A12}$$

which maps each OP (X,O) (in the language of $\mathscr{R}$) into the description of another OP (in the same language) which bears the same relation to the observer $g\mathscr{R}$ as the OP (X,O) does to the observer $\mathscr{R}$. However, it is important to realize that even though $\gamma_{g,\mathscr{R}}$ (together with $\widetilde{\gamma}_{g,\mathscr{R}}$) constitutes a symmetry transformation of the theory [in the sense that an analog of Eq. (A18) below is satisfied], the correspondence $g\rightarrow\gamma_{g,\mathscr{R}}$ does *not* constitute a representation of the relativity group $P^{\uparrow}_{+}$. In fact, from Eq. (A12) it follows that

$$\gamma_{g,\mathscr{R}} = \beta_{g^{-1},\mathscr{R}}, \tag{A13}$$

and therefore we have

$$\gamma_{g_1,\mathscr{R}}\gamma_{g_2,\mathscr{R}} = \gamma_{g_2g_1,\mathscr{R}}. \tag{A14}$$

We now consider the following question: "To what extent is the action of the transformation $\beta_{g,\mathscr{R}}(\widetilde{\beta}_{g,\mathscr{R}})$ which acts on the description of OP (SPP) in the language of $\mathscr{R}$, similar to that of the transformation $\beta_{g,\overline{\mathscr{R}}}(\widetilde{\beta}_{g,\overline{\mathscr{R}}})$ which acts on the descriptions of OP (SPP) in the language of an equivalent observer $\overline{\mathscr{R}}$? First of all it should be noted that if the OP (X,O) (in the language of $\mathscr{R}$) is bodily identical to the OP $(\overline{X},\overline{O})$ in the language of $\overline{\mathscr{R}}$, it does not follow that in general the OP $\beta_{g,\mathscr{R}}(X,O)$ and $\beta_{g,\overline{\mathscr{R}}}(\overline{X},\overline{O})$ are also bodily identical. In fact, if $\overline{\mathscr{R}} = \overline{g}\mathscr{R}$ and it is given that the OP (X,O) and $(\overline{X},\overline{O})$ are bodily identical, then it follows that the OP $\beta_{g,\mathscr{R}}(X,O)$ and $\beta_{\overline{g}g\overline{g}^{-1},\overline{\mathscr{R}}}(\overline{X},\overline{O})$ will also be bodily identical, as is evident from the equation

$$\beta_{\overline{g}g\overline{g}^{-1},\overline{g}\mathscr{R}} = B_{\overline{g}\mathscr{R}\leftarrow\mathscr{R}}\beta_{g,\mathscr{R}}B_{\mathscr{R}\leftarrow\overline{g}\mathscr{R}}, \tag{A15}$$

for all $g,\overline{g}\in P^{\uparrow}_{+}$.

Actually, the actions of the "symmetry transformation" $\beta_{g,\mathscr{R}}$ and $\beta_{g,\overline{\mathscr{R}}}$ are similar only in the following sense: *For any two equivalent observers* $\mathscr{R}$, $\overline{\mathscr{R}}$ *if the* OP (X,O) *and* $(\overline{X},\overline{O})$ *are subjectivley identical, then the* OP $\beta_{g,\mathscr{R}}(X,O)$ *and* $\beta_{g,\overline{\mathscr{R}}}(\overline{X},\overline{O})$ *are also subjectively identical*, as is evident from the equation

$$\beta_{g,\overline{g}\mathscr{R}} = S_{\overline{g}\mathscr{R}\leftarrow\mathscr{R}}\beta_{g,\mathscr{R}}S_{\mathscr{R}\leftarrow\overline{g}\mathscr{R}}, \tag{A16}$$

for all $g,\overline{g}\in P^{\uparrow}_{+}$. Relations similar to (A15) and (A16) also hold between the "symmetry transformations" $\widetilde{\beta}_{g,\mathscr{R}}$ and $\widetilde{\beta}_{g,\overline{\mathscr{R}}}$, from which the corresponding conclusions can be drawn. The above result, that the transformations $\beta_{g,\mathscr{R}}(\widetilde{\beta}_{g,\mathscr{R}})$ transform subjectively identical procedures into subjectively identical procedures, is of considerable significance. For example, it implies that the process by which one empirically verifies whether the requirement of relativistic symmetry [cf. Eq. (A18) below] is satisfied, is (subjectively) the same for equivalent observers.

All our considerations till now have been based only on the requirements I–III. In order to express the requirement IV also in an appropriate form, let us consider a SPP (μ,O_0) and a sequence of OP $\{(X_1,O_1),\ldots,(X_{i_1},O_{i_1}),\ldots,(X_{i_n},O_{i_n}),\ldots,(X_r,O_r)\}$ in the language of an observer $\mathscr{R}$. If $\{O_0,O_1,\ldots,O_r\}$ is a COS of space–time regions (cf. Appendix B), then we shall denote by $\Pr^{(\mu,O_0)}_{(X_1,O_1),\ldots,(X_r,O_r)}\{X_{i_1}(E_{i_1}),\ldots,X_{i_n}(E_{i_n})\}$ the joint probability that the outcomes of the OP $\{(X_{i_k},O_{i_k})\}$ are found to lie in $\{E_{i_k}\}$ $(k=1,2,\ldots,n)$, when an ensemble of systems prepared as per SPP (μ,O_0) is subjected to the sequence of OP $\{(X_1,O_1),\ldots,(X_r,O_r)\}$. Now, if $\overline{\mathscr{R}}=g\mathscr{R}$ is another equivalent observer then the requirements IV(a) and IV(b) can be expressed [in the notation of Eqs. (A2), (A3), and (A4)] as follows:

$$\Pr^{(\mu,O_0)}_{(X_1,O_1),\ldots,(X_r,O_r)}\{X_{i_1}(E_{i_1}),\ldots,X_{i_n}(E_{i_n})\}$$
$$=\Pr^{(\overline{\mu}^B,gO_0)}_{(\overline{X}^B_1,gO_1),\ldots,(\overline{X}^B_r,gO_r)}\{\overline{X}^B_{i_1}(E_{i_1}),\ldots,\overline{X}^B_{i_n}(E_{i_n})\}\qquad\text{(A17a)}$$

$$\Pr^{(\mu,O_0)}_{(X_1,O_1),\ldots,(X_r,O_r)}\{X_{i_1}(E_{i_1}),\ldots,X_{i_n}(E_{i_n})\}$$
$$=\Pr^{(\overline{\mu}^S,O_0)}_{(\overline{X}^S_1,O_1),\ldots,(\overline{X}^S_r,O_r)}\{\overline{X}^S_{i_1}(E_{i_1}),\ldots,\overline{X}^S_{i_n}(E_{i_n})\}\qquad\text{(A17b)}$$

for all $E_{i_k}\in B(S_{i_k})$ $(k=1,2,\ldots,n)$. Since from theorem B1 it follows that $\{gO_0,gO_1,\ldots,gO_r\}$ is also a COS of space–time regions, the right-hand side of Eq. (A17a) is well-defined.

The two sides of Eqs. (A17a) and (A17b) are expressed, respectively, in the languages of the observers $\mathscr{R}$ and $\overline{\mathscr{R}}=g\mathscr{R}$. Hence these equations may be said to express the "*relativistic invariance*" of the laws of nature. From Eqs. (A8) and (A17) we can obtain the following requirement of "*relativistic symmetry*" of the laws of nature, now expressed entirely in the language of a single observer $\mathscr{R}$ [after adopting the notation of Eqs. (A11a), (A11b)]:

Theorem A 2 (*relativistic symmetry*): For each $g\in P^{\uparrow}_{+}$, each SPP (μ,O_0) and an aribitrary sequence $\{(X_1,O_1),\ldots,(X_{i_1},O_{i_1}),\ldots,(X_{i_n},O_{i_n}),\ldots,(X_r,O_r)\}$ of OP where $\{O_0,O_1,\ldots,O_r\}$ is a COS, we have

$$\Pr^{(\mu,O_0)}_{(X_1,O_1),\ldots,(X_r,O_r)}\{X_{i_1}(E_{i_1}),\ldots,X_{i_n}(E_{i_n})\}$$
$$=\Pr^{(\widetilde{\beta}_g\mu,gO_0)}_{(\beta_gX_1,gO_1),\ldots,(\beta_gX_r,gO_r)}\{(\beta_gX_{i_1})(E_{i_1}),\ldots,(\beta_gX_{i_n})(E_{i_n})\},\qquad\text{(A18)}$$

for all $E_{i_k}\in B(S_{i_k})$, where S_{i_k} is the value space of (X_{i_k},O_{i_k}) $(k=1,2,\ldots,n)$.

In summary, our discussion has been mostly concerned with the reformulation of the fundamental requirements (I–IV) of relativistic invariance (which postulate certain relations between the descriptions of physical phenomena and formulations of laws of nature as carried out in the languages of *different* equivalent observers, as certain conditions on the theory as formulated in the language of *any single* observer. This is necessitated by the fact that usually a physical theory is formulated entirely in terms of the language adopted by a single observer—as is for example the case with our formulation of relativistic quantum theory in Sec. 2. In such cases the requirements of relativistic invariance essentially reduce to the following conditions on the theory (as formulated in the language of a single observer):

(i) The OP and SPP of the theory should be local procedures which may be carried out in any translatable set of space–time regions (cf. Sec. 2).

(ii) For each $g\in P^{\uparrow}_{+}$ there are associated transformations β_g and $\widetilde{\beta}_g$ which map, respectively, the set of OP and SPP in accordance with Eqs. (A11a), (A11b); also these transformations carry the physical interpretation that was discussed in connection with Eqs. (A8), (A10), and (A16).

(iii) The conclusions stated in Theorem A1 and Theorem A2 are satisfied.

If the above conditions are satisfied, we may say that the theory is *in conformity* with the requirements of relativistic invariance.

In conclusion we shall make a few remarks concerning the phsyical basis of the particular framework of relativistic quantum theory that was proposed in Sec. 2. The general motivation of course is that relativistic quantum theory should be formulated as an appropriate generalization of nonrelativistic quantum theory (in conformity with the fundamental requirements of relativistic invariance). Therefore, it should also be essentially a theory of successive observations. In order to achieve this (it has been argued in Sec. 2 and Appendix B that) our heuristic ideas of relativistic causality suggest that we should stipulate that the theory should satisfy the following requirement:

V. Relativistic succession principle—stated in Appendix B.

It may be noted that the requirements I–V stated so far, are of a very general nature, and are not in any way specific to a particular mathematical scheme employed in formulating the theory. However, the analogy with nonrelativistic quantum theory suggests that we also adopt the scheme of quantum probability theory in the formulation of the relativistic theory. More specifically we assume the following:

VI. The local OP and local SPP are defined (in analogy with nonrelativistic quantum theory) in terms of a space of local events $\mathscr{O}=\cup\mathscr{O}(O)$.

Finally as regards the relativistic transformation properties of OP and SPP we shall assume that all these transformations are induced by transformations defined on the event space, as per the following requirement:

VII. If $\mathscr{R}$ and $\overline{\mathscr{R}}$ are two equivalent observers, then the transformations $B_{\overline{\mathscr{R}}\leftarrow\mathscr{R}}$ and $\widetilde{B}_{\overline{\mathscr{R}}\leftarrow\mathscr{R}}$ ($S_{\overline{\mathscr{R}}\leftarrow\mathscr{R}}$ and $\widetilde{S}_{\overline{\mathscr{R}}\leftarrow\mathscr{R}}$) are induced by the transformation $\mathscr{B}_{\overline{\mathscr{R}}\leftarrow\mathscr{R}}$ ($\mathscr{S}_{\overline{\mathscr{R}}\leftarrow\mathscr{R}}$) which maps the event space $\mathscr{O}=\cup\mathscr{O}(O)$ in the language of $\mathscr{R}$ onto the event space $\overline{\mathscr{O}}=\cup\overline{\mathscr{O}}(O)$ in the language adopted by $\overline{\mathscr{R}}$. Moreover, both $\mathscr{B}_{\overline{\mathscr{R}}\leftarrow\mathscr{R}}$ and $\mathscr{S}_{\overline{\mathscr{R}}\leftarrow\mathscr{R}}$ map $\mathscr{O}$ somorphically onto $\mathscr{O}$ in the sense that both of them satisfy conditions analogous to (E5bi)–(E5bv) of Sec. 2 [see (D13)].

The framework of relativistic quantum theory outlined in Sec. 2 was essentially based on the postulates (E1)–(E5) (which characterized the structure of the event space), the definitions of OP and SPP and their transformation properties as given by (OP1)–(OP3) and (SPP1)–(SPP3), and the fundamental statistical law as given by the relativistic Wigner formula (2.13). We shall now briefly indicate some of the heuristic arguments that may be provided in justification of these "postulates" on the basis of the above requirements I–VII. First of all it may be noted that (OP1), (OP2), (SPP1), (SPP2), and the relativistic Wigner formula (2.13) are the appropriate generalizations of their nonrelativistic counterparts once we adopt the requirements I–VI. The following remarks suggest that the postulates (E1)–(E4) are also based on the same requirements once we notice that each local event is nothing but the realization of certain outcomes when some local OP is carried out.

The postulate (E1a) is just a consistency requirement which allows us to consider the situation in which no observations are carried out in the space–time region O as being equivalent to the performance of the trivial OP (ξ_I, O), where ξ_I is the trivial operation-valued measure with value space $\{1\}$ such that $\xi_I(\{1\}) = I$. The postulates (E1b) and (E1d) ensure that the definition of a local OP as given by (OP1) and (OP2) is consistent. The postulate (E1c) ensures that every local event $A \in \mathscr{O}(O)$ is associated with some local OP carried out in O. The postulate (E2) enables us to consider every OP carried out in a region O_1 as also a procedure carried out in a larger region $O_2 \supset O_1$. In fact, this has been our viewpoint all through this paper.

The postulate (E3) now provides the appropriate interpretation of conjunction of two local events in relativistic quantum theory and is directly based on the relativistic succession principle. It essentially states that $A_2 \in \mathscr{O}(O_2)$ may be thought of as an event which follows another event $A_1 \in \mathscr{O}(O_1)$ iff $O_2 > O_1$, and then the occurrence of the succession of these events can be considered as the occurrence of the compound event $A_1 \wedge A_2$ in the space–time region $O_1 \cup O_2$.

The postulate (E4), which states that two local events $A_1 \in \mathscr{O}(O_1)$ and $A_2 \in \mathscr{O}(O_2)$ are "compatible" whenever O_1 is spacelike to O_2, plays a crucial role in all our considerations. First of all, this postulate is needed to ensure that the occurrence of two events in mutually spacelike regions O_1, O_2 can be considered as the occurrence of a *unique* compound event in $O_1 \cup O_2$— because, from requirement V, the occurrence of either of these events may be thought of as preceding that of the other. In our view however, the most important reason for adopting the postulate (E4) is that it is essential to ensure that the fundamental statistical law of the theory [Eq. (2.13)] provides an unambiguous prescription for the statistical correlations between the outcomes of successive observations. In fact the proof of our Theorem 2.1 [which established the consistency of the RWF (2.13)], depends crucially on the Eq. (2.9) [which asserts the compatibility of two OP (X_1, O_1) and (X_2, O_2) whenever $O_1 \times O_2$] which in turn is based on the postulate (E4). In other words, *once we adopt the relativistic succession principle* [*on which Eq.* (2.13) *is based*] *and also assume that any arbitrary local* OP (X_2, O_2) *can be conjointly performed with the* OP (X_1, O_1) *whenever* $O_1 \times O_2$ (*which is some form of a local causality principle*), *then* (*E* 4) *will follow as a consistency condition and is in this sense analogous to Eqs.* (1.6) *and* (1.18) *in the nonrelativistic theory*. In this context we may also note that the postulate (E4) can be justified on other independent physical grounds, for example (cf. Refs. 13,18) by demanding that the principle of local causes (say, in its simplest form as stated in Theorem 3.1) be one of the fundamental requirements imposed on the theory.

Finally we shall make a few remarks on the postulates (E5), (OP3), and (SPP3). From the requirement VII it follows that the transformations $B_{\overline{\mathscr{R}} \leftarrow \mathscr{R}}$ and $S_{\overline{\mathscr{R}} \leftarrow \mathscr{R}}$ are to be considered as being induced by the transformations $\mathscr{B}_{\overline{\mathscr{R}} \leftarrow \mathscr{R}}$ and $\mathscr{S}_{\overline{\mathscr{R}} \leftarrow \mathscr{R}}$ relating the event spaces of these observers. Therefore, if we adopt the notation of Eqs. (A2a) and (A4a), we can write

$$\overline{X}^B(E) = \mathscr{B}_{\overline{\mathscr{R}} \leftarrow \mathscr{R}}[X(E)] \tag{A19a}$$

$$\overline{X}^S(E) = \mathscr{S}_{\overline{\mathscr{R}} \leftarrow \mathscr{R}}[X^S(E)] \tag{A19b}$$

for all $E \in B(S)$. Hence, it follows from Eqs. (A2a), (A3), and (A4a) that

$$A \in \mathscr{O}(O) \Rightarrow \mathscr{B}_{\overline{\mathscr{R}} \leftarrow \mathscr{R}} A \in \overline{\mathscr{O}}(gO) \tag{A20a}$$

and

$$A \in \mathscr{O}(O) \Rightarrow \mathscr{S}_{\overline{\mathscr{R}} \leftarrow \mathscr{R}} A \in \overline{\mathscr{O}}(O), \tag{A20b}$$

where $\overline{\mathscr{R}} = g\mathscr{R}$. If we now define the mapping $\alpha_{g,\mathscr{R}}: \mathscr{O} \rightarrow \mathscr{O}$ (denoted as α_g in Sec. 2) by the equation

$$\alpha_{g,\mathscr{R}} = \mathscr{S}_{\mathscr{R} \leftarrow g\mathscr{R}} \mathscr{B}_{g\mathscr{R} \leftarrow \mathscr{R}}, \tag{A21}$$

then it can easily be seen that the postulates (E5), (OP3), and (SPP3) follow directly from the requirements I–VII.

APPENDIX B: CAUSALLY ORDERED SEQUENCES OF SPACE–TIME REGIONS

In Sec. 1 it was seen that the joint probabilities for any sequence of events in nonrelativistic quantum theory, depend crucially on the time-ordering of the corresponding sequence of observations. In the same way, when we go over to a relativistic theory the manner in which a sequence of local observations should be ordered, depends on the causal structure of space–time. We shall devote this appendix to a study of such causally ordered sequences (COS) of local observation procedures (OP) when the space–time is just the ordinary Minkowski space of special relativity.

For each space–time point x in the Minkowski space M, let $J^+(x)$ ($J^-(x)$) denote the closed forward (backward) light cone with the apex at the point x [which is also included in $J^{\pm}(x)$]. Then, for each space–time region $O \subset M$, its *causal future* $J^+(O)$ and the *causal past* $J^-(O)$ are defined by the equations

$$J^+(O) = \bigcup_{x \in O} J^+(x), \tag{B1a}$$

$$J^-(O) = \bigcup_{x \in O} J^-(x). \tag{B1b}$$

The *causal complement* O' of space-time region O is defined as the set of all points of M whch are spacelike to every point in O. We may also define the regions $S^{\pm}(O)$ by the equation

$$S^{\pm}(O) = M \setminus J^{\mp}(O). \tag{B2}$$

We shall refer to $S^+(O)$ ($S^-(O)$) as the *domain of causal suc-*

cession (*precedence*) for O. For each space–time region O, the region O' is disjoint with O, $J^+(O)$ and $J^-(O)$; however we have

$$O \subset J^+(O) \cap J^-(O). \tag{B3}$$

We may also note the following relations:

$$J^+(O) \cup J^-(O) \cup O' = M, \tag{B4a}$$

$$S^+(O) \cup S^-(O) \cup O = M, \tag{B4b}$$

$$J^\pm(O_1 \cup O_2) = J^\pm(O_1) \cup J^\pm(O_2). \tag{B5}$$

If O_1, O_2 are any two space–time regions we shall say that "O_1 *causally precedes* O_2" (to be denoted by $O_1 < O_2$), or equivalently that "O_2 *causally succeeds* O_1" $(O_2 > O_1)$ whenever

$$O_1 \cap J^+(O_2) = \phi. \tag{B6a}$$

Equation (B6a) is also equivalent to each of the following conditions (B6b), (B7a), and (B7b):

$$O_2 \cap J^-(O_1) = \phi, \tag{B6b}$$

$$O_1 \subset S^-(O_2), \tag{B7a}$$

$$O_2 \subset S^+(O_1). \tag{B7b}$$

We shall say that the regions O_1, O_2 are spacelike to each other (to be denoted by $O_1 \times O_2$) whenever $O_1 O'_2$ or equivalently $O_2 O'_1$. It can easily be shown that

$$O_1 \times O_2 \Leftrightarrow O_1 > O_2 \quad \text{and} \quad O_2 > O_1. \tag{B8}$$

The relations $<$ and $\times$ are both invariant under relativistic transformations—i.e., for each $g \in P^{\uparrow}_{+}$, the group of all restricted inhomogeneous Lorentz transformations, we have

$$O_1 < O_2 \Leftrightarrow gO_1 < gO_2, \tag{B9a}$$

$$O_1 \times O_2 \Leftrightarrow gO_1 \times g(O_2), \tag{B9b}$$

where for each $O \subset M$, $g(O) = \{x | g^{-1}(x) \in O\}$. We should also note that both the relations $<$ and $\times$ are *irreflexive and nontransitive* in general.

We have argued in Sec. 2 that in a relativistic theory the way in which a sequence of local OP (and SPP) are to be ordered should be deduced from the following "*relativistic succession principle*."

"A local OP (X_2,O_2) can be thought of as a procedure which follows another OP (X_1,O_1) [or a SPP (μ,O_1)] iff the space–time region O_2 has a null intersection with the causal past $[J^{-1}(O_1)]$ of O_1—i.e., $O_2 > O_1$. The performance of the sequence of OP $\{(X_1,O_1),(X_2,O_2)\}$ can then be thought of as the performance of a composite OP [to be denoted by $(X_1 \circ X_2, O_1 \cup O_2)$] in the space–time region $O_1 \cup O_2$."

Based on the above principle, we can immediately conclude that a sequence of OP $\{(X_1,O_1),\ldots,(X_r,O_r)\}$ can be thought of as a sequence of "successive" observation procedures which follow a SPP (μ,O_0) iff the sequence of space–time regions $\{O_0,O_1,\ldots,O_r\}$ has the following property:

$$O_i > \bigcup_{j=o}^{i-1} O_j \quad (i = 1,2,\ldots,r). \tag{B10}$$

We shall say that a sequence $\{O_0,O_1,\ldots,O_r\}$ is a *causally ordered sequence* (COS) of space–time regions if it satisfies Eq. (B10). We shal also say that a sequence of OP $\{(X_1,O_1),\ldots,(X_r,O_r)\}$ is a COS of OP if the corresponding sequence of space–time regions $\{O_1,\ldots,O_r\}$ is a COS.

From Eq. (B5) we can conclude that

$$\{O_1,\ldots,O_r\} \text{ is a COS iff } O_i > O_j \text{ whenever } i > j \quad (i,j = 1,2,\ldots,r). \tag{B11}$$

Some of the basic properties of COS of space–time regions are collected in the following theorem which can be proved easily from the above definitions:

Theorem B 1: Every COS $S = \{O_1,O_2,\ldots,O_r\}$ has the following properties (C1)–(C6):

(C1) For every subset $\{O_{i_1},O_{i_2},\ldots,O_{i_n}\} \subset S$, $O_{i_1} \not< O_{i_2}$ and

$$O_{i_2} \not< O_{i_3} \text{ and} \cdots \text{and } O_{i_{n-1}} \not< O_{i_n} \Rightarrow O_{i_1} > O_{i_n}, \tag{B12}$$

where $O_\alpha \not< O_\beta$ denotes the negation of $O_\alpha < O_\beta$. In particular if O_α, O_β are distinct elements of S, then at least one of the two relations $O_\alpha < O_\beta$, $O_\beta < O_\alpha$ holds; also, the regions $O_1,O_2,\ldots,O_r$ are all mutually disjoint.

(C2) Every subsequence $\{O_{i_1},O_{i_2},\ldots,O_{i_n}\}$ of S is also a COS whenever $i_1 < i_2 < \cdots < i_n$.

(C3) For any integer $j < r$, the sequence $\{O_1,\ldots,O_{j-1},O_{j+1},O_j,O_{j+2},\ldots,O_r\}$ is also a COS iff $O_j \times O_{j-1}$.

(C4) For each $g \in P^{\uparrow}_{+}$, the sequence $\{gO_1,gO_2,\ldots,gO_r\}$ is also a COS.

(C5) If the integers $\{i_k | k = 1,\ldots,n\}$ are such that $1 \leqslant i_1 < i_2 < \cdots < i_n < r$, then

$$\{\bigcup_{\alpha=1}^{i_1} \mathscr{O}_\alpha, \bigcup_{\alpha=i_1+1}^{i_2} O_\alpha, \ldots, \bigcup_{\alpha=i_n+1}^{r} O_\alpha\}$$

is also a COS.

(C6) For any integer $j \leqslant r$, if the regions $\{O_{j_\alpha} | \alpha = 1,2,\ldots,n\}$ are such that

$$O_{j_1} \cup O_{j_2} \cup \cdots \cup O_{j_n} \subset O_j,$$

then the sequence $\{O_1,\ldots,O_{j-1},O_{j_1},O_{j_2},\ldots,O_{j_n},O_{j+1},\ldots,O_r\}$ is a COS iff $\{O_{j_1},O_{j_2},\ldots,O_{j_n}\}$ is a COS.

Another very important property of causally ordered sequences of space–time regions (which plays a crucial role in the formulation of the fundamental statistical prescription of quantum theory (cf. Theorem 2.1) is the following:

Theorem B 2: If $S' = \{O_{\pi_1},O_{\pi_2},\ldots,O_{\pi_r}\}$ is a permutation of a COS $S = \{O_1,O_2,\ldots,O_r\}$ then S' is a COS iff S' can be obtained from S by successive interchanges between neighboring elements which are spacelike to each other.

Proof: The "if" part of the theorem follows directly from the result (C3) of Theorem B.1. The "only if" part will now be proved by induction on the length "r" of the sequence S.

The theorem is true for any sequence $\{O_1,O_2\}$ of length 2—for, if both $\{O_1,O_2\}$ and $\{O_2,O_1\}$ are causally ordered sequences, then from (B8) it follows that $O_1 \times O_2$. Now let us assume that the theorem is true for every sequence of length $r-1$. Let O_1 be the jth element of S', i.e.,

$$S' = \{O_{\pi_1}, O_{\pi_2}, \ldots, O_{\pi_{j-1}}, O_1, O_{\pi_{j+1}}, \ldots, O_{\pi_r}\}.$$

since S' is a COS it follows that

$$O_1 > O_{\pi_\alpha} (\alpha = 1,2,\ldots,j-1).$$

But, since S itself is a causally ordered sequence, we also have

$$O_1 < O_{\pi_\alpha} (\alpha = 1,2,\ldots,j-1).$$

Therefore from (B8) it follows that

$$O_1 \times O_{\pi_\alpha} (\alpha = 1,2,\ldots,j-1).$$

Therefore the sequence S' is obtained from the COS

$$S'' = \{O_1, O_{\pi_1}, O_{\pi_2}, \ldots, O_{\pi_{j-1}}, O_{\pi_{j+1}}, \ldots, O_{\pi_r}\}$$

by successive interchanges between neighboring elements which are spacelike to each other. The theorem will be proved if we now show that S'' itself is obtained from S by similar successive interchanges. This follows directly from the hypothesis of induction (that the theorem is true for every COS of length $r-1$) when applied to the causally ordered permutation $\{O_{\pi_1}, O_{\pi_2}, \ldots, O_{\pi_{j-1}}, O_{\pi_{j+1}}, \ldots, O_{\pi_r}\}$ of the COS $\{O_2, O_3, \ldots, O_r\}$ of length $r-1$. Hence the theorem has been established.

We shall not comment on the physical significance of the above results, as that will be sufficiently clear from the discussion of Sec. 2. Since the fundamental statistical law of relativistic quantum theory (as formulated in Sec. 2) prescribes the statistical correlations between the outcomes of a COS of OP only, it is important to know when a given set $\{(X_1,O_1),\ldots,(X_r,O_r)\}$ of OP can be ordered into a COS. For this purpose, we now define a sequence $\{O_1, O_2, \ldots, O_r\}$ to be a "*causally orderable sequence*" of space–time regions if there exists a permutation π of the indices $\{1,2,\ldots,r\}$ such that $\{O_{\pi_1}, O_{\pi_2}, \ldots, O_{\pi_r}\}$ is a COS. From (C3) of Theorem B1, it follows that such a permutation (if it exists) will not be unique in general, because we can interchange any two neighboring mutually spacelike elements in any causally ordered form of the sequence and obtain another causally ordered form. We may also note that if $\{O_1, \ldots, O_r\}$ is a causally orderable sequence, so are all its subsequences and also the sequence $\{gO_1, \ldots, gO_r\}$ for any $g \in P^{\uparrow}_{+}$.

It can easily be verified that *every sequence of space–time points is causally orderable*. [See (D14).] In general, a sequence of space–time regions will have to satisfy quite a stringent condition (especially if some of the regions are unbounded) for it to be causally orderable. However, (most often) by decomposing some of the elements of a sequence which is not itself causally orderable into mutually disjoint subregions a new causally orderable sequence may be formed.

From a mathematical point of view, the condition of causal orderability cannot be expressed in a very simple form because the relation $<$ is not transitive and, in addition, satisfies Eq. (B8). In fact, we can easily find examples of sequences $\{O_1, O_2, O_3\}$ which are not causally orderable even though $O_1 < O_2$ and $O_2 < O_3$. In the folowing theorem we shall state a necessary and sufficient condition for a sequence of space–time regions to be causally orderable.

Theorem B 3. A sequence $S = \{O_1, O_2, \ldots, O_r\}$ of space–time regions is causally orderable iff the following condition (C1) is satisfied:

(C1) For every subset $\{O_{i_1}, O_{i_2}, \ldots, O_{i_n}\} \subset S$, $O_{i_1} \nless O_{i_2}$ and $O_{i_2} \nless O_{i_3}$ and- - - - - -and $O_{i_{n-1}} \nless O_{i_n} \Rightarrow O_{i_1} > O_{i_n}$.

The fact that (C1) is a necessary condition follows immediately from Theorem B1. We can show that (C1) is also a sufficient condition for causal orderability by means of an argument based on induction on the number of elements "r" in the sequence S.

APPENDIX C: NOTES

(D1) In the present investigation we shall not consider the important problem of localization in relativistic quantum theory (and the related difficulties in defining a suitable positin observable). See in this connection Refs. 41–45.

(D2) We shall refer to observables characterized this way as observation procedures (OP), if in addition the space–time region of observation is also specified.

(D3) This is a more general (but purely statistical) version of the principle of local causes recently discussed by Stapp.[35–38] In the literature one also comes across references to "Einstein causality,." "relativistic causality," "local compatibility," etc.—which essentially refer to some version of the above principle of local causes.

(D4) It should be noted that this notion "local state of a system in a space–time region O" is quite different from our notion of local SPP and also the notion of the so called "localized states."[41] In the formalism of Sec. 2, these "*local states in O*" may be identified with appropriate linear functionals on $\mathscr{O}(O)$ which satisfy conditions analogous to (SPP1) and (SPP2). From the specification of such a "local state of the system in O" we can only infer the probabilities for the various outcomes of OP carried out in O. Also the local state in any region $\bar{O} > O$ depends on what particular OP has been carried out in O.

(D5) Strictly speaking, a satisfactory solution to this problem can emerge only after we remove the arbitrary restriction imposed in the nonrelativistic theory that all measurements are instantaneous in character (cf. Sec. 2 below).

(D6) *Note added in proof*: A recent study[46] shows that for discussing the collapse associated with observables with continuous spectra it is necessary to formulate the theory in terms of "finitely additive expectation valued measures" which are more general than the "operation valued measures" considered in this paper. However, such a generalization will not affect the main conclusions of this paper.

(D7) The notion of local events (operations) was introduced (as elements of a C^*-algebra) by Haag and Kastler.[46] The local event spaces $\{\mathscr{O}(O)\}$ of our theory can also be characterized (more abstractly, without bringing in Hilbert space considerations) as appropriate subsets of the set of all positive elements in the unit ball of a Banach algebra as indicated in Ref. 24.

(D8) These postulates will be justified on somewhat more general physical grounds in Appendix A. We should also make a note of the fact that topological considerations will not play any signficant role in our investigations and therefore will be mostly ignored in what follows.

(D9) The simplest way is of course that of replacing the postulate (E3) itself by the following one: "If the OP (X_1,O_1), (X_2,O_2) are such that $O_1 < O_2$, the range of the operation-valued measure $X_1 \circ X_2$ is contained entirely in $\mathscr{O}(O_1 \cup O_2)$."

(D10) It may be noted that in order to prove Theorem 3.3 (principle of local causes), the following (weaker) assumption (E6′) is sufficient:

(E6′) For each maximal element $\xi \in \Sigma \cap \mathscr{O}(O)$, where $O = O_1 \cup O_2$ and $O_1 < O_2$, there exist maximal elements $\xi_1 \in \Sigma \cap \mathscr{O}(O_1)$ and $\xi_2 \in \Sigma \cap \mathscr{O}(O_2)$ such that $\xi = \xi_1 \wedge \xi_2$.

(D11) The fact that in a relativistic theory the characterization of OP and SPP should also include a specification of the space–time regions in which these procedures are carried out follows from the fundamental requirement (I) of relativistic invariance (cf. below), as we have argued in Sec. 2.

(D12) These are essentially the well-known requirements studied in the literature[2,4,40,48,49] which have been restated somewhat more precisely in the context of our formulation of relativistic quantum theory.

(D13) The requirement that $\mathscr{S}_{\bar{\mathscr{R}} \leftarrow \mathscr{R}}$ should map the events space $\mathscr{O}$ isomorphically on to the event space $\bar{\mathscr{O}}$, may be thought of to be a natural extension of the earlier requirement (II) that equivalent observers see subjectively identical worlds. Moreover, the assumption that both $\mathscr{S}_{\bar{\mathscr{R}} \leftarrow \mathscr{R}}$ and $\mathscr{B}_{\bar{\mathscr{R}} \leftarrow \mathscr{R}}$ are isomorphisms between the event spaces of $\mathscr{R}$ and $\bar{\mathscr{R}}$, is also essential to ensure that the transformations $S_{\bar{\mathscr{R}} \leftarrow \mathscr{R}}$ and $B_{\bar{\mathscr{R}} \leftarrow \mathscr{R}}$ [which are induced from $\mathscr{S}_{\bar{\mathscr{R}} \leftarrow \mathscr{R}}$ and $\mathscr{B}_{\bar{\mathscr{R}} \leftarrow \mathscr{R}}$—cf. Eqs. (A19), (A21) below] are well defined on all local OP and transform in addition the composition of two OP into the composition of the transformed OP.

(D14) Every set $\{x_1,x_2,\ldots,x_n\}$ of space–time points satisfies the condition $\{x_1\} \nless \{x_2\}$ and $\{x_2\} \nless \{x_3\}$ and $\cdots$ $\{x_{n-1}\} \nless \{x_n\} \Rightarrow \{x_1\} > \{x_n\}$ and $\{x_1\} \nless \{x_n\}$, which is stronger than the condition (C1) (cf. Theorem B3) which ensures the causal orderability of a set. In fact, given any sequence $\{x_1,\ldots,x_n\}$ of space–time points, any time-ordered permutation (say $\{x_{\pi_1},\ldots,x_{\pi_n}\}$, where the time coordinates satisfy $x^0_{\pi_i} \leqslant x^0_{\pi_j}$ whenever $i<j$) will be a COS.

[1]E.P. Wigner, "Epistomological perspective on quantum theory," in *Contemporary Research in the Foundations and Philosophy of Quantum Theory*, edited by C.A. Hooker (Reidel, Dordrecht, 1973).
[2]E.P. Wigner, Nuovo Cimento **3**, 517 (1956).
[3]E.P. Wigner, Am. J. Phys. **31**, 6 (1963) (also reprinted in Ref. 31).
[4]R.M.F. Houtappel, H. Van Dam, and E.P. Wigner, Rev. Mod. Phys. **37**, 595 (1965).
[5]E.P. Wigner, The subject of our discussions, in *Foundations of Quantum Mechanics*, edited by B. d'Espagnat (Academic, New York, 1971).
[6]W.C. Davidon and H. Ekstein, J. Math. Phys. **5**, 1588 (1964).
[7]A.L. Licht, J. Math. Phys. 7, 1656 (1966).
[8]I. Bloch, Phys. Rev. **156**, 1377 (1967).
[9]S. Schlieder, Commun. Math. Phys. **7**, 305 (1968).
[10]A.L. Licht, J. Math. Phys. **9**, 1468 (1968).
[11]K.E. Hellwig and K. Kraus, Phys. Rev. D **1**, 566 (1970).
[12]A.L. Licht, "The relativistic quantum theory of measurement," in the *Proceedings of VIII Winter School in Theoretical Physics in Karpacz, Wroclaw*, 1972.
[13]K. Kraus, Ann. Phys. (N.Y.) **64**, 81 (1971).
[14]J.S. Bell, "Subject and object," in *The Physicist's Conception of Nature*, edited by J. Mehra (Reidel, Dordrecht, 1973).
[15]J.S. Bell, "The Theory of Local Beables," CERN preprint (Ref. TH.2053–CERN), 1975.
[16]Ken-ichi Ono, "On the origin of indeterminacy," in *The Study of Time II*, edited by J.T. Fraser and N. Lawrence (Springer Verlag, Berlin, 1975).
[17]B. d'Espagnat, *Conceptual Foundations of Quantum Mechanics* (Benjamin, New York, 1976).
[18]P. Benioff and H. Ekstein, Phys. Rev. D **15**, 3563 (1977), and references cited therein.
[19]E.B. Davies and J.T. Lewis, Commun. Math. Phys. **15**, 305 (1969).
[20]C.M. Edwards, Commun. Math. Phys. **16**, 207 (1970).
[21]M.D. Srinivas, J. Math. Phys. **16**, 1672 (1975) [to be reprinted in *The Logico-Algebraic Approach to Quantum Theory*, edited by C.A. Hooker (Reidel, Dordrecht, 1978), Vol. II.
[22]E.B. Davies, *Quantum Theory of Open Systems* (Academic, New York, 1976), and references cited therein.
[23]M.D. Srinivas, J. Math. Phys. **19**, 1705 (1978).
[24]M.D. Srinivas, J. Math. Phys. **19**, 1952 (1978).
[25]J. Von Neumann, *Mathematical Foundations of Quantum Mechanics* (Princeton U.P., Princeton, New Jersey, 1955).
[26]B. Misra and E.C.G. Sudarshan, J. Math. Phys. **18**, 756 (1977) and references cited therein.
[27]M.D. Srinivas, J. Math. Phys. **18**, 2138 (1977).
[28]M.D. Srinivas, "Quantum point process model for photodetection," to appear in the *Proceedings of the IV Rochester Conference on Coherence and Quantum Optics*, edited by L. Mandel and E. Wolf (1978).
[29]E.P. Wigner, *Group Theory and Its Application to the Quantum Mechanics of Atomic Spectra* (Academic, New York, 1959).
[30]E.P. Wigner, Ann. Math. **40**, 149 (1939).
[31]Several of the articles reprinted in E.P. Wigner, *Symmetries and Reflections* (Indiana U.P., Bloomington, 1967).
[32]A. Einstein, B. Podolsky, and N. Rosen, Phys. Rev. **47**, 777 (1935).
[33]D. Bohm and Y. Aharanov, Phys. Rev. **108**, 1070 (1957).
[34]J.S. Bell, Physics (N.Y.) **1**, 195 (1964).
[35]H.P. Stapp, "Correlation experiments and the nonvalidity of ordinary ideas about the physical world," Lawrence Berkeley Laboratory Report LBL-5333 (1968).
[36]H.P. Stapp, Phys. Rev. D **3**, 1303 (1971).
[37]H.P. Stapp, Found. Phys. **7**, 313 (1977).
[38]H.P. Stapp, "Whiteheadian approach to quantum theory and the generalised Bell's theorem," Lectures given at the University of Texas, 1977.
[39]P. Benioff and H. Ekstein, Nuovo Cimento B **40**, 9 (1977), and references cited therein.
[40]A.S. Wightman, Suppl. Nuovo Cimento **14**, 81 (1959).
[41]A.J. Kalnay, "The Localization Problem," in *Studies in the Foundations, Methodology and Philosophy of Science, problems in the Foundations of Physics*, edited by M. Bunge (Springer Verlag, Berlin, 1971), Vol. IV, and references cited therein.
[42]E.P. Wigner, "Relativistic Equations in Quantum Mechanics," in *The Physicist's Conception of Nature*, edited by J. Mehra (Reidel, Dordrecht, 1973).
[43]G.C. Hegerfeldt, Phys. Rev. D **10**, 3320 (1974).
[44]J.F. Perez and I.F. Wilde, Phys. Rev. D **16**, 315 (1977).
[45]A.A. Broyles, "Relativistic Quantum Mechanics," University of Florida report, Gainesville (1977).
[46]M.D. Srinivas, "Collapse Postulate for observables with continuous spectra," ICTP preprint (IC 78/23).
[47]R. Haag and D. Kastler, J. Math. Phys. **5**, 848 (1964); see also R. Haag, "Observables and fields," in *Lectures on Elementary Particles and Quantum Field Theory*, edited by S. Deser, M. Grisaru, and H. Pendleton (M.I.T. Press, Cambridge, Massachussetts, 1970).
[48]E. Fabri, Nuovo Cimento **14**, 1130 (1959), and references cited therein.
[49]A.S. Wightman, "L' invariance dans la mécanique quantique relativistique," in *Relations de Dispersion et Particals Elémentaires*, edited by C. De Witt and R. Omnes (Herman, Paris, 1960).