The unitary irreducible representations of SL(2, R) in all subgroup reductions [a)]

Debabrata Basu [b)] and Kurt Bernardo Wolf

Instituto de Investigaciones en Matemáticas Aplicadas y en Sistemas (IIMAS), Universidad Nacional Autónoma de México, Apdo. Postal 20-726 México 20, D. F., Mexico

(Received 23 December 1980; accepted for publication 5 June 1981)

We use the canonical transform realization of SL(2, R) in order to find all matrix elements and integral kernels for the unitary irreducible representations of this group. Explicit results are given for all mixed bases and subgroup reductions. These provide the full multiparameter set of integral transforms and series expansions associated to SL(2, R).

PACS numbers: 02.20.Df

1. INTRODUCTION

The complete classification of the Unitary Irreducible representations (UIRs) of the three-dimensional Lorentz group SO(2, 1) and of its twofold covering group SL(2, R) were given by Bargmann in his classic 1947 article,[1] where one can find the UIR matrix elements—rows and columns classified by the UIRs of the compact subgroup SO(2)—in explicit form. This group, its covering groups $\mathrm{SO}(2,1) \overset{1:2}{\simeq} \mathrm{SU}(1,1) \simeq \mathrm{Sp}(2,R) \approx \mathrm{SL}(2,R) \overset{1:\infty}{\simeq} \overline{\mathrm{SL}(2,R)}$ and its representations were further studied by Barut and Fronsdal,[2] Pukański,[3] Sally, Jr.,[4] and in a book by Lang.[5]

The study of group representations in different bases is of interest both from the mathematical and the physical point of view. The intimate connections between the representations of Lie groups and the special functions of mathematical physics have long been recognized and treated in textbooks.[6] In physics, subgroup reductions corresponding to different bases of the Lorentz and other groups lead to various ways to correlate or interpret data, as in the description of the high-energy scattering dynamics,[7] which requires the reduction SO(2, 1)$\supset$SO(1, 1) among others. This interest coincided with the investigations of Mukunda,[8–11] Barut,[2,12] Lindblad and Nagel,[13] and others, who analyzed this chain in some detail and computed the generalized representation matrices (or integral kernels) of one-parameter subgroups and found the coupling coefficients.

In the study of the role of canonical transformations in quantum mechanics, the work of Moshinsky and Quesne[14,15] started from linear transformations between coordinate and momentum observables and lead to the oscillator (metaplectic) representation of Sp(2, R). In contrast to the realizations given by Bargmann[1] and by Gel'fand *et al.*,[16] in which the group acts as a Lie transformation group on functions of a coset manifold, the group actions in the constructions of Moshinsky,[14,15,17] Seligman, Wolf,[18–23] Burdet, Perrin and Perroud,[24] and present in the work of others,[25–27] is an integral transform realization of SL(2, R) on $\mathscr{L}^2(R)$ Hilbert spaces. This group of integral transforms has been called *canonical* transforms.[18,28] It is unique in that the associated Lie algebra is an algebra of second-order differential operators on a dense common domain in these Hilbert spaces. The action is thus distinct from—although unitarily equivalent[20,21] to—the SL(2, R) action as a Lie transformation group on coset spaces, of the Lie–Bargmann multiplier representations[29] on the unit circle or disk.

The canonical transform realization has provided a degree of uniformity in the treatment of the discrete series[19] of UIRs on the one hand and the continuous series[21] of UIRs on the other. In this article it has enabled us to evaluate, in a straightforward and unified way, the UIR matrix elements and integral kernels of finite SL(2, R) elements. In contrast with some of the previous investigations, this approach deals with the general SL(2, R) group element, rather than with specific one-parameter subgroups. Although Bargmann's results on UIRs of SL(2, R) in the compact subgroup basis[30] are well known, it is also true that other continuous noncompact and mixed-basis reductions have so far not received uniform consideration[2,9,10,12,31–33] and are scattered in the literature. The discrete series of UIRs in all subgroup reductions was undertaken by Boyer and Wolf[34] using canonical transforms. We repeat their results here since the journal is not generally available and the article contains some errata. The mixed-basis matrix elements of the continuous series were treated by Kalnins,[31] who gave expressions for one-parameter subgroups in terms of Whittaker and Laguerre functions of the second kind.[35] All our expressions are given in terms of confluent and Gauss hypergeometric functions, and have uniformity of notation, normalization, and phase conventions. The purpose of this paper is to give a comprehensive derivation and listing of all subgroup reductions.

The plan of the article is as follows. In Sec. 2 we display the needed formulas from the theory of canonical transforms for the general method of construction and, since we want to describe all UIR matrix elements and integral kernels, we organize the notation properly in due accordance with Bargmann's conventions. In Sec. 3 and 4 we give the results for the discrete and continuous (nonexceptional and exceptional) representation series. The first subsection of each lists the subgroup-adapted basis functions, the second treats the mixed-basis expressions, while the third subsection treats the subgroup reductions, i.e., the cases when the row and column variables refer to the same subgroup. These are ex-

a) Work performed under financial assistance from Consejo Nacional de Ciencia y Tecnología (CONACYT) Project ICCBIND 790370.

b) On leave from Department of Physics, Indian Institute of Technology, Kharagpur 721302, India.

pressed as Gauss or confluent hypergeometric functions and, alternatively, as cylinder and Whittaker functions[36,37] of the three independent SL(2, R) parameters. Certain cases of interest are pointed out in a further subsection. Comparison with alternative derivations available in the literature is pointed out whenever we are aware of such results.

The representation matrix elements for the compact subgroup chain were obtained by Bargmann as solutions to differential equations[38] with boundary conditions imposed by the group identity. We come to the evaluation of an integral as the last step to the same end. We make use of a method by Majumdar and Basu[32] on hypergeometric series Mellin expansions to solve three of the six chains in each series. In the special case of the continuous series in the compact subgroup reduction, such an integral (a Gaussian of imaginary width times two Whittaker functions, one with a rescaled argument) is not available in the literature. Through Bargmann's result this is evaluated.

In Sec. 5 we point out that the six different mixed-basis and subgroup-reduced representation matrix elements constitute six families of SL(2, R) integral and discrete transforms, as well as series expansions, of which the set of canonical transforms is but one. The Appendix summarizes some information about the groups SU(1, 1), SL(2, R), and their UIRs as classified by Bargmann. Throughout this article Z and R stand for the set of integers and real numbers. Boldfaced symbols indicate vectors or matrices. For brevity, we shall speak of UIR *matrix elements* encompassing both the ordinary and generalized (i.e., integral transform kernel) cases.

As a general observation, we should remark that the canonical transform realization of SL(2, R) can be regarded as a complementary alternative to Bargmann's treatment of the same group. The latter is simpler in certain respects, particularly when dealing with the compact subgroup chain, while the former seems to be most appropriate for noncompact subgroup chains.

2. CANONICAL TRANSFORMS

A. The construction of SL(2, R) representations

The determination of representation matrices (or integral kernels) for group elements $g\in G$ may proceed as follows: Provided (i) one has a Hilbert space $\mathscr{H}$ of functions $f(r)$, r in some carrier space X, endowed with a sesquilinear positive definite inner product $(\cdot,\cdot)$, where the action of G is well defined and onto,

$$f(r) \overset{g}{\rightarrow} f_g(r) = [\mathbb{C}_g f](r), \quad f, f_g \in \mathscr{H} \tag{2.1}$$

(ii) one has a complete orthonormal, or generalized Dirac-orthonormal basis for $\mathscr{H}$, $\{\psi_\lambda(r)\}_{\lambda\in\Lambda}$ (Λ being the range of the label specifying the basis vectors uniquely), one can build a representation $\mathbf{D}$: $G \rightarrow \mathrm{Hom}\Lambda$ as

$$\mathbf{D}(g) = \|D_{\lambda,\lambda'}(g)\|, \tag{2.2a}$$

$$D_{\lambda,\lambda'}(g) = (\psi_\lambda, \mathbb{C}_g \psi_{\lambda'}). \tag{2.2b}$$

The completeness of the (possibly generalized) basis function set will then guarantee the representation property

$$\underset{\lambda'\in\Lambda}{\mathrm{S}}\, D_{\lambda,\lambda'}(g_1) D_{\lambda',\lambda''}(g_2) = D_{\lambda,\lambda''}(g_1 g_2), \tag{2.2c}$$

where the symbol $\mathrm{S}_{\lambda'\in\Lambda}$ stands for summation in the case of proper, and integration in the case of generalized, bases. The unitarity and irreducibility properties of $\mathbf{D}$ follow from similar requirements for the action (2.1) on $\mathscr{H}$.

The reasons for which this straightforward program often fails to provide a definite result have to do more with knowing the "best" choice of basis functions $\{\psi_\lambda(r)\}_{\lambda\in\Lambda}$ and the problem of explicit computation of the integral in (2.2b), than with matters of principle. The bases are usually chosen as the eigenvectors of one or more operators in the Lie algebra—so that subgroup reductions result—while the space $\mathscr{H}$ is an $\mathscr{L}^2(X)$ space on a coset manifold $X = G/H$ (or $H\setminus G$) with some convenient subgroup $H\subset G$. A closely related approach to part (ii) of evaluation of (2.2b) calls for (ii') finding these functions for various one-parameter subgroups of G as solutions of differential equations obtained from the subgroup generators, subject to the boundary conditions $\mathbf{D}(e) = \mathbf{1}$ at the group identity $e\in G$.

The group G which we consider here is SL(2, R):

$$\left\{ \mathbf{g} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \Big| a,b,c,d\in R, \quad \det \mathbf{g} = 1 \right\}. \tag{2.3}$$

Starting with Bargmann[1] a number of authors have implemented the program (i)–(ii) or (i)–(ii'), using for the supporting space X the coset space provided by the Iwasawa decomposition $NA\setminus NAK = S_1$ (i.e., the circle) and Bargmann's multiplier action.[29] This is unitary in $\mathscr{L}^2(S_1)$ for the continuous nonexceptional representation series[29]; for the continuous exceptional and discrete series it is $\mathscr{L}^2_{\Omega^C}(S_1)$ and $\mathscr{L}^2_{\Omega^D}(S_1)$ with nonlocal measures[39,40] Ω^C and Ω^D. The latter is equivalent[20] to a space of analytic functions on the unit disk[29] or on the complex half-plane.[16] These realizations are very appropriate for finding the SL(2, R) representation matrices reduced with respect to the compact SO(2) subgroup, since, the ensuing analysis makes use of Fourier series on $\mathscr{L}^2(S_1)$ for UIRs belonging to the continuous class, or Hardy spaces for those belonging to the discrete series.[39] When one makes use of the same action and spaces for the reduction under a noncompact subgroup, calculations become awkward.

The Hilbert spaces and SL(2, R) action we use in this article have been developed in Refs. 9, 15, 19, 21, and 22 for Sp(2, R)$\simeq$SL(2, R), as well as the oscillator representation[14,18] of Sp(2N, R) on an N-dimensional carrier space R^N. As we shall see in implementing part (ii) of the program outlined above, these techniques are best suited for noncompact subgroup reduction.

B. The discrete series $D_k^\pm$

The oscillator representation of the subgroup SO(2)$\times$SL(2, R) of Sp(4, R), restricted to a given one-dimensional UIR M of SO(2), $M\in Z$, generates the conjugate SL(2, R) representation[15,19,22,27] belonging to the discrete series D_k^+ with $k = (1 + |M|)/2$. When the two-dimensional carrier space R^2 is parametrized in polar coordinates, this representation is realized as an integral transform group on the radial variable $r \in R^+$ and defines the *k-radial canonical transform* on the Hilbert space $\mathscr{L}^2(R^+)$. The inner product is thus the standard one,

$$(f, h) = \int_0^\infty dr\, f(r)^* h(r), \tag{2.4}$$

and the action of the group element $\mathbf{g}$ is given by

$$[\mathbb{C}_{\mathbf{g}}^k f](r) = \int_0^\infty dr' C_{\mathbf{g}}^k(r, r') f(r'), \quad \mathbf{g} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \tag{2.5a}$$

where the integral kernel $C_{\mathbf{g}}^k(r, r')$ is given by an imaginary Gaussian times a Bessel function:

$$C_{\mathbf{g}}^k(r, r') = e^{-i\pi k} b^{-1} (rr')^{1/2} \exp(i[dr^2 + ar'^2]/2b) J_{2k-1}(rr'/b), \tag{2.5b}$$

$$2k - 1 = 0, 1, 2, ..., \quad \text{i.e., } k = \tfrac{1}{2}, 1, \tfrac{3}{2}, 2, ... \tag{2.5c}$$

When $\mathbf{g}$ is a lower-triangular matrix ($b = 0$) one finds from the asymptotic properties of the Bessel function[41] that Eq. (2.5a) becomes the multiplier action

$$\left[\mathbb{C}^k \begin{pmatrix} a & 0 \\ c & a^{-1} \end{pmatrix} f\right](r) = (\operatorname{sgn} a)^{2k} |a|^{-1/2} \exp(icr^2/2a) f(r/|a|). \tag{2.5d}$$

We shall write $\mathbb{C}^k(\mathbf{g})$ for $\mathbb{C}_{\mathbf{g}}^k$ whenever $\mathbf{g}$ is displayed as a matrix. The k-canonical transform (2.5) is unitary under the inner product (2.4) and a Parseval relation $(f, h) = (\mathbb{C}_{\mathbf{g}}^k f, \mathbb{C}_{\mathbf{g}}^k h)$ holds.

The Lie generators of $\mathbb{C}_{\mathbf{g}}^k$ are *second*-order differential operators[42] given by

$$J_1^\gamma = \frac{1}{4}\left(-\frac{d^2}{dr^2} + \frac{\gamma}{r^2} - r^2\right), \tag{2.6a}$$

$$J_2^\gamma = -\frac{i}{2}\left(r\frac{d}{dr} + \tfrac{1}{2}\right), \tag{2.6b}$$

$$J_0^\gamma = \frac{1}{4}\left(-\frac{d^2}{dr^2} + \frac{\gamma}{r^2} + r^2\right), \tag{2.6c}$$

on a space dense in $\mathscr{L}^2(R^+)$, and γ is related to k through

$$\gamma = (2k-1)^2 - \tfrac{1}{4}, \tag{2.7}$$

so that $\gamma = -\frac{1}{4}, \frac{3}{4}, \frac{15}{4}, \ldots$. These generators close into a Lie algebra sl(2, R) under commutation. We shall also come to use

$$J_+^\gamma = J_0^\gamma + J_1^\gamma = \frac{1}{2}\left(-\frac{d^2}{dr^2} + \frac{\gamma}{r^2}\right), \tag{2.8a}$$

$$J_-^\gamma = J_0^\gamma - J_1^\gamma = \tfrac{1}{2} r^2. \tag{2.8b}$$

The Casimir invariant of sl(2, R) is a multiple of the identity:

$$Q = (J_1^\gamma)^2 + (J_2^\gamma)^2 - (J_0^\gamma)^2 = q\mathbf{1}, \tag{2.9a}$$

$$q = -\tfrac{1}{4}\gamma + \tfrac{3}{16} = k(1-k), \tag{2.9b}$$

i.e., $q = \frac{1}{4}, 0, -\frac{3}{4}, -2, \ldots$.

The association of (2.6)–(2.8) with the one-parameter subgroups of SL(2, R) is as follows

$$\exp(i\alpha J_1) \mapsto \mathbf{M}_1(\alpha) = \begin{pmatrix} \cosh \alpha/2 & -\sinh \alpha/2 \\ -\sinh \alpha/2 & \cosh \alpha/2 \end{pmatrix} \in \mathrm{SO}(1, 1)_1, \tag{2.10a}$$

$$\exp(i\beta J_2) \mapsto \mathbf{M}_2(\beta) = \begin{pmatrix} \exp(-\beta/2) & 0 \\ 0 & \exp(\beta/2) \end{pmatrix} \in \mathrm{SO}(1, 1)_2, \tag{2.10b}$$

$$\exp(i\gamma J_0) \mapsto \mathbf{M}_0(\gamma) = \begin{pmatrix} \cos(\gamma/2) & -\sin(\gamma/2) \\ \sin(\gamma/2) & \cos(\gamma/2) \end{pmatrix} \in \mathrm{SO}(2)_0, \tag{2.10c}$$

$$\exp(ibJ_+) \mapsto \mathbf{M}_+(b) = \begin{pmatrix} 1 & -b \\ 0 & 1 \end{pmatrix} \in \mathrm{E}(1)_+, \tag{2.10d}$$

$$\exp(icJ_-) \mapsto \mathbf{M}_-(c) = \begin{pmatrix} 1 & 0 \\ c & 1 \end{pmatrix} \in \mathrm{E}(1)_-. \tag{2.10e}$$

All nonequivalent one-parameter subgroups of SL(2, R) are present in (2.10): the compact rotation elliptic subgroup SO(2), the noncompact Euclidean parabolic subgroup E(1), and the boost hyperbolic subgroup SO(1, 1). For the latter two we have the following equivalence relations between the equivalent pairs (2.10a)–(2.10b) and (2.10d)–(2.10e):

$$\mathbf{S}\mathbf{M}_2(\zeta)\mathbf{S}^{-1} = \mathbf{M}_1(\zeta), \quad \mathbf{S} = 2^{-1/2}\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}, \tag{2.11a}$$

$$\mathbf{F}\mathbf{M}_-(z)\mathbf{F}^{-1} = \mathbf{M}_+(z), \quad \mathbf{F} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} = \mathbf{S}^{-2}. \tag{2.11b}$$

The spectrum of J_0^γ in (2.6c) for $\gamma \geqslant \frac{3}{4}$ in $\mathscr{L}^2(R^+)$ has a lower bound given by its corresponding $k \geqslant 1$. (For $k = \frac{1}{2}$ or $\gamma = -\frac{1}{4}$ this is also the case for the self-adjoint extension specified in Sec. 3) The k-radial canonical transforms (2.5) thus belong to the lower-bound UIRs D_k^+ of SL(2, R).

The UIRs D_k^- are obtained from the D_k^+ ones through the sl(2, R) outer automorphism[43]

$$J_0^\gamma \leftrightarrow -J_0^\gamma, \quad J_1^\gamma \leftrightarrow -J_1^\gamma, \quad J_2^\gamma \leftrightarrow J_2^\gamma, \quad J_\pm^\gamma \leftrightarrow -J_\pm^\gamma. \tag{2.12a}$$

This exchanges the raising and lowering operators with a change of sign:

$$J_\uparrow^\gamma \leftrightarrow -J_\downarrow^\gamma, \quad J_{\uparrow\downarrow}^\gamma = J_1^\gamma \pm iJ_2^\gamma. \tag{2.12b}$$

The automorphism acts on the SL(2, R) group elements[44] as

$$\mathbf{g} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \leftrightarrow \begin{pmatrix} a & -b \\ -c & d \end{pmatrix} = \mathbf{g}^A. \tag{2.12c}$$

The D_k^- matrix elements can be thus expressed in terms of the corresponding D_k^+ ones, as will be detailed for the various subgroup reductions, at the end of the next section.

C. The continuous nonexceptional series C_q^ϵ

The oscillator representation of Sp(4, R) can also be reduced with respect to an O(1, 1)$\times$SL(2, R) subgroup[11,21,22] by making use of hyperbolic coordinates on the plane. The resulting reduction, on being restricted to a definite UIR $(p, 2s)$ of O(1, 1), $p = \pm 1$, $s \in R$, yields a conjugate reduction of SL(2, R) to one of the continuous series of UIRs C_q^ϵ. The case of vector ($\epsilon = 0$) and spinor ($\epsilon = \frac{1}{2}$) representations correspond to even ($p = +1$) and odd ($p = -1$) parity representations of O(1, 1) with $q = \frac{1}{4} + s^2 \geqslant \frac{1}{4}$. Since hyperbolic coordinates require two coordinate patches to cover the plane, the "hyperbolic radial" carrier space will be $X = R^+ + R^+$ and the Hilbert space correspondingly a two-component $\mathscr{L}^2$ space of functions

$$\mathbf{f}(r) = \begin{pmatrix} f_1(r) \\ f_{-1}(r) \end{pmatrix} = \| f_j(r) \|, \quad j = 1, -1, \quad f_j(r) \in \mathscr{L}^2(R^+). \tag{2.13}$$

The inner product in this Hilbert space $\mathscr{L}^2_{\rm II}(R^+) = \mathscr{L}^2(R^+) \dot{+} \mathscr{L}^2(R^+)$ will be

$$(\mathbf{f}, \mathbf{h}) = \sum_{j=\pm 1} \int_0^\infty dr\, f_j(r)^* h_j(r). \tag{2.14}$$

Calling $k = \frac{1}{2} + is$, this reduction leads to the (ϵ, k)-hyperbolic canonical transform

$$[\mathbb{C}_{\mathbf{g}}^{\epsilon,k} \mathbf{f}]_j(r) = \sum_{j'=\pm 1} \int_0^\infty dr' [\mathbb{C}_{\mathbf{g}}^{\epsilon,k}]_{j,j'}(r,r') f_{j'}(r'). \tag{2.15a}$$

The 2×2 matrix integral kernel $\mathbb{C}_{\mathbf{g}}^{\epsilon,k}(r, r')$ is given by a Gaussian times Hankel and Macdonald functions of imaginary index. For $2k-1 = 2is$, $s \in R$, $p_0 = 1$, $p_{1/2} = -1$, we can write[45]

$$[\mathbb{C}_{\mathbf{g}}^{\epsilon,k}]_{jj'}(r, r') = G_{\mathbf{g}jj'}(r, r') H_{jj'}^{\epsilon,k}(-rr'/b), \tag{2.15b}$$

$$G_{\mathbf{g}jj'}(r, r') = (2\pi|b|)^{-1}(rr')^{1/2}\exp(i[djr^2 + aj'r'^2]/2b), \tag{2.15c}$$

$$\begin{aligned} H_{1,1}^{\epsilon,k}(\zeta) &= p_\epsilon H_{-1,-1}^{\epsilon,k}(\zeta) = p_\epsilon H_{1,1}^{\epsilon,k}(-\zeta) = H_{1,1}^{\epsilon,1-k}(\zeta) \\ &= i\pi[e^{-\pi s} H_{2is}^{(1)}(\zeta + i0^+) - p_\epsilon e^{\pi s} H_{2is}^{(2)}(\zeta - i0^+)] \\ &= 2i\pi(-\operatorname{sgn}\zeta)^{2\epsilon}[-g_{1/2-\epsilon}(k) J_{2is}(|\zeta|) \\ &\quad + i\, g_\epsilon(k) Y_{2is}(|\zeta|)], \end{aligned} \tag{2.15d}$$

$$\begin{aligned} H_{1,-1}^{\epsilon,k}(\zeta) &= p_\epsilon H_{-1,1}^{\epsilon,k}(\zeta) = p_\epsilon H_{1,-1}^{\epsilon,k}(-\zeta) = p_\epsilon H_{1,-1}^{\epsilon,1-k}(\zeta) \\ &= 4(-\operatorname{sgn}\zeta)^{2\epsilon} g_\epsilon(k) K_{2is}(|\zeta|), \end{aligned} \tag{2.15e}$$

$$\epsilon = 0: \begin{cases} k - \frac{1}{2} = is, & s \geqslant 0 \\ k - \frac{1}{2} = \sigma, & 0 < \sigma < \frac{1}{2} \end{cases}, \quad g_0(k) = \sin \pi k = \begin{cases} \cosh \pi s \\ \cos \pi\sigma \end{cases}, \tag{2.15f}$$

$$\epsilon = \tfrac{1}{2}: k - \tfrac{1}{2} = is, \quad s > 0, \quad g_{1/2}(k) = i\cos\pi k = \sinh \pi s. \tag{2.15g}$$

In the last two equations we are defining the function $g_\epsilon(k)$ for values of k which will make it applicable to the exceptional continuous series discussed in the next subsection. Note that for $\zeta < 0$, $\arg(\zeta \pm i0^+) = \pm\pi$, so (2.15d) valuates $H_{2is}^{(1)}$ *above* the branch cut of the function (placed along the negative real half-axis), and $H_{2is}^{(2)}$ is valuated *below* the cut.

When **g** in Eq. (2.3) is lower-triangular ($b = 0$), as for the oscillator radial case (2.5), one finds from the asymptotic properties of the cylinder functions that Eq. (2.15a) becomes the multiplier action

$$\left[\mathbb{C}^{\epsilon k}\begin{pmatrix} a & 0 \\ c & a^{-1} \end{pmatrix} \mathbf{f}\right]_j(r) = (\operatorname{sgn} a)^{2\epsilon}|a|^{-1/2}\exp(ijcr^2/2a) f_j(r/|a|). \tag{2.15h}$$

The (ϵ, k)-hyperbolic canonical transform is unitary under (2.14), and a corresponding Parseval relation holds.

Here too, the Lie generators of the integral transform action are second-order differential operators, but arranged in 2×2 matrix form. In terms of the formal operators (2.6) they are[11,21]

$$\mathbf{J}_1^\gamma = \begin{pmatrix} J_1^\gamma & 0 \\ 0 & -J_1^\gamma \end{pmatrix} = \| j\delta_{jj'} J_1^\gamma \|, \tag{2.16a}$$

$$\mathbf{J}_2^\gamma = \begin{pmatrix} J_2^\gamma & 0 \\ 0 & J_2^\gamma \end{pmatrix} = \| \delta_{jj'} J_2^\gamma \|, \tag{2.16b}$$

$$\mathbf{J}_0^\gamma = \begin{pmatrix} J_0^\gamma & 0 \\ 0 & -J_0^\gamma \end{pmatrix} = \| j\delta_{jj'} J_0^\gamma \|, \tag{2.16c}$$

$$\mathbf{J}_\pm^\gamma = \begin{pmatrix} J_\pm^\gamma & 0 \\ 0 & -J_\pm^\gamma \end{pmatrix} = \| j\delta_{jj'} J_\pm^\gamma \|. \tag{2.16d}$$

Again γ is related to k through (2.7), but now as k is in the range (2.15f) and (2.15g) [instead of (2.5c)], we have $\gamma \leqslant -\frac{1}{4}$. As the subgroup assignments (2.10) are representation-independent statements, they continue to hold here as well. The Casimir invariant of SL(2, R) is now $q \geqslant \frac{1}{4}$, corresponding to the continuous nonexceptional series of UIRs. The one point we must clarify in this regard (See the Appendix) is that for spinor representations ($\epsilon = \frac{1}{2}$) the hyperbolic canonical transforms (2.15) do not include the point $k = \frac{1}{2}$ (i.e., $s = 0$ or $q = \frac{1}{4}$). Indeed, from (2.15e) we can verify that for $k = \frac{1}{2} + is$, $s \to 0^+$ the off-diagonal kernel elements ($j \neq j'$) vanish and hence the two j-component spaces uncouple. The diagonal elements are now $\sim J_0(\zeta)$, that is, they are the $D_{1/2}^+$ ($k = \frac{1}{2}$)-radial canonical transform kernel for the upper component, and the $D_{1/2}^-$ one for the lower component, as is clearly suggested by (2.12a)–(2.16).

D. The continuous exceptional series C_q^0

The oscillator representation of Sp(4, R) does not contain the exceptional continuous representation series of any of its SL(2, R) subgroups. However, there exist unique self-adjoint extensions[46] of the generators (2.16) in $\mathscr{L}^2_{\rm II}(R^+)$, which enable us to reach this series by analytic continuation in the variable k in (2.15f) to values off $k = \frac{1}{2}$, in the range

$$\tfrac{1}{2} < k < 1 \ (\text{i.e., } 0 < 2k - 1 = 2\sigma < 1), \text{ for } \epsilon = 0 \ (p_\epsilon = 1). \tag{2.17}$$

For these UIRs $-\frac{1}{4} < \gamma < \frac{3}{4}$, i.e., $0 < q < \frac{1}{4}$.

The features one must check are that the integral kernels corresponding to these values of k continue to map $\mathscr{L}^2_{\rm II}(R^+)$ functions into functions in the same space, and that the representation property (2.2c) holds. That this is the case follows from the integrability properties of cylinder functions in the range $(-1, 1)$ of the index, in particular their behavior at zero and infinity, and from the completeness relations for the similarly extended basis functions, to be seen in Sec. 4.

Again, as for the $\epsilon = \frac{1}{2}$, $k = \frac{1}{2} + is$, $s \to 0^+$ case seen above, when $\epsilon = 0$ and $k \to 1^-$ the integral kernel matrix (2.13) becomes diagonal and the two j components uncouple. In the limit, the upper and lower-diagonal components become proportional to $J_1(\zeta)$, and belong to the D_1^+ and D_1^- representations.

We have assembled in the last subsections the tools for the calculation of the matrix elements of SL(2, R) in point (i) of our program. In the next two sections we shall implement point (ii) for the discrete and continuous UIRs.

E. Notation

A word about notation: we shall use the eigenbases of J_α^γ, $\alpha = 0, 1, 2, +, -$, generating the discrete UIRs $D_k^\pm$. We denote their eigenfunctions by ${}^\alpha\Phi_\lambda^k(r)$, λ being a function of the eigenvalue. When J_α^γ is in the elliptic orbit ($\alpha = 0$) the eigenvalue set of J_0^γ is discrete and we shall denote its eigenvalues λ by m. The range will be understood by the context. When J_α^γ is in the hyperbolic orbit ($\alpha = 1, 2$) or in the parabolic orbit ($\alpha = +, -$), its eigenvalue set is continuous. In the first case λ will be denoted by $\mu \in R$, the eigenvalue under $J_{1,2}^\gamma$ being μ. In the second case λ will be called $\rho \in R^+$, the eigenvalues of $J_\pm^\gamma$ being $\rho^2/2$. Eigenbases for the D_k^- UIRs will not be needed separately. In the continuous series C_q^ϵ the eigenbases of J_α^γ will be similarly denoted by ${}^\alpha\Psi_\lambda^{\epsilon,k}(r)$, these are two-component functions with elements ${}^\alpha\Psi_{\lambda,j}^{\epsilon,k}$, $j = 1, -1$. We use m again for λ, the eigenvalue under J_0^γ. The multiplicity of the eigenvalues of the generators in the hyperbolic and parabolic orbits is now doubled, however. For the former we use for λ the pair (κ, μ), $\kappa = \pm 1, \mu \in R$, and for the latter ($\operatorname{sgn}\rho$, $|\rho|) = \rho, \rho \in R$, the eigenvalues being again μ and $\rho^2/2$ under the respective J^γ's.

The representations $\mathbf{D}(\mathbf{g})$ constructed in (2.2) have their matrix elements

$$ {}^{\alpha,\beta}D_{\lambda,\lambda'}^k(\mathbf{g}) = ({}^\alpha\Phi_\lambda^k, \mathbb{C}_\mathbf{g}^k\, {}^\beta\Phi_{\lambda'}^k) = [{}^{\beta,\alpha}D_{\lambda',\lambda}^k(\mathbf{g}^{-1})]^*, \tag{2.18a} $$

$$ {}^{\alpha,\beta}D_{\lambda,\lambda'}^{\epsilon,k}(\mathbf{g}) = ({}^\alpha\Psi_\lambda^{\epsilon,k}, \mathbb{C}_\mathbf{g}^{\epsilon,k}\, {}^\beta\Psi_{\lambda'}^{\epsilon,k}) = [{}^{\beta,\alpha}D_{\lambda',\lambda}^{\epsilon,k}(\mathbf{g}^{-1})]^*, \tag{2.18b} $$

in the appropriate inner product. When $\alpha = \beta$ we write ${}^\alpha D_{..}$ for ${}^{\alpha,\alpha}D_{..}$. The cases $\alpha \neq \beta$ and $\alpha = \beta$ in (2.18) will be called *mixed-basis* and *subgroup-reduced* UIR matrix elements. We shall work mostly with the D_k^+ UIRs and use (2.18a). In Sec. 3D, when we express the D_k^- UIRs in terms of the D_k^+ ones, we shall write ${}^{\cdot}D_{..}^{k(-)}$ and ${}^{\cdot}D_{..}^{k(+)}$ to distinguish between them.

3. THE DISCRETE SERIES $D_k^\pm$

In this section we present the evaluation of the matrix elements (or integral kernels) of finite SL(2, R) transformations for the UIRs belonging to the discrete series $D_k^\pm$. The first subsection gives the E(1), SO(1, 1), and SO(2) subgroup-adapted eigenfunctions, while the second and third subsections provide the explicit evaluation of D_k^+ mixed-basis and subgroup-reduced cases respectively. The last subsection relates these results to those of the D_k^- representations.

A. The subgroup-adapted eigenfunctions

i. E(1)$\subset$*SL* (2, *R*). The two operators generating conjugate E(1) subgroups [c.f. Eqs. (2.10d) and (2.10e)] are, as given by (2.8a), and (2.8b), J_+^γ and J_-^γ. They are unitarily equivalent through the Hankel transform (2.11b).

The eigenfunctions of J_+^γ in $\mathscr{L}^2(R^+)$ are, for $\gamma = (2k-1)^2 - \frac{1}{4}$,

$$ {}^+\Phi_\rho^k(r) = e^{i\pi k}(\rho r)^{1/2} J_{2k-1}(\rho r), \quad \rho \in R^+, \quad k = \tfrac{1}{2}, 1, \tfrac{3}{2}, \dots, \tag{3.1} $$

with eigenvalue $\rho^2/2 \in R^+$. The phase has been chosen so that the phase of the ${}^-\Phi_\rho^k$ functions, below, be as simple as possible.

A more convenient operator in the E(1) orbit is J_-^γ, as its eigenfunctions are simply

$$ {}^-\Phi_\rho(r) = \delta(\rho - r) = [\mathbb{C}_\mathrm{F}^k\, {}^+\Phi_\rho^k](r), \quad r \in R^+, \tag{3.2a} $$

with eigenvalue $\rho^2/2$. These are Dirac-orthonormal and complete:

$$ ({}^-\Phi_\rho, {}^-\Phi_{\rho'}) = \delta(\rho - \rho'), $$

$$ \int_0^\infty d\rho\, {}^-\Phi_\rho(r)^*\, {}^-\Phi_\rho(r') = \delta(r - r'), \tag{3.2b} $$

and independent of k.

ii. SO(1, 1)$\subset$*SL* (2, *R*). Here again we have two operators generating conjugate SO(1, 1) subgroups [c.f. Eqs. (2.10a) and (2.10b) and (2.11a)]: J_1^γ and J_2^γ, as given by (2.5a) and (2.5b). The latter is the simpler one, and its eigenfunctions are

$$ {}^2\Phi_\mu(r) = \pi^{-1/2} r^{-1/2 + 2i\mu}, \quad \mu \in R, \tag{3.3a} $$

with eigenvalue μ. They are Dirac-orthonormal and complete:

$$ ({}^2\Phi_\mu, {}^2\Phi_{\mu'}) = \delta(\mu - \mu'), \quad \int_{-\infty}^{\infty} d\mu\, {}^2\Phi_\mu(r)^*\, {}^2\Phi_\mu(r') = \delta(r - r'), \tag{3.3b} $$

and independent of k. The expansion in terms of them is—up to a factor—the positive Mellin transformation,[47] so an appropriate phase choice has been made.

The J_1^γ Dirac-normalized eigenfunctions may be found from (3.3a) and (2.11a) to be

$$ \begin{aligned} {}^1\Phi_\mu^k(r) &= [\mathbb{C}_\mathrm{S}^k\, {}^2\Phi_\mu](r) \\ &= C_\mu^k e^{i\pi k/2} r^{-1/2} M_{i\mu, k-1/2}(-ir^2) \\ &= C_\mu^k r^{2k-1/2} e^{ir^2/2}\, {}_1F_1\left[\begin{matrix} k - i\mu \\ 2k \end{matrix}; -ir^2\right], \end{aligned} \tag{3.4a} $$

$$ C_\mu^k = e^{i\pi k/2} 2^{i\mu} \pi^{-1/2} e^{\pi\mu/2} \Gamma(k + i\mu)/\Gamma(2k). \tag{3.4b} $$

and where $M_{..}(\cdot)$ is one of the Whittaker functions.[48] They correspond to eigenvalue μ under J_1^γ, and are Dirac-orthonormal and complete as in (3.3b).

iii. SO (2)$\subset$*SL* (2,*R*). The compact SO(2) subgroup is generated by J_0^γ as given in Eq. (2.6c). Its normalized eigenfunctions are given by

$$ \begin{aligned} {}^0\Phi_m^k(r) &= [2n!/(2k+n-1)!]^{1/2} r^{2k-1/2} e^{-r^2/2} L_n^{(2k-1)}(r^2) \\ &= [2(2k+n-1)!/n!(2k-1)!]^{1/2} r^{-1/2} M_{m,k-1/2}(r^2) \\ &= [2(2k+n-1)!/n!]^{1/2} [(2k-1)!]^{-1} r^{2k-1/2} e^{-r^2/2} \\ &\quad \times {}_1F_1\left[\begin{matrix} -n \\ 2k \end{matrix}; r^2\right], \end{aligned} $$

$$ m = k + n, \ n = 0, 1, 2, \dots \tag{3.5a} $$

with eigenvalue $m = k, k+1, \cdots$. The phase of these functions has been chosen following Bargmann's convention,[49] namely, such that the raising and lowering operators $J_1^\gamma \pm iJ_2^\gamma$ have real, positive, matrix elements. They are orthonormal and complete (dense) in $\mathscr{L}^2(R^+)$:

$$ ({}^0\Phi_m^k, {}^0\Phi_{m'}^k) = \delta_{m,m'}, \quad \sum_{m=k}^{\infty} {}^0\Phi_m^k(r)^*\, {}^0\Phi_m^k(r') = \delta(r - r'). \tag{3.5b} $$

B. The mixed-basis matrix elements

i. E (1)$\subset$*SL* (2,*R*)$\supset$*SO* (2). For all $\mathbf{g} \in$ SL(2, R) we may perform the Iwasawa decomposition

$$\mathbf{g}=\begin{pmatrix}a&b\\c&d\end{pmatrix}=\begin{pmatrix}\bar a&0\\ \bar c&\bar a^{-1}\end{pmatrix}\begin{pmatrix}\cos\theta/2&-\sin\theta/2\\ \sin\theta/2&\cos\theta/2\end{pmatrix},\tag{3.6a}$$

where

$$e^{i\theta}=(a-ib)/(a+ib),\quad \bar a=(a^2+b^2)^{1/2},\quad \bar a\,\bar c=ac+bd.\tag{3.6b}$$

Application of $\mathbb{C}^k_{\mathbf{g}}$ decomposed as above, multiplies the J^{γ}_0 eigenfunction by $e^{im\theta}$, followed subsequently by a multiplier Lie transformation, Eq. (2.5d). Thus

$$[\mathbb{C}^k_{\mathbf{g}}\,{}^0\Phi^k_m](r)=[(a-ib)/(a+ib)]^m(a^2+b^2)^{-1/4}\times\exp(ir^2[ac+bd]/2[a^2+b^2])\times{}^0\Phi^k_m(r/[a^2+b^2]^{1/2}).\tag{3.7}$$

Since the J^{γ}_- eigenfunctions are simple Dirac deltas, we immediately obtain

$$\begin{aligned}{}^{-,0}D^k_{\rho m}\begin{pmatrix}a&b\\c&d\end{pmatrix}&=\left({}^-\Phi_\rho,\mathbb{C}^k\begin{pmatrix}a&b\\c&d\end{pmatrix}{}^0\Phi^k_m\right)\\&=\left[\mathbb{C}^k\begin{pmatrix}a&b\\c&d\end{pmatrix}{}^0\Phi^k_m\right](\rho)={}^{+,0}D^k_{\rho m}\begin{pmatrix}c&d\\-a&-b\end{pmatrix}\\&=\left[{}^{0,-}D^k_{m\rho}\begin{pmatrix}d&-b\\-c&a\end{pmatrix}\right]^*\\&=\left(\frac{a-ib}{a+ib}\right)^m\left[\frac{2\Gamma(k+m)}{(m-k)!}\right]^{1/2}\frac{(a^2+b^2)^{-k}}{\Gamma(2k)}\\&\quad\times\rho^{2k-1/2}\exp\left(-\frac{\rho^2}{2}\frac{d-ic}{a+ib}\right)\\&\quad\times{}_1F_1\left[\begin{matrix}-(m-k)\\2k\end{matrix};\frac{\rho^2}{a^2+b^2}\right].\end{aligned}\tag{3.8}$$

The overlap coefficient between the $\mathrm{E}(1)_-$ and $\mathrm{SO}(2)_0$ subgroup chains is obtained by setting $\mathbf{g}=\mathbf{1}$, i.e., $a=1=d$, $b=0=c$ in Eq. (3.8). This is ${}^0\Phi^k_m(\rho)$, i.e., this change of basis is basically the Laguerre series expansion of functions of $\rho\in R^+$.

ii. $SO(1,1)\subset SL(2,R)\supset SO(2)$. This mixed basis element is essentially the Mellin transform of Eq. (3.8), and is given by[50]

$$\begin{aligned}&{}^{2,0}D^k_{\mu m}\begin{pmatrix}a&b\\c&d\end{pmatrix}\\&=\left({}^2\Phi_\mu,\mathbb{C}^k\begin{pmatrix}a&b\\c&d\end{pmatrix}{}^0\Phi^k_m\right)\\&={}^{1,0}D^k_{\mu m}\begin{pmatrix}2^{-1/2}(a-c)&2^{-1/2}(b-d)\\2^{-1/2}(a+c)&2^{-1/2}(b+d)\end{pmatrix}\\&=2^{k-i\mu}\left[\frac{\Gamma(k+m)}{2\pi(m-k)!}\right]^{1/2}\frac{\Gamma(k-i\mu)}{\Gamma(2k)}\\&\quad\times(a+ib)^{-m}(a-ib)^{m-k+i\mu}(d-ic)^{-k+i\mu}\\&\quad\times{}_2F_1\left[\begin{matrix}-(m-k),k-i\mu\\2k\end{matrix};\frac{2}{(a-ib)(d-ic)}\right]\\&=(-1)^{m-k}2^{m-i\mu}[2\pi(m-k)!\Gamma(k+m)]^{-1/2}\Gamma(m-i\mu)\\&\quad\times(a+ib)^{-m}(a-ib)^{i\mu}(d-ic)^{-m+i\mu}\\&\quad\times{}_2F_1\left[\begin{matrix}-(m-k),1-k-m\\1-m+i\mu\end{matrix};\tfrac{1}{2}(a-ib)(d-ic)\right].\end{aligned}\tag{3.9}$$

In all power-function factors, the principal branch of this function is to be taken in an obvious way. The hypergeometric function is a polynomial of degree $m-k=n$ so no multivaluation problems occur on its account.

The overlap coefficient between these two chains in the discrete series is obtained by setting $\mathbf{g}=\mathbf{1}$. Using an identity for the hypergeometric function[51] we find

$$\begin{aligned}({}^2\Phi_\mu,{}^0\Phi^k_m)&={}^{2,0}D^k_{\mu m}(\mathbf{1})\\&=(-1)^{m-k}2^{k-i\mu}\frac{\Gamma(m-i\mu)}{[2\pi(m-k)!\Gamma(k+m)]^{1/2}}\\&\quad\times{}_2F_1\left[\begin{matrix}-(m-k),k+i\mu\\1-m+i\mu\end{matrix};-1\right].\end{aligned}\tag{3.10a}$$

Correspondingly

$$\begin{aligned}&({}^1\Phi^k_\mu,{}^0\Phi^k_m)\\&={}^{2,0}D^k_{\mu m}\left(2^{-1/2}\begin{pmatrix}1&1\\-1&1\end{pmatrix}\right)=e^{-im\pi/2}({}^2\Phi_\mu,{}^0\Phi^k_m),\end{aligned}\tag{3.10b}$$

which may be compared with prior results.[52]

iii. $E(1)\subset SL(2,R)\supset SO(1,1)$. The application of $\mathbb{C}^k_{\mathbf{g}}$ to ${}^2\Phi_\mu$ in Eq. (3.3a) is up to a factor the Mellin transform of the k-canonical transform kernel (2.5b) with respect to the second argument r'. Although integrals of this type appear in the standard tables,[53] if we want to have expressions valid for all group parameters, positive as well as negative, care must be taken to choose the appropriate parameter products and ratios so that the ensuing complex power function be evaluated in a definite way: We choose here the principal sheet (with the branch cut along the negative real axis). Following the general method of finding the Mellin transforms of hypergeometric functions due to Majumdar and Basu,[39] which will be explained in some detail in the next section, we find the value of the integral to be

$$\begin{aligned}{}^{-,2}D^k_{\rho\mu}\begin{pmatrix}a&b\\c&d\end{pmatrix}&=\left({}^-\Phi_\rho,\mathbb{C}^k\begin{pmatrix}a&b\\c&d\end{pmatrix}{}^2\Phi_\mu\right)=\left[\mathbb{C}^k\begin{pmatrix}a&b\\c&d\end{pmatrix}{}^2\Phi_\mu\right](\rho)\\&=e^{-i\pi k}2^{-k+i\mu}\pi^{-1/2}\frac{\Gamma(k+i\mu)}{\Gamma(2k)}\\&\quad\times b^{-2k}(-ia/b)^{-k-i\mu}\\&\quad\times\rho^{2k-1/2}\exp(id\rho^2/2b)\\&\quad\times{}_1F_1\left[\begin{matrix}k+i\mu\\2k\end{matrix};\frac{-i\rho^2}{2ab}\right].\end{aligned}\tag{3.11}$$

The complex-power function argument $-ia/b$ lies, for all signs of a and b on the imaginary axis.[54] Valuation on the principal sheet means that the phase of $-ia/b$ is $-\pi/2$ for $\mathrm{sgn}\,ab=1$ and $\pi/2$ for $\mathrm{sgn}\,ab=-1$.

The overlap coefficient between these two chains may be obtained as the limit $\mathbf{g}\to\mathbf{1}$ in Eq. (3.11), or directly, as

$$({}^-\Phi_\rho,{}^2\Phi_\mu)={}^{-,2}D^k_{\rho\mu}(\mathbf{1})=\pi^{-1/2}\rho^{-1/2+2i\mu},\tag{3.12}$$

which is[47] $2^{1/2}$ times the positive Mellin transform kernel, of argument 2μ, between a function of $\rho\in R^+$ and its transform function of $\mu\in R$.

C. The matrix elements in the subgroup bases

i. $E(1)\subset SL(2,R)$. In this generalized basis the integral kernel is the simplest to obtain, as no integrations need be performed:

$$ {}^{-}D^{k}_{\rho\rho'}\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \left({}^{-}\Phi_{\rho}, \mathbb{C}^{k}\begin{pmatrix} a & b \\ c & d \end{pmatrix} {}^{-}\Phi_{\rho'} \right) = C^{k}\begin{pmatrix} a & b \\ c & d \end{pmatrix}(\rho,\rho') $$
$$ = e^{-i\pi k} b^{-1}(\rho\rho')^{1/2}\exp(i[d\rho^2 + a\rho'^2]/2b) J_{2k-1}(\rho\rho'/b) $$
$$ = 2(2ib)^{-2k}[\Gamma(2k)]^{-1}(\rho\rho')^{2k-1/2} $$
$$ \times\exp(i[d\rho^2 - 2\rho\rho' + a\rho'^2]/2b)\, {}_1F_1\left[\begin{matrix} 2k-\frac{1}{2} \\ 4k-1 \end{matrix}; \frac{2i\rho\rho'}{b}\right]. \tag{3.13} $$

For $g \in E(1)$, the subgroup generated by J^{γ}_{-} [c.f. Eq. (2.10e)], the kernel becomes diagonal. In fact, it is diagonal for the two-parameter subgroup generated by the first-order differential operators, for which (2.13) converges weakly to

$$ {}^{-}D^{k}_{\rho\rho'}\begin{pmatrix} a & 0 \\ c & a^{-1} \end{pmatrix} = (\operatorname{sgn} a)^{2k}|a|^{-1/2} \times\exp(ic\rho^2/2a)\delta(\rho' - \rho/|a|). \tag{3.14} $$

From this form it is manifest that ${}^{-}D^{k}_{\rho\rho'}(\mathbf{1}) = \delta(\rho - \rho')$, the unit operator in $\mathscr{L}^2(R^{+})$, while ${}^{-}D^{k}_{\rho\rho'}(-\mathbf{1}) = (-1)^{2k}\delta(\rho - \rho')$. The composition property is satisfied, i.e., Eq. (2.2c) under $\int_{R^+} d\rho..$, as under this measure the eigenbasis is Dirac-orthonormal and complete.

The matrix elements between the J^{γ}_{+} eigenfunctions can now be immediately computed:

$$ {}^{+}D^{k}_{\rho\rho'}\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \left({}^{+}\Phi^{k}_{\rho}, \mathbb{C}^{k}\begin{pmatrix} a & b \\ c & d \end{pmatrix} {}^{+}\Phi^{k}_{\rho'} \right) = {}^{-}D^{k}_{\rho\rho'}\begin{pmatrix} d & -c \\ -b & a \end{pmatrix}. \tag{3.15} $$

The matrix elements (3.14) and (3.15) are manifestly unitary. This is a direct consequence of the unitarity of the canonical transforms.

The E(1) reduction shows in particular that the Bessel functions in ${}^{+}\Phi^{k}_{\rho}(r)$ are self-reciprocating[55] under the k-radial canonical transforms, i.e., the $\mathbb{C}^{k}_{\mathbf{g}}$-transform of ${}^{+}\Phi^{k}_{\rho}$ may be written as a multiplier function times a function of the transformed argument:

$$ \left[\mathbb{C}^{k}\begin{pmatrix} a & b \\ c & d \end{pmatrix} {}^{+}\Phi^{k}_{\rho}\right](r) $$
$$ = \left[\mathbb{C}^{k}\begin{pmatrix} a & 0 \\ c & a^{-1} \end{pmatrix} \exp(-iba^{-1}J^{\gamma}_{+})\, {}^{+}\Phi^{k}_{\rho}\right](r) $$
$$ = |a|^{-1/2}\exp(-ib\rho^2/2a)\exp(icr^2/2a)\, {}^{+}\Phi^{k}_{\rho}(r/|a|). \tag{3.16} $$

Here we have made use of the decomposition of $\mathbf{g}$ as a lower-triangular matrix times $\mathbf{M}_{+}(b/a)$ [c f. Eqs. (2.10d) and (2.10e)]; the latter factor gives rise to the phase $\exp(-ib\rho^2/2a)$ while the former is the point transformation as given by Eq. (2.5c). Similar self-reciprocation formulas hold for other subgroup-reduced matrix elements throughout this article.

ii. $SO(1,1) \subset SL(2,R)$. This matrix element[56] is essentially the Mellin transform of Eq. (3.11) with respect to the argument ρ. Again, as the general method for evaluating Mellin transforms of hypergeometric functions[39] is presented in the next section, we simply quote here the result:

$$ {}^{2}D^{k}_{\mu\mu'}\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \left({}^{2}\Phi_{\mu}, \mathbb{C}^{k}\begin{pmatrix} a & b \\ c & d \end{pmatrix} {}^{2}\Phi_{\mu'} \right) $$
$$ = {}^{1}D^{k}_{\mu\mu'}\begin{pmatrix} (a-b-c+d)/2 & (a+b-c-d)/2 \\ (a-b+c-d)/2 & (a+b+c+d)/2 \end{pmatrix} $$
$$ = e^{-i\pi k}2^{i(\mu'-\mu)}\frac{\Gamma(k-i\mu)\Gamma(k+i\mu')}{2\pi\Gamma(2k)} $$
$$ \times b^{-2k}\left(\frac{-ia}{b}\right)^{-k-i\mu'}\left(\frac{-id}{b}\right)^{-k+i\mu} $$
$$ \times\, {}_2F_1\left[\begin{matrix} k-i\mu, k+i\mu' \\ 2k \end{matrix}; \frac{1}{ad}\right]. \tag{3.17} $$

As in (3.11), we give this expression in terms of complex power functions, taking care that these variables be evaluated for points along the imaginary axis, in the principal sheet of the power functions, where the cut is chosen along the negative real half-axis.[57] An alternative expression in terms of the absolute values of a, b, and d may be written through

$$ b^{-2k}(-ia/b)^{-k-i\mu'}(-id/b)^{-k+i\mu'} $$
$$ = (\operatorname{sgn} b)^{2k}\exp(i\tfrac{1}{2}\pi[k+i\mu']\operatorname{sgn} ab) $$
$$ \times\exp(i\tfrac{1}{2}\pi[k-i\mu]\operatorname{sgn} bd)|a|^{-k-i\mu'} $$
$$ \times|b|^{i(\mu'-\mu)}|d|^{-k+i\mu}. \tag{3.18} $$

One can obtain from these expressions the diagonal and anti-diagonal cases

$$ {}^{2}D^{k}_{\mu\mu'}\begin{pmatrix} a & 0 \\ 0 & a^{-1} \end{pmatrix} = (\operatorname{sgn} a)^{2k}|a|^{-2i\mu}\delta(\mu - \mu'), \tag{3.19} $$

$$ {}^{2}D^{k}_{\mu\mu'}\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} $$
$$ = e^{-i\pi k}2^{-2i\mu}[\Gamma(k-i\mu)/\Gamma(k+i\mu)]\delta(\mu + \mu') $$
$$ = \exp i(-\pi k - 2\mu\ln 2 + 2\arg[k-i\mu])\delta(\mu+\mu'). \tag{3.20} $$

From (3.19) we verify that ${}^{2}\mathbf{D}^{k}(\pm\mathbf{1}) = (\pm 1)^{2k}\mathbf{1}$, while (3.20) is the Fourier–Hankel transform in the Mellin basis. The representations are unitary in all cases. The direct evaluation of (3.20) allows us to give alternative forms for (3.17) through

$$ {}^{2}D^{k}_{\mu,\mu'}\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \int_{-\infty}^{\infty} d\mu''\, {}^{2}D^{k}_{\mu,\mu''}\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} {}^{2}D^{k}_{\mu'',\mu'}\begin{pmatrix} -c & -d \\ a & b \end{pmatrix} $$
$$ = e^{-i\pi k}2^{-2i\mu}[\Gamma(k-i\mu)/\Gamma(k+i\mu)]\, {}^{2}D^{k}_{-\mu,\mu'}\begin{pmatrix} -c & -d \\ a & b \end{pmatrix} $$
$$ = e^{-i\pi k}2^{2i\mu'}[\Gamma(k+i\mu')/\Gamma(k-i\mu')]\, {}^{2}D^{k}_{\mu,-\mu'}\begin{pmatrix} b & -a \\ d & -c \end{pmatrix}. \tag{3.21} $$

iii. $SO(2) \subset SL(2,R)$. This matrix element is the inner product of Eq. (3.7) with ${}^{0}\Phi^{k}_{m}$. The resulting integral in available from the tables.[58] It is

$$
\begin{aligned}
{}^0D^k_{mm'}\begin{pmatrix} a & b \\ c & d \end{pmatrix} &= \left({}^0\Phi^k_m, \mathbb{C}^k\begin{pmatrix} a & b \\ c & d \end{pmatrix}{}^0\Phi^k_{m'}\right) \\
&= 2^{2k}\Gamma(m+m')[\Gamma(k+m)\Gamma(1-k+m)\Gamma(k+m')\Gamma(1-k+m')]^{-1/2} \\
&\quad \times [(d-a)-i(b+c)]^{m-k}[(a-d)-i(b+c)]^{m'-k}[(a+d)+i(b-c)]^{-m-m'} \\
&\quad \times {}_2F_1\left[\begin{matrix} -(m-k), -(m'-k) \\ 1-m-m' \end{matrix}; \frac{a^2+b^2+c^2+d^2+2}{a^2+b^2+c^2+d^2-2}\right] \\
&= (-1)^{m-k}\Gamma(m+m')[\Gamma(k+m)\Gamma(1-k+m)\Gamma(k+m')\Gamma(1-k+m')]^{1/2} \\
&\quad \times \alpha^{*-m-m'}\beta^{m-k}\beta^{*m'-k}\,{}_2F_1\left[\begin{matrix} -(m-k), -(m'-k) \\ 1-m-m' \end{matrix}; \frac{|\alpha|^2}{|\beta|^2}\right]. \qquad (3.22)
\end{aligned}
$$

In the last expression we have given the SL(2, R) representation matrix elements in terms of the complex SU(1, 1) parameters of Bargmann through (A3). The hypergeometric function appearing above is actually a polynomial of degree $\min(m-k, m'-k)$. One also checks easily that ${}^0\mathbf{D}^k(\pm\mathbf{1}) = (\pm 1)^{2k}\mathbf{1}$ and that the representation matrix is unitary.

The expression (3.22) for the UIR matrix elements gives the value of the group unit at the point at infinity of the hypergeometric function. We can bring[59] (3.22) to coincide with the form given by Bargmann,[60] which values the group unit at the zero of the hypergeometric function, taking care to distinguish the cases $m \geqslant m'$ from $m \leqslant m'$.

D. The D_k^- representations

The discrete representation series D_k^- is obtained from the D_k^+ series through the group automorphism (2.12c), i.e., $\mathbf{D}^{k(-)}(\mathbf{g}) = \mathbf{D}^{k(+)}(\mathbf{g}^A)$. The basis functions ${}^\alpha\Phi^k_\lambda(r)$ are now to be taken as eigenfunctions of the algebra generators $\tilde\sigma_\alpha J^\gamma_\alpha$, where $\tilde\sigma_\alpha = -1$ for $\alpha = 0, 1, +, -$ and $\tilde\sigma_\alpha = 1$ for $\alpha = 2$, with eigenvalue $\tilde\sigma_\alpha$ times the eigenvalue of the J^γ_α representation generator. In addition, for the SO(2) subgroup chain, if we are to follow Bargmann's phase convention[49] of having the raising and lowering operators represented by matrices with positive elements, (2.12b) implies that the phase of the basis functions ${}^0\Phi^k_m(r)$ must be multiplied by a sign factor $\tau_0^m = (-1)^{m-k}$ [recall (3.5b)]. For convenience we set $\tau_\alpha^\lambda = 1$ for all other $\alpha \neq 0$. We can then write all D_k^- mixed-basis and subgroup-reduced matrix elements in terms of the D_k^+ expressions given above in this section as

$$
{}^{\alpha,\beta}D^{k(-)}_{\lambda,\lambda'}\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \tau^\lambda_\alpha \tau^{\lambda'}_\beta\, {}^{\alpha,\beta}D^{k(+)}_{\sigma_\alpha\lambda,\sigma_\beta\lambda'}\begin{pmatrix} a & -b \\ -c & d \end{pmatrix}, \qquad (3.23a)
$$

$$
\sigma_\alpha = \begin{cases} 1, & \alpha = 2, +, - \\ -1, & \alpha = 0, 1 \end{cases},
\qquad
\tau^\lambda_\alpha = \begin{cases} 1, & \alpha = 1, 2, +, - \\ (-1)^{m-k}, & \alpha = 0 \end{cases}. \qquad (3.23b)
$$

4. THE CONTINUOUS SERIES C_q^ϵ

In this section we follow the same general strategy in finding the unitary irreducible matrix elements (or integral kernels) corresponding to the continuous series C_q^ϵ. The difference is that here we use the hyperbolic canonical transforms of Sec. 2C, rather than the radial ones employed above. The function space has now two components, the inner product is given by Eq. (2.14), the group action by (2.15), and the subgroup generators by Eqs. (2.16). The noncompact subgroup generators $\mathbf{J}_-$ and $\mathbf{J}_2$ of $\mathrm{E}(1)_-$ and $\mathrm{SO}(1,1)_2$ are just as simple as those in the last section—although their spectra are doubly degenerate. The eigenfunctions of $\mathbf{J}_0$ and $\mathbf{J}_1$ are in general less simple: linear combinations of the first and second solutions of the confluent hypergeometric differential equation. Although the $\mathbf{J}_0$ eigenfunctions sum up to a Whittaker function,[61] the $\mathbf{J}_1$ eigenfunctions do not.

A. The subgroup-adapted eigenfunctions

i. $E(1) \subset SL(2, R)$. The simplest operator in the parabolic orbit, as for its discrete counterpart, is $\mathbf{J}_-$, given by (2.16c). Its generalized eigenfunctions are

$$
{}^-\mathbf{\Psi}_\rho(r) = \begin{cases} \begin{pmatrix} \delta(\rho - r) \\ 0 \end{pmatrix}, & \rho \geqslant 0 \\ \begin{pmatrix} 0 \\ \delta(|\rho| - r) \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} {}^-\mathbf{\Psi}_{|\rho|}(r), & \rho < 0, \end{cases} \qquad (4.1a)
$$

with eigenvalue $(\operatorname{sgn}\rho)\rho^2/2$. The spectrum of $\mathbf{J}_-$ in the continuous series UIRs thus ranges over R, rather than over R^+ as in the discrete ones. In (4.1a) a definite choice of phase has been made. The set of functions (4.1a) is Dirac-orthonormal and complete in $\mathscr{L}^2_{\mathrm{II}}(R^+)$:

$$
({}^-\mathbf{\Psi}_\rho, {}^-\mathbf{\Psi}_{\rho'}) = \delta(\rho - \rho'), \quad \int_{-\infty}^{\infty} d\rho\, \Psi_{\rho,j}(r)^* \Psi_{\rho,j'}(r') = \delta_{j,j'}\delta(r - r'). \qquad (4.1b)
$$

From Eqs. (4.1a) and the hyperbolic inverse Fourier canonical transform [Eqs. (2.15) for $\mathbf{F}^{-1}$ as given in (2.11b)] we find the $\mathbf{J}^\gamma_+$ generalized eigenfunctions to be

$$
{}^+\mathbf{\Psi}^{\epsilon k}_\rho(r) = \frac{(\rho r)^{1/2}}{2\pi}\begin{pmatrix} H^{\epsilon k}_{1,1}(-\rho r) \\ H^{\epsilon k}_{-1,1}(-\rho r) \end{pmatrix} = \begin{pmatrix} (2\pi)^{-1/2}[e^{-i\pi/4}W_{0,2k-1}(2i\rho r) + p_\epsilon e^{i\pi/4}W_{0,2k-1}(-2i\rho r)] \\ (2/\pi)^{1/2}p_\epsilon g_\epsilon(k) W_{0,2k-1}(2\rho r) \end{pmatrix}, \quad \rho \geqslant 0, \qquad (4.2a)
$$

$$ {}^{+}\Psi_{\rho}^{\epsilon k}(r) = \frac{(|\rho| r)^{1/2}}{2\pi} \begin{pmatrix} H_{1,-1}^{\epsilon k}(\rho r) \\ H_{-1,-1}^{\epsilon k}(\rho r) \end{pmatrix} = p_\epsilon \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} {}^{+}\Psi_{|\rho|}^{\epsilon k}(r), \quad \rho \leqslant 0, \tag{4.2b} $$

where the $H_{jj'}^{\epsilon k}(\zeta)$ are given in (2.15d)–(2.15e). We have expressed the Hankel and Macdonald functions in terms of Whittaker functions[61] of argument phase 0 and $\pm\pi/2$. As in (4.1a), (4.2) correspond to the eigenvalue $(\operatorname{sgn}\rho)\rho^2/2 \in R$. Recall that for the continuous nonexceptional series $2k-1=2is$, $s \geqslant 0$ for $\epsilon=0$ and $s>0$ for $\epsilon=\frac{1}{2}$, while for the exceptional interval $\epsilon=0$, $2k-1=2\sigma$, $0<\sigma<\frac{1}{2}$.

ii. $SO(1,1) \subset SL(2,R)$. The simplest operator in the hyperbolic orbit is $\mathbf{J}_2$, as given by (2.16b). Notice that the signs of the entries are the same. The spectrum of J_2 covers R once in $\mathscr{L}^2(R^+)$, while that of $\mathbf{J}_2$ does so twice in $\mathscr{L}^2_{\text{II}}(R^+)$. The normalized eigenfunctions ${}^2\Psi_{\kappa,\mu}(r)$ thus require an extra dichotomic index $\kappa = \pm 1$, and are

$$ {}^2\Psi_{\kappa,\mu}(r) = (2\pi)^{-1/2} \begin{pmatrix} 1 \\ \kappa \end{pmatrix} r^{-1/2+2i\mu}, \quad \kappa = \pm 1, \mu \in R, \tag{4.3a} $$

belonging to the eigenvalue μ under $\mathbf{J}_2$. The dichotomic index κ has been introduced by Mukunda and Radhakrishnan[11]; it can be seen as the eigenvalue of ${}^2\Psi_{\kappa,\mu}(r)$ under a transformation in $\mathscr{L}^2_{\text{II}}(R^+)$ given by $A: f_j(r) \to f_{-j}(r)$, which may be represented[62] as

$$ \mathbf{A} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}. $$

The statement of Dirac orthonormality and completeness is

$$ ({}^2\Psi_{\kappa,\mu}, {}^2\Psi_{\kappa',\mu'}) = \delta_{\kappa,\kappa'}\delta(\mu-\mu'), $$
$$ \sum_{\kappa=\pm 1} \int_{-\infty}^{\infty} d\mu \, {}^2\Psi_{\kappa,\mu,j}(r)^* \, {}^2\Psi_{\kappa,\mu,j'}(r') = \delta_{jj'}\delta(r-r'). \tag{4.3b} $$

The eigenfunctions ${}^1\Psi_{\kappa,\mu}^{\epsilon,k}(r)$ of $\mathbf{J}_1^\gamma$ [Eq. (2.16a)], on the other hand, using (2.11a) are given by[63]

$$ \begin{aligned} {}^1\Psi_{\kappa,\mu,j}^{\epsilon,k}(r) &= [\mathbb{C}_{\mathbf{S}}^{\epsilon,k}\, {}^2\Psi_{\kappa,\mu}]_j(r) = (-1)^{2\epsilon}(2\pi)^{-3/2}2^{i\mu+1}g_\epsilon(k) \\ &\times [e^{-ij\pi(k+i\mu)/2}\{p_\epsilon G_{\mu,j}^{k}(r) + G_{\mu,j}^{1-k}(r)\} \\ &+ \kappa e^{ij\pi(k+i\mu)/2}\{G_{\mu,j}^{k}(r) + G_{\mu,j}^{1-k}(r)\}], \end{aligned} \tag{4.4a} $$

$$ \begin{aligned} G_{\mu,j}^{k}(r) &= \Gamma(1-2k)\Gamma(k+i\mu)r^{2k-1}e^{ijr^2/2} \\ &\times {}_1F_1(k-i\mu; 2k; -ijr^2). \end{aligned} \tag{4.4b} $$

They are obtained from Eqs. (4.17)–(4.18), below.

iii. $SO(2) \subset SL(2,R)$. For the continuous series C_q^ϵ of UIRs belonging to the nonexceptional or exceptional series, the eigenfunctions of the compact generator $\mathbf{J}_0^\gamma$ are given by

$$ {}^0\Psi_m^{\epsilon,k}(r) = \frac{g_\epsilon(k)}{\pi r^{1/2}} \times \begin{pmatrix} (-1)^{m-\epsilon}[2\Gamma(k-m)\Gamma(1-k-m)]^{1/2}W_{m,k-1/2}(r^2) \\ [2\Gamma(k+m)\Gamma(1-k+m)]^{1/2}W_{-m,k-1/2}(r^2) \end{pmatrix}. \tag{4.5a} $$

These eigenfunctions belong to the eigenvalue m under $\mathbf{J}_0^\gamma$. We have chosen the phase in accordance with Bargmann's convention,[64] i.e., such that the raising and lowering operators have positive matrix elements. They are orthonormal and complete in $\mathscr{L}^2_{\text{II}}(R^+)$:

$$ ({}^0\Psi_m^{\epsilon,k}, {}^0\Psi_{m'}^{\epsilon,k}) = \delta_{m,m'}, $$
$$ \sum_{m\in Z} {}^0\Psi_{m,j}^{\epsilon,k}(r)^* \, {}^0\Psi_{m,j'}^{\epsilon,k}(r') = \delta_{j,j'}\delta(r-r'). \tag{4.5b} $$

B. The mixed-basis matrix elements

i. $E(1) \subset SL(2,R) \supset SO(2)$. Application of $\mathbb{C}_{\mathbf{g}}^{\epsilon k}$ decomposed as in (3.6) gives

$$ \begin{aligned} [\mathbb{C}_{\mathbf{g}}^{\epsilon,k}\, {}^0\Psi_m^{\epsilon,k}]_j(r) &= \left(\frac{a-ib}{a+ib}\right)^m (a^2+b^2)^{-1/4} \\ &\times \exp\left(\frac{ijr^2[ac+bd]}{2[a^2+b^2]}\right) \\ &\times {}^0\Psi_{m,j}^{\epsilon,k}(r/[a^2+b^2]^{1/2}). \end{aligned} \tag{4.6} $$

This formula displays the Whittaker functions (4.5a) as self-reciprocating under the corresponding hyperbolic canonical transforms.[65] Since the $\mathbf{J}_-$ eigenfunctions are simple Dirac deltas, we obtain[66]

$$ \begin{aligned} {}^{-,0}D_{\rho,m}^{\epsilon,k}\begin{pmatrix} a & b \\ c & d \end{pmatrix} &= \left({}^{-}\Psi_\rho, \mathbb{C}^{\epsilon,k}\begin{pmatrix} a & b \\ c & d \end{pmatrix} {}^0\Psi_m^{\epsilon,k}\right) = \left[\mathbb{C}^{\epsilon,k}\begin{pmatrix} a & b \\ c & d \end{pmatrix} {}^0\Psi_m^{\epsilon,k}\right]_{\operatorname{sgn}\rho}(|\rho|) = {}^{+,0}D_{\rho,m}^{\epsilon,k}\begin{pmatrix} c & d \\ -a & -b \end{pmatrix} \\ &= (-\operatorname{sgn}\rho)^{m-\epsilon}\left(\frac{a-ib}{a+ib}\right)^m \frac{g_\epsilon(k)}{\pi|\rho|^{1/2}} \exp\left(\frac{\rho^2[a-ib\operatorname{sgn}\rho][d-ic\operatorname{sgn}\rho]}{2(a^2+b^2)}\right) \\ &\times \left[\left\{\Gamma(1-2k)\left[\frac{2\Gamma(k-m_\rho)}{\Gamma(1-k-m_\rho)}\right]^{1/2}\left[\frac{\rho^2}{a^2+b^2}\right]^k \right.\right. \\ &\left.\left.\times {}_1F_1\left[\begin{matrix} k-m_\rho \\ 2k \end{matrix}; \frac{\rho^2}{a^2+b^2}\right]\right\} + \{k \leftrightarrow 1-k\}\right], \\ & m_\rho = m\operatorname{sgn}\rho. \end{aligned} \tag{4.7} $$

The overlap coefficient between the $E(1)_-$ and $SO(2)_0$ subgroup chains is easily found from (4.7) for $\mathbf{g} = \mathbf{1}$ and is ${}^0\Psi^{\epsilon,k}_{m,\operatorname{sgn}\rho}(|\rho|)$. This change of basis thus represents basically the Whittaker series expansion ($m \in Z$) of a function of $\rho \in R$.

ii. $SO(1,1) \subset SL(2,R) \supset SO(2)$. The evaluation of this mixed-basis matrix element will be given in some detail because the method presented here has been used to obtain all the matrix elements carrying SO(1, 1) reductions, both in the continuous and in the discrete series in the last section, where its discussion was postponed. The method[32] essentially consists of a Taylor expansion of $[\mathbb{C}^k_{\mathbf{g}}\, {}^0\mathbf{\Psi}^{\epsilon,k}_m](r)$ followed by a Mellin–Barnes transformation.

The Taylor expansion of the Gaussian and ${}_1F_1$ functions appearing in (4.7) [for $|\rho| \mapsto r$ and $\operatorname{sgn}\rho \mapsto j$] yields, after an exchange of summations which allows us to recognize one of them as a ${}_2F_1$ series,

$$[\mathbb{C}^k_{\mathbf{g}}\, {}^0\mathbf{\Psi}^{\epsilon,k}_m]_j(r)$$

$$= (-j)^{m-\epsilon}\left(\frac{a-ib}{a+ib}\right)^m \frac{g_\epsilon(k)}{\pi}\,[2\Gamma(k-jm)\Gamma(1-k-jm)]^{1/2} \times [X^j_k + X^j_{1-k}], \tag{4.8}$$

where

$$X^j_k = \left(-\frac{q_j}{t}\right)^{1/2-jm}\left(\frac{r}{|\bar{a}|}\right)^{1/2} \times \sum_{n=0}^{\infty}\frac{(-1)^n\Gamma(1-2k-n)(-q_j r^2)^{k-1/2+n}}{n!\,\Gamma(1-k-jm-n)} \times {}_2F_1\left[\begin{matrix}k-jm,\ 1-k-jm\\ 1-k-jm-n\end{matrix};\ 1+\frac{q_j}{t}\right], \tag{4.9a}$$

and where we are using the abbreviations from (3.6b) for $\bar{a}$ and $\bar{c}$, and

$$q_j = -(1-ij\bar{a}\,\bar{c})/2\bar{a}^2, \quad t - 1/\bar{a}^2, \quad j = \pm 1. \tag{4.9b}$$

The terms in the sum over n are now recognized as the residues, at $z = z_n = -k-n,\ -1+k-n$, $(n = 0, 1, 2, \ldots)$ of the following meromorphic function:

$$\chi^j(z) = \left(-\frac{q_j}{t}\right)^{1/2-jm}\left(\frac{r}{|\bar{a}|}\right)^{1/2}(-q_j r^2)^{-1/2-z} \times \frac{\Gamma(k+z)\Gamma(1-k+z)}{\Gamma(1+z-jm)} \times {}_2F_1\left[\begin{matrix}k-jm,\ 1-k-jm\\ 1+z-jm\end{matrix};\ 1+\frac{q_j}{t}\right]. \tag{4.10}$$

Since for fixed ζ, $\Gamma(c)^{-1}\,{}_2F_1(a,b;c;\zeta)$ is an entire function of the parameters, $\chi^j(z)$ is a meromorphic function falling to zero rapidly as $|z| \to \infty$ in the region $\operatorname{Re} z < 0$. The singularities of $\chi^j(z)$ are simple poles arising from the Gamma functions in the factor $\Gamma(k+z)\Gamma(1-k+z)$ and are located at the points $z = z_n$.

For the nonexceptional UIRs, $k - \frac{1}{2}$ is pure imaginary and the poles lie symmetrically with respect to the real axis. For the exceptional UIRs k is real, but no two pole points z_n are coincident.

If we now choose a closed contour $\mathscr{C}$ consisting of the infinite semicircle $\mathscr{S}$ on the left, and the imaginary axis, we obtain

$$\frac{1}{2\pi i}\oint_{\mathscr{C}} dz\,\chi(z) = \sum_{n=0}^{\infty}\operatorname{Res}[\chi(z)]_{z=-k-n} + \sum_{n=0}^{\infty}\operatorname{Res}[\chi(z)]_{z=-1+k-n}. \tag{4.11}$$

The first and second terms on the right-hand side, by our previous analysis, are respectively equal to X^j_k and X^j_{1-k} and hence the integral in (4.11) vanishes on $\mathscr{S}$, as can be easily verified. We obtain

$$X^j_k + X^j_{1-k} = \frac{1}{2\pi}\int_{-\infty}^{\infty} d\lambda\,\chi^j(-i\lambda). \tag{4.12}$$

This expression, replaced in (4.8), represents the solution of the problem of finding the integral of ${}^2\mathbf{\Psi}_{\kappa,\mu}(r)$ with it, since the latter integral is essentially the Mellin transform of (4.8), integrated over r for the value $-\mu$; we note that (4.12) is expressed as an inverse Mellin transform of the coefficient (function of λ) of the $r^{-1/2+2i\lambda}$ factor in (4.10). The value of this coefficient for $z = -\mu$ and summed over the two j components will be the inner product of ${}^2\mathbf{\Psi}_{\kappa,\mu}$ with (4.8). We thus obtain[67]

$$\begin{aligned}
{}^{2,0}D^{\epsilon,k}_{\kappa,\mu;m}\begin{pmatrix}a & b\\ c & d\end{pmatrix} &= \left({}^2\mathbf{\Psi}_{\kappa,\mu},\ \mathbb{C}^{\epsilon,k}\begin{pmatrix}a & b\\ c & d\end{pmatrix}{}^0\mathbf{\Psi}^{\epsilon,k}_m\right)\\
&= {}^{1,0}D^{\epsilon,k}_{\kappa,\mu;m}\begin{pmatrix}2^{-1/2}(a-c) & 2^{-1/2}(b-d)\\ 2^{-1/2}(a+c) & 2^{-1/2}(b+d)\end{pmatrix}\\
&= g_\epsilon(k)\pi^{-3/2}\Gamma(k-i\mu)\Gamma(1-k-i\mu)(a+ib)^{-m-i\mu}(a-ib)^{m-i\mu}\\
&\quad\times\sum_{j=\pm 1}(-j)^{m-\epsilon}\kappa^{(1-j)/2}\frac{[\Gamma(k-jm)\Gamma(1-k-jm)]^{1/2}}{\Gamma(1-i\mu-jm)}\\
&\quad\times{}_2F_1\left[\begin{matrix}k-i\mu,\ 1-k-i\mu\\ 1-i\mu-jm\end{matrix};\ \tfrac{1}{2}(a-ijb)(d-ijc)\right].
\end{aligned} \tag{4.13}$$

The overlap coefficient between these two chains[68] in the continuous series is obtained by setting $\mathbf{g} = \mathbf{1}$:

$$({}^2\mathbf{\Psi}_{\kappa,\mu},\ {}^0\mathbf{\Psi}^{\epsilon,k}_m) = g_\epsilon(k)\pi^{-3/2}\Gamma(k-i\mu)\Gamma(1-k-i\mu)\sum_{j=\pm 1}(-j)^{m-\epsilon}\kappa^{(1-j)/2}\frac{[\Gamma(k-jm)\Gamma(1-k-jm)]^{1/2}}{\Gamma(1-i\mu-jm)} \times {}_2F_1\left[\begin{matrix}k-i\mu,\ 1-k-i\mu\\ 1-i\mu-jm\end{matrix};\ \tfrac{1}{2}\right]. \tag{4.14}$$

iii. $E(1)\subset SL(2,R)\supset SO(1,1)$. As in all cases involving E(1), the calculation here consists in applying $\mathbf{C}_{\mathbf{g}}^{\epsilon,k}$ on ${}^2\boldsymbol{\Psi}_{\kappa,\mu}$, that is, performing the integral in

$$[\mathbf{C}_{\mathbf{g}}^{\epsilon,k}\,{}^2\boldsymbol{\Psi}_{\kappa,\mu}]_j(r) = \sum_{j'=\pm 1}\int_0^\infty dr'\,[\mathbf{C}_{\mathbf{g}}^{\epsilon,k}]_{j,j'}(r,r')\,{}^2\Psi_{\kappa,\mu,j'}(r'), \tag{4.15}$$

of the kernel $[\mathbf{C}_{\mathbf{g}}^{\epsilon,k}]_{j,j'}(r,r')$ with the Mellin basis function. We resort to the expansion of the hyperbolic canonical transform kernel in Taylor series and to the Mellin–Barnes contour deformation presented above. We obtain

$$[\mathbf{C}_{\mathbf{g}}^{\epsilon,k}\,{}^2\boldsymbol{\Psi}_{\kappa,\mu}]_j(r) = A^{\,\epsilon,k}_{\,\mathbf{g};\kappa,\mu,j}(r) + \kappa B^{\,\epsilon,k}_{\,\mathbf{g};\kappa,\mu,j}(r), \tag{4.16}$$

where

$$A\begin{pmatrix} a & b\\ c & d\end{pmatrix}^{\epsilon,k}_{\kappa,\mu,1}(r) = \kappa p_\epsilon A\begin{pmatrix} -a & b\\ c & -d\end{pmatrix}^{\epsilon,k}_{\kappa,\mu,-1}(r) = \frac{(\mathrm{sgn}b)^{2\epsilon}g_\epsilon(k)}{(2\pi)^{3/2}|b|}\left(\frac{-ia}{2b}\right)^{-1/2-i\mu} r^{1/2}\exp\left(\frac{idr^2}{2b}\right)$$
$$\times\left[p_\epsilon\left\{\Gamma(1-2k)\Gamma(k+i\mu)\left(\frac{ir^2}{2ab}\right)^{k-1/2}{}_1F_1\left[\begin{matrix}k+i\mu\\ 2k\end{matrix};\frac{-ir^2}{2ab}\right]\right\} + \{k\leftrightarrow 1-k\}\right], \tag{4.17}$$

$$B\begin{pmatrix} a & b\\ c & d\end{pmatrix}^{\epsilon,k}_{\kappa,\mu,1}(r) = \kappa p_\epsilon B\begin{pmatrix} -a & b\\ c & -d\end{pmatrix}_{\kappa,\mu,-1}(r) = \frac{(\mathrm{sgn}\,b)^{2\epsilon}g_\epsilon(k)}{(2\pi)^{3/2}|b|}\left(\frac{ia}{2b}\right)^{-1/2-i\mu} r^{1/2}\exp\left(\frac{idr^2}{2b}\right)$$
$$\times\left[\left\{\Gamma(1-2k)\Gamma(k+i\mu)\left(\frac{-ir^2}{2ab}\right)^{k-1/2}{}_1F_1\left[\begin{matrix}k+i\mu\\ 2k\end{matrix};\frac{-ir^2}{2ab}\right]\right\} + \{k\leftrightarrow 1-k\}\right]$$
$$= \frac{(\mathrm{sgn}b)^{2\epsilon+1}g_\epsilon(k)}{(2\pi)^{3/2}}\left(\frac{ia}{2b}\right)^{-i\mu} r^{-1/2}\exp\left(\frac{ir^2}{4ab}[ad+bc]\right)\Gamma(k+i\mu)\Gamma(1-k+i\mu)W_{-i\mu,k-1/2}(-ir^2/2ab), \tag{4.18}$$

which come, respectively, from the Mellin transforms of the on- and off-diagonal integral kernel elements. We remind the reader again that the complex power functions are to be evaluated inthe principal sheet.

Since the $E(1)_-$ basis has simple Dirac deltas, we immediately obtain[69]

$$^{-,2}D^{\,\epsilon,k}_{\rho;\kappa,\mu}\begin{pmatrix} a & b\\ c & d\end{pmatrix} = \left({}^-\boldsymbol{\Psi}_\rho,\,\mathbf{C}^{\epsilon,k}\begin{pmatrix} a & b\\ c & d\end{pmatrix}{}^2\boldsymbol{\Psi}_{\kappa,\mu}\right)$$
$$= A\begin{pmatrix} a & b\\ c & d\end{pmatrix}^{\epsilon,k}_{\kappa,\mu,\mathrm{sgn}\rho}(|\rho|) + B\begin{pmatrix} a & b\\ c & d\end{pmatrix}^{\epsilon,k}_{\kappa,\mu,\mathrm{sgn}\rho}(|\rho|). \tag{4.19}$$

The overlap coefficient between these two chains may be obtained upon letting $\mathbf{g}\rightarrow\mathbf{1}$, or directly as

$$({}^-\boldsymbol{\Psi}_\rho,\,{}^2\boldsymbol{\Psi}_{\kappa,\mu}) = {}^2\Psi_{\kappa,\mu,\mathrm{sgn}\rho}(|\rho|). \tag{4.20}$$

C. The matrix elements in the subgroup bases

i. $E(1)\subset SL(2,R)$. The integral kernel representations of SL(2, R) in this chain are given by the hyperbolic canonical transform integral kernel, which we may rewrite in terms of the confluent hypergeometric function as follows:

$$^-D^{\,\epsilon,k}_{\rho,\rho'}\begin{pmatrix} a & b\\ c & d\end{pmatrix} = \left({}^-\boldsymbol{\Psi}_\rho,\,\mathbf{C}^{\epsilon,k}\begin{pmatrix} a & b\\ c & d\end{pmatrix}{}^-\boldsymbol{\Psi}_{\rho'}\right)$$
$$= C^{\,\epsilon,k}\begin{pmatrix} a & b\\ c & d\end{pmatrix}_{\mathrm{sgn}\rho,\mathrm{sgn}\rho'}(|\rho|,|\rho'|)$$
$$= (\mathrm{sgn}b)^{2\epsilon}p_\epsilon^{(1+\mathrm{sgn}\rho')/2}(\pi|b|)^{-1}g_\epsilon(k)$$
$$\times|\rho\rho'|^{1/2}\exp(i[dj\rho^2 - 2\eta\rho\rho' + aj'\rho'^2]/2b)$$
$$\times\left[\left\{\Gamma(1-2k)\left|\frac{\rho\rho'}{2b}\right|^{2k-1}\right.\right.$$
$$\left.\times{}_1F_1\left[\begin{matrix}2k-1/2\\ 4k-1\end{matrix};\frac{2i\rho\rho'}{\eta b}\right]\right\}$$
$$\left.+\{k\leftrightarrow 1-k\}\right], \tag{4.21}$$

where $\eta = 1$ for $\mathrm{sgn}\rho = \mathrm{sgn}\rho'$, and $\eta = -i$ for $\mathrm{sgn}\rho\neq\mathrm{sgn}\rho'$. In particular, for the $b = 0$ subgroup we have, as from Eqs. (2.15h),

$$^-D^{\,\epsilon,k}_{\rho,\rho'}\begin{pmatrix} a & 0\\ c & a^{-1}\end{pmatrix}$$
$$= (\mathrm{sgn}a)^{2k}|a|^{-1/2}\exp(i(\mathrm{sgn}\rho)c\rho^2/2a)\delta(\rho'-\rho/|a|). \tag{4.22}$$

In the $E(2)_+$ reduction, as in (3.17),

$$^+D^{\,\epsilon,k}_{\rho,\rho'}\begin{pmatrix} a & b\\ c & d\end{pmatrix} = \left({}^+\boldsymbol{\Psi}^{\epsilon,k}_\rho,\,\mathbf{C}^{\epsilon,k}\begin{pmatrix} a & b\\ c & d\end{pmatrix}{}^+\Psi^{\,\epsilon,k}_{\rho'}\right)$$
$$= {}^-D^{\,\epsilon,k}_{\rho,\rho'}\begin{pmatrix} d & -c\\ -b & a\end{pmatrix}. \tag{4.23}$$

ii. $SO(1,1)\subset SL(2,R)$. These matrix elements are essentially the Mellin transforms of (4.16)–(4.18), and can be obtained by the same technique[32] of Taylor expansion and Mellin–Barnes contour deformation. The Taylor expansion of, for example, the function (4.17) yields

$$A\begin{pmatrix} a & b\\ c & d\end{pmatrix}^{\epsilon,k}_{\kappa,\mu,1}(r)$$
$$= \frac{(-\mathrm{sgn}b)^{2\epsilon}g_\epsilon(k)}{(2\pi)^{3/2}|b|}\left(\frac{-ir^2}{2ab}\right)^{-1/2-i\mu} r^{1/2}[Y_k + p_\epsilon Y_{1-k}], \tag{4.24}$$

with

$$Y_k = \exp(i\pi[2k-1][\alpha+\beta]/4)$$
$$\times\Gamma(1-2k)\Gamma(k+i\mu)|ad|^{1/2-k}$$
$$\times\sum_{n=0}^\infty\frac{(-1)^n}{n!}\left(\frac{-idr^2}{2b}\right)^{-1/2+k+n}{}_2F_1\left[\begin{matrix}-n,\,k+i\mu\\ 2k\end{matrix};\frac{1}{ad}\right], \tag{4.25}$$

where we denote for brevity $\alpha = \mathrm{sgn}(ab)$, $\beta = \mathrm{sgn}(bd)$. The terms in this series can be identified as the residues of the meromorphic function

$\nu_k(z)$

$$= \Gamma(k+z)\left(\frac{-idr^2}{2b}\right)^{-1/2-z} {}_2F_1\left[\begin{matrix} k+z, k+i\mu \\ 2k \end{matrix}; \frac{1}{ad}\right] \quad (4.26)$$

at the simple poles at $z = z_n = -k-n$. Through the same argument as in (4.9)–(4.12), we may express

$$Y_k = \exp(i\pi[2k-1][\alpha+\beta]/4)\Gamma(1-2k) \times \Gamma(k+i\mu)|ad|^{1/2-k}\frac{1}{2\pi}\int_{-\infty}^{\infty} d\lambda\, \nu_k(-i\lambda). \quad (4.27)$$

As before, the function $\nu_k(z)$ on the integration contour in (4.27) contains the kernel $r^{-1/2+2i\lambda}$, so (4.27) is the inverse Mellin transform of the coefficient of that term in (4.26). The corresponding Mellin transform of B term (4.18) follows (4.24)–(4.27) with the same meromorphic function (4.26), but with different linear combination coefficients which originate from the corresponding coefficients in the two summands of (4.17) vs (4.18). We consequently find[70]

$$\begin{aligned} {}^2D^{\epsilon,k}_{\kappa,\mu,\kappa',\mu'}\begin{pmatrix} a & b \\ c & d\end{pmatrix} &= \left({}^2\Psi_{\kappa,\mu}, \mathbb{C}^{\epsilon k}\begin{pmatrix} a & b \\ c & d\end{pmatrix} {}^2\Psi_{\kappa',\mu'}\right) \\ &= (-\operatorname{sgn} b)^{2\epsilon}(2\pi)^{-2}g_\epsilon(k) \\ &\quad \times [(\tau_k + \kappa\kappa' p_\epsilon \tau_k^{-1} + \kappa'\theta_\kappa + \kappa p_\epsilon \theta_k^{-1})T_k \\ &\quad + (p_\epsilon \tau_{1-k} + \kappa\kappa'\tau_{1-k}^{-1} + \kappa'\theta_{1-k} + \kappa p_\epsilon \theta_{1-k}^{-1})T_{1-k}], \end{aligned} \quad (4.28a)$$

$$T_k = \Gamma(1-2k)\Gamma(k-i\mu)\Gamma(k+i\mu')|a|^{-k-i\mu'}|2b|^{i(\mu'-\mu)} |d|^{-k+i\mu}\, {}_2F_1\left[\begin{matrix} k-i\mu, k+i\mu' \\ 2k \end{matrix}; \frac{1}{ad}\right], \quad (4.28b)$$

$$\tau_k = \exp(i\tfrac{1}{2}\pi[\{k+i\mu\}\operatorname{sgn} ab + \{k-i\mu\}\operatorname{sgn} bd]), \quad (4.28c)$$

$$\theta_k = \exp(i\tfrac{1}{2}\pi[-\{k+i\mu'\}\operatorname{sgn} ab + \{k-i\mu\}\operatorname{sgn} bd]). \quad (4.28d)$$

Whereas in the discrete series we are able to express the ${}^2D$ function as a meromorphic function in b, $-ia/b$, and $-id/b$ [c f. Eq. (3.19)] the corresponding continuous series functions do not have this property, and must be written in terms of powers of $|a|$, $|b|$, and $|d|$, with phase factors (4.28c) and (4.28d). This stems from the corresponding lack of meromorphicity of the hyperbolic canonical transform kernel (2.15d) and (2.15e), where the two Hankel functions are to be evaluated in the upper and lower half-planes, vis-á-vis the radial canonical transform kernel (2.5b), which is meromorphic in the group parameters. It has been pointed out before[21] that the continuous series UIRs cannot be subject to analytic continuation to a unitarizable representation of a subsemigroup of SL(2, C), such as may be done for the discrete series.[19]

Finally, it is easy to verify that our result is consistent with the expected behavior near the identity, namely

$${}^2D^{\epsilon,k}_{\kappa,\mu;\kappa',\mu'}\begin{pmatrix} a & 0 \\ 0 & a^{-1}\end{pmatrix} = (\operatorname{sgn} a)^{2\epsilon}|a|^{-2i\mu}\delta_{\kappa,\kappa'}\delta(\mu-\mu'), \quad (4.29)$$

which acts as a reproducing kernel when we sum over κ and integrate over μ as in (4.3b). The Fourier transform case is

$${}^2D^{\epsilon k}_{\kappa,\mu;\kappa',\mu'}\begin{pmatrix} 0 & 1 \\ -1 & 0\end{pmatrix} = p_\epsilon \frac{g_\epsilon(i\mu)+\kappa g_\epsilon(k)}{\sin(\pi[k+i\mu])}2^{-2i\mu}\frac{\Gamma(k-i\mu)}{\Gamma(k+i\mu)}\delta_{\kappa,p_\epsilon\kappa'}\delta(\mu+\mu'), \quad (4.30)$$

Remarks similar to those made on Eq. (4.28) apply here.

iii. SO(2)⊂*SL*(2, *R*). This matrix element should be obtained in the same way as the discrete series case given in Eq. (3.25a), with the basis functions which are now ${}^0\Psi^{\epsilon,k}_m(r)$ as given in (4.5a) [instead of the simpler ones ${}^0\Phi^k_m(r)$ in (3.5a)], and the inner product which is now the $\mathscr{L}^2_{11}(R^+)$ given in (2.14) [in place of the $\mathscr{L}^2(R^+)$ inner product (2.4)]. The application of the hyperbolic canonical transform $\mathbf{C}^{\epsilon,k}_g$ to ${}^0\Psi^{\epsilon,k}_m(r)$ is the exact analog of (3.6)–(3.7), namely, these functions are self-reciprocating[65] under $\mathbf{C}^{\epsilon,k}_g$. We can thus write

$$\begin{aligned} {}^0D^{\epsilon,k}_{m,m'}\begin{pmatrix} a & b \\ c & d\end{pmatrix} &= \left({}^0\Psi^{\epsilon,k}_m, \mathbb{C}^{\epsilon,k}\begin{pmatrix} a & b \\ c & d\end{pmatrix} {}^0\Psi^{\epsilon,k}_{m'}\right) \\ &= [(a-ib)/(a+ib)]^{m'}(a^2+b^2)^{-1/4}\sum_{j=\pm1}\int_0^\infty dr\, {}^0\Psi^{\epsilon,k}_{mj}(r)^* \\ &\quad \times \exp(ir^2[ac+bd]/2[a^2+b^2])\, {}^0\Psi^{\epsilon,k}_{m'j}(r/[a^2+b^2]^{1/2}) \\ &\left\{\begin{aligned} &= 2^{2m'}(m'!)^{-1}[\Gamma(k+m)\Gamma(1-k+m)/\Gamma(k+m')\Gamma(1-k+m')]^{1/2} \\ &\quad \times [(a+d)+i(b-c)]^{-m-m'}[(a-d)+i(b+c)]^{m-m'} \\ &\quad \times {}_2F_1(k-m', 1-k-m'; 1+m-m'; -\tfrac{1}{4}[a^2+b^2+c^2+d^2-2]), \quad m \geqslant m' \\ &= (-1)^{m'-m}2^{2m}(m!)^{-1}[\Gamma(k+m')\Gamma(1-k+m')/\Gamma(k+m)\Gamma(1-k+m)]^{1/2} \\ &\quad \times [(a+d)+i(b-c)]^{-m-m'}[(a-d)-i(b+c)]^{m'-m} \\ &\quad \times {}_2F_1(k-m, 1-k-m; 1+m'-m; -\tfrac{1}{4}[a^2+b^2+c^2+d^2-2]), \quad m \leqslant m' \end{aligned}\right\}. \end{aligned} \quad (4.31)$$

The right-hand term has been taken from Bargmann's work,[71] rewriting his phases and normalization constants, and using (A3) for the parameters. We have not been able to solve the integral in (4.31) directly: When we replace ${}^0\Psi^{\epsilon,k}_m(r)$ from (4.5a), we are confronted with a solution of a sum of two integrals whose integrands are each a product of two Whittaker functions, one of them with a rescaled argument, times an oscillating Gaussian function. This type of integral does not appear in the standard tables nor, apparently, does it yield easily to reduction to simpler forms. Bargmann's method of evaluation[38] of (4.31) does not make use of any explicit form of the basis functions $\Psi_m^{\epsilon,k}$. Instead, the function ${}^0D^{\epsilon,k}_{m,m'}(\mathbf{g})$ is shown to factorize into two exponentials of the first and third Euler angles, and the Bargmann-d function of the second Euler angle. The latter is subject to the differential relation stemming from (2.9) with $\mathbf{J}_\alpha$ expressed as operators on the group manifold. The condition ${}^0D^{\epsilon,k}_{m,m'}(\mathbf{1}) = \delta_{m,m'}$ provides the normalization and boundary conditions. This line of reasoning applies to any operator realization of the group belonging to that representation and subgroup reduction. The result provided by Bargmann[71] thus evaluates (4.31) and gives the solution for the integral. We can set $b = 0$, $a > 0$, and $r^2 = x$ in thus writing[72]

$$\sum_{j=\pm 1} |\Gamma(\tfrac{1}{2}+is-jm)\Gamma(\tfrac{1}{2}+is-jn)| \int_0^\infty dx\, x^{-1}\exp(icx/2a) W_{jm,is}(x) W_{jn,is}(x/a^2)$$

$$= \left\{ \begin{array}{l} = \left(\dfrac{\pi}{g_\epsilon(k)}\right)^2 \dfrac{2^{2n}}{n!} \left|\dfrac{\Gamma(\frac{1}{2}+is+m)}{\Gamma(\frac{1}{2}+is+n)}\right| (a+a^{-1}-ic)^{-m-n}(a-a^{-1}+ic)^{m-n} \\ \quad \times {}_2F_1(\tfrac{1}{2}+is-n, \tfrac{1}{2}-is-n; 1+m-n; -\tfrac{1}{4}[a^2+a^{-2}+c^2-2]), \quad m \geqslant n \\ = \left(\dfrac{\pi}{g_\epsilon(k)}\right)^2 \dfrac{2^{2m}}{m!} \left|\dfrac{\Gamma(\frac{1}{2}+is+n)}{\Gamma(\frac{1}{2}+is+m)}\right| (a+a^{-1}-ic)^{-m-n}(a-a^{-1}+ic)^{-m+n} \\ \quad \times {}_2F_1(\tfrac{1}{2}+is-m, \tfrac{1}{2}-is-m; 1-m+n; -\tfrac{1}{4}[a^2+a^{-2}+c^2-2]), \quad m \leqslant n \end{array} \right\} \tag{4.32}$$

where $\epsilon = 0(\frac{1}{2})$ for m, n integer (odd-half-integer), $g_\epsilon(k\,(s))$ is given by (2.15f) and (2.15g) and the range of s is, as above, $s > 0$ and $s = -i\sigma$, $0 < \sigma < \frac{1}{2}$ for $\epsilon = 0$.

D. The limits of continuous to discrete representations

i. $C_q^{1/2} \underset{q\to\frac{1}{4}}{\longrightarrow} D^+_{1/2} \dotplus D^-_{1/2}$. At the end of Sec. 2C we noted that the continuous series integral kernel $[\mathbf{C}_\mathbf{g}^{1/2,k}]_{j,j'}(r, r')$, for $k = \frac{1}{2} + is$, $s \to 0^+$, uncoupled in the sense of having its off-diagonal ($j \neq j'$) terms vanish. The hyperbolic canonical transform kernel becomes the direct sum of the $D^+_{1/2}$ radial canonical transform for the $j = 1$ component, and the $D^-_{1/2}$ one for the $j = -1$ component. In terms of the E(1) representation integral kernels,

$${}^-D^{1/2,\,1/2+is}_{\rho,\rho'}(\mathbf{g}) \underset{s\to 0^+}{\longrightarrow} \delta_{\mathrm{sgn}\rho,\mathrm{sgn}\rho'}\, {}^-D^{1/2(\mathrm{sgn}\,\rho)}_{|\rho|,|\rho'|}(\mathbf{g}), \tag{4.33}$$

as can be verified using (4.21) for the $C_q^{1/2}$ representation, (2.5b) for the $D^+_{1/2}$, and (3.23) for the $D^-_{1/2}$ representations. The SO(2)$\subset$SL(2, R) UIR matrices found by Bargmann follow (4.33) (replacing ρ, ρ' by m, m', and $-$ by 0). Indeed, after (4.7) we remarked that the E(1)$\subset$SL(2, R)$\supset$SO(2) overlap coefficient in the continuous series is ${}^0\Psi^{\epsilon,k}_{m,\mathrm{sgn}\rho}(|\rho|)$. From its functional form (4.5a) we can see that

$${}^0\Psi^{1/2,\,1/2+is}_{jm,j}(r) \underset{s\to 0^+}{\longrightarrow} j^{m-1/2}\, {}^0\Phi^{1/2}_m(r), \tag{4.34a}$$

$${}^0\Psi^{1/2,\,1/2+is}_{jm,-j}(r) \underset{s\to 0^+}{\longrightarrow} 0, \quad m = \tfrac{1}{2} + n, \quad n = 0, 1, 2, \ldots. \tag{4.34b}$$

The continuous series UIR in the SO(2) basis thus also separates in block-diagonal form into the $D^+_{1/2}$ and $D^-_{1/2}$ representations:

$${}^0D^{1/2,\,1/2+is}_{m,m'}(\mathbf{g}) \underset{s\to 0^+}{\longrightarrow} \delta_{\mathrm{sgn}m,\mathrm{sgn}m'}\, {}^0D^{1/2(\mathrm{sgn}m)}_{|m|,|m'|}(\mathbf{g}). \tag{4.35}$$

The SO(1,1) subgroup-reduced integral kernels do separate, although not in block-diagonal form as in the former cases. The E(1)$\subset$SL(2, R)$\supset$SO(1,1) overlap coefficient in the continuous series (4.20) for $\mathbf{g} = \mathbf{1}$ are, in terms of those of the discrete series (3.14),

$$({}^-\Psi_\rho, {}^2\Psi_{\kappa,\mu}) = {}^2\Psi_{k,\mu,\mathrm{sgn}\rho}(|\rho|) = \begin{cases} 2^{-1/2}({}^-\Phi_{|\rho|}, {}^2\Phi_\mu), & \rho \geqslant 0 \\ \kappa 2^{-1/2}({}^-\Phi_{|\rho|}, {}^2\Phi_\mu), & \rho < 0, \end{cases} \tag{4.36}$$

and hence we obtain a sum of the $D^+_{1/2}$ and $D^-_{1/2}$ representations:

$${}^2D^{1/2,\,1/2+is}_{\kappa,\mu;\kappa',\mu'}(\mathbf{g}) \underset{s\to 0^+}{\longrightarrow} \frac{1}{2} \sum_{\tau=\pm 1} (\kappa\kappa')^{(1-\tau)/2}\, {}^2D^{1/2(\tau)}_{\mu,\mu'}(\mathbf{g}). \tag{4.37}$$

From this and the remark following (4.18) on the bilateral Mellin transform, it may appear more convenient to use $\mathbf{J}_2$ eigenfunctions whose dichotomic index label functions with upper or lower components only, instead of those used in (4.3a). This may be a useful alternative in some contexts, such as matching the two components of the bilateral Mellin transform kernel.[47] In some other cases, as in the study of an (uncoupled) hyperbolic Fourier transform class,[73] still another linear combination of the two ${}^-\Psi_\rho$ rows proves to be useful, as it diagonalizes the 2×2 kernel matrix.

ii. $C_q^0 \underset{q\to 0}{\longrightarrow} D_1^+ \dotplus D_1^-$. We also remarked at the end of Sec. 2D that the exceptional continuous series integral kernel $[\mathbf{C}_\mathbf{g}^{0k}]_{j,j'}(r, r')$ for $k = \frac{1}{2} + \sigma$, $\sigma \to (\frac{1}{2})^-$ also uncoupled into the D_1^+ and D_1^- radial canonical transform kernels:

$${}^-D^{0,\,1/2+\sigma}_{\rho,\rho'}(\mathbf{g}) \underset{\sigma\to(1/2)^-}{\longrightarrow} \delta_{\mathrm{sgn}\,\rho,\mathrm{sgn}\,\rho'}\, {}^-D^{1(\mathrm{sgn}\,\rho)}_{|\rho|,|\rho'|}(\mathbf{g}). \tag{4.38}$$

The significance of this limit is the same as for (4.33), and equations parallel to (4.34)–(4.37) follow for all other overlap coefficients and subgroup reductions. In particular, ${}^0\Psi_0^{0,1/2+\sigma}(r)$ vanishes as $\sigma \to (\frac{1}{2})^-$.

5. SL (2, R) TRANSFORMS AND SERIES

In Sec. 2 we introduced the SL(2, R) group of unitary k-canonical integral transforms for all UIR series of this group. The ensuing developments in Secs. 3 and 4 have detailed three families of bases for these spaces, associated with the E(1), SO(1, 1), and SO(2) families of subgroup reductions, and have given their overlap coefficients. These define as many families of integral transforms and series expansions.

A. The discrete series

i. $E(1)\subset SL(2,R)\supset E(1)$. For the discrete series, we can write in terms of the $\mathscr{L}^2(R^+)$ inner product and E(1) basis functions (3.2)

$$({}^-\Phi_r, f) = f(r), \quad r\in R^+. \tag{5.1a}$$

The k-radial canonical transform may be thus implemented as a change of coordinates

$$f(r) \xrightarrow{g} f_g(r) = [\mathbb{C}_g^k f](r) = ({}^-\Phi_r, \mathbb{C}_g^k f)$$
$$= (\mathbb{C}_{g^{-1}}^k\, {}^-\Phi_r, f) = \int_0^\infty dr'\, {}^-D^k_{r,r'}(g) f(r'), \tag{5.1b}$$

from the Dirac-orthonormal E(1) eigenbasis $\{{}^-\Phi_r\}_{r\in R^+}$ to a similar family of bases $\{\mathbb{C}_{g^{-1}}^k\, {}^-\Phi_r\}_{r\in R^+}$ of generalized eigenfunctions of $\mathbb{C}_g^k\, {}_1J_-\mathbb{C}_g^k$, for every fixed $g\in SL(2,R)$. The UIR matrix elements are the radial canonical transform kernels, as has been noted before. The transform inverse to (5.1b) has a kernel ${}^-D^k_{r,r'}(g^{-1}) = [D^k_{r',r}(g)]^*$. The unitarity of the transform implies the Parseval identity $(f, h) = (f_g, h_g)$. In particular, it contains the Hankel transform of $g = \mathbf{F}$ [Eq. (2.11b)].

ii. $E(1)\subset SL(2,R)\supset SO(1,1)$. In the point of view we are developing in this section, the coordinates of f in the $SO(1,1)_2$ eigenbasis $\{{}^2\Phi_\mu\}_{\mu\in R}$ are

$$\hat{f}(\mu) = ({}^2\Phi_\mu, f)$$
$$= \int_0^\infty dr ({}^2\Phi_\mu, {}^-\Phi_r)({}^-\Phi_r, f)$$
$$= \int_0^\infty dr\, \pi^{-1/2} r^{-1/2-2i\mu} f(r)$$
$$= 2^{1/2} f_+^M(2\mu), \tag{5.2a}$$

where f_+^M is the positive Mellin transform[47] of f. The family of SL(2, R)-similar Dirac bases $\{\mathbb{C}_{g^{-1}}^k\, {}^2\Phi_\mu\}_{\mu\in R}$ defines a corresponding SL(2, R)-parametrized family of integral transforms between $\mathscr{L}^2(R^+)$ and $\mathscr{L}^2(R)$,

$$f(r) \xrightarrow{(M)g} \hat{f}_g^k(\mu) = ({}^2\Phi_\mu, \mathbb{C}_g^k f)$$
$$= (\mathbb{C}_{g^{-1}}^k\, {}^2\Phi_\mu, f) = \int_0^\infty dr\, {}^{2,-}D^k_{\mu,r}(g) f(r), \tag{5.2b}$$

whose kernel (3.11) contains in general a confluent hypergeometric function, with μ in one index and r in the argument. In particular, it contains the positive Mellin transform (5.2a) for $g = 1$. The transform inverse to (5.2b) has a kernel ${}^{-,2}D^k_{r,\mu}(g^{-1}) = [{}^{2,-}D^k_{\mu,r}(g)]^*$ and the integration is performed over $\mu\in R$. An obvious Parseval identity holds between (f, h) and $\hat{f}_g^k(\mu)^*\hat{h}_g^k(\mu)$ integrated over μ.

iii. $E(1)\subset SL(2,R)\supset SO(2)$. The coordinates of f in the $SO(2)\subset SL(2,R)$-similar eigenbases $\{\mathbb{C}_{g^{-1}}^k\, {}^0\Phi_m^k\}_{m=k}^\infty$ define a mapping between $\mathscr{L}^2(R^+)$ and l^2_+ (lower-bound square-summable sequences):

$$f(r) \xrightarrow{(L)g} f_{g,m}^k = ({}^0\Phi_m^k, \mathbb{C}_g^k f)$$
$$= (\mathbb{C}_{g^{-1}}^k\, {}^0\Phi_m^k, f) = \int_0^\infty dr\, {}^{0,-}D^k_{m,r}(g) f(r), \tag{5.3}$$

which contains, essentially, the normalized Laguerre series analysis [in $L_{m-k}^{(2k-1)}(r^2)$] of $f(r)$ for $g = \mathbf{1}$. The series synthesis is provided by the functions ${}^{-,0}D^k_{r,m}(g^{-1}) = [{}^{0,-}D^k_{m,r}(g)]^*$ and a corresponding Parseval identity holds.

iv. $SO(1,1)\subset SL(2,R)\supset SO(2)$. We may also use the overlap coefficients between the SO(1, 1) and SO(2) bases to define the expansion of an $\mathscr{L}^2(R)$ function $\hat{f}(\mu)$ in a series of hypergeometric functions of argument $\frac{1}{2}$, as given by (3.10a), or its generalization for any fixed argument as given by (3.9), through the analysis

$$\hat{f}(\mu) \xrightarrow{(H)g} \hat{f}_{g,m}^k = \int_{-\infty}^\infty d\mu\, {}^{0,2}D^k_{m,\mu}(g)\hat{f}(\mu) \tag{5.4}$$

and the corresponding synthesis with $[{}^{0,2}D^k_{m,\mu}(g)]^*$, with an appropriate Parseval identity.

v. $SO(1,1)\subset SL(2,R)\supset SO(1,1)$. The SO(1, 1) subgroup decomposition of the discrete UIR series provides an SL(2, R)-parametrized family of unitary integral transforms between $\mathscr{L}^2(R)$ and itself,

$$\hat{f}(\mu) \xrightarrow{(F),g} \hat{f}_g^k(\mu) = \int_{-\infty}^\infty d\mu'\, {}^2D^k_{\mu,\mu'}(g)\hat{f}(\mu'), \tag{5.5}$$

with a kernel involving hypergeometric functions of fixed argument, as given by (3.17). This is basically the Mellin transform of the k-radial canonical transform family (5.1).

vi. $SO(2)\subset SL(2,R)\supset SO(2)$. The SO(2) subgroup decomposition, finally, provides an SL(2, R)-parametrized family of mappings of discrete unitary transforms between l^2_+ and l^2_+ which repesents the well-known action of the group—for a fixed element $\mathbf{g}$ and k—on the space of sequences $\{f_m\}_{m=k}^\infty$.

The SL(2, R) D_k^+ UIR matrix elements of the discrete series thus provide six different SL(2, R)-parametrized families of integral or discrete transforms, or series expansions between $\mathscr{L}^2(R^+)$, $\mathscr{L}^2(R)$, and l^2_+, of which the k-canonical radial transforms given in Sec. 2 are but one family.

B. The continuous series

The same pattern of six families of transforms hold for the continuous series of SL(2, R) UIRs, between spaces $\mathscr{L}^2_{\rm II}(R^+)$ [extendable to $\mathscr{L}^2(R)$ through $f(\rho) = f_{{\rm sgn}\rho}(|\rho|)$], $\mathscr{L}^2_{\rm II}(R)$ and l^2. These families include the k-hyperbolic canonical transforms given in Sec. 2, bilateral Mellin transforms, Whittaker and hypergeometric series and transforms.

C. Further extensions

Since these six families of transforms have a group-theoretical origin and parametrization, pairs of transforms belonging to one or two families (with the same k) may be applied in succession, respecting the mixed-basis transitivity properties, to give another transform of the same or of a different family. These are transforms which are all associated with the SL(2, R) group and its representations, so we would like to close our account of these with some comments on further extensions to this set, which have been published in the literature, and to other sets as yet not fully explored.

The first extension pertains consideration of the covering group $\overline{\mathrm{SL}(2,R)}$. Indeed, the oscillator (metaplectic) representation is the two-fold covering of SL(2, R) [four-fold covering of SO(2, 1)] provided by $D^+_{1/4} \dot{+} D^+_{3/4}$. The case D^+_k, for real $k > 0$, has been described in Refs. 19, 20, and 34, but as yet it has not been as thoroughly analyzed as would be desirable. The continuous series of SL(2, R) have not been treated, although partial results exist. The subject of complex extensions of SL(2, R) to a semigroup of integral transforms,[17,19,28] possible for the discrete series—which includes the bilateral Laplace, Gauss–Weierstrass (heat diffusion), Bargmann[74] and Barut-Girardello[75] transforms—and the extension of SL(2, R) to W $\wedge$ SL(2, R) (W being the Heisenberg–Weyl group), has not been touched upon in this work, as it falls outside the scope of the title. Parts of it have appeared in various articles by one of the authors,[76] but the description of this last extension in various subgroup—and mixed bases is still wanting. Finally, the subject of nonsubgroup decompositions[77] in this context is still open.

ACKNOWLEDGMENTS

We would like to thank Dr. Alberto Alonso y Coria and Dr. Antonmaría Minzoni for several useful conversations. One of us (D. B.) gratefully acknowledges the hospitality extended by IIMAS.

APPENDIX: THE UNITARY IRREDUCIBLE REPRESENTATIONS OF SL(2, R)

Bargmann[1] classified all UIRs of SU(1, 1) $\approx$ SL(2, R) $\approx$ Sp(2, R) $\overset{1:2}{\approx}$ SO(2, 1). We give here a summary of the results, nomenclature, and notation followed in this article.

We denote by SL(2, R) the special linear group in two dimensions over the real field, i.e., the group of 2×2 matrices

$$\mathbf{g} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \; a, b, c, d \in R, \quad \det \mathbf{g} = ad - bc = 1. \tag{A1}$$

Due to the unimodularity condition, (A1) also satisfy $\mathbf{g}\boldsymbol{\sigma}_p \mathbf{g}^T = \boldsymbol{\sigma}_p$, $\mathbf{g}^T$ being the transpose of $\mathbf{g}$, with the symplectic metric matrix

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

The elements of the real symplectic group Sp(2, R) are thus also given by $\mathbf{g}$ as in (A1). The "1 + 1" unimodular pseudounitary group SU(1, 1), on the other hand, is the set of unimodular 2×2 complex matrices $\mathbf{u}$ satisfying $\mathbf{u}\boldsymbol{\sigma}_3\mathbf{u}^\dagger = \boldsymbol{\sigma}_3$, $\mathbf{u}^\dagger$ being the adjoint (transpose, complex conjugate) of $\mathbf{u}$, with the metric matrix

$$\boldsymbol{\sigma}_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

It is easy to show that the most general form of $\mathbf{u}$ is

$$\mathbf{u} = \begin{pmatrix} \alpha & \beta \\ \beta^* & \alpha^* \end{pmatrix}, \; \alpha, \beta \in C, \; \det \mathbf{u} = |\alpha|^2 - |\beta|^2 = 1. \tag{A2}$$

The link between SL(2, R) and SU(1, 1) matrices which relates the results of this article with those of Bargmann is given by the similarity transformation

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \mathbf{W}\begin{pmatrix} \alpha & \beta \\ \beta^* & \alpha^* \end{pmatrix}\mathbf{W}^{-1} = \begin{pmatrix} \mathrm{Re}\alpha + \mathrm{Re}\beta & -\mathrm{Im}\alpha + \mathrm{Im}\beta \\ \mathrm{Im}\alpha + \mathrm{Im}\beta & \mathrm{Re}\alpha - \mathrm{Re}\beta \end{pmatrix}, \tag{A3a}$$

$$\mathbf{W} = 2^{-1/2}\begin{pmatrix} \omega^{-1} & \omega^{-1} \\ -\omega & \omega \end{pmatrix}, \quad \omega = e^{i\pi/4}. \tag{A3b}$$

Other isomorphisms found in the literature are determined by W's such as

$$2^{-1/2}\begin{pmatrix} 1 & -i \\ -i & 1 \end{pmatrix}, \; 2^{-1/2}\begin{pmatrix} 1 & -1 \\ -i & -i \end{pmatrix},$$

and

$$2^{-1/2}\begin{pmatrix} 1 & i \\ i & 1 \end{pmatrix}.$$

The latter yields the complex conjugate of (A3a). The 2:1 homomorphism between SU(1, 1) and the Lorentz group SO(2, 1) is often exploited through parametrizing the former in terms of Euler angles,

$$\begin{pmatrix} \alpha & \beta \\ \beta^* & \alpha^* \end{pmatrix} = \begin{pmatrix} e^{-i\mu} & 0 \\ 0 & e^{i\mu} \end{pmatrix}\begin{pmatrix} \cosh\zeta & \sinh\zeta \\ \sinh\zeta & \cosh\zeta \end{pmatrix}\begin{pmatrix} e^{-i\nu} & 0 \\ 0 & e^{i\nu} \end{pmatrix}. \tag{A4}$$

Our favored set of parameters are those in (A1), and in terms of those we express the UIR matrix elements. Of particular interest to many authors are the representations of the hyperbolic rotation (boost) subgroup in the middle factor of (A4). This is given by $\mathbf{M}_2(-2\zeta)$ in (2.10b).

Out of the matrix realization (A1)–(A2) Bargmann[1] finds the sl(2, R) Lie algebra. Without having to realize the algebra elements through differential operators, but only under the assumption of the existence of a Hilbert space endowed with a sesquilinear positive-definite inner product, one can find the self-adjoint irreducible representations of the algebra classified through the eigenvalues q of the Casimir operator (2.9), and through the usual raising- and lowering-operator techniques, the SO(2) representations m contained in any one SL(2, R) UIR are found.

The following are all nonequivalent single-valued representations of SL(2, R).

Discrete series $q = k(1-k)$ for $k = \frac{1}{2}, 1, \frac{3}{2}, 2, \ldots$ containing:

D^+_k positive discrete UIRs, $m = k, k+1, k+2, \ldots$

D^-_k negative discrete UIRs, $m = -k, -k-1, -k-2, \ldots$.

Continuous series

C^0_q the vector nonexceptional continuous UIRs $q = k(1-k) \geqslant \frac{1}{4}$; $k = \frac{1}{2} + is$, $s \geqslant 0$,

C^0_q the (vector) exceptional continuous UIRs $0 < q = k(1-k) < \frac{1}{4}$; $k = \frac{1}{2} + \sigma$, $0 < \sigma < \frac{1}{2}$,

$C^{1/2}_q$ the spinor (nonexceptional) continuous UIRs $q = k(1-k) > \frac{1}{4}$; $k = \frac{1}{2} + is$, $s > 0$.

Values of k other than these give rise to nonunitary and/or multivalued representations of SL(2, R).

[1]V. Bargmann, Ann. Math. **48**, 568 (1947).
[2]A. O. Barut and C. Fronsdal, Proc. R. Soc. (London) **A287**, 532 (1965).
[3]L. Pukański, Math. Anal. **156**, 96 (1964).
[4]P. J. Sally, Jr., Bull. Am. Math. Soc. **72**, 269 (1966).
[5]S. Lang, $sl_2(R)$ (Addison-Wesley, Reading, MA, 1975).
[6]N. Ya. Vilenkin, *Special Functions and the Theory of Group Representations*, Trans. Math. Soc. **22** (1968); W. Miller, Jr., *Lie Theory and Special Functions* (Academic, New York, 1968); J. D. Talman, *Special Functions, a Group Theoretic Approach* (Benjamin, New York 1968); W. Miller, Jr., *Symmetry Groups and Separation of Variables*, Encyclopedia of Mathematics, Vol. 4, edited by G.-C. Rota (Addison-Wesley, Reading, MA, 1977).
[7]M. Andrews and J. Gunson, J. Math. Phys. **5**, 1391 (1964); M. Toller, Nuovo Cimento **37**, 631 (1965); A. Sciarrino and M. Toller, J. Math. Phys. **8**, 1252 (1967); J. F. Boyce, J. Math. Phys. **8**, 675 (1967); M. Toller, Nuovo Cimento **53**, 671 (1968); D. A. Akeyampong, J. F. Boyce, and M. A. Rashid, Nuovo Cimento **53**, 737 (1968); G. Sollani and M. Toller, Nuovo Cimento **15**, 430 (1973).
[8]N. Mukunda, J. Math. Phys. **8**, 2210 (1967); **9**, 50, 417 (1968); J. G. Kuriyan, N. Mukunda, and E. C. G. Sudarshan, *ibid.* **9**, 2100 (1968).
[9]N. Mukunda, J. Math. Phys. **10**, 2068 (1969).
[10]N. Mukunda, J. Math. Phys. **10**, 2092 (1969).
[11]N. Mukunda and B. Radhakrishnan, J. Math. Phys. **14**, 254 (1973).
[12]A. O. Barut and E. C. Phillips, Commun. Math. Phys. **8**, 52 (1968).
[13]G. Lindblad and B. Nagel, Ann. Inst. H. Poincaré **13**, 27 (1970).
[14]M. Moshinsky and C. Quesne, J. Math. Phys. **12**, 1772, 1780 (1971); M. Moshinsky, SIAM, J. Appl. Math. **25**, 193 (1973).
[15]M. Moshinsky, T. H. Seligman, and K. B. Wolf, J. Math. Phys. **13**, 1634 (1972).
[16]I. M. Gel'fand, M. I. Graev, and N. Ya. Vilenkin, *Generalized Functions* (Academic, New York, 1966), Vol. 5, Chap. VII.
[17]P. Kramer, M. Moshinsky, and T. H. Seligman, "Complex Extensions of Canonical Transformations in Quantum Mechanics," in *Group Theory and its Applications*, Vol. III, edited by E. M. Loebl (Academic, New York, 1975).
[18]K. B. Wolf, J. Math. Phys. **15**, 1295 (1974).
[19]K. B. Wolf, J. Math. Phys. **15**, 2102 (1974).
[20]C. P. Boyer and K. B. Wolf, J. Math. Phys. **16**, 1493 (1975).
[21]K. B. Wolf, J. Math. Phys. **21**, 680 (1980).
[22]K. B. Wolf, in *Proceedings of the IX International Colloquium on Group Theoretical Methods in Physics*, Lecture Notes in Physics **135** (Springer, New York, 1980).
[23]S. Steinberg and K. B. Wolf, Nuovo Cimento **53A**, 149 (1979).
[24]G. Burdet and M. Perrin, J. Math. Phys. **16**, 1692, 2172 (1975); M. Perroud, Helv. Phys. Acta **50**, 233 (1977); G. Burdet, M. Perrin, and M. Perroud, Commun. Math. Phys. **58**, 241 (1978).
[25]L. Infeld and J. Pelbański, Acta Phys. Polon. **14**, 41 (1955); C. Itzykson, J. Math. Phys. **10**, 1109 (1969).
[26]E. G. Kalnins and W. Miller, Jr., J. Math. Phys. **15**, 1263 (1974).
[27]M. Kashiwara and M. Vergne, Inv. Math. **44**, 1 (1978).
[28]K. B. Wolf, *Integral Transforms in Science and Engineering* (Plenum, New York, 1979), Part 4.
[29]Reference 1, Secs. 6–9.
[30]Reference 1, Sec. 10.
[31]E. G. Kalnins, J. Math. Phys. **14**, 654 (1973).
[32]D. Basu, J. Math. Phys. **19**, 1667 (1978).
[33]D. Basu and D. Mitra, J. Math. Phys. **22**, 946 (1981).
[34]C. P. Boyer and K. B. Wolf, Rev. Mex. Fís. **25**, 31 (1976).
[35]Ya. I. Azimov, Sov. J. Nucl. Phys. **4**, 469 (1967).
[36]A. Erdelyi, W. Magnus, F. Oberhettinger,a nd F. G. Tricomi, *Higher Transcendental Functions* (3 Vols.) (McGraw-Hill, New York, 1953–1955).
[37]I. S. Gradshteyn and I. M. Ryzhik, *Tables of Integrals, Sums and Products* (Academic, New York, 1965).
[38]Ref. 1, Secs. 10a–10d.
[39]C. P. Boyer, Comunicaciones Técnicas CIMAS, No. 107 (1975) (unpublished).
[40]K. B. Wolf, "Topics in Noncompact Group Representations," in *Lecture Notes of the 1980 Latin American School of Physics*, edited by T. H. Seligman, American Institute of Physics, Conference Proceedings **71** (AIP, New York, 1981), Sec. 3.17.
[41]Ref. 37, Eq. 8.451.1.
[42]For example, see Ref. 28, Eqs. (9.75)–(9.77).
[43]Ref. 1, Eqs. (9.16) and (9.21).
[44]On the SU(1, 1) matrices of Bargmann, (A2), it acts through complex conjugation: See Ref. 1, Eqs. (9.1) and (9.20).
[45]Expressions (2.15) differ for $b<0$ from those presented in Ref. 11, Eqs. (2.15) and (2.16) and those in Ref. 21, Eq. (3.11) due to the fact that the integrals leading to the Hankel functions must take account of the appropriate Sommerfeld contour deformation: $0<\arg z<\pi$ for $H_\nu^{(1)}(z)$ and $-\pi<\arg z<0$ for $H_\nu^{(2)}(z)$. This implies that one should approach the real axis from above and below, respectively. This point was overlooked in the quoted references, and is quite crucial for subsequent calculations with these kernels.
[46]M. Reed and B. Simon, *Methods of Modern Mathematical Physics*, Vol. 2 (Academic, New York, 1975).
[47]Ref. 28, Sec. 8.2.
[48]Ref. 37, Eq. 9.220.2.
[49]Ref. 1, Eqs. (6.22), valid for the discrete series, as remarked after Eq. (9.18).
[50]Ref. 37, Eq. 7.414.7.
[51]Ref. 37, Eq. 9.131.1.
[52]Ref. 12, Eqs. (3.10), (3.24), and (3.25) for $\Phi \mapsto -k$ and $\epsilon\lambda \mapsto i\mu$.
[53]See, e.g., Ref. 37, Eq. 6.621.1.
[54]In Ref. 34 we used a decomposition analogous to (3.6a) but with $\mathbf{M}_1(\theta)$ in place of the second factor. The decomposition followed there, in contrast with the present one, is not global and a patch matching must follow.
[55]K. B. Wolf, J. Math. Phys. **18**, 1046 (1977).
[56]Compare with Ref. 12, Eq. (3.18').
[57]In Ref. 34, Eq. (3.6), not enough care was taken to ensure the correctness of the phases. The temptation to collect powers of these variables must be avoided if the formulae are to retain their validity for the full range of group parameters.
[58]Ref. 37, Eq. 7.414.4. See also Ref. 34, Eq. (2.5), where the expression is derived in detail. There, the argument of the hypergeometric function contains an erratum.
[59]Ref. 37, Eq. 9.132.1.
[60]Ref. 1, Eqs. (10.28) with normalization (10.11).
[61]Ref. 37, Eq. 9.220.4.
[62]Its meaning in the Sp(4, R) parent group, for the continuous UIR series, can be seen in Ref. 21, Eq. (2.9).
[63]Two of the four ${}_1F_1$ functions sum to a Whittaker W function [cf, Eq. (4.18b)]; the other two do not due to a phase difference. In Ref. 21, Eq. (4.19) the claim to reduce ${}^1\psi^{\epsilon,k}_{\kappa,\mu,j}(r)$ to a single W function is thus incorrect. This is one consequence of the imprecision in the phases of the hyperbolic canonical transform kernel of Eq. (3.11a) in Ref. 21 versus Eq. (2.15d) here.
[64]Ref. 1, Eqs. (6.22)–(6.26), (7.10)–(7.11), and (8.10)–(8.15).
[65]These results extend those presented in Ref. 55.
[66]Compare with Eq. (2.5) of Ref. 31 for the $\mathrm{SO}(1,1)_2$ subgroup.
[67]Compare with Eq. (2.16) of Ref. 31 for the $\mathrm{SO}(1,1)_2$ subgroup.
[68]This is the method and result of Ref. 32, Eq. (2.24). One can compare this result with Ref. 12, Eqs. (3.10), (3.24), and (3.25) after a ${}_2F_1$ transformation is used. The cases $\epsilon=0$ and $\epsilon=\frac{1}{2}$ are not distinguished there.
[69]Compare with Eqs. (2.26)–(2.27) of Ref. 31 for the $\mathrm{SO}(1,1)_2$ subgroup. It should be noted that the symmetry $k \leftrightarrow 1-k$ (i.e., $\rho \leftrightarrow -\rho$ there) is not apparent. Caution should be excised as the dichotomic index in Ref. 31, and our κ are not the same.
[70]Compare with Ref. 12, Eqs. (3.18')–(3.25). There is no indication of whether the $\epsilon=0$ or $\epsilon=\frac{1}{2}$ continuous series is being discussed.
[71]Ref. 1, Eq. (10.27).
[72]This result generalizes that of Kalnins, Ref. 31, Eq. (2.12), in allowing for a Gaussian factor in the integrand.
[73]Ref. 22, Eq. (14). Transforms of this kind have been studied by E. C. Titchmarsh, *Introduction to the Theory of Fourier Integrals* (Clarendon, Oxford, 1937), p. 217. The result presented there corresponds to $\mathbf{g}=\mathbf{F}^{\pm 1}$ for the point $k=\frac{1}{2}$ in the continuous $\epsilon=0$ series. For the whole continu-

ous series, we obtain a kernel with a linear combination of Bessel, Neumann, and Macdonald functions.

[74]V. Bargmann, Commun. Pure Appl. Math. **14**, 187 (1961); **20**, 1 (1967).
[75]A. O. Barut and L. Girardello, Commun. Math. Phys. **21**, 41 (1971).
[76]K. B. Wolf, J. Math. Phys. **17**, 601 (1976); K. B. Wolf, SIAM J. Appl. Math. **40**, 419 (1981).
[77]J. Patera and P. Winternitz, J. Math. Phys. **14**, 1130 (1973); E. G. Kalnins and W. Miller Jr., J. Math. Phys. **15**, 1728 (1974).

Geometric quantization and representations of semisimple Lie groups: spin (2,1) and spin (2,2)

B. R. Sitaram and K. C. Tripathy
Department of Physics, University of Delhi, Delhi-110007, India

(Received 30 July 1980; accepted for publication 26 November 1980)

The UIR's of spin (2,1) and spin(2,2) are studied in the light of Kostant–Auslander induction scheme, and compared with those obtained earlier by Harish–Chandra and Schmid.

PACS numbers: 02.20.Rt

I. INTRODUCTION

Ever since the advent of relativistic symmetries in particle physics, the utility of noncompact groups has been increasingly felt. The spin (2,1) group was the attack of physicists and mathematicians as well since it provided a prototype building block of a hierchy of symmetries.[1] The remarkable success of Auslander–Kostant work[2–4] on simply connected solvable Lie groups, and equally astounding success of the work of Harish-Chandra and Schmid and others on semisimple Lie groups,[5–7] sufficiently warrants an in depth study of the links between the two. In recent times, the representation theory of some of the semisimple Lie groups have been studied in the light of the Auslander–Kostant program with encouraging success. In the present paper we make a systematic study of two physically relevant noncompact groups: spin(2,1) and spin(2,2). It is interesting to note that both these groups possess discrete series representations and provide a basis for generalization to the case of arbitrary semisimple Lie groups. We hope to report on the latter in a forthcoming paper.

Our paper is arranged as follows: In Sec. II, we give a brief resumé of (a) the Auslander–Kostant induction scheme, (b) the theory of nondiscrete UIR's as semisimple Lie groups, and (c) Harish-Chandra and Schmid's work on discrete series representaions.

In Sec. III we compute the orbits and polarizations for the afore said groups.

The identification of polarizations associated with noncompact orbits with parabolic subalgebras in displayed in Sec. IV. Further, we show that the representations obtained for these polarizations coincide with the principal and degenerate series representations.

In Sec. V, we construct the representations associated with compact orbits and show their equivalence with discrete series representations of Harish-Chandra and Schmid.

We shall use the following notations throughout the paper. Lie groups will be denoted by capital Roman letters, the corresponding Lie algebra being denoted by lower case Roman letters. For a group G, $\hat{G}$, will denote the set of equivalence classes of UIR's. Finally, all direct sums are to be taken as vector space direct sums and not necessarily Lie algebra direct sums.

II. RESUMÉ OF THE AUSLANDER–KOSTANT THEORY AND THE CONVENTIONAL THEORY OF SEMISIMPLE LIE GROUPS

A. The Auslander–Kostant theory

Let G be a semisimple Lie group with $\mathbf{g}$ as its Lie algebra and let $\mathbf{g}^*$ be the real dual of $\mathbf{g}$. We use the Cartan–Killing isomorphism to identify $\mathbf{g}^*$ with $\mathbf{g}$. Let O_X represent the orbit of a point $X\in\mathbf{g}$ under the adjoint transformation

$$Ad_s X = sXs^{-1}, \quad s\in G. \tag{II.1}$$

Let G_X be the isotropy group at X such that

$$G_X = \{s\in G: sXs^{-1} = X\} \tag{II.2}$$

and let $\mathbf{g}_X$ be the corresponding Lie algebra. An element $X\in\mathbf{g}$ defines a mapping

$$2\pi iX: \mathbf{g}_X \to i\mathbb{R}, \quad Y \mapsto 2\pi iB(X,Y)\,\forall Y\in\mathbf{g}_X, \tag{II.3}$$

where B is the canonical Cartan–Killing form. We call X, or equivalently O_X, quantizable if the mapping $2\pi iX$ is the differential of some character,

$$\chi: G_X \to S^1. \tag{II.4}$$

A polarization at X is defined to be a Lie subalgebra $\mathbf{s}$ of $\mathbf{g}_c = \mathbf{g} + i\mathbf{g}$ which satisfies

(i) $\langle B,[\mathbf{s},\mathbf{s}]\rangle = 0$, where B is extended to $\mathbf{g}_c \times \mathbf{g}_c$ by complex linearity, and $\mathbf{s}$ is maximal with respect to (w.r.t.) this condition,

(ii) $\dim_C \mathbf{s} = \frac{1}{2}(\dim_R \mathbf{g} + \dim \mathbf{g}_x)$,

(iii) $\mathbf{s} + \bar{\mathbf{s}}$ is a Lie subalgebra of $\mathbf{g}_c$, where the bar indicates complex conjugation,

(iv) $\mathbf{g}_x \subseteq \mathbf{s}$ and $\mathbf{s}$ is ad_{G_x}-stable: $\mathrm{ad}_s\, \mathbf{s} \subseteq \mathbf{s}\ \forall\, s\in G_X$. $\mathbf{s}$ is said to be positive if $iB(x,[\bar{z},z]) \geqslant 0\ \forall z\in\mathbf{s}$.

Let O_x be a quantizable orbit with χ being the corresponding character and let $\mathbf{s}$ be a positive polarization at X. Let us define

$$\boldsymbol{\delta} = \mathbf{s} \cap \mathbf{g}, \quad \mathbf{e} = (\mathbf{s} + \bar{\mathbf{s}}) \cap \mathbf{g}, \tag{II.5a}$$

$$D = D_0 G_x, \quad E = E_0 G_X, \tag{II.5b}$$

where D_0 and E_0 are the analytic subgroups corresponding to $\boldsymbol{\delta}$, $\mathbf{e}$ respectively. We extend χ form G_X to D. The Auslander–Kostant induction scheme can be written as

$$\sigma = \mathrm{ind}_E^G(\mathrm{ind}_D^E \chi), \tag{II.6}$$

where ind_A^B denotes induction from A to B where ind_D^E is holomorphic induction. For semisimple groups, our problems can be stated as follows:

(i) Is σ a UIR of G?

(ii) Does every UIR of G arise in the above manner?

B. Nondiscrete UIR's of semisimple Lie groups

Let $\mathbf{g} = \mathbf{k} + \mathbf{p}$ be the Cartan decomposition of $\mathbf{g}$; let $\mathbf{a}$ be a maximal abelian subalgebra of $\mathbf{p}$. Let $\mathbf{n}$ be the set of root vecotrs of $\mathbf{a}$.

$$\mathbf{n} = \{X\in\mathbf{g}: [A,X] = \lambda(A)X\ \forall A\in\mathbf{a} \text{ and for some } \lambda:\mathbf{a}\to\mathbb{C},\ \ker\lambda \neq \mathbf{a}\}. \tag{II.7}$$

Let n^+, n^- be the set of positive and negative root vectors relative to some ordering. Let **m** be normalizer of **a** in **k**:

$$\mathbf{m} = \{M\in\mathbf{k}:[M,\mathbf{a}]\subseteq\mathbf{a}\}. \tag{II.8}$$

then, $\mathbf{p}_0 = \mathbf{m} + \mathbf{a} + \mathbf{n}^+$ is known as a minimal parabolic subalgebra of **g**. A parabolic subalgebra of **g** is then defined to be any subalgebra of **g** containing $\mathbf{p}_0$. The set of all parabolic subalgebras of **g** can be explicitly constructed as follows[8]: Let Σ be a set of roots for **n** and let Ψ be the set of positive simple roots in Σ. Let θ be a subset of Ψ. Let $\langle\theta\rangle$ denote the set of roots in Σ which arises as linear combinations of roots in θ. We define

$$\langle\theta\rangle_+ = \Sigma_+\cap\langle\theta\rangle, \quad \langle\theta\rangle_- = \Sigma_-\cap\langle\theta\rangle, \tag{II.9}$$

where Σ_+, Σ_- denote positive and negative roots in Σ. Let $\mathbf{n}_+(\theta)$, $\mathbf{n}_-(\theta)$, $\mathbf{n}_\theta$ denote the subspaces of **n** corresponding to $\langle\theta\rangle_+$, $\langle\theta\rangle_-$ and $\{\Sigma_+ - \langle\theta\rangle_+\}$.

Define

$$\mathbf{a}_\theta = \{A\in\mathbf{a}: \lambda(A) = 0\ \forall\lambda\in\theta\} \tag{II.10}$$

and let $\mathbf{a}(\theta)$ be the orthogonal complement of $\mathbf{a}_\theta$ in **a** w.r.t. the Cartan–Killing form. Then $\mathbf{p}_\theta = \mathbf{m}_\theta + \mathbf{a}_\theta + \mathbf{n}_\theta$ is a parabolic subalgebra of **g** where

$$\mathbf{m}_\theta = \mathbf{m} + \mathbf{n}_+(\theta) + \mathbf{n}_-(\theta) + \mathbf{a}(\theta). \tag{II.11}$$

A Cartan subalgebra **h** of **g** is said to be invariant w.r.t the cartan decomposition if

$$\mathbf{h} = (\mathbf{h}\cap\mathbf{k}) + (\mathbf{h}\cap\mathbf{p}). \tag{II.12}$$

A parabolic subalgebra $\mathbf{p}_\theta$ is said to be cuspidal if there exists an invariant Cartan subalgebra **h** such that

$$\mathbf{a}_\theta = \mathbf{h}\cap\mathbf{p}. \tag{II.13}$$

Let M_θ, A_θ, N_θ be the Lie groups corresponding to $\mathbf{p}_\theta = \mathbf{m}_\theta + \mathbf{a}_\theta + \mathbf{n}_\theta$. P_θ is called a parabolic subgroup. These subgroups define the following series of representations.

(i) $\rho = \mathrm{ind}^G_{P_\theta}(\sigma\times\tau)$, $\sigma\in\widehat{M}_\theta$, $\tau\in\widehat{A}_\theta$, $P_\theta = M_\theta A_\theta N_\theta$, a cuspidal parabolic subgroup, defines the principal P_θ series of representation. If P_θ is minimal parabolic and σ is the trivial representation then ρ is irreducible.

(ii) $\rho = \mathrm{ind}^G_{P_\theta}(\sigma\times\tau)$, $\sigma\in\widehat{M}_\theta$, $\tau:A_\theta\to\mathbb{C}^*(=\mathbb{C}-\{0\})$, a nonunitary character; $P_\theta = M_\theta A_\theta N_\theta$ cuspidal parabolic defines the complementary P_θ series of representations.

(iii) $\rho = \mathrm{ind}^G_{P_\theta}(\sigma\times\tau)$, $\sigma\in\widehat{M}_\theta$, $\tau\in\widehat{A}_\theta$, $P_\theta = M_\theta A_\theta N_\theta$, a noncuspidal parabolic, defines the degenerate P_θ series of representations.

C. Discrete UIR's of semisimple Lie groups

A UIR of G is said to belong to the discrete series if it is square integrable. Let H be a maximal compact abelian subgroup of G.

Theorem: G has discrete series representation iff it is also a Cartan subgroup, or equivalently, if rank G = rank K, $G = K\cdot P$ being the Cartan decomposition. Let G satisfy the condition of the above theorem and let H be a compact Cartan subalgebra. As usual, every $\chi\in\widehat{H}$ determines a linear map $\lambda:\mathbf{h}\to i\mathbb{R}$ through

$$\langle\chi,\exp\mathbf{h}\rangle = \exp[\lambda(\mathbf{h})]. \tag{II.14}$$

We emphasize this relation by writing χ as e^λ. The set of all λ which arise in this manner defines a lattice L. Let

$$g = \mathbf{h} + (\oplus_{\alpha\in\Delta^+}\mathbf{g}_\alpha) + (\oplus_{\alpha\in\Delta^-}g_\alpha) \tag{II.15}$$

be the triangular decomposition of **g** w.r.t. **h**, where Δ^+ is the set of positive roots relative to some ordering. We define a map $i:L\to\mathbf{h}_c$, $\lambda\mapsto H_\lambda$ by

$$B(H_\lambda,H) = \lambda(H)\ \forall H\in\mathbf{h}_c. \tag{II.16}$$

This induces a scalar product in L by

$$\langle\lambda_1,\lambda_2\rangle = B(H_{\lambda_1}, H_{\lambda_2}). \tag{II.17}$$

Let $\omega(\lambda) = \prod_{\alpha\in\Delta^+}\langle\lambda,\alpha\rangle$ and $L' = \{\lambda\in L: \omega(\lambda)\neq 0\}$.

Let $D = G/H$. D has a complex structure inherited from G_c/B, where B is a Borel (i.e., maximally solvable) subgroup of G_c. Then

$$H\to G\to D \tag{II.18}$$

defines a principal bundle. Every $\lambda\in L'$ defines an associated line bundle $L_\lambda\to D$ through the character $\chi = e^\lambda$. Let $A^i_0(L_\lambda)$ be the space of C^∞, L_λ valued forms of degree i with compact support. Explicitly, these forms can be written as

$$\omega^{(r)} = \Sigma f_{i_1 i_2\cdots i_r} d\bar{z}_{i_1}\wedge d\bar{z}_{i_2}\cdots\wedge d\bar{z}_{i_r}, \tag{II.19}$$

where the Z_i's are suitable local complex coordinates for D and the f's are sections of the line bundle $L_\lambda\to D$. We know that

$$A^0_0(L_\lambda)\sim\{f\in c^\infty(p^{-1}U)f(gh) = e^\lambda(h^{-1})f(g) \times\forall g\in p^{-1}U,\ h\in H,\ U \text{ open in } D\}, \tag{II.20}$$

where $p:G\to D$ is the canonical projection. Define ∂: $A^i_0(L_\lambda)\to A^{i+1}_0(L_\lambda)$ by

$$\partial\omega^{(\gamma)} = \sum\partial f_{i_1 i_2\cdots i_r}\wedge d\bar{z}_{i_1}\wedge d\bar{z}_{i_2}\cdots\wedge d\bar{z}_{i_r}, \tag{II.21}$$

where

$$\partial f = \sum\frac{\partial f}{\partial\bar{z}_i}d\bar{z}_i.$$

Define a scalar product on $\mathbf{n}\sim T_e(G/H)$,

$$\begin{aligned}(x,y) &= -B(x,\bar{y}) \quad (x,y \text{ compact})\\ &= B(x,\bar{y}) \quad (x,y \text{ noncompact})\\ &= 0 \quad (\text{otherwise}).\end{aligned} \tag{II.22}$$

This induces a scalar product on $T(G/H)$ (by left translation) and hence, by duality on $A^i_0(L_\lambda)$. Let ∂^* be the adjoint of ∂ w.r.t. this scalar product, denote the closure of $A^i_0(L_\lambda)$ by $\mathrm{CL}\{A^i_0(L_\lambda)\}$, and define

$$H^i(L_\lambda) = \{\omega\in cl\{A^i_0(L_\lambda)\},\ \omega \text{ is square integrable},\ \partial\omega = 0,\ \partial^*\omega = 0\} \tag{II.23}$$

$H^i(L_\lambda)$ is known as the ith L^2-cohomology group of L_λ. Define

$$k(\lambda) = \mathrm{card}\{\alpha\in\Delta^+\cap\mathbf{k}|(\lambda,\alpha)<0\} + \mathrm{card}\{\alpha\in\Delta^+\cap\mathbf{p}|(\lambda,\alpha)>0\}. \tag{II.24}$$

Then one has

Theorem: $\exists b$: if $\lambda\in L'$ and $|(\lambda,\alpha)|>b\ \forall\alpha\in D$, then

(i) $H^i(L_\lambda) = 0\ \forall i\neq k(\lambda)$, (ii)$H^i(L_\lambda)$ bears a UIR of G if $i = k(\lambda)$.

Corollary: If $k(\lambda) = 0$, $H^0(L_\lambda)$ bears a UIR which is equivalent to $\text{ind}_H^G e^\lambda$, the induction being in spaces of totally holomorphic functions on G/H, i.e., $\partial f/\partial \bar{z}_i = 0 \ \forall i$.

III. ISOTROPY ALGEBRAS AND POLARIZATIONS FOR SPIN(2,1) AND SPIN (2,2)

A. Spin (2,1)

We describe the spin (2,1) Lie algebra by

$$[J_{ij}, J_{kl}] = g_{ik}J_{jl} + g_{jl}J_{ik} - g_{il}J_{jk} - g_{ik}J_{il},$$
$$J_{ij} = -J_{ji}, \quad i,j,k,l = 1,2,3,4. \tag{III.1}$$

and

$$g_{ij} = \text{diag}\{+1, +1, -1\}.$$

The Cartan–Killing form is defined as

$$-B(J_{12},J_{12}) = B(J_{13},J_{13}) = B(J_{23},J_{23}) = 1. \tag{III.2}$$

There exists one Casimir invariant

$$C = J_{12}^2 - J_{13}^2 - J_{23}^2.$$

There exist four classes of orbits:

Class I: generated by λJ_{12} each λ generating a distinct orbit.

Class II: generated by μJ_{13}, each $\mu > 0$ generating a distinct orbit.

Class III: two orbits generated by $J_{13} \pm J_{12}$.

Class IV: the origin $\{0\}$.

Let $X = \lambda J_{12}$. Then, $[y,x] = 0 \Longrightarrow y = \gamma x$, i.e., $\mathbf{g}_x = \{J_{12}\}$. To compute $\mathbf{s}_x$, we write the isotropicity condition: If $y_1 = a_1J_{12} + a_2J_{13} + a_3J_{23}$, $Y_2 = b_1J_{12} + b_2J_{13} + b_3J_{23}$, then $B(X,[Y_1,Y_2]) = 0 \Longrightarrow a_2b_3 - a_3b_2 = 0$, i.e., $a_2/a_3 = b_2/b_3 = \alpha$ (say). The condition $[\mathscr{S}_x, S_x] \subseteq \mathbf{s}_X$ then implies that $\alpha = \pm i$; finally the positivity condition yields $\alpha = -i\text{sgn}\lambda$. Thus,

$$\mathscr{S}_x = \{J_{12}, J_{13} - i(\text{sgn}\lambda)J_{23}\}. \tag{III.3}$$

This yields $\boldsymbol{\delta} = \mathbf{g}_X$, $\mathbf{e} = \mathbf{g}$. One can also see that X is quantizable iff $2\pi\lambda \in \mathbb{Z}$, the corresponding character being given by

$$\chi(\exp \alpha J_{12}) = \exp(2\pi i \lambda \alpha). \tag{III.4}$$

Similarly, one gets the following results:

Class II: $\mathbf{g}_X = \{J_{13}\}$, $\mathbf{s}_X = \{J_{13}, J_{12} \pm J_{23}\}$, $\boldsymbol{\delta} = \mathbf{e} = \mathbf{s}_X$. All orbits are quantizable.

Class III:
$\mathbf{g}_X = \{J_{13} \pm J_{12}\}$, $\mathbf{s}_X = \{J_{23}, J_{13} \pm J_{12}\} = \boldsymbol{\delta} = \mathbf{e}$. Since $2\pi i x$: $\mathbf{g}_X \to i\mathbb{R}$ is the trivial map, X is quantizable.

B. Spin (2,2)

We describe the spin (2,2) Lie algebra by

$$[J_{ij}, J_{kl}] = g_{ik}J_{jl} + g_{jl}J_{ik} - g_{il}J_{jl} - g_{jk}J_{il},$$
$$J_{ij} = -J_{ji}, \quad i,j,k,l = 1,2,3,4, \tag{III.5}$$

where

$$g_{ij} = \text{diag}\{1,1,-1,-1\}.$$

We define the Cartan–Killing form by

$$B(J_{ij},J_{kl}) = g_{il}g_{jk} - g_{ik}g_{jl}. \tag{III.6}$$

The two Casimir invariants are

$$C_{\text{I}} = J_{12}^2 + J_{34}^2 - J_{13}^2 - J_{23}^2 - J_{14}^2 - J_{24}^2,$$
$$C_{\text{II}} = J_{12}\cdot J_{34} - J_{13}\cdot J_{24} + J_{14}\cdot J_{23}. \tag{III.7}$$

The orbits can be written down as:

Class I: generated by $\lambda J_{12} + \mu J_{34}$, each pair (λ,μ) generating a distinct orbit.

Class II: generated by $\lambda J_{13} + \mu J_{24}$; the pairs (λ,μ) and $(-\lambda,-\mu)$ generate the same orbit.

Class III: generated by $\lambda J_{12} + \mu J_{34} + \nu J_{13}$, $\nu^2 = \lambda^2 + \mu^2$. ν, $-\nu$ generates the same orbit while (λ,μ,ν) and $(-\lambda,-\mu,-\nu)$ generate distinct orbits.

Class IV: two orbits generated by $J_{13} \pm J_{12}$.

Class V: the origin $\{0\}$.

Let $X = \lambda J_{12} + \mu J_{34}$, and let $\mathbf{h} = \{J_{12}, J_{34}\}$.

Let

$$\mathbf{g}_{\pm\alpha_1} = \{J_{13} \pm iJ_{23} \pm iJ_{14} - J_{24}\}, \tag{III.8}$$
$$\mathbf{g}_{\pm\alpha_2} = \{J_{13} \pm iJ_{23} \mp iJ_{14} + J_{24}\}. \tag{III.9}$$

The two-dimensional space spanned by λ,μ can be conveniently divided into six subspaces:

(i)$\lambda + \mu > 0, \lambda - \mu > 0$: $\mathbf{g}_X = \mathbf{h}$,
$\mathbf{s}_X = \mathbf{g}_X + \mathbf{g}_{-\alpha_1} + \mathbf{g}_{-\alpha_2}$.
(ii)$\lambda + \mu < 0, \lambda - \mu < 0$: $\mathbf{g}_X = \mathbf{h}$,
$\mathbf{s}_X = \mathbf{g}_X + \mathbf{g}_{\alpha_1} + \mathbf{g}_{\alpha_2}$.
(iii)$\lambda + \mu < 0, \lambda - \mu > 0$: $\mathbf{g}_X = \mathbf{h}$,
$\mathbf{s}_X = \mathbf{g}_X + \mathbf{g}_{-\alpha_1} + \mathbf{g}_{\alpha_2}$.
(iv)$\lambda + \mu > 0, \lambda - \mu < 0$,
$$\mathbf{g}_X = \mathbf{h}, \mathbf{s}_X = \mathbf{g}_X + \mathbf{g}_{-\alpha_1} + \mathbf{g}_{-\alpha_2}. \tag{III.10}$$

For the above four subspaces, $\boldsymbol{\delta} = \mathbf{g}_X$, $\mathbf{e} = \mathbf{g}$. The characters are given by

$$\chi[\exp(\alpha J_{12} + \beta J_{34})] = \exp i(\lambda\alpha + \mu\beta),\ 2\pi\lambda \in \mathbb{Z},\ 2\pi\mu \in \mathbb{Z}.$$

(v) $\lambda = \mu$, and

(vi) $\lambda = -\mu$: $\mathbf{g}_X = \{J_{13} \mp J_{24}, J_{14} \pm J_{23}\} + \mathbf{h}$. (III.11)

$$\mathbf{s}_X = \mathbf{h} + \mathbf{g}_{-\alpha_1} + \mathbf{g}_{\alpha_2} + \mathbf{g}_{-\alpha_2} \quad (\lambda = \mu)$$

and

$$\mathbf{s}_X = \mathbf{h} + \mathbf{g}_{\alpha_1} + \mathbf{g}_{-\alpha_1} + \mathbf{g}_{-\alpha_2} \quad (\lambda = -\mu).$$

For both cases $\boldsymbol{\delta} = \mathbf{g}_X$, $\mathbf{e} = \mathbf{g}$. The characters are

$$\chi\{\exp[\alpha_1 J_{12} + \alpha_2 J_{34} + \alpha_3(J_{13} \mp J_{24}) + \alpha_4(J_{14} \pm J_{23})]\}$$
$$= \exp i\lambda(\alpha_1 \pm \alpha_2),\ 2\pi\lambda \in \mathbb{Z}. \tag{III.12}$$

Similarly, for the other classes, we have the following results:

Class II: Let $\mathbf{a} = \{J_{13}, J_{24}\}$;
$$\tilde{\mathbf{g}}_{\pm\alpha_1} = J_{12} + J_{34} \pm (J_{14} - J_{23}).$$
$$\tilde{\mathbf{g}}_{\pm\alpha_2} = J_{12} - J_{34} \pm (J_{14} - J_{23}). \tag{III.13}$$

We divide Class II into two subclasses:

(i) $|\lambda| \neq |\mu|$, $\mathbf{g}x = \mathbf{a}$. There are four polarizations:

$$\mathbf{s}_1 = \mathbf{a} + \mathbf{g}_{\alpha_1} + \mathbf{g}_{\alpha_2},$$
$$\mathbf{s}_2 = \mathbf{a} + \mathbf{g}_{\alpha_1} + \mathbf{g}_{-\alpha_2},$$
$$\mathbf{s}_3 = \mathbf{a} + \mathbf{g}_{-\alpha_1} + \mathbf{g}_{\alpha_2},$$
$$\mathbf{s}_4 = \mathbf{a} + \mathbf{g}_{-\alpha_1} + \mathbf{g}_{-\alpha_2}. \tag{III.14}$$

All the four polarizations are positive and $\boldsymbol{\delta} = \mathbf{e} = \mathbf{s}$, where $\mathbf{s}$

can be any of the four polarizations given above.

(ii) $\lambda = \pm\mu$: $\mathbf{g}_X = \mathbf{a} + \{J_{12} \mp J_{34}, J_{14} \pm J_{23}\}$ respectively. There exist two polarizations for each sign:

$$\left.\begin{array}{l}\mathbf{s}_1 = \mathbf{a} + \mathbf{g}_{\alpha_1} + \mathbf{g}_{\alpha_2} + \mathbf{g}_{-\alpha_2}\\ \mathbf{s}_2 = \mathbf{a} + \mathbf{g}_{-\alpha_1} + \mathbf{g}_{\alpha_2} + \mathbf{g}_{-\alpha_2}\end{array}\right\}, \quad \lambda = \mu, \tag{III.15a}$$

$$\left.\begin{array}{l}\mathbf{s}_1 = \mathbf{a} + \mathbf{g}_{\alpha_1} + \mathbf{g}_{-\alpha_1} + \mathbf{g}_{\alpha_2}\\ \mathbf{s}_2 = \mathbf{a} + \mathbf{g}_{\alpha_1} + \mathbf{g}_{-\alpha_1} + \mathbf{g}_{-\alpha_2}\end{array}\right\}, \quad \lambda = -\mu. \tag{III.15b}$$

Once again $\boldsymbol{\delta} = \mathbf{e} = \mathbf{s}$, where $\mathbf{s}$ is either of the two polarizations given for each case. The Kostant integrality condition imposes no conditions on λ.

Class III: Define

$$\lambda' = \lambda(\lambda^2 + \mu^2)^{1/2}/(\lambda^2 - \mu^2),$$
$$\mu' = \mu(\lambda^2 + \mu^2)^{1/2}/(\lambda^2 - \mu^2). \tag{III.16}$$

Then,

$$\mathbf{g}_X = \{J_{12} + \mu' J_{13} - \lambda' J_{24}, J_{34} - \lambda' J_{13} + \mu' J_{24}\}, \tag{III.17}$$

and

$$\mathbf{s}_X = \{J_{12} - (\lambda' + \mu')J_{24}, J_{13} + J_{24}, J_{14} + J_{23}, J_{34} + (\lambda' + \mu')J_{24}\}; \tag{III.18}$$

$\boldsymbol{\delta} = \mathbf{e} = \mathbf{s}_X$. The Kostant integrality condition imposes no conditions on λ,μ.

Class IV: We have,

$$\mathbf{g}_X = \{J_{12} - J_{13}, J_{24} - J_{34}\}, \tag{III.19}$$
$$\mathbf{s}_X = \{J_{14}, J_{23}, J_{12} - J_{13}, J_{34} - J_{24}\},$$

$\boldsymbol{\delta} = \mathbf{s}_X = \mathbf{e}$. The Kostant integrality condition is vacuous.

IV. PARABOLIC SUBALGEBRAS AND POLARIZATIONS

In this section, we show the connection between parabolic subalgebras and some of the polarizations given above.

A. Spin (2,1)

We choose the Cartan decomposition as

$$\mathbf{k} = \{J_{12}\}, \quad \mathbf{p} = \{J_{13}, J_{23}\}. \tag{IV.1}$$

Choose $\mathbf{a} = \{J_{13}\}$. Trivially, $\mathbf{m} = 0$. $\mathbf{n}$ can be easily calculated and shown to be

$$\mathbf{n} = \{J_{13} + J_{23}, J_{13} - J_{23}\}. \tag{IV.2}$$

One concludes that the polarization given in Class II is a minimal parabolic subalgebra. The Kostant induction scheme reduces to $\text{ind}_{P_1}^G(\chi)$ where P_1 is the corresponding minimal parabolic subgroup of spin(2,1). Note that $\chi \in \hat{A}$ and hence the Kostant induction corresponds to the conventional induction scheme with $\sigma \sim 1_M$. The representations, by Kostant's theorem, are UIR's. In identical fashion, we see that the polarizations for Class III are also minimally parabolic with Langland's decomposition

$$\mathbf{a} = \{J_{12} \pm J_{23}\}, \quad m = \{0\}, \quad \mathbf{n}^+ = \{J_{13}\}. \tag{IV.3}$$

Once again $\chi \sim 1_M \times \tau$, $\tau \in \hat{A}$ and hence, the representations are UIR's. These sets of polarizations give rise to the principal series of UIR's of spin(2,1).

B. Spin (2,2)

The Cartan decomposition is

$$\mathbf{k} = \{J_{12}, J_{34}\}, \quad \mathbf{p} = \{J_{13}, J_{14}, J_{23}, J_{24}\}. \tag{IV.4}$$

Choose $\mathbf{a} = \{J_{13}, J_{24}\}$; we have $m = \{0\}$ and

$$\mathbf{n} = \tilde{\mathbf{g}}_{\alpha_1} + \tilde{\mathbf{g}}_{\alpha_2} + \tilde{\mathbf{g}}_{-\alpha_1} + \tilde{\mathbf{g}}_{-\alpha_2}, \tag{IV.5}$$

using the notation of Sec. III B. Then, all the polarizations given in Class II (i) are minimal parabolic with $\chi \sim 1_M \times \tau, \tau \in \hat{A}$. These polarizations yield the principal series of UIR's of spin (2,2). For Class III, we choose

$$\mathbf{a} = \{J_{12} + \mu' J_{13} - \lambda' J_{24}, J_{34} - \lambda' J_{13} + \mu J^{24}\},$$
$$m = \{0\}, \tag{IV.6}$$
$$\mathbf{n} = \{J_{13} + J_{24} + (\lambda' + \mu')J_{34} - (\lambda' + \mu')J_{12} \pm \rho_1(J_{23} + J_{14}),$$
$$J_{13} + J_{24} + (\lambda' + \mu')J_{34} - (\lambda^1 + \mu^1)J_{12} \pm \rho_2(J_{23} + J_{14})\}$$

where $\rho_1 = |[(\lambda' + \mu')^2 - 1]^{1/2}|$,

$\rho_2 = -|[(\lambda' - \mu')^2 - 1]^{1/2}|$.

$\mathbf{s}_x$ can then be shown to be equal to $\mathbf{m} + \mathbf{a} + \mathbf{n}^+$, hence it is minimal parabolic.

Once again, we get the principal series of UIR's. For class IV, the Langland decomposition of the polarization is

$$\mathbf{a} = \{J_{12} - J_{13}, J_{24} - J_{34}\}, m = \{0\},$$
$$n^+ = \{J_{14}, J_{23}\}. \tag{IV.7}$$

This class of orbits yield the principal series of UIR's. Finally, in class II (ii), it is possible to show that the polarizations are parabolic, with Langland's decomposition

$$\mathbf{m}_\theta = \{J_{12} + J_{34}, J_{14} - J_{23}, J_{13} + J_{24}\}, \tag{IV.8a}$$
$$\mathbf{a}_\theta = \{J_{13} - J_{24}\},$$
$$\mathbf{n}_\theta = \{J_{12} - J_{34} + J_{14} + J_{23}\} \quad (\lambda = \mu),$$

and

$$\mathbf{m}_\theta = \{J_{12} - J_{34}, J_{14} + J_{23}, J_{13} - J_{24}\}, \tag{IV.8b}$$
$$\mathbf{a}_\theta = \{J_{13} + J_{24}\},$$
$$\mathbf{n}_\theta = \{J_{12} - J_{34} + J_{14} - J_{23}\} \quad (\lambda = -\mu),$$

corresponding to the choices $\theta = \{\alpha_1\}, \{\alpha_2\}$, respectively. These subalgebras are noncuspidal, and hence, generate the degenerate series of representations.

V. DISCRETE SERIES

A. Spin (2,1)

We have a compact Cartan subalgebra

$$\mathbf{h} = \mathbf{k} = \{J_{12}\}$$

The corresponding root is $\alpha(J_{12}) = i$, with

$$g_{\pm\alpha} = \{J_{13} \pm iJ_{23}\} \subseteq \mathbf{p}, \quad H_\alpha = -iJ_{12}. \tag{V.1}$$

The characters of H, as noted in Sec. III, are given by

$$\chi(\exp aJ_{12}) = \exp(ina), \quad n \in \mathbb{Z}. \tag{V.2}$$

Hence,

$$L = \{\lambda: \mathbf{h} \to \mathbb{C}, \lambda = n\alpha, n \in \mathbb{Z}\}, \tag{V.3}$$

with

$$H_\lambda = -inJ_{12}.$$

Since $(\lambda,\alpha) \equiv B(H_\lambda, H_\alpha) = n$, $L' = \{\lambda \in L: \lambda = n\alpha, n \neq 0\}$.

We note that there exist an isomorphism between the set of Kostant integrable elements of $\mathbf{h}$, denoted $\mathbf{h}_K$, and L. Let $X \in \mathbf{h}$ and let λ_X be the element of $\mathbf{h}^*$ associated with X by the Cartan–Killing map. Then

$$\lambda_X \in L \Longleftrightarrow X/2\pi \in \mathbf{h}_K. \tag{V.4}$$

We choose a canonical complex structure on G/H by

$$T_e(G/H)_c = \mathbf{g}_\alpha, \Delta^+{}_c = \{-\alpha\}, \tag{V.5}$$

and choose a corresponding complex coordinate z. Note that $(\lambda,\alpha) = n$.

(i) $n > 0$: It follows from Eq. (II.24) that $k(\lambda) = 0$ with respect to Δ_c^+. Hence, the Schmidt theory yields us UIR's in terms of functions of G which are holomorphic in Z. The Kostant induction scheme can be written as $\operatorname{ind}_H^G e^\lambda$.

Choose a complex structure and a positive subspace of $\mathbf{n}$ in the following way:

$$\mathbf{n}' = \{\mathbf{g}_\alpha : \mathbf{g}_{-\alpha} \subseteq \mathbf{s}_X\}, \tag{V.6}$$

where $\mathbf{s}_X$ is the corresponding positive polarization and define,

$$T_e(G/H)_K = \mathbf{n}^-. \tag{V.7}$$

It is obvious that in this case, the complex structure defined by $T_e(G/H)_c$ and $T_e(G/H)_K$ are identical, and hence, so are the corresponding representations.

(ii) $n < 0$: Once again, it is easy to see that $k(\lambda) = 1$ and hence the Schmid theory yield UIR's in spaces of 1-forms. Note, however, that the prescription yields $T_e(G/H)_K = g_{-\alpha}, \Delta_K^+ = \{\alpha\}$ with respect to which $k(\lambda) = 0$. Hence, the Kostant induction yields representations in spaces of functions which are holomorphic w.r.t. $T_e(G/H)_K$ or equivalently, antiholomorphic w.r.t. $T_e(G/H)_c$.

B. Spin (2,2)

The calculations for spin (2,2) are entirely analogous and yield the following:

$$\begin{aligned} &\mathbf{h} = \mathbf{k} = \{J_{12}, J_{34}\}, \\ &L = \{\lambda \in \mathbf{h}^*, \lambda = n_1\alpha_1 + n_2\alpha_2, n_1,n_2 \in \mathbb{Z}\}, \\ &L' = \{\lambda \in L : |n_1| \neq |n_2|\}, \end{aligned} \tag{V.8}$$

where α_1,α_2 are the roots associated with $\mathbf{g}_{\alpha_1},\mathbf{g}_{\alpha_2}$ introduced in Sec. II. We once again have an isomorphism between L and the set of Kostant-integral elements of $\mathbf{h}$. Define

$$T_e(G/H)_c = \mathbf{g}_{\alpha_1} + \mathbf{g}_{\alpha_2}. \tag{V.9}$$

and denote the corresponding complex coordinates by Z_1,Z_2 respectively. Let, as above $T_e(G/H)_K$ be the complex structure defined by the corresponding polarization. Then, one has for Class I,

(i) $n_1 + n_2 > 0$, $n_1 - n_2 > 0$,

$$T_e(G/H)_K = \mathbf{g}_{\alpha_1} + \mathbf{g}_{\alpha_2}, \quad \Delta^+{}_K = \{-\alpha_1, -\alpha_2\}. \tag{V.10}$$

Kostant-induced representation is in the space functions holomorphic in both Z_1 and Z_2.

(ii) $n_1 + n_2 < 0$, $n_1 - n_2 > 0$,

$$T_e(G/H)_K = \mathbf{g}_{\alpha_1} + \mathbf{g}_{-\alpha_2}; \quad \Delta_K^+ = \{-\alpha_1,\alpha_2\}. \tag{V.11}$$

Representation is in the space of functions holomorphic in Z_1 and antiholomorphic in Z_2.

(iii) $n_1 + n_2 < 0$ $n_1 - n_2 < 0$,

$$T_e(G/H)_K = \mathbf{g}_{-\alpha_1} + \mathbf{g}_{-\alpha_2}; \quad \Delta_K^+ = \{\alpha_1,\alpha_2\}. \tag{V.12}$$

A space of functions antiholomorphic in both Z_1 and Z_2 is to be used.

(iv) $n_1 + n_2 > 0$, $n_1 - n_2 < 0$,

$$T_e(G/H)_K = \mathbf{g}_{\alpha_1} + \mathbf{g}_{-\alpha_2}; \quad \Delta_K^+ = \{-\alpha_1,\alpha_2\}. \tag{V.13}$$

Representation space is the space of functions antiholomorphic in Z_1 and holomorphic in Z_2. In all the above cases $K(\lambda) = 0$ w.r.t. the corresponding. Δ_K^+ and hence, the Kostant representation and the Schmid representation coincide. Note, however, that case (v) and (vi) of Class I are outside the purview of Schmid theory (they belong to Kerω).

VI. CONCLUSION

Our results can be summarized as follows:

(1) The Auslander–Kostant theory yields representation which coincide with those given by the conventional theory of principal and degenerate series.

(2) Further, the Auslander–Kostant theory, through the concept of a complex polarization provides a natural complex structure w.r.t. which $K(\lambda) = 0$, and hence, yields UIR's in the space of functions which are holomorphic w.r.t. this complex structure.

(3) However, the present investigation does not throw any light on the complementary series of representation. This could possibly be done by considering orbits in $\mathbf{g}_c^*$ rather than $\mathbf{g}^*$, and inducing from nonunitary characters.

[1]J. G. Kuriyan, N. Mukunda, and E. C. G. Sudarshan, Commun. Math. Phys. **8**, 204 (1963); A. O. Barut and C. Fronsdal, Proc. R. Soc. London, Ser. A **287**, 532 (1965); C. Fronsdal in Proceedings of the IAEA Conference, Trieste, 1965, IAEA, Vienna, 1965; A. O. Barut and R. N. Philips, Commun. Math. Phys. **8**, 52 (1968); E. C. G. Sudarshan, in *Proceedings of the third Corals Gables Conference* 1966 edited by A. Perlmutter, J. Wojtsazek, E. C. G. Sudarshan, and B. Kursunonglu (Freeman, San Francisco, 1966).
[2]L. Auslander and B. Kostant, Invet. Math. **14**, 255 (1971).
[3]A. Kirillov, Elements of the Theory of Representations, *Grundlehren der Mathematischen Wissenchaften*, Vol. **220** (Springer-Verlag, Berlin, 1976).
[4]J. H. Rawnsley, Proc. Cambridge Philos. Soc. **76**, 473 (1974).
[5]W. Schmid, Ann. Math. **93**, 1 (1971).
[6]Harish-Chandra, Acta Math. **113**, 241 (1965); **116**, 1 (1966).
[7]R. L. Lipsmann, *Group Representations, Lecture Notes in Mathematics*, Vol. 388 (Springer, New York, 1974).
[8]J. F. Cornwall, J. Math. Phys. **16**, 1992 (1975); **20**, 547 (1979).

The Riccati equation and Hamiltonian systems [a]

F. Guil Guerrero
Departamento de Métodos Matemáticos de la Física, Facultad de Ciencias Físicas, Universidad Complutense, Madrid-3, Spain

(Received 9 December 1980; revised manuscript accepted for publication 10 August 1981)

The involutive system of functionals associated by Gel'fand and Dikii to a nth-order scalar differential operator is obtained from a set of solutions of a generalized Riccati equation. These solutions allow us to explain the involutive character of the system of functionals in terms of the Riccati equation properties only.

PACS numbers: 02.30. + g

INTRODUCTION

In this paper we discuss some well-known facts concerning the Hamiltonian systems constructed by Gel'fand and Dikii. In a series of papers [1–3] devoted to the construction of infinite-dimensional examples of completely integrable Hamiltonian systems, these authors stressed the important role played by the resolvent of differential operators to obtain an infinite set of functionals in involution.

In the last years much work has been done in the present subject of Lax equations, constants of motion, and involuteness. One can cite as representatives the papers of Manin[4], Wilson,[5] and Adler.[8]

In Ref. 3 the Riccati equation is presented in connection with the theory of the resolvent of a scalar differential operator of order n. We shall prove here that the set of first integrals mentioned above are the coefficients of the power series solutions of this Riccati equation. Such a Riccati equation is associated directly to the differential operator of order n.

Section I is concerned with preliminary aspects. In Sec. II we introduce the Riccati equation for the scalar nth-order operator (2.1). Some formulas involving determinants are used to get the necessary collection of functionals for the construction of the integrable systems. The variational derivatives of the solutions of the Riccati equation characterized in Sec. II are calculated in Sec. III. Finally, in Sec. IV it is proved that the Lax-type equations admit as constants of motion the functionals related to the Riccati equation mentioned above. That is all that one needs to prove that the system of functionals so constructed is in involution.

I. THE RING OF DIFFERENTIAL POLYNOMIALS

We shall summarize here some aspects of the algebra of differential polynomials which will be used later on. For more information see Refs. 1–3.

Let $A(u)$ denote the ring of polynomials in the letters u_1, $u_2, \ldots, u_N, u_1', u_2', \ldots, u_N', u_1'', u_2'', \ldots$, where the u_k are functions on the real variable x and $u_k^{(j)} = \partial^j u_k$. By ∂ we denote the total derivative with respect to x; $\partial f = \Sigma u_k^{(j+1)} \partial f / \partial u_k^{(j)}$, $f \in A(u)$ and also we shall put $f^{(k)} = \partial^k f$

The space $\tilde{A}(u)$ of functionals is defined as the set of the equivalence classes $\tilde{f} = \{ f_1, f_2, \ldots, \in A : f_i - f_j \in \partial A(u) \}$, where $\partial A(u)$ is the set of elements of $A(u)$ which can be written as total derivatives. Then $\tilde{A}(u) = A(u)/\partial A(u)$ and the notation $\tilde{f} = \int dx\, f$ will be used for the class $\tilde{f}$ which contains f.

[a] Partially supported by the J. E. N., Instituto de Estudios Nucleares, Madrid.

$\tilde{f} = \int dx\, f$ will be used for the class $\tilde{f}$ which contains f.

The variational derivative operator $\delta / \delta u_k, k = 1, \ldots, N$ acts on $A(u)$ by the formula

$$\frac{\delta f}{\delta u_k} = \sum_l (-\partial)^l \frac{\partial f}{\partial u_k^{(l)}}, \quad k = 1,2,\ldots,N. \tag{1.1}$$

With the variational derivative one characterizes $\partial A(u)$ in the following manner:

$$f \in \partial A(u) \Leftrightarrow \frac{\delta f}{\delta u_k} = 0, \quad k = 1,2,\ldots,N \tag{1.2}$$

and this result tells us that $\delta f / \delta u_k = \delta \tilde{f} / \delta u_k$ and allows us to obtain, by means of integration by parts, the formula

$$\frac{d\tilde{f}}{dt} = \int dx \sum_k \frac{\delta \tilde{f}}{\delta u_k} \dot{u}_k \tag{1.3}$$

when the u_k are left to depend on some parameter t, $\dot{u}_k = du_k/dt$.

Differential 1-forms are introduced as finite sums $\omega = \Sigma \omega_{kj} \delta u_k^j$, where $\omega_{kj} \in A(u)$ and the δu_k^j are new independent variables.

We define δf by $\delta f = \Sigma \delta u_k^{(j)} \partial f / \partial u_k^{(j)}$. By using the relation[3] $\partial \delta u_k^{(j)} = \delta u_k^{(j+1)}$, δf may be written as $\delta f = \Sigma_k d_k f \delta u_k$, where $d_k f = \Sigma_j \partial f / \partial u_k^{(j)} \partial^j$ is the Gateaux differential[6] of f with respect to u_k,

$$d_k f a = \left. \frac{d}{d\epsilon} \right|_{\epsilon = 0} f(u_1, \ldots, u_k + \epsilon a, \ldots, u_N). \tag{1.4}$$

Here a is some arbitrary function.

One can prove that δf can be written uniquely in the form

$$\delta f = \sum_k g_k \delta u_k + \partial \omega \tag{1.5}$$

with $g_k \in A(u)$, ω is an appropriate 1-form, and $g_k = \delta f / \delta u_k$.

Let now z be a complex variable. We construct $A(u, z^{-1})$ as the set of formal power series $\xi = \Sigma \alpha_r f_r / z^r$. The first coefficient f_{r_0} is constant and the α_r are complex numbers, the other terms $f_r \in A(u)$. We shall use the notation $\xi = \xi_0 / z^{r_0} + O(z^{-r_0 - 1})$. The set $A(u, z^{-1})$ inherits the operations ∂, $\delta / \delta u_k$, and δ introduced for $A(u)$. In the same way as we did for $A(u)$ one gets $\tilde{A}(u, z^{-1})$ from $\tilde{A}(u)$. To $\xi = \xi_{r_0} / z^{r_0} + O(z^{-r_0 - 1})$ we make correspond

$$\tilde{\xi} = \int dx \left(\frac{\xi_{r_0}}{z^{r_0}} - \xi \right)$$

 0022-2488/82/020211-06$02.50

to remove the constant term ξ_{r_0} in ξ.

II. THE RICCATI EQUATION

Let us consider a differential operator

$$L = \sum_{k=0}^{n} u_k \partial^k \tag{2.1}$$

for which $u_n = 1$, $u_{n-1} = 0$. We assume that u_k, $k = 0,1...,n-2$ are free generators of the ring $A(u)$. In order to obtain the system of functionals mentioned above, we examine the differential equation $(L - z^n)\varphi = 0$ for the solutions φ which can be represented in the form $\varphi(x,z) = \exp\int^x \chi(x,z)dx$. We shall denote the primitive of $\chi(x,z)$ by $\hat{\chi}(x,z) = \int^x \chi(x,z)dx$.

It is easy to see that

$$(L - z^n)e^{\hat{\chi}} = 0. \tag{2.2}$$

is equivalent to the Riccati equation

$$\sum_{k=0}^{n} u_k P_k(\chi) = z^n, \tag{2.3}$$

where

$$P_k(\chi) = e^{-\hat{\chi}}(\partial^k e^{\hat{\chi}}) \tag{2.4}$$

is a differential polynomial in χ. Thus $P_0 = 1$, $P_1 = \chi$, $P_2 = \chi' + \chi^2$,.... One can prove by induction that P_k admits also the expression $P_k(\chi) = (\partial + \chi)^k 1$ and that $P_k(\chi) = \chi^{(k-1)} + ... + \chi^k, k = 2,3,...$ .The following proposition holds for Eq. (2.3).

Proposition 2.1: Equation (2.3) has n solutions $\chi^{[i]} \in A(u,z^{-1})$, $i = 0,1,...n-1$ with $\chi^{[i]} = \epsilon_i z + O(z^{-1})$. The ϵ_i, $i = 0,...,n-1$ are the nth roots of unity.

Proof: We define $\sigma^{[i]} = \chi^{[i]} - \epsilon_i z = \Sigma_{r\geqslant 1}\chi_r^{[i]}/z^r$. For $\sigma^{[i]}$ we have the differential equation

$$\sum_{l=0}^{n-1} (\epsilon_i z)^l \sum_{k=l}^{n} \binom{k}{l} u_k P_{k-l}(\sigma^{[i]}) = 0$$

since from (2.4)

$$P_k(\chi^{[i]}) = \sum_{l=0}^{k} \binom{k}{l} (\epsilon_i z)^l P_{k-l}(\sigma^{[i]}).$$

The equation for $\sigma^{[i]}$ can be written in the form

$$\sigma^{[i]} = -\frac{1}{n}\sum_{l=0}^{n-2} \frac{\epsilon_i^{l+1}}{z^{n-1-l}} \sum_{k=l}^{n} \binom{k}{l} u_k P_{k-l}(\sigma^{[i]})$$

in which we introduce the power series

$$P_k(\sigma^{[i]}) = \sum_{r\geqslant 1} \frac{1}{z^r}(P_k^{[i]})_r, \qquad k = 1,2,...$$

and $P_0 = 1$, to get the relation

$$\chi_r^{[i]} = -\frac{1}{n}\sum_{l=0}^{n-2} \epsilon_i^{l+1} \sum_{k=l}^{n} \binom{k}{l} u_k (P_{k-l}^{[i]})_{r+l+1-n}. \tag{2.5}$$

By virtue of the polynomial character of $P_k(\sigma^{[i]}),(P_k^{[i]})_r$ is a differential polynomial on the first r coefficients $\chi_1^{[i]},\chi_2^{[i]},..., \chi_r^{[i]}$. Then, (2.5) is a recurrence formula which allows us to calculate the coefficients $\chi_r^{[i]}$ as differential polynomials on $(u_0,...,u_{n-2})$.

Corollary 2.1: If we take $\epsilon_k = e^{2\pi i k/n}, k = 0,1,...,n-1$ we have the relations $\chi^{[k+1]}(x,z) = \chi^{[k]}(x,\epsilon_1 z)$, for $k = 0,1,...,n-2$. $\chi^{[n-1]}(x,\epsilon_1 z) = \chi^{[0]}(x,z)$.

Proof: If $\chi(x,z)$ is a solution of (2.3) so is $\chi(x,\epsilon_k z)$, $k = 0,1,...,n-1$. But $\chi^{[k]}(x,\epsilon_1 z) = \epsilon_{k+1} z + O(z^{-1})$ is just $\chi^{[k+1]}(x,z)$. At this point we observe that all the solutions $\chi^{[k]}$ are equivalent with respect to their dependence in $(u_0, u_1,..,u_{n-2})$; in fact one has

$$\chi_r^{[k+1]} = \frac{1}{\epsilon_1^r}\chi_r^{[k]}.$$

Corollary 2.2:

$$\frac{1}{n}\sum_{k=0}^{n-1}\chi^{[k]} = \sum_{r\geqslant 1}\frac{\chi_{rn}^{[0]}}{(z^n)^r}.$$

As a matter of fact from Eq. (2.5) we have for the first coefficients

$$\chi_1^{[0]} = -\frac{1}{n}u_{n-2}, \quad \chi_2^{[0]} = \frac{n-1}{2n}u'_{n-2} - \frac{1}{n}u_{n-3},$$

$$\chi_3^{[0]} = -\frac{1}{n}\partial\left[\binom{n}{2}\chi_2^{[0]} + \binom{n}{3}\partial\chi_1^{[0]}\right] - \frac{1}{n}\left(u_{n-4} + \frac{3-n}{2n}u_{n-2}^2\right). \tag{2.6}$$

Now we look at the solutions of the equation

$$(L^* - z^n)\psi = 0,$$

$L^* = \Sigma_{k=0}^n(-\partial)^k u_k$ is the adjoint operator of the operator L [(2.1)]. We form the Wronskian determinant for the solutions $\varphi_0, \varphi_1,...,\varphi_{n-1}$ of $(L - z^n)\varphi = 0$ obtained from the $\chi^{[i]}$ of Proposition 2.1, that is

$$W = \begin{vmatrix} \varphi_0^{(n-1)} & \varphi_1^{(n-1)} & \cdots & \varphi_{n-1}^{(n-1)} \\ \varphi_0^{(n-2)} & \varphi_1^{(n-2)} & \cdots & \varphi_{n-1}^{(n-2)} \\ \vdots & \vdots & \cdots & \vdots \\ \cdot & \cdot & \cdots & \cdot \\ \varphi_0 & \varphi_1 & \cdots & \varphi_{n-1} \end{vmatrix}. \tag{2.7}$$

By (2.4) we can write the derivatives of φ_i as

$$\varphi_i^{(k)} = P_k(\chi^{[i]})e^{\hat{\chi}^{[i]}},$$

which after the substitution in (2.7) allows us to write W in the form

$$W = \Omega \exp\sum_{i=0}^{n-1}\hat{\chi}^{[i]}, \tag{2.8}$$

where Ω is the determinant

$$\Omega = \begin{vmatrix} P_{n-1}(\chi^{[0]}) & P_{n-1}(\chi^{[1]}) & \cdots & P_{n-1}(\chi^{[n-1]}) \\ P_{n-2}(\chi^{[0]}) & P_{n-2}(\chi^{[1]}) & \cdots & P_{n-2}(\chi^{[n-1]}) \\ \vdots & \vdots & \cdots & \vdots \\ \cdot & \cdot & \cdots & \cdot \\ \chi^{[0]} & \chi^{[1]} & \cdots & \chi^{[n-1]} \\ 1 & 1 & \cdots & 1 \end{vmatrix}. \tag{2.9}$$

We note that $\Omega \in A(u,z^{-1}),\Omega$ being an element of the form

$$\Omega = |\epsilon| z^{n(n-1)/2} + O(z^{(n(n-1)/2)-1}), \tag{2.10}$$

where $|\epsilon|$ is the Vandermonde determinant

$$|\epsilon| = \begin{vmatrix} \epsilon_0^{n-1} & \epsilon_1^{n-1} & \cdots & \epsilon_{n-1}^{n-1} \\ \epsilon_0^{n-2} & \epsilon_1^{n-2} & \cdots & \epsilon_{n-1}^{n-2} \\ \vdots & \vdots & \cdots & \vdots \\ \epsilon_0 & \epsilon_1 & \cdots & \epsilon_{n-1} \\ 1 & 1 & \cdots & 1 \end{vmatrix} = \prod_{k<l}(\epsilon_k - \epsilon_l),$$

which guarantees that $W \neq 0$ since $\epsilon_k \neq \epsilon_l$ for $k \neq 1$. We introduce further the cofactors W_k of the elements in the first row of W. Concerning the W_k we have the following.

Proposition 2.2:

$$(L^* - z^n)W_k = 0, \quad k = 0,1,\ldots,n-1. \tag{2.11}$$

Proof: We consider W_0. From (2.7)

$$W_0 = \begin{vmatrix} \varphi_1^{(n-2)} & \varphi_2^{(n-2)} & \cdots & \varphi_{n-1}^{(n-2)} \\ \varphi_1^{(n-3)} & \varphi_2^{(n-3)} & \cdots & \varphi_{n-1}^{(n-3)} \\ \vdots & \vdots & \cdots & \vdots \\ \varphi_1' & \varphi_2' & \cdots & \varphi_{n-1}' \\ \varphi_1 & \varphi_2 & \cdots & \varphi_{n-1} \end{vmatrix}$$

may be written in the form

$$W_0 = \det(\Phi^{(n-2)}, \Phi^{(n-3)}, \ldots, \Phi', \Phi)$$

to denote the determinant which has as row vectors $\Phi^{(n-2)}$, $\ldots, \Phi'$, Φ, where $\Phi^{(k)} = (\varphi_1^{(k)}, \varphi_2^{(k)}, \ldots, \varphi_{n-1}^{(k)})$. According to the equation $(L - z^n)\varphi_i = 0$ one has

$$\Phi^{(n)} = (z^n - u_0)\Phi - u_1\Phi' - \cdots - u_{n-2}\Phi^{(n-2)}.$$

By using this expression for $\Phi^{(n)}$, the rule for the derivation of determinants, and the fact that a determinant vanishes when two rows are repeated, we get the formula

$$\sum_{i=0}^{k}(-1)^i(u_{n-i}W_0)^{(k-i)} = \det(\Phi^{(n-1)}, \Phi^{(n-2)}, \ldots, \mathring{\Phi}^{(n-1-k)}, \ldots, \Phi', \Phi), \quad 0 \leqslant k \leqslant n-1,$$

where $\mathring{\Phi}^{(k)}$ denotes the absence of the row $\Phi^{(k)}$. To see that this formula is correct take the derivative on both sides of it, to obtain

$$\begin{aligned}\sum_{i=0}^{k}(-1)^i(u_{n-i}W_0)^{(k+1-i)} &= \det(\Phi^{(n)}, \Phi^{(n-2)}, \ldots, \mathring{\Phi}^{(n-1-k)}, \ldots, \Phi', \Phi) \\ &\quad + \det(\Phi^{(n-1)}, \Phi^{(n-2)}, \ldots, \mathring{\Phi}^{(n-k-2)}, \ldots, \Phi', \Phi) \\ &= \det(-u_{n-1-k}\Phi^{(n-1-k)}, \Phi^{(n-2)}, \ldots, \mathring{\Phi}^{(n-1-k)}, \ldots, \Phi', \Phi) \\ &\quad + \det(\Phi^{(n-1)}, \Phi^{(n-2)}, \ldots, \mathring{\Phi}^{(n-k-2)}, \ldots, \Phi', \Phi) \\ &= (-1)^k u_{n-1-k}W_0 \\ &\quad + \det(\Phi^{(n-1)}, \Phi^{(n-2)}, \ldots, \mathring{\Phi}^{(n-k-2)}, \ldots, \Phi', \Phi).\end{aligned}$$

Thus, we have

$$\sum_{i=0}^{n-1}(-1)^i(u_{n-i}W_0)^{(n-1-i)} = \det(\Phi^{(n-1)}, \Phi^{(n-2)}, \ldots, \Phi')$$

and the derivative of this equation yields

$$\begin{aligned}\sum_{i=0}^{n-1}(-1)^i(u_{n-i}W_0)^{(n-i)} &= \det(\Phi^{(n)}, \Phi^{(n-2)}, \ldots, \Phi') \\ &= \det((z^n - u_0)\Phi, \Phi^{(n-2)}, \ldots, \Phi') = (-1)^n(z^n - u_0)W_0,\end{aligned}$$

that is (2.11) up to a factor $(-1)^n$. The proof for $W_1, \ldots, W_{n-1}$ is the same.

The solutions W_k, of Eq. (2.11) also admit the factorization

$$W_k = \Omega_k \exp\sum_{i \neq k}\widehat{\chi}^{[i]}, \quad k = 0,1,\ldots,n-1. \tag{2.12}$$

Here Ω_k is the cofactor of $P_{n-1}(\chi^{[k]})$ in (2.9). Since $u_{n-1} = 0$ the formula of Liouville[7] for the function $W(x)$

$$W(x_2) = W(x_1)\exp\left(-\int_{x_1}^{x_2} u_{n-1}(x')dx'\right)$$

implies here that $W =$ constant and hence we will have

$$(L^* - z^n)(W_k/W) = 0.$$

For the solutions W_k/W of Eq. (2.11) one finds the expression $(W_k/W) = S^{[k]}\exp(-\chi^{[k]})$ with $S^{[k]}$ defined by

$$S^{[k]} = \Omega_k/\Omega, \quad k = 0,1,\ldots,n-1. \tag{2.13}$$

The constant coefficient of $S^{[k]} \in A(u, z^{-1})$ is found to be ϵ_k/nz^{n-1} considering (2.10) and a similar formula for Ω_k,

$$\Omega_k = |\epsilon_k| z^{(n-1)(n-2)/2} + O(z^{((n-1)(n-2)/2)-1}),$$

where $|\epsilon_k|$ is the cofactor of ϵ_k^{n-1} in the Vandermonde determinant $|\epsilon|$.

To see that $|\epsilon_k|/|\epsilon|$ is just equal to ϵ_k/n one uses the relation $|\epsilon_k| = \epsilon_1^k|\epsilon_0|$, $\epsilon_1 = \exp(2\pi i/n)$, in the expansion of $|\epsilon|$ in terms of the $|\epsilon_k|$:

$$|\epsilon| = \sum_{k=0}^{n-1}\epsilon_k^{n-1}|\epsilon_k| = \sum_{k=0}^{n-1}|\epsilon_k|/\epsilon_1^k = n|\epsilon_0|$$

(we choose $\epsilon_k = \epsilon_1^k$). This immediately implies that

$$|\epsilon_k|/|\epsilon| = \epsilon_k|\epsilon_0|/|\epsilon| = \epsilon_k/n$$

$$S^{[k]} = \epsilon_k/nz^{n-1} + O(z^{-n}). \tag{2.14}$$

Now, we go on to the interpretation of the equation

$$(L^* - z^n)(S^{[k]}e^{-\hat{\chi}^{[k]}}) = 0, \tag{2.15}$$

which is another version of (2.11) in terms of $S^{[k]}$. Keeping in mind Eq. (1.4) and the definition (2.4) of $P_k(\chi)$ it is easy to see that

Lemma 2.1: The Gateaux differential of $P_k(\chi)$ is the $(k-1)$–order differential operator given by

$$dP_k(\chi) = \sum_{l=1}^{k}\binom{k}{l}P_{k-l}(\chi)\partial^{l-1}. \tag{2.16}$$

Lemma 2.2: If $\sum_{k=0}^{n}u_kP_k(\chi) = z^n$ the differential operator $e^{-\hat{\chi}}\circ(L - z^n)\circ e^{\hat{\chi}}$ can be expressed in terms of the operators dP_k according to the formula

$$e^{-\hat{\chi}}\circ(L - z^n)\circ e^{\hat{\chi}} = \sum_{k=1}^{n}u_k dP_k(\chi)\partial. \tag{2.17}$$

Proof: By definition

$$dP_k a = \left.\frac{d}{d\epsilon}\right|_{\epsilon=0}P_k(\chi + \epsilon a),$$

then

$$dP_k a = \frac{d}{d\epsilon}\Big|_{\epsilon = 0}(e^{-(\hat{\chi}+\epsilon\hat{a})}\partial^k e^{(\hat{\chi}+\epsilon\hat{a})})$$
$$= -P_k(\chi)\hat{a} + e^{-\hat{\chi}}\partial^k(\hat{a}e^{\hat{\chi}})$$
$$= -P_k(\chi)\hat{a} + \sum_{l=0}^{k}\binom{k}{l}P_{k-l}(\chi)\partial^{l-1}\hat{a}$$
$$= \sum_{l=1}^{k}\binom{k}{l}P_{k-l}(\chi)\partial^{l-1}a$$

according to (1.4) and the Leibniz rule for the derivative of a product. This proves Lemma 2.1. For (2.17)

$$e^{-\hat{\chi}}\circ\partial^k\circ e^{\hat{\chi}} = \sum_{l=0}^{k}\binom{k}{l}e^{-\hat{\chi}}(\partial^{k-l}e^{\hat{\chi}})\partial^l$$
$$= \sum_{l=0}^{k}P_{k-l}(\chi)\partial^l = dP_k\partial + P_k$$

by (2.16). Mulitply both sides on the left by u_k

$$e^{-\hat{\chi}}\circ\sum_{k=0}^{n}u_k\partial^k\circ e^{\hat{\chi}} = \sum_{k=1}^{n}u_k dP_k\partial + \sum_{k=0}^{n}u_k P_k$$

and use the fact that $\Sigma_{k=0}^{n}u_k P_k(\chi) = z^n$ to get (2.17).

Of course, it is also true that

$$e^{\hat{\chi}}\circ(L^* - z^n)\circ e^{-\hat{\chi}} = -\partial\circ\sum_{k=1}^{n}d^*P_k(\chi)\circ u_k, \tag{2.18}$$

where

$$d^*P_k = \sum_{l=1}^{k}(-1)^{l-1}\binom{k}{l}\partial^{l-1}\circ P_{k-l}$$

is the adjoint operator of the operator (2.16). With that, we are in a position to prove that

Proposition 2.3: For each $i = 0,1,\ldots,n-1$, $S^{[i]}$, as given by (2.13), satisfies the differential equation

$$\sum_{k=1}^{n}d^*P_k(\chi^{[i]})(u_k S^{[i]}) = 1. \tag{2.19}$$

Proof: Take (2.18) with the solution $\chi^{[i]}$ of (2.3). Since by Eq. (2.15) $(L^* - z^n)(S^{[i]}\exp(-\hat{\chi}^{[i]})) = 0$, then $\partial\Sigma_{k=1}^{n}d^*P_k(\chi^{[i]})(u_k S^{[i]}) = 0$. Therefore $\Sigma_{n}^{k=1}d^*P_k(\chi^{[i]})(u_k S^{[i]}) =$ constant. that this constant should be equal to one can be seen as follows. In the expression

$$\sum_{k=1}^{n}d^*P_k(\chi^{[i]})(u_k S^{[i]})$$
$$= \sum_{k=1}^{n}\sum_{l=1}^{k}(-1)^{l-1}\binom{k}{l}\partial^{l-1}(P_{k-l}(\chi^{[i]})u_k S^{[i]})$$

the only term from which a constant term can be obtained is $nP_{n-1}(\chi^{[i]})S^{[i]}$ since the $(u_0,\ldots,u_{n-2})$ are assumed to be free generators of the ring A there does not exist differential relations between them. But $P_{n-1}(\chi^{[i]}) = \epsilon_i^{n-1}z^{n-1} + O(z^{n-2})$ and $S^{[i]}$ from (2.14) is $S^{[i]} = \epsilon_i/nz^{n-1} + O(z^{-n})$ so the constant term of $nP_{n-1}(\chi^{[i]})S^{[i]}$ is just equal to one and (2.19) holds.

Equation (2.19) was considered in the Gel'fand-Dikii paper.[3] It was motivated by different reasons as these exposed here, namely, the representation as products of the inexact components of the resolvent of L. Similarly to what they did there one can prove that for each $\chi^{[i]}$, $S^{[i]}$ is in fact the only element of $A(u,z^{-1})$ which satisfies (2.19). Hence, we have constructed in terms of the $\chi^{[i]}$ all the solutions of (2.19) (formula 2.13) belonging to $A(u,z^{-1})$

III. VARIATIONAL DERIVATIVES

In the foregoing section we have obtained two systems of elements of the set $A(u,z^{-1})$, namely the solutions $(\chi^{[0]},\chi^{[1]},\ldots,\chi^{[n-1]})$ of Eq. (2.3) characterized by Proposition 2.1 and those of Eq. (2.19) $(S^{[0]}, S^{[1]},\ldots,S^{[n-1]})$ constructed in terms of the $\chi^{[i]}$ by means of the formula (2.13). We now search for the expressions of their variational derivatives as they were defined in Sec. I.

Theorem 3.1: For each $\chi^{[i]}$, $i = 0,\ldots,n-1$, one has

$$\frac{\delta\chi^{[i]}}{\delta u_k} = -S^{[i]}P_k(\chi^{[i]}). \tag{3.1}$$

Proof: Putting $\chi^{[i]}$ in Eq. (2.3) one has identically $\Sigma_{k=0}^{n}u_k P_k(\chi^{[i]}) = z^n$ we carry out the variation of such relation

$$\sum_{k=0}^{n-2}P_k\delta u_k + \sum_{k=1}^{n}u_k dP_k\delta\chi^{[i]} = 0$$

and multiply both sides on the left by $S^{[i]}$, we obtain after integration by parts in the second summand

$$\sum_{k=0}^{n-2}S^{[i]}P_k\delta u_k + \delta\chi^{[i]}\sum_{k=1}^{n}d^*P_k(\chi^{[i]})(u_k S^{[i]}) = \partial\omega$$

where ω is an appropriate 1-form. But by (2.19) $\Sigma_{k=1}^{n}d^*P_k(\chi^{[i]})(u_k S^{[i]}) = 1$ and in view of the unicity of the solution $S^{[i]}$ and the formula (1.5) our statement is true.

Theorem 3.2: The variational derivatives of the solutions $S^{[i]}$ of Eq. (2.19) are given by

$$\frac{\delta S^{[i]}}{\delta u_k} = -\frac{d}{d(z^n)}(S^{[i]}P_k(\chi^{[i]})). \tag{3.2}$$

Proof: We calculate the derivative of $\Sigma_{k=0}^{n}u_k P_k(\chi^{[i]}) = z^n$ with respect to z^n,

$$\sum_{k=1}^{n}u_k dP_k\frac{d\chi^{[i]}}{d(z^n)} = 1,$$

multiply both sides on the left by $S^{[i]}$ and integrate by parts the left-hand side to obtain, by taking into account Eq. (2.19) for $S^{[i]}$,

$$\frac{d\chi^{[i]}}{d(z^n)} = S^{[i]} + \partial\omega.$$

Since $(\delta/\delta u_k)\partial$ is always equal to zero, Eq. (3.2) follows after application of $\delta/\delta u_k$ to both sides and the introduction of (3.1).

We remark here that the formula (3.2) was obtained by Gel'fand and Dikii[1,3] in the context of the resolvent methods for the operator L. Here it appears as a simple consequence of Theorem 3.1 in the present context of the Riccati equation. Formula (3.2) will guarantee that the Gel'fand–Dikii Hamiltonian systems coincide with those which will be obtained in the next section.

It is interesting to note that all the terms $\chi_n^{[i]}, \chi_{2n}^{[i]},\ldots$ are total derivatives. To see that, consider

$$\sum_{i=0}^{n-1} \frac{\delta \chi^{[i]}}{\delta u_k} = - \sum_{i=0}^{n-1} S^{[i]} P_k(\chi^{[i]}), \quad k = 0,1,\ldots,n-2$$

as follows from (3.1). Remember that $S^{[i]} = \Omega_i/\Omega$, Eq. (2.13), and that the Ω_i are the cofactors of the elements on the first row of Ω. Then

$$\frac{\delta}{\delta u_k} \sum_{i=0}^{n-1} \chi^{[i]} = - \frac{1}{\Omega} \sum_{i=0}^{n-1} P_k(\chi^{[i]})\Omega_i = 0,$$
$$k = 0,1,\ldots,n-2$$

since $\Sigma_{i=0}^{n-1} P_k(\chi^{[i]})\Omega_i$ is just the determinant Ω (2.9) when we substitute its first row by $(P_k(\chi^{[0]}), P_k(\chi^{[1]}),\ldots,P_k(\chi^{[n-1]}))$ which is contained in Ω for $k = 0,1,\ldots,n-2$ and hence is equal to zero. From the corollaries 2.1 and 2.2 $\delta\chi_{rn}^{[i]}/\delta u_k = 0$, $r = 1,2,\ldots$, and this is equivalent (see Sec. I) to our initial assertion.

4. THE INVOLUTIVE SYSTEMS RELATED TO THE RICCATI EQUATION

We start by considering evolution equations for which the functional

$$H^{[i]}[u,z] = \int dx\,(\epsilon_i z - \chi^{[i]}) \tag{4.1}$$

is a constant of motion.

Now let the elements of the ring $A(u)$ depend on a new parameter t. We shall prove that the Lax equations preserve (4.1)

Theorem 4.1: For every linear operator K, for which the equation

$$\dot{L} = [L,K] \tag{4.2}$$

becomes equivalent to an evolution system for the $u_0, u_1,\ldots,u_{n-2}$ the functionals $H^{[i]} \in \tilde{A}(u,z^{-1})$ are constants of motion for this evolution system.

Proof: Apply (4.2) to $e^{\hat{\chi}^{[i]}}$ and multiply on the left by $S^{[i]} e^{-\hat{\chi}^{[i]}}$. From the left-hand side one obtains, after integration, that

$$\int dx S^{[i]} e^{-\hat{\chi}^{[i]}} \dot{L} e^{\hat{\chi}^{[i]}} = \int dx \sum_{k=0}^{n-2} S^{[i]} P_k(\chi^{[i]}) \dot{u}_k$$
$$= \frac{d}{dt} H^{[i]}$$

is the derivative with respect to t of $H^{[i]}$ [see Eq. (1.3)].

Going over the commutator on the right-hand side, we have

$$\int dx\, e^{-\hat{\chi}^{[i]}} S^{[i]} [L,K] e^{\hat{\chi}^{[i]}}$$
$$= \int dx (L^*(S^{[i]} e^{\hat{\chi}^{[i]}}) K e^{\hat{\chi}^{[i]}} - \int dx\, e^{-\hat{\chi}^{[i]}} S^{[i]} K (L e^{\hat{\chi}^{[i]}}) = 0$$

by (2.2) and (2.15). Therefore, we have proved that the operators K which one needs in Theorem 4.1 have been constructed for the first time by Gel'fand and Dikii. They gave[3] a form of such operators which is particularly useful here. The operator

$$\mathscr{P}^{[i]} = (\partial - \chi^{[i]})^{-1} \circ S^{[i]}, \quad i = 0,1,\ldots,n-1 \tag{4.3}$$

can be expanded in powers of z with coefficients which are differential operators of increasing order. For each of these coefficients Theorem 4.1 is satisfied.

To prove that, we shall use the operator (4.3) in a slightly different fashion which will be more convenient here. One can put

$$\mathscr{P}^{[i]} = (e^{\hat{\chi}^{[i]}} \circ \partial \circ e^{-\hat{\chi}^{[i]}})^{-1} \circ S^{[i]} = e^{\hat{\chi}^{[i]}} \circ \partial^{-1} \circ e^{-\hat{\chi}^{[i]}} \circ S^{[i]}$$

to calculate the commutator $[L,\mathscr{P}^{[i]}]$. We observe that by (2.2) and (2.15) $\mathscr{P}^{[i]}$ is an operator of the form $\varphi \circ \partial^{-1} \circ \psi$, where $(L - z^n)\varphi = 0$ and $(L^* - z^n)\psi = 0$, which is all that one needs to have $[L,\mathscr{P}^{[i]}]$ as a differential operator.

Proposition 4.1: Let $\mathscr{P}^{[i]}$ be the operator (4.3) and set $\varphi_i = \exp\hat{\chi}^{[i]}$, $\psi_i = S^{[i]}\exp(-\hat{\chi}^{[i]})$ then one has the identities

$$(L - z^n) \circ \mathscr{P}^{[i]} = \sum_{k=0}^{n-1} \sum_{l=0}^{n-k-1} \sum_{r=0}^{n-k-l-1} \binom{k+r}{r}$$
$$\times u_{k+l+r+1} (\psi_i \varphi_i^{(l)})^{(r)} \partial^k, \tag{4.4a}$$

$$\mathscr{P}^{[i]} \circ (L - z^n) = \sum_{k=0}^{n-1} \sum_{l=0}^{n-k-1} \sum_{r=0}^{n-k-l-1} (-1)^r \binom{l+r}{r}$$
$$\times (u_{k+l+r+1} \psi_i \varphi_i^{(l)})^{(r)} \partial^k. \tag{4.4b}$$

With the help of (4.4a) and (4.4b) and the definition (4.1) of $H^{[i]}$, by substituting φ_i and ψ_i in terms of $\exp\hat{\chi}^{[i]}$ and $S^{[i]}\exp(-\hat{\chi}^{[i]})$, it is easy to see that

$$[L,\mathscr{P}^{[i]}] = [L - z^n, \mathscr{P}^{[i]}] = \sum_{k=0}^{n-2} X_k^{[i]} \partial^k, \tag{4.5}$$

where

$$X_k^{[i]} = \sum_{l=0}^{n-2-k} \mathscr{T}_{kl} \frac{\delta H^{[i]}}{\delta u_l} \tag{4.6}$$

and $\mathscr{T} = (\mathscr{T}_{kl})$ is the symplectic operator of Gel'fand–Dikii,[1,3,8]

$$\mathscr{T}_{kl} = \sum_{r=0}^{n-k-l-1} \left[\binom{k+r}{r} u_{k+l+r+1} \partial^r - \binom{l+r}{r} (-\partial)^r \circ u_{k+l+r+1} \right]. \tag{4.7}$$

We return now to $\mathscr{P}^{[i]}$ and write them in the form

$$\mathscr{P}^{[i]} = e^{-\hat{\sigma}^{[i]}} \circ e^{\epsilon_i z x} \circ \partial^{-1} \circ e^{-\epsilon_i z x} \circ (e^{\hat{\sigma}^{[i]}} S^{[i]})$$
$$= - e^{-\hat{\sigma}^{[i]}} \circ (\epsilon_i z - \partial)^{-1} \circ (e^{\hat{\sigma}^{[i]}} S^{[i]}),$$

where $\sigma^{[i]}$ is the part of $\chi^{[i]}$ which contains only negative powers of z, $\sigma^{[i]} = \chi^{[i]} - \epsilon_i z$. We also note that the expansion of $(\epsilon_i z - \partial)^{-1}$ in power of ∂,

$$(\epsilon_i z - \partial)^{-1} = \sum_{r \geqslant 0} \frac{1}{(\epsilon_i z)^{r+1}} \partial^r,$$

immediately implies that of $\mathscr{P}^{[i]}$ as

$$\mathscr{P}^{[i]} = - \sum_{r \geqslant 0} \sum_{s=0}^{r} \frac{1}{(\epsilon_i z)^{r+1}} \binom{r}{s} P_{r-s}(\sigma^{[i]}) \partial^s \circ S^{[i]}, \tag{4.8}$$

where $P_{r-s}(\sigma^{[i]})$ is the differential polynomial (2.4) with $\sigma^{[i]}$ as argument. Now we note that (4.8) enables us to write

$$\mathscr{P}^{[i]} = \sum_{r \geqslant 0} \frac{1}{z^{n+r}} \mathscr{P}_r^{[i]}, \tag{4.9}$$

in view of the fact that $\sigma^{[i]}$ and $S^{[i]} = \epsilon_i / nz^{n-1} + O(z^{-n})$ contains only negative powers of z. The operator $\mathscr{P}_r^{[i]}$ is a differential operator whose commutator with L is an operator of $(n-2)$ order with coefficients $X_{k,r+n}^{[i]}$ as follows from (4.5) and (4.6) when $H^{[i]}$ is expressed as a power series in z.

We summarize this in the following theorem.

Theorem 4.2: For each opeator $\mathscr{P}_r^{[i]}$, $i = 0,\ldots,n-1, r \geqslant 0$, the Lax equation

$$\dot{L} = [L, \mathscr{P}_r^{[i]}] \tag{4.10}$$

becomes equivalent to the Hamiltonian system

$$\dot{u}_k = \sum_{l=0}^{n-k-2} \mathscr{T}_{kl} \frac{\delta H_{n+r}^{[i]}}{\delta u_k}. \tag{4.11}$$

Define now on $\tilde{A}(u)$ the Poisson bracket

$$\{F_1, F_2\} = \int dx \sum_{k,l} \frac{\delta F_1}{\delta u_k} \left(\mathscr{T}_{kl} \frac{\delta F_2}{\delta u_l} \right), \tag{4.12}$$

which extends in a natural way to $\tilde{A}(u, z^{-1})$. Under (4.11) we can write

$$\frac{dH^{[j]}}{dt} = \{H^{[j]}, H_{n+r}^{[i]}\}, \tag{4.13}$$

but since for (4.11) Theorem 4.1 holds $dH^{[j]}/dt = 0$ and hence

$$\{H_s^{[j]}, H_{n+r}^{[i]}\} = 0.$$

In view of (4.5), (4.6), and (4.9)

$$\sum_{l=0}^{n-k-2} \mathscr{T}_{kl} \frac{\delta H_s^{[i]}}{\delta u_l} = 0$$

for $s \leqslant n$. Thus $(\delta H_s^{[i]} / \delta u_l)_{l=0}^{n-2}$ belongs to the kernel[9] of $\mathscr{T}$ (remember that $H_n^{[i]}$ is always zero because $\chi_n^{[i]}$ is a total derivative). This assures that all the coefficients of $H^{[i]}$ are in involution with respect to (4.12).

$$\{H_s^{[j]}, H_r^{[i]}\} = 0, \quad r,s \geqslant 1.$$

And this was the property claimed for the Riccati equation (2.3)

ACKNOWLEDGMENTS

The author wishes to thank L. Abellanas, F. J. Chinea, A. Galindo, and L. Martínez Alonso for conversations.

[1] I. M. Gel'fand and L. A. Dikii, Funct. Anal. Appl. **10**, 13 (1976).
[2] I. M. Gel'fand and L. A. Dikii, Funct. Anal. Appl. **11**, 11 (1977).
[3] I. M. Gel'fand and L. A. Dikii, Funct. Anal. Appl. **12**, 8 (1978).
[4] Yu. I. Manin, J. Sov. Math. **11**, 1 (1979).
[5] G. Wilson, Math. Proc. Cambridge Philos. Soc. **86**, 131 (1979).
[6] M. M. Vainberg, *Variational Methods for the Study of Nonlinear Operators* (Holden-Day, San Francisco, 1964).
[7] L. S. Pontryagin, *Equations Differentielles Ordinaires* (Mir, Moscow, 1975).
[8] M. Adler, Inventions Math. **50**, 219 (1979).
[9] A. P. Veselov, Funct. Anal. Appl. **13**, 1 (1979).

Factorization of operators and completely integrable Hamiltonian systems [a)]

F. Guil Guerrero
Departmento de Métodos Matemáticos de la Física, Facultad de Ciencias Físicas, Universidad Complutense, Ciudad Universitaria, Madrid-3, Spain

(Received 9 December 1980; revised manuscript accepted for publication 10 August 1981)

Generalized Miura transformations induced by factorization of an nth-order scalar operator are used to characterize a set of Hamiltonian systems by requiring the conservation of the Gel'fand–Dikii first integrals sequence. The second symplectic structure for the Gel'fand–Dikii equations is obtained in connection with the previous Hamiltonian systems. Bäcklund transformations are also analyzed.

PACS numbers: 02.30. + g

1. INTRODUCTION

Since the discovery by Gel'fand and Dikii[1] of the completely integrable Hamiltonian systems related to the resolvent of differential operators, a number of interesting properties exhibited by such systems has been studied. One can cite the elegant form given by Adler, Lebedev, and Manin to the Gel'fand–Dikii theory in terms of the Kirillov symplectic structure.[2]

There was mentioned a series of problems concerning the existence of a second symplectic structure and the Lenard relations. Such relations were constructed by Adler and Symes by means of the fractional powers of the symbols of differential operators.[2,3]

On the other hand, the celebrated Miura transformation[4] between the Korteweg de Vries (KdV) and the modified KdV equations was found to be connected with the problem of the second symplectic structure for the KdV equation; it may be regarded as a canonical transformation between the symplectic structure of the modified KdV equation and the second one for the KdV equation.[5] This guarantees that the second operator is in fact symplectic by constuction. As was noted by Adler and Moser,[6] the Miura transformation is induced by factorization of the Schrödinger operator into two first-order differential operators. The same factorization procedure was put forward by Jaulent and Miodek[7] in the context of energy-dependent Schrödinger operators. Also in this energy-dependent case, one can see that there exists a canonical map between two symplectic structures.[8]

The generalization of these facts to an arbitrary nth-order differential scalar operator is presented in the paper of Sokolov and Shabat,[9] where the construction of the Lax equations for the modified systems is given. Such modified equations are developed even for nth-order differential matrix operators by Kupershmidt and Wilson.[10] The proof of the symplectic character of the second operator in the Gel-'fand–Dikii equations is also given by Kuperschmidt and Wilson. Another proof is made Ref. 11. More applications of the factorization method can be found in Ref. 12.

In this paper we arrive at these results for scalar operators in the following way.

The system of first integrals of Gel'fand–Dikii equations as they were found in the Riccati equation context[13] are used to characterize the modified equations (the analog of the modified KdV equation). These modified equations are constructed as the Lax equations which preserve the system of first integrals mentioned above. The relevant operator to be considered in the Lax representation is determined by the form of the variational derivatives of the first integrals with respect to the new variables. From the Lax equations we obtain completely integrable Hamiltonian systems with a symplectic operator which is a first-order differential operator. The Lax representation of the new Hamiltonian systems is found to be particularly useful to prove that in fact such systems are connected by the transformation with those of Gel'fand–Dikii. If one transforms directly the Hamiltonian form of modified equations, then the Gel'fand–Dikii equations written in terms of the second symplectic operator are obtained. Finally canonical invariance maps for the modified equations as well as the Bäcklund transformations induced in the Gel'fand–Dikii ones, are analyzed. The Bäcklund transformations were found by Kupershmidt in the paper of Ref. 14.

Throughout this paper we shall use the results and notations of the previous work of Ref. 13. For completeness we include two Appendices with some results which are used in Sec. 6.

We would like to thank Boris Kupershmidt for information about his own results in this field when this work was in preparation during the Workshop on Dynamical Systems held in Crete (July 1980).

2. VARIATIONAL DERIVATIVES

We shall introduce a new parametrization of the ring A of differential polynomials[13] with free generators $v_0, v_1,...,v_{n-2}$ which will be related to the standard collection $u_0, u_1,...u_{n-2}$ by the factorization procedure described in Sec. 1. To do that, we define

$$V_\alpha = \sum_{k=0}^{n-2} b_{\alpha k} v_k, \quad \alpha = 0,1,...,n-1 \tag{2.1}$$

where the $n\times(n-1)$ matrix $b = (b_{\alpha k})$, $\alpha = 0,1,..,n-1, k = 0,1,..,n-2$ should satisfy the following two conditions.

$$\sum_{\alpha=0}^{n-1} b_{\alpha k} = 0, \quad k = 0,1,...,n-2 \Rightarrow \sum_{\alpha=0}^{n-1} V_\alpha = 0, \tag{2.2a}$$

$$\operatorname{rank} b = n-1. \tag{2.2b}$$

[a)]Partially supported by the J. E. N. (Madrid).

Let us now define the differential nth-order operator

$$\Lambda = (\partial + V_0)(\partial + V_1)\cdots(\partial + V_{n-1}), \tag{2.3}$$

which can be written in the form $\Lambda = \Sigma_{k=0}^{n} Q_k \partial^k$. The coefficients Q_k are differential polynomials in $v_0,...v_{n-2}$, $Q_n = 1$ and $Q_{n-1} = \Sigma_{\alpha=0}^{n-1} V_\alpha = 0$ according to (2.2a). By setting $u_k = Q_k(v)$ one has the relation

$$\sum_{k=0}^{n} u_k \partial^k = (\partial + V_0)(\partial + V_1)\cdots(\partial + V_{n-1}) \tag{2.4}$$

or $L(u,\partial) = \Lambda(v,\partial)$.

The power series $\chi^{[i]}$ and $S^{[i]}$ introduced in the context of the operator $L = \Sigma u_k \partial^k$ (see Ref. 13) as solutions of

$$\sum_{k=0}^{n} u_k P_k(\chi) = z^n, \quad P_k(\chi) = e^{-\hat{\chi}}\partial^k e^{\hat{\chi}}, \quad \hat{\chi} = \int^x \chi dx, \tag{2.5a}$$

$$\sum_{k=1}^{n} d^* P_k(u_k S) = 1 \tag{2.5b}$$

are found in terms of $v_0,...,v_{n-2}$ by substituting $u_k = Q_k(v)$.

We retain the same notation $\chi, S,...$ for $\chi(u,z), S(u,z),...$ as for $\chi(Q(v),z) S(Q(v),z),...$ In order to calculate the variational derivatives of the solutions $\chi^{[i]}$ with respect to the $v_0,...,v_{n-2}$ we introduce the following conventions.

$$\Lambda_{0,\alpha} = (\partial + V_0)(\partial + V_1)\cdots(\partial + V_\alpha), \quad \Lambda_{0,1} \equiv 1, \quad \alpha = 0,1,...,n-1$$
$$\Lambda_{\alpha,n-1} = (\partial + V_\alpha)(\partial + V_{\alpha+1})\cdots(\partial + V_{n-1}), \quad \Lambda_{n,n-1} \equiv 1, \quad \alpha = 0,1,...,n-1. \tag{2.6}$$

By Λ we understand $\Lambda_{0,n-1}$ and $\Lambda^*, \Lambda^*_{0,\alpha},...$ will denote the adjoint operators.

Theorem 2.1: The variational derivatives of the solution $\chi^{[i]}, i = 0,1,...,n-1$ of (2.5a) are given by the expressions

$$\frac{\delta \chi^{[i]}}{\delta v_k} = -\sum_{\alpha=0}^{n-1} b_{\alpha k} (\Lambda^*_{0,\alpha-1} S^{[i]} e^{-\hat{\chi}^{[i]}})(\Lambda_{\alpha+1,n-1} e^{-\hat{\chi}^{[i]}}). \tag{2.7}$$

Note the polynomial character of $\delta\chi^{[i]}/\delta v_k$ since the exponentials cancel.

Proof: Rewrite (2.5a) to keep explicitly the operator Λ [(2.3)],

$$e^{-\hat{\chi}^{[i]}}(\Lambda e^{\hat{\chi}^{[i]}}) = z^n.$$

We perform the variation of this identity to get

$$e^{-\hat{\chi}^{[i]}}(\delta\Lambda e^{\hat{\chi}^{[i]}}) + e^{\hat{\chi}^{[i]}}\Lambda(e^{\hat{\chi}^{[i]}}\delta\hat{\chi}^{[i]}) - z^n\delta\hat{\chi}^{[i]} = 0.$$

If we take into account the relation[13]

$$e^{-\hat{\chi}^{[i]}}\Lambda e^{\hat{\chi}^{[i]}} = z^n + \sum_{k=1}^{n} u_k dP_k(\chi^{[i]})\partial$$

we arrive at

$$e^{-\hat{\chi}^{[i]}}(\delta\Lambda e^{\hat{\chi}^{[i]}}) + \sum_{k=1}^{n} u_k dP_k \delta\chi^{[i]} = 0.$$

Multiply both sides on the left by $S^{[i]}$, apply (2.5b) after integration by parts to deduce, with an appropriated 1-form ω

$$\delta\chi^{[i]} + \sum_{k=0}^{n-2}\sum_{\alpha=0}^{n-1} b_{\alpha k}(\Lambda^*_{0,\alpha-1} S^{[i]} e^{-\hat{\chi}^{[i]}})(\Lambda_{\alpha+1,n-1} e^{\hat{\chi}^{[i]}})\delta v_k = \partial\omega$$

as it follows from the formula

$$\delta\Lambda = \sum_{\alpha=0}^{n-1}\sum_{k=0}^{n-2} b_{\alpha k}\Lambda_{0,\alpha-1}\delta v_k \Lambda_{\alpha+1,n-1},$$

easily provable by induction over n. This concludes the proof of (2.7).

From Now on we shall drop the index on $\chi^{[i]}, S^{[i]},...$ and write simply $\chi, S,...$.

We set

$$\xi_\alpha = z^\alpha \Lambda_{\alpha+1,n-1} e^{\hat{\chi}}, \tag{2.8a}$$

$$\zeta_\alpha = z^{-\alpha}\Lambda^*_{0,\alpha-1}(Se^{-\hat{\chi}}) \tag{2.8b}$$

and define the column vectors $\xi = (\xi_\alpha), \zeta = (\zeta_\alpha)$ which enable us to write (2.7) in the form

$$\frac{\delta\chi}{\delta v_k} = -\sum_{\alpha=0}^{n-1} b_{\alpha k}\xi_\alpha\zeta_\alpha. \tag{2.9}$$

Let us now define the matrix

$$a = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ & & & & \\ 0 & 0 & 0 & \cdots & 1 \\ 1 & 0 & 0 & \cdots & 0 \end{pmatrix}. \tag{2.10}$$

The columns of a are the vectors of the standard basis of $\mathbb{R}^n(e_{n-1}, e_0, e_1,...,e_{n-2})$. Denote by $\tilde{a}$ the transposed of a, then

$$a\tilde{a} = \tilde{a}a = \mathbb{1}, \quad \det a = (-1)^{n-1}, \quad a^n = \mathbb{1} \tag{2.11}$$

and a is a unitary matrix. Let V be the column vector $V = (V_\alpha)$ with components V_α given by (2.1); we set $D(V) = \text{diag}(v_0,...,v_{n-1})$ and similarly for $D(\xi), D(\zeta),...$.

Proposition 2.1: The vectors ξ and ζ as defined by (2.8) satisfy the linear differential equations

$$(T - z)\xi = 0, \tag{2.12a}$$

$$(\tilde{a}T^+ a - z)\zeta = 0. \tag{2.12b}$$

Here T is the first-order differential operator

$$T = a(\partial + D(V)) \tag{2.13}$$

and $T^+ = (-\partial + D(V))\tilde{a}$.

Proof: From the definition (2.6) for $\Lambda_{0,\alpha}$ one sees that $(\partial + V_\alpha)\xi_\alpha = z^\alpha \Lambda_{\alpha,n-1} e^{\hat{\chi}} = z\xi_{\alpha-1}$ according to (2.8a), $\alpha = 1,...,n-1$. Moreover one has $(\partial + V_0)\xi_0 = \Lambda e^{\hat{\chi}} = z^n e^{\hat{\chi}}$ or $(\partial + V_0)\xi_0 = z\xi_{n-1}$. We have proved therefore that $(\partial + D(V))\xi = z\tilde{a}\xi$, which immediately implies (2.12a). In the same way one obtains (2.12b).

We shall cite here some properties about the linear application D of $\mathbb{R}^n$ in the set of $n\times n$ diagonal matrices. For every vector $\theta \in \mathbb{R}^n$ it is easily seen that

$$D(a\theta) = aD(\theta)\tilde{a}, \qquad D(\tilde{a}\theta) = \tilde{a}D(\theta)a \tag{2.14}$$

and for diagonal matrices $D(\theta)D(\rho) = D(D(\theta)\rho)$.

III. LAX PAIRS

Suppose that there exists some linear operator J for which the Lax equation

$$\frac{d}{dt} T = [T,J] \tag{3.1}$$

makes sense as a system of differential equations for the functions $v_0,...,v_{n-2}$ which are left to depend on a new parameter t. The precise set of conditions on the J-operators will be stipulated later. Now, we claim that concerning the functionals $\mathscr{H}(v,z) = \int dx(\epsilon z - \chi)$ [in short for $\mathscr{H}^{(i)}(v,z) = \int dx(\epsilon_i z - \chi^{(i)})$] one has

Theorem 3.1: The functional $\mathscr{H}$ is a constant of motion for the evolution equations which issue from the Lax equation (3.1), with an admissible operator J.

Proof: In like manner as we did in the Theorem 4.1 of the previous work[13], we multiply on the left of (3.1) by $(\widetilde{a\zeta})$ and apply this operator to the vector ξ, where ξ and ζ are the solutions (2.8) of Eqs. (2.12). After integration we have

$$\int dx\, \zeta D(\dot{V})\xi = \int dx(\widetilde{a\zeta})[T,J]\xi, \quad \dot{V} = \frac{dV}{dt}.$$

From the left one obtains

$$\int dx\, \tilde{\zeta} D(\dot{V})\xi = \int dx \sum_{k=0}^{n-2} \dot{v}_k \sum_{\alpha=0}^{n-1} b_{\alpha k}\xi_\alpha \zeta_\alpha = \frac{d\mathscr{H}}{dt}$$

by (2.9). Now we examine the right-hand side,

$$\int dx(\widetilde{a\zeta})[T,J]\xi = \int dx((T^+ a\zeta)^\sim J\xi - (\widetilde{a\zeta})JT\xi),$$

which becomes identically equal to zero by virtue of (2.12a) and (2.12b). Therefore we have proved that $(d\mathscr{H}/dt) = 0$ as it was claimed.

Theorem 3.1 tells us that the appropriate Lax equation which one does select in order to have evolution equations with an infinite sequence of first integrals, should be attached to a certain differential operator T which in this case does not coincide with the starting operator Λ.

The precise set of conditions on J to make Eq. (3.1) meaningful are

(i) $T,J]$ should be a matrix multiplicative operator,

(ii) $\tilde{a}[T,J]$ remains to be a diagonal matrix with null trace.

These conditions are obtained by examining the form of $dT/dt = aD(\dot{V})$ and the assumption $\sum_{\alpha=0}^{n-1} V_\alpha = 0$ for the V_0, $V_1,...,V_{n-1}$.

In order to find operators J with the required properties, we translate Eqs. (2.12a) and (2.12b) to equations for the diagonal matrices $D(\xi)$ and $D(a\zeta)$.

Proposition 3.1: Let the vectors ξ and ζ satisfy (2.12a) and (2.12b), then

$$TD(\xi) = zD(\xi)a, \tag{3.2a}$$

$$T^+D(a\zeta) = zD(a\zeta)\tilde{a} \tag{3.2b}$$

for the diagonal matrices $D(\xi)$ and $D(a\zeta)$.

Proof: By (2.12a) $D(T\xi) = zD(\xi)$. But (2.14) implies that

$$D(T\xi) = D(a(\partial + D(v))\xi) = a(\partial + D(v))D(\xi)\tilde{a}$$

just equal to $TD(\xi)\tilde{a}$. From (2.11) (3.2a) follows. In the same way the Eq. (3.2b) is proved.

Let us now construct the operator

$$\mathscr{M} = z(D(\xi)\partial^{-1}D(a\zeta) \tag{3.3}$$

to be understood as a power series of differential operators when we write them in the form

$$\mathscr{M} = zD(e^{-\hat{x}}\xi)(\partial - \chi)^{-1}D(ae^{\hat{x}}\zeta)$$

in which the exponentials' factor do not appear and $(\partial - \chi)^{-1}$ expanded in powers of z (see Ref. 13) to write

$$\mathscr{M} = \sum_{r\geqslant 0} \frac{1}{z^r}\mathscr{M}_r, \tag{3.4}$$

the $\mathscr{M}_r$ being differential operators of increasing order.

Proposition 3.2:

$$[T,\mathscr{M}] = z[a,D(\xi)D(a\zeta)]. \tag{3.5}$$

Proof: One has the identity

$$[T,\mathscr{M}] = z\{(TD(\xi))\partial^{-1}D(a\zeta) - D(\xi)\partial^{-1}(T^+D(a\zeta))^\sim + [a,D(\xi)D(a\zeta)]\}$$

in which we take into account Eqs. (3.2) to see that they reduce to (3.5).

We also note that

$$\tilde{a}[T,\mathscr{M}] = zD(\xi)D(a\zeta) - zD(\tilde{a}\xi)D(\zeta)$$

is manifestly a diagonal matrix with null trace.

4. THE MODIFIED HAMILTONIAN SYSTEMS

Our next goal will be to express the commutator (3.5) in terms of the variational derivatives of the functional

$$\mathscr{H}(v,z) = \int dx(\epsilon z - \chi) = \sum_{r\geqslant 1} \frac{\mathscr{H}_r[v]}{z^r}, \tag{4.1}$$

that by Theorem 3.1 is a constant of motion for the Lax equations defined by the operators $\mathscr{M}_r$ in the power series of $\mathscr{M}$. By doing so, we shall obtain the Hamiltonian systems, the symplectic operator, and an involutive system of functionals with respect to the Poisson bracket given by the new symplectic operator.

We shall need the following two identities,

$$z[a,D(\xi)D(a\zeta)] = -\partial(D(a\zeta)aD(\xi)), \tag{4.2}$$

$$\tilde{a}D(a\zeta)aD(\xi) = D(\zeta)D(\xi). \tag{4.3}$$

The relation (4.3) is an obvious consequence of (2.14), since $\tilde{a}D(a\zeta)a = D(\tilde{a}a\zeta) = D(\zeta)$. Formula (4.2) is found by multiplication of (3.2a) on the left by $D(a\zeta)$ and the transpose of (3.2b) on the right by $D(\xi)$ [take care of the commutativity of the diagonal matrices $D(\xi)$ and $D(a\zeta)$], then

$$z(aD(\xi)D(a\zeta) - D(\xi)D(a\zeta)a) = D(a\zeta)\tilde{T}^+D(\xi) - D(a\zeta)TD(\xi)$$

yields (4.2).

Proposition 4.1: Let $\mathscr{M}$ be the operator defined by (3.3), then

$$\tilde{a}[T,\mathscr{M}] = -\partial D(\xi)D(\zeta). \tag{4.4}$$

This relation comes directly from (3.5), (4.2), and (4.3).

It proves to be convenient to introduce the vector $\pi = (\pi_\alpha)$ defined by the product $D(\xi)D(\zeta)$ which we write as the diagonal matrix $D(\pi)$ with $\pi_\alpha = \xi_\alpha \zeta_\alpha, \alpha = 0,1,...,n-1$.

Moreover, let us define the vector $\kappa = (\kappa_\alpha)$ with components

$$\kappa = -\frac{\delta\chi}{\delta v_k}, \quad k = 0,1,...,n-2$$

$$\kappa_{n-1} = 0.$$

Keeping in mind Eq. (2.9) one finds $\kappa_k = \sum_{\alpha=0}^{n-1} b_{\alpha k}\pi_\alpha$ or $\partial\kappa_k = \sum_{k=0}^{n-1} b_{\alpha k}\partial\pi_\alpha$, $k = 0,...,n-2$. In (4.4), the traceless character of $\tilde{a}[T,\mathscr{M}]$ implies that $\partial\sum_{\alpha=0}^{n-1}\pi_\alpha = 0$ [as one can check directly in Eqs. (2.12)]. We summarize the above relations in the formula

$$\partial\kappa = c\partial\pi, \tag{4.5}$$

where the matrix c is

$$c = \begin{pmatrix} b_{00} & b_{10} & \cdots & b_{n-1,0} \\ b_{01} & b_{11} & \cdots & b_{n-1,1} \\ b_{0,n-2} & b_{1,n-2} & \cdots & b_{n-1,n-2} \\ 1 & 1 & \cdots & 1 \end{pmatrix}, \tag{4.6}$$

the transpose of b (2.1), bordered by the row (1, 1,...,1).

By conditions (2.2) concerning b one sees that c is invertible [rank $b = n-1$ implies that at least one of the cofactors of one of the elements, say the $c_{n-1,i}$ in the last row of c, is non-null. To calculate the determinant of c we add up to the column i the restant ones and expand det c by the elements of that column i. By (2.2a) $\Sigma_{\alpha=0}^{n-1} b_{\alpha k} = 0, k = 0,...,n-2$, thus det c is proportional to the non-null cofactor of $c_{n-1,i}$].

From (4.5) $\partial\pi = c^{-1}\partial\kappa$, and this enables us to express Eq. (4.4) in terms of $\delta\chi/\delta v_k, k = 0,1,...,n-2$.

To go over the motion equations, we pick out a term $\mathscr{M}_r$ in the series (3.4) to examine Eq. (3.1) with $J = \mathscr{M}_r$, which may be written as

$$D(\dot{V}) = \tilde{a}[T,\mathscr{M}_r]. \tag{4.7}$$

From Proposition 4.1 and Eq. (4.5) it easily follows that

$$D(\dot{V}) = D(-c^{-1}\partial\kappa_r)$$

turns out to be equivalent to (4.7), or

$$\dot{V} = -c^{-1}\partial\kappa_r,$$

to be expressed in terms of the independent variables $v_0,...,v_{n-2}$ as follows. Keeping in mind Eq. (2.1) $V = bv, v = (v_k)_{k=0}^{n-2}$, since b is a constant matrix

$$b\dot{v} = -c^{-1}\partial\kappa_r.$$

Multiply on the left by c to obtain $cb\dot{v} = -\partial\kappa_r$. Take care of the fact that by (4.6) and the property (2.2)

$$cb = \begin{pmatrix} \tilde{b}b \\ \hline 00\cdots0 \end{pmatrix}$$

is the matrix $\tilde{b}b$ boardered by the row (0,...,0). Remember that by definition $\kappa_{n-1} = 0$ to see that one has simply

$$\tilde{b}b\dot{v} = -\partial\nabla\mathscr{H}_r,$$

where we have introduced the vector $\nabla\mathscr{H}_r = (-\delta\chi_r/\delta v_k)_{k=0}^{n-2}$ according to the expression (4.1) of the functional $\mathscr{H}$, related to the H introduced in Ref. 13 by

$$\mathscr{H} = H\circ Q. \tag{4.8}$$

Here Q is the transformation defined by (2.4).

We must prove again that a determinant does not vanish; here is now det $\tilde{b}b$.

This comes in this case by considering the product $c\tilde{c}$, which by (4.6) and (2.2) equals

$$c\tilde{c} = \begin{pmatrix} & & & 0 \\ & \tilde{b} & b & \vdots \\ & & & 0 \\ 0 & \cdots & 0 & n \end{pmatrix};$$

together the above proved property det $c \neq 0$.

These considerations allow us to write finally the motion equations (4.7) in terms of $v_0,...,v_{n-2}$ in the form

$$\dot{v} = \mathscr{S}\partial\nabla\mathscr{H}_r, \quad \mathscr{S} = -(\tilde{b}b)^{-1} \tag{4.9}$$

in which we want them.

The matrix $\mathscr{S}$ which appears in (4.9) is manifestly a constant, nonsingular, symmetrical matrix. We define the operator

$$\mathscr{k} = \mathscr{S}\partial \tag{4.10}$$

that turns out to be a symplectic operator. (Here the Jacobi identity follows directly from the self-adjoint character of the Gateaux differential for a gradient.[14] So we have

Theorem 4.1: The Hamiltonian systems (4.9)

$$\dot{v} = \mathscr{k}\nabla\mathscr{H}_r, \quad r = 1,2,... \tag{4.11}$$

admit a Lax representation

$$\dot{T} = [T,\mathscr{M}_r]$$

in terms of the first-order differential operator T [(2.13)] and the pairing operators $\mathscr{M}_r$ [(3.4)]. This result is contained in Ref. 10.

Let us now introduce the Poisson bracket

$$(F_1,F_2) = \int dx \sum_{k,l=0}^{n-2} \mathscr{S}_{kl}\frac{\delta F_1}{\delta v_k}\,\delta\,\frac{\delta F_2}{\delta v_l} \tag{4.12}$$

associated with the symplectic operator $\mathscr{k}$. The Hamiltonian systems (4.11) may be also written as

$$\dot{v} = (v,\mathscr{H}_r), \tag{4.13}$$

and morever one has

Corollary: The functionals $\mathscr{H}_r[v]$ in the series (4.1) are in involution with respect to the Poisson bracket (4.12). In consequence the Hamiltonian systems (4.11) are completely integrable.

Proof: By Theorems 3.1 and 4.1 $\mathscr{H}_s$ is a constant of motion for Eq. (4.11), but $d\mathscr{H}_s/dt = (\mathscr{H}_s,\mathscr{H}_r)$ and this implies that

$$(\mathscr{H}_s,\mathscr{H}_r) = 0, \quad r,s = 1,2,... . \tag{4.14}$$

We shall call Eqs. (4.11) the modified Hamiltonian systems. Note the simple form of the symplectic operator $\mathscr{k}$[(4.10)] and that (4.11) depends on some free parameters through the matrix b [(2.2)].

5. MODIFIED EQUATIONS AND GEL'FAND–DIKII SYSTEMS

Firstly we take into consideration the Lax representation of (4.11). We further introduce the following two matrix differential operators:

$$D_+ = \mathrm{diag}(1, \Lambda_{00}, \Lambda_{01}, \dots, \Lambda_{\cdot, n-2}), \tag{5.1a}$$

$$D_- = \mathrm{diag}(\Lambda_{1,n-1}, \Lambda_{2,n-1}, \dots, \Lambda_{n-1,n-1}, 1), \tag{5.1b}$$

where the $\Lambda_{0,\alpha}$ and $\Lambda_{\alpha,n-1}$ are defined in (2.6).

Multiply the Lax equation $\dot T = [T, \mathscr{M}_r]$ on the left by $D_+\tilde a$ and on the right by D_- to form the expression

$$D_+\tilde a \dot T D_- = D_+\tilde a [T, \mathscr{M}_r] D_-.$$

Both sides in this equation are diagonal matrices. The trace of the matrix in the left turns out to be

$$S_p D_+ \tilde a \dot T D_- = S_p D_+ D(\dot V) D_- = \sum_{\alpha=0}^{n-1} \Lambda_{0,\alpha-1} \dot V_\alpha \Lambda_{\alpha+1,n-1} = \frac{d}{dt}\Lambda,$$

where according to (2.6) $\Lambda_{0,\ -1} = \Lambda_{n,n-1} = 1$ and Λ is the operator [(2.3)].

To calculate the right-hand side, we examine the expression

$$D_+\tilde a [T, \mathscr{M}] D_-$$

with the complete operator $\mathscr{M}$ [(3.3)] which we write in the form

$$\mathscr{M} = \mathrm{diag}(\mathscr{M}_0, \mathscr{M}_1, \dots, \mathscr{M}_{n-1}).$$

The definition (2.13) of T gives us the formula

$$S_p D_+ \tilde a [T, \mathscr{M}] D_- = \left(\sum_{\alpha=0}^{n-2} \Lambda_{0,\alpha} \mathscr{M}_\alpha \Lambda_{\alpha+1,n-1} + \Lambda \mathscr{M}_{n-1}\right) - \left(\sum_{\alpha=0}^{n-2} \Lambda_{0,\alpha} \mathscr{M}_\alpha \Lambda_{\alpha+1,n-1} + \mathscr{M}_{n-1}\Lambda\right) = \Lambda \mathscr{M}_{n-1} - \mathscr{M}_{n-1}\Lambda.$$

By (3.3) and (2.8) $\mathscr{M}_{n-1}$ is found to be

$$\mathscr{M}_{n-1} = z\xi_{n-1}\partial^{-1}\zeta_0 = z^n e^{\hat x}\partial^{-1} S e^{-\hat x} = z^n \mathscr{P},$$

where $\mathscr{P}$ is the operator introduced in connection with L to pair with it in the Lax equations problem.[13] Thus we have

$$S_p D_+ \tilde a [T, \mathscr{M}] D_- = z^n [\Lambda, \mathscr{P}].$$

In this way we arrive at the following result.[9]

Theorem 5.1. The Lax equation

$$\dot T = [T, \mathscr{M}_r], \quad r = 1, 2, \dots$$

implies that Λ evolves according to the equation

$$\dot \Lambda = [\Lambda, \mathscr{P}_r].$$

6. THE SECOND HAMILTONIAN STRUCTURE

In this section we examine the operator $\mathscr{G}$ defined by (A5) (see Appendix A). The property of $\mathscr{G}$ to be symplectic was conjectured by Adler[2] and finally proved in the works of Refs. 10, 11, and 14. We shall prove that by the same method followed by Kupershmidt and Wilson.

To deal with the functions $u_k = Q_k(v)$ [(2.4)] in the context of the Hamiltonian systems associated to the Lax equations of $L = \Sigma_k u_k \partial^k$ the property to be free generators of the ring $A(u)$ should be proved. This result follows from some preliminary considerations.

Lemma 6.1: Let $U_\alpha = N_\alpha(V_1, V_2, \dots, V_n)$ $\alpha = 1, 2, \dots, n, n \geqslant 2$, where the polynomials N_α define the set of variables $U_1, \dots, U_n$ in terms of the differentially independent (∂-indep.) ones $V_1, V_2, \dots, V_n$. If the set U is not ∂-indep. there exists a non-null polynomial F such that $F \circ N$ vanishes identically and $\nabla F \circ N$ is a non-null vector.

Proof: Let $F \in A(y_1, y_2, \dots, y_n)$ with $y_1, y_2, \dots, y_n, n \geqslant 2$ ∂-indep., thus $\nabla F = (\delta F/\delta y_\alpha)^n_{\alpha-1}$. From the definition of differential dependence,[10] a non-null polynomial F exists such that $F \circ N \equiv 0$. We consider here polynomials F which are not total derivatives, equivalently $\nabla F \neq 0$. If $F = \partial G$ we take G instead of F. Then $\nabla F \circ N \neq 0$ or $\nabla F \circ N = 0$. In the first case the lemma is proved. If $\nabla F \circ N = 0$ we take as initial polynomial F one of the polynomials $\delta F/\delta y_\alpha$; if F was of degree M as polynomial in $y_1, \dots, y_n$, $\delta F/\delta y_\alpha$ is of degree $M - 1$. By continuing this descending process we arrive in the limit case to a polynomial F which has degree 1 for which $\nabla F = \nabla F \circ N = \text{const.} \neq 0$.

Example: Consider

$$U_1 = V_1 + V_2,$$

$$U_2 = V_1' + V_2',$$

and $F(U_1, U_2) = \frac{1}{2}(U_2 - U_1')^2$. Then $F \circ N$ and $\nabla F \circ N$ vanish identically,

$$\nabla F = \begin{pmatrix} U_2' - U_1'' \\ U_2 - U_1' \end{pmatrix}.$$

If we consider the first component $U_2' - U_1'' = \partial(U_2 - U_1')$ we retain $U_2 - U_1'$, for which the lemma works.

Lemma 6.2: Let $U_\alpha = N_\alpha(V_1, \dots, V_n)$ define the variables $U_1, \dots, U_n$ by means of the polynomials N_α in terms of the differentially independent ones $V_1, \dots, V_n$. Then, if the kernel of $d_V^+ N$ on $A(V_1, \dots, V_n)$ (the ring of differential polynomials in $V_1, \dots, V_n$) contains no other vector than zero the variables $U_1, \dots, U_n$ are ∂-indep. ($d_V^+ N$ denotes the adjoint of the Gateaux differential $d_V N$[15]).

Proof: If the $U_1, \dots, U_n$ were differentially dependent variables, by Lemma 6.1 we will have a polynomial F such that $F \circ N = 0$ and $\nabla F \circ N \neq 0$. From the identity $(F \circ N)(V_1, V_2, \dots, V_n) = 0$ we deduce

$$\nabla_V (F \circ N) = d_V^+ N(\nabla F \circ N) = 0,$$

a contradiction.

We apply this result to the transformation induced by

$$\partial^n + U_1\partial^{n-1} + U_2\partial^{n-2} + \cdots + U_n = (\partial + V_n)(\partial + V_{n-1}) \cdots (\partial + V_1). \tag{6.1}$$

Proposition 6.1: Let $U_\alpha = N_\alpha(V_1, \dots, V_n), \alpha = 1, 2, \dots, n, n \geqslant 2$ be the transformation defined by (6.1), where the $V_1, \dots V_n$ are assumed to be ∂-indep. variables. Then the set $U_1, \dots, U_n$ is also ∂-indep.

Proof: By induction on n.

For $n = 2$, from $\partial^2 + U_1\partial + U_2 = (\partial + V_2)(\partial + V_1)$ we deduce

$$d_V N = \begin{pmatrix} 1 & 1 \\ \partial + V_2 & V_1 \end{pmatrix}, \quad d_V^+ N = \begin{pmatrix} 1 & -\partial + V_2 \\ 1 & V_1 \end{pmatrix}.$$

From

$$d_V^+ Na = \begin{pmatrix} a_1 + (-\partial + V_2)a_2 \\ a_1 + V_1 a_2 \end{pmatrix} = 0,$$

$$a = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \in A(V_1, V_2)$$

we get $a_2' + (V_1 - V_2)a_2 = 0$, which implies $a_2 = 0$ and $a_1 = 0$ (if a_2 contains derivatives of maximal order m_1 and m_2 in V_1 and V_2, respectively, the coefficients of $V_1^{(m_1+1)}$ and $V_2^{(m_2+1)}$ which are $\partial a_2/\partial V_1^{(m_1)}$ and $\partial a_2/\partial V_2^{(m_2)}$ in the equation for a_2 should be equal to zero, a contradiction).

Now, we write (6.1) in the form

$$\partial^n + U_1\partial^{n-1} + \cdots + U_n = (\partial + V_n)(\partial^{n-1} + W_1\partial^{n-2} + \cdots + W_{n-1})$$

to examine the transformation of the ∂-indep. set $(W_1,\ldots,W_{n-1},W_n) W_n \equiv V_n$ into the variables $U_1,\ldots,U_n$. Then

$$d_W^+ N = \begin{pmatrix} 1 & -\partial + W_n & 0\cdots 0 & 0 \\ 0 & 1 & -\partial + W_n \cdots 0 & 0 \\ \cdot & \cdot & \cdots\cdots & \cdot \\ 0 & 0 & 0\cdots 1 & -\partial + W_n \\ 1 & W_1 & W_2 \cdots W_{n-2} & W_{n-1} \end{pmatrix}$$

and from $d_W^+ Na = 0$ we deduce $a = (a_\alpha)_1^N = 0$ by the same argument used in the case $n = 2$ which we apply here to the equation

$$[(\partial - W_n)^{n-1} + W_1(\partial - W_n)^{n-2} + \cdots + W_{n-1}]a_n = 0;$$

then

$$a_\alpha = (\partial - W_n)^{n-\alpha} a_n = 0.$$

Another proof of this result can be found in Ref. 10. The differential independence of the set $(u_0, u_1,\ldots,u_{n-2})$ defined in terms of $(v_0, v_1,\ldots,v_{n-2})$ by (2.4) follows from

Proposition 6.2: The variables $(u_0, u_1,\ldots,u_{n-2}), n\geqslant 3$ defined by (2.4) are ∂-independent.

Proof: Write the formula (2.4) in the form

$$\partial^n + u_{n-2}\partial^{n-2} + \cdots u_1\partial + u_0 = (\partial + V_0)(\partial^{n-1} + W_{n-2}\partial^{n-2} + \cdots + W_1\partial + W_0) \tag{6.2}$$

to see that

(a) $V_0 = -\sum_{i=0}^{n-2} W_i$, to have $u_{n-1} \equiv 0$,

(b) The set $(W_0, W_1,\ldots,W_{n-2})$ is ∂-indep. The ∂-independence of this set follows from Proposition 6.1 applied to the transformation induced by

$$\partial^{n-1} + W_{n-2}\partial^{n-2} + \cdots + W_1\partial + W_0 = (\partial + V_1)(\partial + V_2)\cdots(\partial + V_{n-1}).$$

We observe that $(V_1, V_2,\ldots,V_{n-1})$ are ∂-indep. since they are obtained from the ∂-indep. set $(V_0, V_1,\ldots,V_{n-2})$ through the transformation

$$V_i = \sum_{k=0}^{n-2} b_{ik} v_k, \quad i = 1, 2,\ldots,n-1$$

and the matrix b_{ik} obtained from the $b_{\alpha k}$ [(2.1)] is nonsingular [to see that, use the matrix c [(4.6)] which was nonsingular, and property (2.2a) of b to get the formula $0 \neq \det c = (-1)^{n-1} n \det b_{ik}$]. That the ∂-independence is preserved by the composition of transformations is easily seen from Lemma 6.2.

Thus we can apply Lemma 6.2 to (6.2) written in the form

$$\partial^n + u_{n-2}\partial^{n-2} + \cdots + u_1\partial + u_0 = \Big(\partial - \sum_0^{n-2} W_i\Big)(\partial^{n-1} + W_{n-2}\partial^{n-2} + \cdots + W_1\partial + W_0)$$

to confirm the ∂-indep.of $(u_0, u_1,\ldots,u_{n-2})$.

Now, we are in a position to prove that $\mathscr{G}$ is symplectic.[10,11,14]

Theorem 6.1: Let $\mathscr{G}$ be the operator induced on $A(u_0, u_1,\ldots,u_{n-2})$ by (AI.5). $\mathscr{G}$ is symplectic.

Proof: Theorem 5.1 tells us that Eq. (4.11),

$$\dot{v} = \mathscr{K}\nabla_v \mathscr{H}_r, \tag{6.3}$$

goes into the equation (see Theorem 4.2 in Ref. 13)

$$\dot{Q}(v) = \mathscr{T}(\nabla_u H_{r+n} \circ Q) \tag{6.4}$$

if $u = Q(v)$[(2.4)]. On the other hand, from (4.8) $\mathscr{H}_r = H_r \circ Q$,

$$\nabla_v \mathscr{H}_r = d_v^+ Q(\nabla_u H_r \circ Q),$$

and

$$\dot{Q}(v) = d_v Q \dot{v} = d_v Q \mathscr{K} d_v^+ Q(\nabla_u H_r \circ Q) \tag{6.5}$$

according to (6.3). By subtracting (6.4) from (6.5) we obtain

$$d_v Q \mathscr{K} d_v^+ Q(\nabla_u H_r \circ Q) - \mathscr{T}(\nabla_u H_{r+n} \circ Q) = 0, \quad r = 1, 2,\ldots.$$

But we have from (A 5)

$$\mathscr{G}\nabla_u H_r - \mathscr{T}\nabla_u H_{r+n} = 0$$

and from these last two equations

$$(\mathscr{G} - d_v Q \mathscr{K} d_v^+ Q)(\nabla_u H_r \circ Q) = 0.$$

To see that $\mathscr{G}$ coincides with $d_v Q \mathscr{K} d_v^+ Q$ we use the following lemma due to Kuperschmidt and Wilson,[10] we shall prove it below:

Lemma 6.3: Let $E = \Sigma_{g=0}^M E_g(v)\partial^g$ be an operator which cancels all the vectors $\nabla_u H_r \circ Q, r = 1, 2,\ldots$. Then E is the null operator. In this way the operation $\{F_1, F_2\}_{\mathscr{G}}$ defined by

$$\{F_1, F_2\}_{\mathscr{G}} = \int dx \sum_{k,e} \frac{\delta F_1}{\delta u_k} \mathscr{G}_{kl} \frac{\delta F_2}{\delta u_l}$$

is a Poisson bracket, due to the relation[16,17]

$$\{F_1, F_2\}_{\mathscr{G}} \circ Q = (F_1 \circ Q, F_2 \circ Q),$$

and $\mathscr{G}$ is symplectic. We observe that the functionals H_r are in involution also with respect to this Poisson bracket,

$$\{H_r, H_s\} \circ Q = (H_r \circ Q, H_s \circ Q) = (\mathscr{H}_r, \mathscr{H}_s) = 0$$

[see Eqs. (4.8) and (4.14)].

Proof of Lemma 6.3: In view of the homogeneity properties[1] of the functionals $H_r, r = 1, 2,\ldots$, the highest-order derivatives contained in H_r appear in the linear part of H_r. By using (B1)

$$\frac{\delta H_{r+n}}{\delta u_k} = \sum_{l=0}^{n-2} \alpha_{kl} u_l^{(r+k+l+1-n)} + \cdots,$$

with α_{kl} nonsingular for r prime to n. We need the linear part of $(\delta H_{r+n}/\delta u_k)\circ Q$, for which we obtain from (2.4)

$$Q_k = \sum_{\alpha=n-1-k}^{n-1} \sum_{l=0}^{n-2} \binom{\alpha}{n-1-k} b_{\alpha l} v_l^{(n-k-1)} + \cdots,$$

where we have calculated $(d/d\lambda)|_{\lambda=0} Q(\lambda v)$. Then $u_k = \Sigma_{l=0}^{n-2} p_{kl} v_l^{(n-k-1)}$ + (lower terms), and

$$\frac{\delta H_{r+n}}{\delta u_k} \circ Q = \sum_{l=0}^{n-2} \gamma_{kl} v_l^{(r+k)} + \cdots.$$

The matrix γ is nonsingular if the same is true for

$$p_{kl} = \sum_\alpha \binom{\alpha}{n-1-k} b_{\alpha l}$$

and this is seen from the relation

$$\begin{pmatrix} \mu \\ 1\ 1 \cdots 1 \end{pmatrix} \begin{pmatrix} b & \begin{matrix} 1 \\ \vdots \\ 1 \end{matrix} \end{pmatrix} = \begin{pmatrix} & & \vdots \\ & p & \sum_\alpha \mu_{l\alpha} \\ & & \vdots \\ 0 & \cdots\ 0 & n \end{pmatrix}.$$

The first matrix on the left-hand side is

$$\mu_{l\alpha} = \binom{\alpha}{n-1-l}$$

bordered by the row (1, 1,...,1); this matrix is triangular with non-null entries on the diagonal and hence nonsingular. The second matrix is $\tilde{c}$ [(4.6)] nonsingular, therefore p is nonsingular.

We take r prime to n such that the coefficients $E_q(v)$ contain derivatives of strictly lower orders than $r+M$. We make the change $w=\gamma v$: if $E(\nabla_u H_{r+n}\circ Q)=0$ each coefficient of $w_k^{(r+M+k)}$ should vanish separately, then $(E_M)_{ik}=0, i,k=0.1,...,n-2$ for arbitrary M and this implies $E=0$.

VII. CANONICAL MAPS AND BÄCKLUND TRANSFORMATIONS

The results of this section were considered in the Kuperschmidt work of Ref. 14.

Let us examine the Lax representation (4.7) of the modified equations (4.11):

$$\dot{T} = [T, \mathscr{M}_r]. \tag{7.1}$$

Note the following property due to the definition (2.13) of $T(V,\partial) = a(\partial + D(V))$:

$$aT(V,\partial)\tilde{a} = T(aV,\partial), \tag{7.2}$$

as follows from (2.14) and the unitary character of a. The invariance of the operator T under the transformation defined by the unitary matrix a suggests to us to investigate the behavior of Eq. (7.1) in this transformation. It is easily seen that (7.1) is left invariant if the operator $\mathscr{M}_r$ is itself an invariant operator. We shall prove later that it also happens for the transformation defined by a.

Let us denote by $\bar{V}=aV$ the action of a on V. The resulting vector $\bar{V}$ is obtained from V by a cyclic permutation of the V-components $\bar{V}_0 = V_1$, $\bar{V}_1 = V_2$,..., $\bar{V}_{n-2} = V_{n-1}$, $\bar{V}_{n-1} = V_0$ and this guarantees that $\Sigma_{\alpha=0}^{n-1} \bar{V}_\alpha = \Sigma_{\alpha=0}^{n-1} V_\alpha = 0$ according to (2.1).

But the V-components are not independent components. We shall define a transformation in terms of the v-variables for which the motion equations are formulated.

Proposition 7.1: Let l be the $(n-1)\times(n-1)$ matrix

$$l = -\mathscr{S}\tilde{b}ab \tag{7.3}$$

with $\mathscr{S}$ defined by (4.9), b in (2.1), and a by (2.10). Then

$$\bar{v} = lv \tag{7.4}$$

implies that the vector $\bar{V}$, defined in terms of $\bar{v}$ by $\bar{V}=b\bar{v}$, is related to V by

$$\bar{V} = aV. \tag{7.5}$$

Proof: From (7.3) and (7.4) one has $-\mathscr{S}^{-1}\bar{v} = \tilde{b}abv$. If we take into account that $-\mathscr{S}^{-1} = \tilde{b}b$, then $\tilde{b}(b\bar{v}-abv) = \tilde{b}(\bar{V}-aV)=0$. Thus $\bar{V}-aV=K$, with a certain K such that $\tilde{b}K=0$ and $\Sigma_{\alpha=0}^{n-1}K_\alpha = 0$. These two conditions about K may be summarized in a single condition $cK=0$ with with the matrix c defined by (4.6). As it was proved that $\det c \neq 0$, K must be equal to zero, and this finishes the proof.

Concerning the matrix l we have

Proposition 7.2: The matrix l has an inverse given by

$$l^{-1} = -\mathscr{S}\tilde{b}\tilde{a}b. \tag{7.6}$$

Proof: We take advantage of the identity

$$ca\tilde{c} = \begin{pmatrix} & & 0 \\ \tilde{b}ab & & \vdots \\ & & 0 \\ 0 & \cdots\ 0 & n \end{pmatrix} \tag{7.7}$$

that comes from the definition (4.6) of c. Since $\det c \neq 0$, this shows us that $\det 1 = -(\det\mathscr{S})\det(\tilde{b}ab) \neq 0$. Now, from (7.5) $b\bar{v} = abv$ or $\tilde{b}\tilde{a}b\bar{v} = \tilde{b}bv$. The introduction of $\mathscr{S} = -(\tilde{b}b)^{-1}$ completes the proof.

Corollary: The transformation (7.4) is a canonical invariance map for the symplectic operator $\mathscr{K}$ [(4.10)].

Proof: From (7.3) we get the expression

$$\tilde{l} = -\tilde{b}\tilde{a}b\mathscr{S}$$

from the transpose of l (remember that $\mathscr{S} = \tilde{\mathscr{S}}$). Formula (7.6) yields $l^{-1} = \mathscr{S}\tilde{l}\mathscr{S}^{-1}$ or

$$\mathscr{S} = l\mathscr{S}\tilde{l}. \tag{7.8}$$

Multiply on both sides by ∂ to obtain $\mathscr{K} = l\mathscr{K}\tilde{l}$.

The invariance of $\mathscr{K}$ under the transformation (7.4) suggests to us another way to prove the invariance of the modified equations. We shall give two different, but equivalent proofs of this fact: one for the Hamiltonian form (4.11), another for the Lax representation (7.1).

Proposition 7.3: Let $\chi(v,z)$, $\chi(\bar{v},z)$ be the solutions of (2.5) corresponding to $(v_0, v_1,...,v_{n-2})$ and $(\bar{v}_0, \bar{v}_1,...,\bar{v}_{n-2})$, respectively. Let $\bar{v}$ and v be connected by [(7.4)] $\bar{v}=lv$, then $\chi(v,z)$ is related to $\chi(\bar{v},z)$ by

$$\chi(v,z) = \chi(\bar{v},z) + \partial \ln[\chi(\bar{v},z) + \bar{V}_{n-1}]. \tag{7.9}$$

Proof: Denote by $\Lambda(V) = (\partial + V_0)\cdots(\partial + V_{n-1})$, $\Lambda(\bar{V}) = (\partial + \bar{V}_0)\cdots(\partial + \bar{V}_{n-1})$. From Eq. (7.5) one finds the relation

$$\Lambda(V)(\partial + \bar{V}_{n-1}) = (\partial + \bar{V}_{n-1})\Lambda(\bar{V}). \tag{7.10}$$

Take the solution[13] $\chi(\bar{v})$ of (2.5a), $\chi(\bar{v}) = \epsilon z + O(z^{-1})$; then (see Ref. 13) $(\Lambda(\bar{V}) - z^n)\exp\hat{\chi}(\bar{v}) = 0$ implies that $\alpha = (\partial + \bar{V}_{n-1})\exp\hat{\chi}(\bar{v}) = (\chi(\bar{v}) + \bar{V}_{n-1})\exp\hat{\chi}(\bar{v})$ satisfies the equation $(\Lambda(V) - z^n)\alpha = 0$ as follows from (7.10). But α is a solution of the form

$$\alpha = \alpha_0 \exp\int^x (\chi(\bar{v}) + \partial \ln[\chi(\bar{v}) + \bar{V}_{n-1}])dx,$$

where α_0 does not depend on x, as follows from the first-order differential equation

$$\alpha' = (\chi(\bar{v}) + \partial \ln[\chi(\bar{v}) + \bar{V}_{n-1}])\alpha.$$

But $\chi(\bar{v}) + \partial \ln[\chi(\bar{v}) + \bar{V}_{n-1}]$ is a power series of z which does satisfy (2.5a),being $\chi(\bar{v}) + \partial \ln[\chi(\bar{v}) + \bar{V}_{n-1}] = \epsilon z + O(z^{-1})$. Thus, (by proposition 2.1 of Ref. 13) formula (7.9) holds.

Corollary. The functional $\mathscr{H}$ [(4.8)] is invariant, that is

$$\mathscr{H}(v,z) = \mathscr{H}(\bar{v},z). \tag{7.11}$$

Theorem 7.1: The modified Hamiltonian systems (4.11) are invariant under (7.4). That is

$$\dot{\bar{v}} = \mathscr{K}\nabla_{\bar{v}}\mathscr{H}_r$$

if $\dot{v} = \mathscr{K}\nabla_v\mathscr{H}_r$ and $\bar{v} = lv$.

Proof: For the proof take the derivative of (7.4) with respect to t having in mind (4.11), (7.8), and (7.11).

On the other hand, one can obtain another proof of Theorem 7.1 by considering the invariance of the operator $\mathscr{M}_r$ in the Lax representation (7.1) as was announced.

From Proposition (7.3) one has the relation

$$e^{\hat{\chi}(v)} = \frac{1}{\epsilon z}(\partial + \bar{V}_{n-1})e^{\hat{\chi}(\bar{v})} \tag{7.12}$$

in which the constant factor $1/\epsilon z$ comes out by examining the power series expansion of $\exp[\hat{\chi}(v) - \hat{\chi}(\bar{v})]$ and that of $\chi(\bar{v}) + \bar{V}_{n-1}$

We shall also need to use the corresponding formula to (7.9) for the solutions S of (2.5b) (see Ref. 13).

$$S(\bar{v}) = S(v) - \partial\left(\frac{S(\bar{v})}{\chi(\bar{v}) + \bar{V}_{n-1}}\right),$$

which is proved in same manner as (7.9) was. This relation enables us to write

$$S(\bar{v})e^{-\hat{\chi}(\bar{v})} = \frac{1}{\epsilon z}(\partial + V_0)^* S(v)e^{-\hat{\chi}(v)} \tag{7.13}$$

the analog of (7.12) for the "adjoint problem".

Now, we are in a position to formulate

Theorem 7.2: The operator $\mathscr{M}$ (3.3) is left invariant by (7.4). Moreover, the Lax equation (7.1) is invariant.

Proof: From (7.5), (7.12), and (7.13), keeping in mind definition (2.8a) and (2.8b) of ξ and ζ, it is easy to see that

$$\xi(\bar{v}) = \frac{1}{\epsilon}a\xi(v), \quad \zeta(\bar{v}) = \epsilon a\zeta(v).$$

Thus, one finds

$$a\mathscr{M}(v)\tilde{a} = aD(\xi)\tilde{a}\partial^{-1}aD(a\zeta)\tilde{a} = \mathscr{M}(\bar{v}),$$

where (2.14) has been used. This formula, together with Eq. (7.2), guarantees the invariance of (7.1).

As we can repeat the transformation (7.4) with $\bar{v}$ as starting solution, $(l) = \{1, l, l^2, \dots, l^{n-1}\}$ are canonical invariance maps for the modified Hamiltonian systems (4.11). That $l^n = 1$ follows from (7.5) and the property $a^n = 1$ [(2.11)].

The same argument used in the Miura transformation for the KdV equation allows us to bring Bäcklund transformations for the Gel'fand–Dikii equations (4.2).

By Theorem 5.1 $u = Q(v)$ satisfied Eq. (6.4), and the same is true for $\bar{u} = Q(lv)$. We take advantage of relation (7.10) in the form

$$L(u)(\partial + \bar{V}_{n-1}) = (\partial + \bar{V}_{n-1})L(\bar{u}) \tag{7.14}$$

according to the transformation law $L(u) = \Lambda(V)$.

In terms of the variables $\partial\alpha_k = \bar{u}_k + u_k$, $\partial\beta_k = \bar{u}_k - u_k$, Eq. (7.14) reads

$$\alpha_k'' = \sum_{l \geqslant k+1} \frac{1}{n}\binom{l}{k}\beta_{n-2}{}^{(l-k)}(\alpha_l' - \beta_l') - \beta_k'' - 2\beta_{k-1}' - 2\beta_k'\left(\frac{1}{n}\beta_{n-2} + \lambda_0\right), \tag{7.15}$$

$$k = 0, 1, \dots, n-2,$$

where $\lambda_0 =$ constant and $\bar{V}_{n-1}$ has been expressed as $\bar{V}_{n-1} = (1/n)\beta_{n-2} + \lambda_0$ as follows from (7.14). We can rewrite (7.15) in the form $\alpha_k' = g_k(\beta)$ but only in the cases $n = 2, 3$ are the g_k differential polynomials in β.

In fact one has

$$\alpha_{n-2}' = (n-2)\beta_{n-2}' - \frac{1}{n}\beta_{n-2}^2 - 2\lambda_0\beta_{n-2} - 2\beta_{n-3} + \lambda_1, \lambda_1 = \text{const}, \tag{7.16a}$$

$$\begin{aligned}\alpha_{n-3}'' = {} & \frac{(n-2)(n-3)}{n}(\beta_{n-2}')^2 \\ & - \frac{2}{n}(\beta_{n-2}\beta_{n-3}' + (n-2)\beta_{n-2}'\beta_{n-3}) \\ & + \frac{(n-1)(n-2)}{3n}\beta_{n-2}''' - \frac{n-2}{(n^2)}\beta_{n-2}^2\beta_{n-2}' \\ & - \frac{2(n-2)}{n}\lambda_0\beta_{n-2}\beta_{n-2}' \\ & + \frac{n-2}{n}\lambda_1\beta_{n-2}' - \beta_{n-3}'' - 2\lambda_0\beta_{n-3}' - 2\beta_{n-4}'.\end{aligned} \tag{7.16b}$$

Note[14] that the first two terms on the right-hand side of (7.16b) are total derivatives only for $n = 3$.

VIII. EXAMPLES

We shall give here some explicit constructions for the symplectic operator and Bäcklund transformations previously considered.

$n = 2$.

$$b = \begin{pmatrix} \gamma \\ -\gamma \end{pmatrix}, \quad \gamma = \text{const}, \quad \mathscr{S} = -(\tilde{b}b)^{-1} = -1/2\gamma^2.$$

The symplectic operator k [(4.10)] is $k = (1/2\gamma^2)\partial$ and Miura transformation $u_0 = Q_0(v_0) = -(\gamma v_0' + \gamma^2 v_0^2)$. The operator $\mathscr{G}$ is found to be

$$\begin{aligned}\mathscr{G} &= d_{v_0}Qkd_{v_0}^+Q = \frac{1}{2\gamma^2}(\gamma\partial + 2\gamma^2 v_0)\partial(\gamma\partial - 2\gamma^2 v_0) \\ &= \tfrac{1}{2}\partial^3 + 2u_0\partial + u_0'.\end{aligned}$$

To formulate the Bäcklund transformation, Eq. (7.16a) is here

$$\alpha_0' = -\tfrac{1}{2}\beta_0^2 - 2\lambda_0\beta_0 + \lambda_1$$

or $\mathscr{T}\alpha_0 = (\delta/\delta\beta_0)F$, with $F = \int dx(-\tfrac{1}{3}\beta_0^3 - 2\lambda_0\beta_0^2 + 2\lambda_1\beta_0)$ being $\mathscr{T} = 2\partial$, the operator of Gel'fand–Dikii.

$n = 3$.

We select for simplicity the matrix

$$b = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \\ -2 & 0 \end{pmatrix},$$

which gives us

$$\mathscr{S} = -\frac{1}{3}\begin{pmatrix} 1 & 0 \\ 0 & 3 \end{pmatrix}.$$

The transformation is induced here by

$$\Big(\partial + \frac{1}{\sqrt{2}}(v_0 + v_1)\Big)\Big(\partial + \frac{1}{\sqrt{2}}(v_0 - v_1)\Big)(\partial - \sqrt{2}v_0) = \partial^3 + u_1\partial + u_0.$$

The operator $\mathscr{G} = d_v Q \not k d_v^+ Q$,

$$\mathscr{G} = \begin{pmatrix} \mathscr{G}_{00} & \mathscr{G}_{01} \\ \mathscr{G}_{10} & \mathscr{G}_{11} \end{pmatrix}$$

with

$$\begin{aligned} \mathscr{G}_{00} = & -\tfrac{1}{3}[2\partial^5 + 4u_1\partial^3 + 6u_1'\partial^2 \\ & + (6u_1'' - 6u_0' + 2u_1^2)\partial] \\ & + 2u_1''' - 3u_0'' + 2u_1u_1'], \\ \mathscr{G}_{01} = & -\mathscr{G}_{10}^* = \partial^4 + u_1\partial^2 + 3u_0\partial + u_0', \\ \mathscr{G}_{11} = & \, 2\partial^3 + 2u_1\partial + u_1', \end{aligned}$$

which are found after some calculations. The operator of Gel'fand–Dikii is here

$$\mathscr{T} = \begin{pmatrix} 0 & 3\partial \\ 3\partial & 0 \end{pmatrix}. \tag{8.1}$$

For the Bäcklund transformations, formulas (7.16a) and (7.16b) yield

$$\begin{aligned} \alpha_1' &= \beta_1' - \tfrac{1}{3}\beta_1^2 - 2\lambda_0\beta_1 - 2\beta_0 + \lambda_1, \\ \alpha_0' &= -\tfrac{2}{3}\beta_0\beta_1 + \tfrac{2}{3}\beta_1'' - \tfrac{1}{27}\beta_1^3 - \tfrac{1}{3}\lambda_0\beta_1^2 \\ & \quad + \tfrac{1}{3}\lambda_1\beta_1 - \beta_0' - 2\lambda_0\beta_0 + \lambda_2 \end{aligned}$$

that may be written in the form

$$\mathscr{T}\alpha = \nabla_\beta F \tag{8.2}$$

with the Gel'fand–Dikii operator (8.1) and the functional

$$F(\beta) = \int dx\Big(3\beta_0\beta_1' - \beta_0\beta_1^2 - 6\lambda_0\beta_0\beta_1 - 3\beta_0^2 + 3\lambda_1\beta_0 - \tfrac{1}{3}(\beta_1')^2 - \tfrac{1}{36}\beta_1^4 - \frac{\lambda_0}{3}\beta_1^3 + \lambda_1\beta_1^2 + 3\lambda_2\beta_1\Big).$$

When a Hamiltonian system admits a symplectic operator which does not contain the variables (the u_k in this case), the Bäcklund transformations which may be written in the form (8.2) prove to be canonical in variance maps for such system.[5] Here, that is the case only for $n = 2, 3$.

APPENDIX A

We shall construct here an equation[2,3,10] for the power series $\nabla_u H[u,z]$ [Eq. (4.1) of Ref. 13].

Consider the formal series

$$\rho(x,\xi,z) = \sum_{-\infty}^{\infty} \frac{S_k(x,z)}{\xi^{k+1}}, \quad S_k = (\chi - \partial)^k S,$$

χ and S being the solutions of (2.5a) and (2.5b). For negative k, $(\chi - \partial)^k$ is defined as a formal series in ∂. From Eqs. (2.2) and (2.15) of Ref. 13 we get the relations

$$\sum_{k=0}^{n}\sum_{\alpha=0}^{k}\binom{k}{\alpha}u_k S_{k+l-\alpha}^{(\alpha)} = z^n S_l, \tag{A1}$$

$$\sum_{k=0}^{n}\sum_{\alpha=0}^{\infty}(-1)^\alpha\binom{k+l}{\alpha}S_{k+l-\alpha}u_k^{(\alpha)} = z^n S_l, \tag{A2}$$

which are obtained in the Gel'fand–Dikii papers cited in Ref. 1. By introducing the multiplication law (see for example the work of Adler in Ref. 2)

$$a_1(x,\xi)\circ a_2(x,\xi) = \sum_{r\geqslant 0}\frac{1}{r!}(\partial_\xi^r a_1)(a_x^r a_2)$$

(here $\partial_x \equiv \partial$) one sees that (A 1) and (A 2) are equivalent to

$$L\circ\rho = z^n\rho,$$

$$\rho\circ L = z^n\rho,$$

with $L(x,\xi) = \Sigma_0^n u_k\xi^k = e^{-\xi x}L(\partial)e^{\xi x}$. Then ρ commutes with L and the decomposition

$$\rho = \rho_+ + \rho_-, \quad \rho_+ = \sum_{k\geqslant 0}S_{-k-1}\xi^k, \quad \rho_- = \sum_{k\geqslant 0}\frac{S_k}{\xi^{k+1}}$$

gives us

$$L\circ\rho_+ - \rho_+\circ L = \rho_-\circ L - L\circ\rho_-.$$

Since the left-hand side is positive in ξ so is the right-hand side and hence it contains the coefficients $S_0, S_1,\ldots,S_{n-2}$ only.

The equations for ρ yield

$$L\circ(\rho_+ + \rho_-) = z^n\rho, \quad (\rho_+ + \rho_-)\circ L = z^n\rho$$

and

$$L\circ\rho_+ + (L\circ\rho_-)_+ = z^n\rho_+, \quad \rho_+\circ L + (\rho_-\circ L)_+ = z^n\rho_+$$

to get finally

$$L\circ(\rho_-\circ L)_+ - (L\circ\rho_-)_+\circ L = z^n(L\circ\rho_+ - \rho_+\circ L). \tag{A3}$$

On the other hand ρ_- may be written in the form

$$\rho_- = \sum_{k\geqslant 0}\frac{S_k}{\xi^{k+1}} = \sum_{k\geqslant 0}\frac{1}{(\xi + \partial_x)^{k+1}}R_k.$$

Here $1/(\xi + \partial_x)^{k+1}$ should be understood as a power series in ∂_x. It is easily seen that

$$R_k = \sum_{\alpha=0}^{k}\binom{k}{\alpha}\partial_x^\alpha S_{k-\alpha} = SP_k(\chi), \quad k = 0,1,2\ldots \tag{A4}$$

by using the identity

$$S_k = (\chi - \partial_x)^k S = \sum_{\alpha=0}^{k}(-1)^\alpha\binom{k}{\alpha}\partial_x^\alpha(SP_{k-\alpha}),$$

which is proved by induction.

The left-hand side of (A 3) depends on $(R_0,R_1,\ldots,R_{n-1})$. To obtain an equation for $\nabla_u H[u,z] = (R_k)_{k=0}^{n-2}$ [see Eq. (3.1) of Ref. 13], R_{n-1} should be expressed in terms of $\nabla_u H[u,z]$. To do that we use Eqs. (2.5a) and (2.5b). From the Riccati equation we get

$$\sum_{k=0}^{n} u_k R_k = z^n R_0$$

(we have multiplied it by S). The equation for S gives us

$$R_n - n\partial_x R_{n-1} = z^n R_0 - \sum_{i=0}^{n-2} \sum_{k>i} (-1)^{k-i}\binom{k}{i}\partial_x^{k-i}(u_k R_i)$$

(see the Proposition 2.3 in Ref. 13). If we eliminate R_n between these two equations we deduce for R_{n-1} the expression

$$R_{n-1} = \frac{1}{n}\sum_{i=0}^{n-2}\sum_{k>i+1}(-1)^{k-i}\binom{k}{i}\partial_x^{k-i-1}(u_k R_i)$$

in terms of $(R_0, R_1, ..., R_{n-2})$. Therefore (A3) induces the equation

$$(\mathscr{G} - z^n\mathscr{T})\nabla_u H[u,z] = 0, \tag{A5}$$

where $\mathscr{T}$ is the symplectic operator of Gel'fand and Dikii.

More detailed versions of these facts are available in the literature cited.

APPENDIX B

In the work of Veselov (cited in Ref. 13) it is proved that for r prime to n, the matrix α_{kl}, $k,l = 0,1,...,n-2$ in the linear part of $\delta H_{r+n}/\delta u_k$

$$\frac{\delta H_{r+n}}{\delta u_k} = \sum_{l=0}^{n-2}\alpha_{kl}u_l^{(r+k+l+1-n)} + \cdots \tag{B1}$$

is nonsingular (the Hamiltonians $H_1, H_2, ..., H_n$ are not of interest here). The following proof is due to Kuperschmidt and Wilson.[10]

We consider the series ρ introduced in the Appendix I for which the equation

$$L\circ\rho - \rho\circ L = 0$$

holds. Let ρ_{r+n} denote the coefficient of $1/z^{r+n}$ obtained from the expansion of $S_k(x,z)$ in powers of z and $\overline{\rho}_{r+n} = (d/d\lambda)|_{\lambda=0}\rho_{r+n}(\lambda u)$ the linear part of it. From $L\circ\rho_{r+n} - \rho_{r+n}\circ L = 0$ we deduce

$$\xi^n\circ\overline{\rho}_{r+n} - \overline{\rho}_{r+n}\circ\xi^n + L_1\circ\rho_{r+n}(0) - \rho_{r+n}(0)\circ L_1 = 0,$$

where $L_1 = \overline{L} = \sum_{k=0}^{n-2} u_k \xi^k$ and $\rho_{r+n}(0) = \rho_{r+n}(u)|_{u=0}$. But

$$\rho(0) = \sum_k \frac{S_k(0)}{\xi^{k+1}} = \sum_k \frac{(z-\partial_x)^k}{\xi^{k+1}}\frac{1}{nz^{n-1}} = \sum_k \frac{z^k}{nz^{n-1}}\frac{1}{\xi^{k+1}}$$

and

$$\rho_{r+n}(0) = \frac{1}{n}\xi^r,$$

as follows from $\chi = z + O(z^{-1})$, $S = \frac{1}{nz^{n-1}} + O(z^{-n})$, and the homogeneity properties of χ and S (we consider $\epsilon_i = 1$ only; the proof for $\chi^{[i]}, S^{[i]}$ arbitrary is the same). If we take into account the relations $\xi^n\circ\overline{\rho}_{r+n} = (\xi+\partial)\overline{\rho}_{r+n}$, $\overline{\rho}_{r+n}\circ\xi^n = \xi^n\overline{\rho}_{r+n}$ we get

$$\xi^n\left[\left(\frac{\partial_x}{\xi}+1\right)^n - 1\right]\overline{\rho}_{r+n} + \frac{1}{n}\xi^r L_1(\xi) - (\xi+\partial_x)^r L_1(\xi) = 0$$

or

$$\overline{\rho}_{r+n} = \frac{1}{n}\sum_{l=0}^{n-2}\frac{(1+\partial/\xi)^r - 1}{(1+\partial/\xi)^n - 1}\frac{u_l}{\xi^{n-r-l}},$$

from which

$$(\overline{S}_k)_{r+n} = \sum_{l=0}^{n-2}\beta_{kl}u_l^{(r+k+l+1-n)},$$

where β_{kl} is the coefficient of $\zeta^{r+k+l+1-n}$ in the series

$$\frac{1}{n}\frac{(1+\zeta)^r - 1}{(1+\zeta)^n - 1}$$

and for r prime to n this Hankel matrix is nondegenerate.[10] The 1–1 relation between $(R_0,...,R_{n-2})$ and $(S_0,...,S_{n-2})$,

$$\overline{S}_k = \sum_{\alpha=0}^{k}(-1)^\alpha\binom{k}{\alpha}\partial^\alpha R_{k-\alpha},$$

the inverse of (B4), finishes the proof.

[1] See references given in the previous paper of Ref. 13.
[2] M. Adler, Invent. Math. **50**, 219 (1979); D. R. Labedev and Yu. I. Manin, Funct. Anal. Appl. **13**, 40 (1979).
[3] W. Symes, J. Math. Phys. **20**, 721 (1979).
[4] R. M. Miura, J. Math. Phys. **9**, 1202 (1968).
[5] F. Guil Guerrero and L. Martínez Alonso, "Hamiltonian Techniques and Backlund Transformations", J. E. N. Report (Madrid, 1979) (Spanish) (unpublished).
[6] M. Adler and J. Moser, Commun. Math. Phys. **61**, 1 (1978).
[7] M. Jaulent and I. Miodek, Lett. Nuovo Cimento **20**, 655 (1977).
[8] F. Guil Guerrero and L. Martínez Alonso, Lett. Nuovo Cimento **27**, 85 (1980).
[9] V. V. Sokolov and A. B. Shabat, Funct. Anal. Appl. **14**, 79 (1980).
[10] B. A. Kupershmidt and G. Wilson, Invent. Math. **62**, 403 (1981).
[11] I. M. Gel'fand and I. Ya. Dorfman, Funct. Anal. Appl. **14**, 71 (1980) (Russian).
[12] A. P. Fordy and J. Gibbons, J. Math. Phys. **21**, 2508 (1980); Commun. Math. Phys. **77**, 21 (1980).
[13] F. Guil Guerrero, "The Riccati equation and Hamiltonian Systems," J. Math. Phys. **23**, 211 (1982).
[14] B. A. Kupershmidt, "Deformations of Integrable Systems" (to be published).
[15] M. M. Vainberg, *Variational Methods for the study of Nonlinear Operators* (Holden-Day, San Francisco, 1964).
[16] R. Jost, Rev. Mod. Phys. **36**, 572 (1964).
[17] V. I. Arnold, *Mathematical Methods of Classical Mechanics* (Springer, New York, 1979).

A global isometry approach to accelerating observers in flat space-time

John R. Urani
Department of Physics, University of Missouri–Kansas City, Kansas City, Missouri 64110
Marilyn H. Kemp
Department of Mathematics, University of Missouri–Kansas City, Kansas City, Missouri 64110

(Received 23 June 1981; accepted for publication 28 August 1981)

A global two-point diffeomorphic extension of Lorentz transformations is constructed which preserves the global Lorentzian metric structure of flat R^4. This global mapping induces, as a tangent-space mapping, instantaneous Lorentz transformations parametrized by interframe velocity functions. The elimination of pseudoterms from particle and electromagnetic field equations leads to an exact analytic expression for the affine connection needed for covariant differentiation. Examination of invariant particle equations gives an obvious proof of the equivalence principle in terms of the symmetric part of the acceleration-group connection. Transformation properties of the connection coefficients are shown to be in accord with general covariance requirements. The specific case of the rotating observer is treated exactly where it is seen that the affine connection merely supplies the exact Thomas precession term. Recent work by DeFacio *et al.* is found to be especially convenient for comparison with the present work. The results of the two approaches agree precisely. A summary of results indicates that the global isometry approach gives results consistent with those obtained via presymmetry arguments.

PACS numbers: 02.40. + m

I. INTRODUCTION

A number of recent papers[1–4] have dealt with accelerating observers in flat Lorentzian R^4 using modern differential geometry and the concept of presymmetry. Local differentiability and covariance arguments are shown to lead to extensions of previous results in such areas as Fermi–Walker transport.[5]

Recent work has also appeared[6–8] treating the accelerating/rotating observer in arbitrary space-times via general geometrical methods. Another approach, taken in this paper, is to construct extensions of global invariance groups [such as O (3) for Euclidean R^3 or the full Lorentz group L_F, for Lorentzian R^4] on flat metric spaces and then to investigate the local (tangent-space) mapping induced by the extension of the global group.

In this work we extend proper Lorentz transformations on Lorentzian R^4 to nonlinear transformations of relative coordinates parametrized by arbitrary C^1 time-like interfame velocity functions. By requiring that the global metric $\eta = (+\ \ +\ \ +\ \ -)$ be invariant and that Galilean and Lorentz transformations result in low-speed and zero-acceleration cases, respectively, a relative-coordinate map is found which induces a tangent-space isometry consisting of instantaneous Lorentz transformations.

Particle four-momenta, defined in terms of tangents to nonspacelike curves, then instantaneously boost to the accelerating reference frame, and the elimination of pseudoterms from particle equations relative to noninertial frames gives an analytic expression for affine connection coefficients applicable in all cases of C^2 time-like observer world lines.

Particle equations lead immediately to a local equivalence principle for the symmetric part of the affine connection.

Applications of the local theory are presented including the simple rotating observer (where the affine connection supplies the Thomas precesion term) and a comparison with the results of DeFacio *et al.*[2] (which are identical with those obtained via the methods shown herein).

Pseudoterms appearing in Maxwell's equations are also eliminated via the same connection and covariant Maxwell equations are written down. The pseudoterms (pseudocurrents) are briefly discussed because of their implications concerning charge conservation.

Finally, a summary indicates that the results of this paper are consistent with those obtained from general presymmetry arguments.[9]

Concerning notation, since some manipulations encountered in this work may be rather novel, many results can most easily be obtained in coordinate form. However, covariant coordinate-free results will be indicated where appropriate.

II. GLOBAL COORDINATE MAP

Covering Lorentzian R^4 with a single orthonormal Cartesian coordinate patch, $(x, y, z, ct) = (x^1, x^2, x^3, x^4)$, the view is taken that the causal structure of the C^∞ manifold is absolute in that all time-like observers must obtain the same causal relationships between any pair of events. More exactly, any mapping of coordinates $\phi: R^4 \to R^4$, parametrized by a future-directed time-like C^2 curve, must preserve the Lorentzian metric structure globally. The mapping is then naturally cast as a relative coordinate map due to its intrinsic nonlinearity (even in the Galilean limit) and due to the necessity of keeping $\eta(X_2 - X_1, X_2 - X_1)$ invariant, where $X_2, X_1 \in R^4$. By hindsight, the relative-coordinate map is convenient, too, in that it is translation invariant and easily defines a tangent-space mapping.

Almost all results may be obtained by considering a two-dimensional (2-D) example. Let (x_1, ct_1) and (x_2, ct_2) be two events relative to an origin O. Let an origin O' move with

 0022-2488/82/020227-04$02.50

velocity $v(t)$ relative to O. The relative coordinate transformation is, in the Galilean limit,

$$(x_2 - x_1)' = x_2 - x_1 - \int_{t_1}^{t_2} v(t)\,dt, \tag{2.1}$$

and $(t_2 - t_1)' \equiv (t_2 - t_1)$.

Similarly, relative coordinates may be taken as the basis for special relativity (SR), to no great advantage since the theory is linear.

We then take as the general form for the mapping, in the 2-D case,

$$(x_2 - x_1)' = \lambda\,[(x_2 - x_1) - \delta c(t_2 - t_1)],$$

and

$$c(t_2 - t_1)' = \lambda\,[c(t_2 - t_1) - \delta(x_2 - x_1)].$$

For an isometry, one immediately requires $\lambda^2 = (1 - \delta^2)^{-1}$, and to obtain the Galilean limit one requires $\delta \to 0$ for $|v(t)| \ll c$. Then, Eq. (2.1) is obtained, for low speeds, if

$$\delta = \left(\int_{t_1}^{t_2} v(t)\,dt\right)\Big/ c(t_2 - t_1). \tag{2.2}$$

Interestingly, the δ factor generalizes the rapidity β of SR to the average velocity of O' relative to O during the $t_2 - t_1$ interval. Clearly, if $v(t) = v_0$ (a constant), $\delta = \beta_0 = v_0/c$.

Even in this 2-D example it is obvious that the global map reduces to identity for pairs of events such that $\delta = 0$. This does not, however, eliminate accumulative tangent-space effects such as proper time intervals and spin precession.

To generalize to arbitrary velocities, $\mathbf{v}(t)$, of O' relative to O, the $\boldsymbol\beta$ parameter of SR is replaced by

$$\boldsymbol\delta = \int_{t_1}^{t_2} \mathbf{v}(t)\,dt / c(t_2 - t_1), \tag{2.3}$$

again the average O' velocity is relative to O. Defining $\lambda = (1 - |\boldsymbol\delta|^2)^{-1/2}$, the λ and $\boldsymbol\delta$ factors are then inserted in an arbitrary rotation-free Lorentz transformation matrix,[10] $\Lambda^{\mu}{}_{\nu}$ and the desired isometric map is

$$(X_2 - X_1)'^{\mu} = \Lambda^{\mu}{}_{\nu}(\boldsymbol\delta)(X_2 - X_1)^{\nu}, \tag{2.4}$$

where the summation convention applies but indices are not tensor indices. Due simply to the structure of $\Lambda^{\mu}{}_{\nu}$ we obtain that $(X_2 - X_1)^2 = |\mathbf{r}_2 - \mathbf{r}_1|^2 - c^2(t_2 - t_1)^2$ is invariant.

The $\boldsymbol\delta$ parameter is well behaved, even as $t_2 \to t_1$, since $\mathbf{v}(t)$ is C^1 and the mapping Eq. (2.4) is differentiable as well since $|\mathbf{v}(t)| < c$ is also assumed. Consequently, the mapping of relative coordinates defined by Eq. (2.4) is the diffeomorphic isometry, $\phi: R^4 \to R^4$, which was sought.

Clearly, for time intervals during which $\mathbf{v}(t)$ is constant, ordinary Lorentz transformations results and if $|\mathbf{v}(t)| \ll c$ during an interval, Galilean transformations result.

While being fundamental in guaranteeing an invariant causal structure on the Lorentzian space-time manifold, the global map is not particularly important for calculations. The induced tangent-space map of Sec. 3 is of prime importance for applications.

III. TANGENT SPACE MAP

For events in R^4 whose t labels are equal, the $\boldsymbol\delta$ and λ factors of Sec. 2 become time-dependent $\boldsymbol\beta$ and γ factors. In other words, equal-time events define relative coordinates which instantaneously Lorentz transform. This will be important in Sec. 4.

Similarly, infinitesimally separated events, for which $(X_2 - X_1)^{\mu} = dx^{\mu}$, will also suffer instantaneous boosts to the accelerating observer's frame via[11]

$$dx^{\mu'} = \Lambda^{\mu}{}_{\nu}(\boldsymbol\beta(t))dx^{\nu}. \tag{3.1}$$

Hence, the global map induces a tangent-space isometry at each point in R^4 according to the time-dependent $\boldsymbol\beta(t)$ of the accelerating observer.

The natural parametrization of time-like curves via an invariant proper time is again possible since, if dx^{μ} represents an infinitesimal displacement along such a curve, $d\tau^2 = -\eta_{\mu\nu}dx^{\mu}dx^{\nu}$ is invariant ($c = 1$). Hence the covariant velocity or normalized tangent vector has components $v^{\mu} = dx^{\mu}/d\tau$ and a massive particle's four-momentum is covariantly defined as $p^{\mu} = mv^{\mu}$. With a quantum hypothesis appended to interpret massless-particle four-momenta in terms of null-curve tangents, the tangent-space map implies an instantaneous Doppler shift, as expected.

Then, (1,0) tensor fields on Lorentzian R^4 simply boost to the accelerating frame as

$$T'^{\mu}(x') = \Lambda^{\mu}{}_{\nu}(x)T^{\nu}(x),$$

with tensor products of vector fields and 1-forms transforming via $\Lambda^{\mu}{}_{\nu}$ and $\Lambda^{-1\mu}{}_{\nu}$ on contravariant and covariant indices, respectively. In particular, the components of the electromagnetic 2-form $F_{\mu\nu}$ transform as

$$F'_{\mu\nu} = \Lambda^{-1\alpha}{}_{\mu}\Lambda^{-1\beta}{}_{\nu}F_{\alpha\beta}.$$

Using Eq. (3.1), transformations of coordinate velocity, $\mathbf{v} = d\mathbf{r}/dt$, and coordinate acceleration, $\mathbf{a} = d\mathbf{v}/dt$, may be easily obtained.[12]

The coordinate 4-acceleration, $a^{\mu} = dv^{\mu}/d\tau$, will illustrate the need for a set of connection coefficients for covariant differentiation. Relative to an accelerating observer,

$$a'^{\mu} = \frac{dv^{\mu}}{d\tau} = \Lambda^{\mu}{}_{\nu}\frac{dv^{\nu}}{d\tau} + \frac{d\Lambda^{\mu}{}_{\nu}}{d\tau}v^{\nu}. \tag{3.2}$$

Clearly Eq. (3.2) contains a pseudoterm and that term is Galilean for small $|\boldsymbol\beta(t)|$ in the matrix Λ. The pseudoterm is absent for inertial observers ($d\Lambda/d\tau = 0$).

If the transformation is to the self-frame of the accelerating observer, $v' = (\mathbf{0},c)$ and $a' = (\mathbf{0},0)$ from Eq. (3.2). Hence, self-4-acceleration has been eliminated all along the accelerating observer's world line but the 4-acceleration is not a covariant quantity. In Section 4, connection coefficients are obtained to eliminate the pseudoterm in $a' = dv'/d\tau$.

IV. THE AFFINE CONNECTION

Let O be an inertial frame and let $\bar{O}$ be noninertial relative to O. Rewriting Eq. (3.2) as

$$\frac{d\bar{v}^{\mu}}{d\tau} - \frac{d\Lambda^{\mu}{}_{\nu}}{d\tau}v^{\nu} = \Lambda^{\mu}{}_{\nu}\frac{dv^{\nu}}{d\tau}, \tag{4.1}$$

the left-hand side of Eq. (4.1) may be written entirely in terms of bar quantities, via $v^\nu = \Lambda^{-1\nu}{}_\alpha \bar{v}^\alpha$ and $\frac{d}{d\tau}(\Lambda^\mu{}_\alpha \Lambda^{-1\alpha}{}_\nu) = 0$, to become

$$\frac{d\bar{v}^\mu}{d\tau} + \Lambda^\mu{}_\alpha\Big(\frac{d}{d\tau}\Lambda^{-1\alpha}{}_\nu\Big)\bar{v}^\nu = \Lambda^\mu{}_\nu\frac{dv^\mu}{d\tau},$$

or

$$\frac{d\bar{v}^\mu}{d\tau} + \Lambda^\mu{}_\alpha\Big(\frac{\partial}{\partial\bar{x}^\beta}\Lambda^{-1\alpha}{}_\nu\Big)\bar{v}^\beta\bar{v}^\nu = \Lambda^\mu{}_\nu\frac{dv^\nu}{d\tau}.$$

We define the affine connection for $\bar{O}$ as

$$\bar{\Gamma}^\mu_{\beta\nu} = \Lambda^\mu{}_\alpha\Big(\frac{\partial}{\partial\bar{x}^\beta}\Lambda^{-1\alpha}{}_\nu\Big). \tag{4.2}$$

Recall that the Λ matrices appearing in Eq. (4.2) involve the variable $\bar{\beta}$ corresponding to $\bar{O}$ motion relative to O (which is an inertial frame with constant $\boldsymbol{\beta}$ factor).

Hence for inertial frames the Γ coefficients will be zero. However, formally adding a (zero-valued) connection Γ on the right in Eq. (4.1) we obtain the convariant equation of motion,

$$\frac{\bar{K}^\mu}{m} = \Big(\frac{\nabla\bar{v}}{d\tau}\Big)^\mu = \Lambda^\mu{}_\nu\Big(\frac{\Delta v}{\partial\tau}\Big)^\nu = \Lambda^\mu{}_\nu\frac{K^\nu}{m}, \tag{4.3}$$

where

$\frac{\nabla}{\partial\tau} = \frac{d}{d\tau} + \Gamma(v,\quad)$ and K is the convariant 4-force.

[Note: $\Gamma(v,\quad)$ means $\Gamma^\alpha_{\beta\gamma}v^\beta$ *not* $\Gamma^\alpha_{\beta\gamma}v^\gamma$ since Γ is *not* symmetric.]

If the acceleration $a^\mu = 0$ relative to O (inertial), Eq. (4.3) is identically the geodesic equation

$$\frac{d\bar{v}}{d\tau} + \bar{\Gamma}(\bar{v},\bar{v}) = 0$$

relative to $\bar{O}$, as required. It is to be noted also that the covariant particle equation [Eq. (4.3)] picks out only the symmetric part of the connection. Hence we have a statement of the equivalence principle that the gravitational connection is locally equivalent to an acceleration connection's symmetric part. The acceleration connection has zero Riemann–Christoffel tensor so the equivalence is strictly local.

The antisymmetric part of the connection, arising from antisymmetric rotation generators, is the source of Thomas precesssion, which is obtained in Sec. 6.

V. TRANSFORMING Γ

The transformation properties of the affine connection are easily shown to satisfy the general requirements imposed by covariance.

For all inertial frames, $\Gamma = 0$, from Eq. (4.2). Let O be inertial, let O' be noninertial with rapidity $\boldsymbol{\beta}(t)$ relative to O, and let O'' be noninertial with rapidity $\boldsymbol{\beta}'(t')$ relative to O'. The motion of O' induces the tangent-space map $\Lambda^\alpha{}_\nu[\boldsymbol{\beta}(t)]$, whereas the motion of O'' induces the the map $\Lambda'^\mu{}_\alpha[\boldsymbol{\beta}'(t')]$. Under composition the combined tangent-space map is $\Lambda'^\mu{}_\alpha(t')\Lambda^\alpha{}_\nu(t)$ and the connection for O'' is given by

$$\Gamma'' = \Lambda'\Lambda\frac{\partial}{\partial x''}(\Lambda'\Lambda)^{-1},$$

where indices have been suppressed for simplicity. Using $\partial/\partial x'' = (\Lambda')^{-1}\partial/\partial x'$ and simply differentiating, one obtains in full-component form,

$$\Gamma''^\alpha{}_{\beta\gamma} = \Lambda'^\alpha{}_\nu\frac{\partial}{\partial x''}\beta(\Lambda')^{-1\nu}{}_\gamma + \Lambda'^\alpha{}_\nu(\Lambda')^{-1\nu}{}_\beta(\Lambda')^{-1\sigma}{}_\gamma\Gamma'^\mu{}_{\nu\sigma}, \tag{5.1}$$

where

$$\Gamma'^\mu{}_{\nu\sigma} = \Lambda^\mu{}_\alpha\frac{\partial}{\partial x'}\nu\Lambda^{-1\alpha}{}_\sigma$$

is the connection for O'. It is seen that Eq. (5.1) is equivalent to the general law for transforming connections.[13]

VI. APPLICATIONS

In this section we treat two problems, the rotating frame[8] and the comparison of present methods with recent presymmetry formulations of accelerating observers.[2]

The rotating observers, $\bar{O}$, relative to an inertial, O, will be taken to have $|\boldsymbol{\beta}| = \text{const}$ and $\boldsymbol{\beta}\cdot d\boldsymbol{\beta}/dt = 0$. For simplicity we confine the motion to the xy plane. Connection coefficients are calculated using Eq. (4.2) in the form

$$\bar{\Gamma}^\mu_{\beta\nu} = \Lambda^\mu{}_\alpha\Big(\Lambda^{-1_4}{}_\beta\frac{\partial}{\partial ct}\Lambda^{-1\alpha}{}_\nu\Big),$$

since Λ depends on $\boldsymbol{\beta}(t)$ explicitly.

Nonzero connection coefficients are

$$\bar{\Gamma}^1_{12} = -\bar{\Gamma}^2_{11} = \frac{\beta_x}{c}\frac{(\gamma-1)}{\beta^2}(\beta_y\dot{\beta}_x - \beta_x\dot{\beta}_y),$$

$$\bar{\Gamma}^1_{22} = -\bar{\Gamma}^2_{21} = \frac{\beta_y}{c}\frac{(\gamma-1)}{\beta^2}(\beta_y\dot{\beta}_x - \beta_x\dot{\beta}_y),$$

$$\bar{\Gamma}^1_{42} = -\bar{\Gamma}^2_{41} = \frac{(\gamma-1)}{c\beta^2}(\beta_y\dot{\beta}_x - \beta_x\dot{\beta}_y),$$

$$\bar{\Gamma}^1_{14} = \bar{\Gamma}^4_{11} = \frac{\gamma}{c}\beta_x\dot{\beta}_x,\quad \bar{\Gamma}^1_{24} = \bar{\Gamma}^4_{21} = \frac{\gamma}{c}\beta_y\dot{\beta}_x,$$

$$\bar{\Gamma}^1_{44} = \bar{\Gamma}^4_{41} = \frac{\gamma}{c}\dot{\beta}_x,\quad \bar{\Gamma}^2_{14} = \bar{\Gamma}^4_{12} = \frac{\gamma}{c}\beta_x\dot{\beta}_y,$$

$$\bar{\Gamma}^2_{24} = \bar{\Gamma}^4_{22} = \frac{\gamma}{c}\beta_y\dot{\beta}_y,\ \text{and}\ \bar{\Gamma}^2_{44} = \bar{\Gamma}^4_{42} = \frac{\gamma}{c}\dot{\beta}_y,$$

where $\dot{\boldsymbol{\beta}} = d\boldsymbol{\beta}/d\tau = \gamma\, d\boldsymbol{\beta}/dt$ and $\bar{\Gamma}^\alpha_{\beta\gamma} \neq \bar{\Gamma}^\alpha_{\gamma\beta}$.

These connection coefficients satisfy the general requirements[4,5] of an acceleration connection.

Consider a classical spin vector, S, satisfying $S^\mu v_\mu = 0$. Let $dS^\mu/d\tau$ be given in the inertial frame, O. Relative to $\bar{O}$, $\bar{S} = (\bar{\mathbf{S}},0)$ and $\bar{v} = (0,c)$ if the spin is taken to be at $\bar{O}$.

After lengthy summations and manipulations, one finds, using

$$\frac{\nabla\bar{S}^\mu}{\partial\tau} = \Lambda^\mu{}_\nu\frac{\nabla S^\nu}{\partial\tau} = \Lambda^\mu{}_\nu\frac{dS^\nu}{d\tau}, \tag{6.1}$$

that

$$\frac{d\bar{S}^i}{d\tau} - (\boldsymbol{\omega}_T\times\bar{\mathbf{S}})^i = \Lambda^i{}_\nu\frac{dS^\nu}{d\tau}, \tag{6.2}$$

and $d\bar{S}^4/d\tau = 0$, where $\boldsymbol{\omega}_T$ is the expected Thomas precession angular velocity,

$$\omega_\tau = \frac{-(\gamma-1)}{\beta^2}(\boldsymbol{\beta}\times\dot{\boldsymbol{\beta}}).$$

A second application is to compare results of the present method with recent exact calculations of accelerating-frame observations of free (inertial) particles.[2] Two cases are considered in Ref. 2, those of spatial and null simultaneity. It is to be noted that in both cases, the authors choose to define $X - z(\tau)$ in terms of a basis of tangent vectors at $z(\tau)$, namely in terms of $u = dz/d\tau$ and a space-like orthornormal basis for $S_{z(\tau)} = u^\perp$. While it is generally not possible to equate manifold and tangent-space coordinates or displacements, expressing $X - z(\tau)$ as $\mathbf{r}\in u^\perp$ for the spatial simultaneous case and as $y = -ru + \mathbf{r}$ in the null simultaneous case has the advantage of being realistic in that an observer makes observations via the local tangent space. We agree that spatial simultaneity has mathematical deficits as well as describing a physically, impossible measurement process.

Accepting the expressions for $X - z(\tau)$ in either case, the calculation of $\dot{\mathbf{r}}$ and $\ddot{\mathbf{r}}$ leads directly to the results given in Ref. 2. Also, we are in complete argreement regarding null simultaneity of two events as an invariant statement and hence as an equivalence relation for all observers. The invariance of the null simultaneity statement is obvious whether manifold coordinates and the global map are used or the tangent-space expression and the induced local map are used.

VII. ELECTRODYNAMICS

The contravariant components of the electromagnetic 2-form, $F^{\mu\nu}$, transform as a (2,0) tensor via

$$F'^{\mu\nu} = \Lambda^\mu{}_\alpha \Lambda^\nu{}_\beta F^{\alpha\beta};$$

however, the Lorentz covariant Maxwell equation $\partial_\mu F^{\mu\nu} = -(4\pi/c)J^\nu$ will take on pseudoterms from the time-dependent boost to an accelerating frame. Specifically, we obtain

$$\partial'_\mu F'^{\mu\nu} = [\partial'_\mu(\Lambda^\mu{}_\alpha \Lambda^\nu{}_\beta)]F^{\alpha\beta} - (4\pi/c)J'^\nu, \tag{7.1}$$

where $J^{\nu'} = \Lambda^\nu{}_\alpha J^\alpha$. Due to the symmetry of the Λ matrices and the antisymmetry of F one still obtains a continuity equation

$$\partial_\nu{}'\partial_\mu{}'F'^{\mu\nu} = -(4\pi/c)\partial'_\nu J^{\nu'} = 0,$$

but Eq. (7.1) may be interpreted as containing pseudocurrent terms. It is easily demonstrated, using $\partial'_\mu = \Lambda^{-1}{}_\mu{}^4 \partial/\partial t$, that the pseudoterms of Eq. (7.1) have vanishing fourth component (no pseudocharge) which is physically reasonable. Also, the boosted Maxwell equation of Eq. (7.1) may be rearranged, whereby the pseudoterms are expressed in terms of the connection coefficients of Sec. 4, to obtain a covariant Maxwell equation.

$$\partial'_\alpha F'^{\alpha\beta} + \Gamma'^\alpha{}_{\alpha\gamma}F'^{\gamma\beta} + \Gamma'^\beta{}_{\alpha\gamma}F'^{\alpha\gamma} = -(4\pi/c)J'^\beta$$

or

$$F'^{\alpha\beta}{}_{;\alpha} = -(4\pi/c)J'^\beta.$$

In noncovariant form, keeping pseudoterms as inhomogeneous pseudocurrent terms, a Green's function integral equation solution for F is easily constructed.[12]

VIII. SUMMARY

The construction of a relative-coordinate-defined nonlinear extension of Lorentz transformations as a diffeomorphic isometry on Lorentzian R^4 leads to a number of pleasing results. Galilean and special relativity can be reproduced as limiting cases of the global-manifold map. The tangent-space mapping is entirely in agreement with presymmetry arguments in that it is parametrized by interframe velocity functions and acceleration covariance can be explicitly demonstrated with a closed form for the affine connection. The equivalence principle follows immediately from invariant particle equations. Finally, covariant electrodynamic equations were constructed based upon the connection for particle equations found in the present work.

Subsequent work has shown that the induced tangent-space isometry found for flat R^4 may be mapped onto the usual $D^{(1/2,0)} \oplus D^{(0,1/2)}$ spinor representation of the full Lorentz group thereby generalizing the Dirac equation to accelerating frames for which the appropriate connection coefficients have been found.

Working, via similar isometry techniques, on arbitrary space-times it has, furthermore, been possible to find a globally isometric tangent-space map induced by arbitrary time-like observers. Acceleration covariance has then been shown for general space-times. Careful analysis of the mappings involved in defining a general spinor theory has resulted in a generally covariant Dirac formalism as well.

ACKNOWLEDGMENTS

J. R. U. thanks J. C. Huang of the University of Missouri–Columbia Physics Department and both authors thank G. P. Barker of the University of Missouri-Kansas City Mathematics Department for useful discussions.

[1] Y. Avishai and H. Ekstein, Commun. Math. Phys. **37**, 193 (1974).
[2] B. DeFacio, P. W. Dennis, and D. G. Retzloff, Phys. Rev. D **18**, 2813 (1978).
[3] B. DeFacio, P. W. Dennis, and D. G. Retzloff, Phys. Rev. D **20**, 570 (1979).
[4] B. DeFacio and D. G. Retzloff, J. Math. Phys. **21**,751 (1980).
[5] C. Misner, K. Thorne, and J. Wheeler, *Gravitation* (Freeman, San Francisco, 1973).
[6] W. -T. Ni and M. Zimmerman, Phys. Rev. D **17**, 1473 (1978).
[7] W. -Q. Li and W. -T. Ni, J. Math. Phys. **20**, 1473 (1979); **20**, 1925 (1979).
[8] P. Burghart, Acta Phys. Austr. **48**, 181 (1978).
[9] Y. Avishai and H. Ekstein, Found. Phys. **2**, 257 (1972).
[10] C. Moller, *The Theory of Relativity* (Clarendon Press, Oxford, 1972), 2nd ed., p. 41.
[11] This may be obtained from the exact differential of Eq. (2.4) by taking the full differential in all eight variables and then letting $X_2 \to X_1$.
[12] J. R. Urani and G. D. Gale, Synthese **50**, No. 3 (1982).
[13] F. Rohrlich, *Classical Charged Particles* (Addison-Wesley, Reading, 1965), p. 288.

Clifford algebra approach to twistors

R. Abłamowicz [a)]
Department of Mathematics, Southern Illinois University at Carbondale, Carbondale, Illinois 62901
Z. Oziewicz
Institute of Theoretical Physics, Wrocław University, Wrocław, Poland
J. Rzewuski [a)]
Department of Physics, Southern Illinois University at Carbondale, Carbondale, Illinois 62901

(Received 24 April 1980; accepted for publication 1 August 1980)

Local particle interpretation or, equivalently, an enlargement of a structure group to the Poincaré group at each point of a Riemannian space-time manifold naturally results in a complexification of the Clifford algebra for the tangent Minkowski space. Following Crumeyrolle, twistor space is identified with an appropriate one-sided ideal of this algebra. Every antiautomorphism of the latter provides a unique projection from the complexified Clifford algebra onto the affine complex Minkowski space. This projection commutes with the action of the Poincaré group. Using the above approach, three projections (the cases of symmetric, antisymmetric, and Hermitian tensors) are derived. The projection in terms of the antisymmetric, decomposable tensors is shown to give the Penrose projection.

PACS numbers: 02.40.Ky

I. INTRODUCTION

The technique of Clifford algebra calculus has been proven useful in the theory of spinors,[1–5] especially in the curved twistor theory.[6,7] In the present paper it is shown how this technique used locally over a complex space-time manifold can be used to derive two new projections from a tensor product of twistor spaces. The very well-known projection given by Penrose[7–9] has been rederived for a purpose of completeness. In Sec. II we give a short historical survey on the problem of accomodation of translations in a spinorial fiber bundle and the logistic of our calculations. Section III starts with a brief review on an abstract Clifford algebra and its representation in a spinor space. The conformal group $\mathscr{C}(1,3)$ is introduced and its elements are expressed in terms of the Clifford numbers. Section IV deals with certain representations of $\mathscr{C}(1,3)$ on a complex Dirac–Clifford algebra defined through the antiautomorphisms of the latter. Main results of this paper are derived in the coordinate-free language. In Sec. V following the well-known approach of Refs. 1 and 10 we introduce a twistor space $\mathscr{T}$ as an ideal in the Dirac–Clifford algebra. Hermitian and symplectic forms invariant under U (2,2) and Sp (4,$\mathbb{C}$) are defined on $\mathscr{T}$. In Sec. VI we define explicitly the isomorphism between a complex Dirac–Clifford algebra and a tensor product of twistor spaces and restate our results from Sec. IV in terms of symmetric and Hermitian tensors. In Appendix A we list all formulas from a Clifford algebra calculus necessary to obtain our results whereas Appendix B explains the notation used in this paper.

II. MOTIVATION AND SHORT SURVEY

Many geometric ideas can be expressed algebraically using the well-known algebra introduced by Clifford one hundred years ago. The advantages of geometry based calculus from the point of view of simplicity and potentially rich physical applications have been stressed by many authors. In particular, in Refs. 11 and 12 such geometric formalism is fully developed.

In the present paper we are interested in exploring the possibility of using exclusively Clifford algebra calculus for the old idea[13] (see also Refs. 14–16) to describe the physical phenomena in spinor space rather than in real affine Minkowski space M^a.

A Clifford algebra gives very natural and clear geometric interpretation to constructions involving affine Minkowski space and twistor spaces. Perhaps the crucial point for all results presented here is the link between the Witt decomposition of tangent Minkowski space, basic in certain approach to spinors and twistors[17] (cf. Refs. 3, 4, 6, 10, and 18) and the Poincaré group as the structure group (cf. Refs. 19–21) of the fiber bundle over a real Riemannian space-time M.

We express all notions in abstract (coordinate-free) Clifford algebraic language and show how all conclusions result from the Clifford associative multiplication and regarding each point of M as having attached to it a 16-dimensional real Dirac–Clifford algebra $\mathscr{D}$. Thus we have a bundle of Clifford algebras over M, which is natural from the viewpoint of geometry and will provide a natural physical interpretation for our results. When we restrict ourselves to the tangent bundle (instead of the Clifford bundle mentioned above) with the Minkowski bilinear form η_p, [signature (1,3)] defined for each tangent space T_pM, we have in mind the geometry of tangent Minkowski space. It is the Clifford algebra $\mathscr{D}$ at $p \in M$ in terms of which the latter can be expressed in full. Having defined a Clifford algebra it is essentially the same as having a symmetric nondegenerate form (cf. Refs 1, 22, 23).

We believe that the above point of view would be also convenient while investigating a nonrelativistic Galilei–Newton space-time N for which T_pN has a singular metric (0 1 1 1). It would be interesting to study the structure of Clif-

[a)]On leave from Wrocław University, Wrocław, Poland.

ford algebra and its representations for this Galilei form.[5] However we are not going to consider this problem here.

In the present paper a twistor space is defined as cetrain ideal in Dirac–Clifford algebra (cf. Ref. 24 and more recently Refs. 10 and 22). The isomorphism between the latter and the tensor product of the twistor space with itself makes it possible to express the coordinates of any algebra element in terms of its tensor coordinates (see Sec. VI).

Three projections from tensor products of twistor spaces on complex affine Minkowski space are derived for symmetric, antisymmetric, and Hermitian tensors respectively. It is proved that the conformal group, while acting linearly on the Dirac–Clifford algebra induces, via twistors and the above isomorphism, correct conformal transformations in the complex Minkowski space. It turns out then that the tensor product of two twistor spaces is the smallest complex manifold[25] from which one can project.

In the case of decomposable tensors the above projections define the same complex vector (this was shown in Ref. 26). Projection involving the antisymmetric forms was found by Penrose and discussed in some detail in Refs. 26–28.

We will now make a few remarks about possible deeper motivation for the present considerations. One is the attempt[13] (cf. also Ref. 16 and the references therein) to describe physical phenomena using fields over the twistor numerical space. This space has more dimensions than Minkowski space which allows more conservation laws, for example, the conservation law of charge, to be given a geometric interpretation. On the other hand, it makes it possible to describe the internal degrees of freedom of elementary particles with the help of the parameters of twistor spaces tensor product, a typical fiber over real Minkowski space being considered as a base. An interesting point is that a Minkowski vector needs to be constructed from at least two twistors (tensor fields of rank two), that may provide a geometrical interpretation for quark confinement.

One can also view twistor space as the classical relativistic phase space (cf. Refs. 9, 28, and 29). The present paper could be one of the steps in the above program to treat the twistor coordinates as the base-space coordinates for the corresponding vector bundles over twistors.

The connection between M^a and twistor spaces (which are the Cartesian products of Lorentz spinor spaces) has its own history. Several different solutions were suggested to the problem of accomodation of translations in a spinor space so that the space-time Poincaré group (or its extension to the conformal group) could be retained. In 1936 Dirac noticed the possibility of representing translation generators by $\partial_\mu + \lambda\frac{1}{2}(1 + i\gamma_5)\gamma_\mu$ (i.e., with an intrinsic part), contrary to the usual ∂_μ for ordinary Dirac bispinors. This point of view was further investigated in Ref. 30, where a close relationship was shown between the complex Dirac–Clifford algebra and conformal group (in this respect cf. Refs. 18, 31, and 32). Thus Dirac[33] (cf. also Ref. 30) had already distinguished the ordinary Dirac bispinors from fields (rediscovered later in Ref. 8 and there called twistors), belonging to the representation with intrinsic translations. Evidently Penrose first found one of the connections between twistors and M^a (cf. Refs. 9, 14, and 28 and the references therein) by means of the complex Plücker coordinates of bitwistor [a skewsymmetric tensor of the conformal group SU(2,2)].

In the present paper, however, we would like to emphasize the possible geometric interpretation of twistors through the real Clifford algebra for de Sitter space rather than the holomorphic aspect of the complex manifolds advocated by Penrose (cf. Ref. 7, 14, and 28.)

A quite different approach to the connection of twistors with space-time was suggested in Ref. 34 (see also Ref. 35) where real space-time coordinates were expressed as bilinear functions of twistors. In particular, we will show in the last section that such a representation is in contradiction to the desired form for linear transformations of the twistor space. We also show that the corrected representation of this type coincides with Penrose's formula in the case of simple Clifford numbers (decomposable tensors).

Different solutions were found in Refs. 36 and 37, where the author abandons the assumption of linear representation of the conformal group in spinor space (this representation is obviously linear when restricted to the Lorentz subgroup). In this case, the following representation of the space-time coordinates, $x^\mu = \bar{\eta}\gamma^\mu\xi$ can be retained.

One may ask, of course, if the formulation in terms of Clifford algebra is only an elegant restatement of known ideas and how much insight and new substance can be obtained using this approach. New light has recently been thrown on this question in Ref. 38, where a very elegant and transparent treatment of the Riemann curvature tensor and its properties was given. We also think that it would be a great deal more difficult to derive these projections without using the concepts of Clifford algebra.

III. PRINCIPAL CONFORMAL BUNDLE

No global properties of a space-time manifold M are investigated in the present paper. $\mathscr{D}$ always denotes a real Dirac–Clifford algebra of a tangent Minkowski space T_pM for some fixed point $p \in M$.

Recall some of the most important notions concerning a general Clifford algebra $C(Q)$ of a linear, finite-dimensional, real vector space V, $\dim V = s$, with a bilinear form $B: V \times V \to R$, and its associated quadratic, nondegenerate form Q. For details see Refs. 1, 2, 5, 10, 22, 23, and 39. The elements of $C(Q)$, denoted here by $m, n, \cdots$, are called "Clifford numbers." Any Clifford algebra $C(Q)$ can be decomposed into a direct sum of k-vector spaces $C_k = \overset{k}{\wedge} V$, $k = 0, 1, 2, \cdots, s$, $C(Q) = \sum_{k=0}^{s} \oplus C_k$, where $\wedge$ denotes the exterior vector space product. Evidently, $C_1 \equiv V$, $C_0 \equiv R$ (real numbers), and $\dim C_k = \binom{s}{k}$. Thus any element $m \in C(Q)$ can be written as $m = \sum_{k=0}^{s} \langle m \rangle_k, \langle m \rangle_k \in C_k, k = 0, 1, 2, \cdots, s$, where $\langle m \rangle_k$ is called a k-vector part of m. At this point we adopt notation from Ref. 12.

The Clifford algebra $C(Q)$ is a Z_2 graded alegebra. This important gradation is due to the linear automorphism α, called also a principal automorphism of $C(Q)$, which is just the reversal of space and time tangent velocities, i.e., PT transformation. Let $m, n \in C(Q)$; then

(1) $\alpha(mn) = \alpha(m)\alpha(n)$,

(2) $\alpha(\langle m\rangle_k) = (-1)^k \langle m\rangle_k \in C_k$, $k = 0, 1, 2, \cdots, s$.

With respect to α, for distinguished V as the subspace of $C(Q)$, there is the following eigenspace decomposition (for a matter of convenience, when Q is not specified we put $C(Q) \equiv C$): $C = C^+ \oplus C^-$ where C^+ = eigenspace for $+1$ $(\alpha|_{C^+} = id)$ and C^- = eigenspace for -1 $(\alpha|_{C^-} = -id)$. C^+ is also a subalgebra of C.

Let β denote the unique linear antiautomorphism on C, an identity when restricted to V. β is commonly called a principal antiautomorphism of $C(Q)$. Explicitly we have

(i) $\beta(mn) = \beta(n)\beta(m)$ for any $m, n, \in C(Q)$,

(ii) $\beta(\langle m\rangle_k) = (-1)^{k(k-1)/2}\langle m\rangle_k$ for $\langle m\rangle_k \in C_k$, $k = 0, 1, 2, \cdots, s$.

The antiautomorphism β allows us to introduce the spinor norm $N(m)$: $= \beta(m)m$ for an arbitrary Clifford number $m \in C(Q)$. Let $K = \{m \in C, mV = Vm\}$. Then the spinor norm of any element from K belongs to the center Z of $C(Q)$ (see also Sec. III).

It is known that if $\dim V = s = p + q$ is even, i.e., Q has signature (p,q), then $Z = R$ and for s odd, $Z = R \oplus C_s$, where C_s is a one-dimenstional space of pseudoscalars (cf. Refs. 1 and 10).

Let $*(p,q)$ be a unit pseudoscalar and

$$*(p,q)^2 = (-1)^{s(s-1)/2}\det Q, \ \det Q = (-1)^q. \qquad (3.1)$$

Then any element of center Z for s odd can be represented in the form $Z \ni a + b*$ for $a, b \in R$. Let Z^* be the set of all invertible elements of Z. Then the Clifford group G of Q is just the subset of K such that $N(G) \in Z^*$ and $N(g_1g_2) = N(g_1)N(g_2)$ for $g_1, g_2 \in G$. The homomorphism $A: G \to O(p,q)$ onto an orthogonal group of quadratic form Q with signature (p,q) is then given by

$$A(g)v = gvg^{-1} \text{ for } v \in V, g \in G \text{ where } \ker A \simeq R^*.$$

One also defines Pin (Q) as the set of elements of G for which the spinor norm is ± 1 valued. The orientation-preserving (pseudoscalar $*$) special Clifford group is $G^+ := G \cap C^+$ and $A(G^+) = SO(p,q)$ for which $N(G^+) = R^*$. The spin group Spin(Q) is defined as the even part of Pin(Q), i.e., Spin$(Q) \equiv [\text{Pin}(Q)]^+ \subset G^+$.

Any 2^s-dimensional associative Clifford algebra is at the same time a Lie algebra with Lie multiplication $[m,n] = mn - nm$ for any Clifford numbers m, n. The Lie structure is uniquely induced by the fundamental Clifford geometric multiplication.

Using β it is easy to see that the space of 2-vectors C_2 forms a Lie algebra $o(p,q)$ of dimension $n(n-1)/2$, the Lie subalgebra of C. Let us decompose $\mathscr{D}$ as follows:

$$\mathscr{D}^+ = S \oplus T \oplus P$$

and

$$\mathscr{D}^- = V \oplus A, \quad \mathscr{D} = \mathscr{D}^+ \oplus \mathscr{D}^-,$$

where from now we adopt new notation uniquely for $\mathscr{D}_k$ $(k = 0, 1, 2, 3, 4)$, namely, $\mathscr{D}_0 = S$ (scalars), $\mathscr{D}_1 = V$ (vectors), $\mathscr{D}_2 = T$ (tensors), $\mathscr{D}_3 = A$ (axial vectors), $\mathscr{D}_4 = P$ (pseudoscalars), while the respective k-vector parts of $m \in \mathscr{D}$ we denote by s, v, t, a, and p. Then the Lie structure of $\mathscr{D}/Z$ is given by

$$[P,A] \subset V, \ [T,\mathscr{D}_k] \subset \mathscr{D}_k, \ [V,V] \subset T,$$
$$[A,A] \subset T, \ [V,A] \subset P, \ [P,V] \subset A, \qquad (3.2)$$

where $\mathscr{D}_k = \overset{k}{\wedge} V$. Here V is identified with T_pM. Equation (3.2) shows that the algebra $\mathscr{D}$ can be also considered as a Lie algebra of the group of inner automorphisms of $\mathscr{D}$.

The following isomorphisms take place:

$$\text{Lie}\{\mathscr{D}/Z\} \simeq \text{so}(1,5) \simeq \text{su}^*(4). \qquad (3.3)$$

The group SU*(4) will be introduced in the next section as the group of linear mappings of a twistor space $\mathscr{T}$ commuting with a particular semilinear map $\mathscr{A}: \mathscr{T} \to \mathscr{T}$ (cf. Ref. 40).

We would like to enlarge the Lorentz algebra of tensors to that of Poincaré group. If one looks for the translation generator as a linear combination of vector and axial vectors there is no solution for the hyperbolic form η unless one complexifies $\mathscr{D}$. Alternatively, one can think in terms of the mapping so(1,5)→so(2,4) via complexification su*(4)→su(2,2) (see next sections). However, formal algebraic complexification without an underlying geometric basis is contrary to the philosophy of the geometrical calculus (compare, for example, Ref. 12). Therefore one would like to have a geometrical interpretation of the imaginary scalar by means of the new "hidden" dimension. This is due to the role of i as the unit pseudoscalar of Clifford algebra Σ for a real, five-dimensional de Sitter space (the center Z is then two-dimensional). From (3.1) for $s = 4$ or 5 we have $*(Q)^2 = \det Q$ and one demands that $\det Q = -1$, which means that the desired de Sitter-Clifford algebra Σ is simple. Therefore, there are three possibilities: $V_{4,1}$—de Sitter space with bilinear form of signature $(+ + + + -)$, $V_{2,3}$ and $V_{0,5}$. However, one would like to consider the Dirac–Clifford algebra $\mathscr{D}$ as an even subalgebra of Σ. This is the case if $V_s \equiv V_{4,1}$. Complex conjugation now in $\mathscr{D}^c$ has a clear , geometric interpretation as the principal automorphism α_s in Σ. That is why algebra $\mathscr{D}$, stable under complex conjugation, can be identified with the even subalgebra of Σ

$$\Sigma^+ \equiv \mathscr{D} = S \oplus T_s \oplus A_s \text{ and } \Sigma^- = V_s \oplus \tilde{T}_s \oplus P_s,$$

where $\Sigma_0 = S$ (scalars), $\Sigma_1 = V_s$ (de Sitter vector space), $\Sigma_2 = T_s$ (tensors), $\Sigma_3 = \tilde{T}_s$ (dual tensors), $\Sigma_4 = A_s$ (axial vectors), and $\Sigma_5 = P_s = i \otimes R$ (pseudoscalars). Moreover,

$$A_s = i \otimes V_s = A + P \text{ or } V + P,$$
$$T_s = i \otimes \tilde{T}_s = V + T \text{ or } A + T,$$

and i is the unit pseudoscalar for the simple algebra Σ, hence $i \in Z_s$ (center of Σ).

Let $\{e_\mu\}$ be an orthonormal basis for T_pM, i.e., $e_\mu \cdot e_\nu = \eta_{\mu\nu}$, $\mu, \nu = 0, 1, 2, 3$ and $\{f_a\}$, $a, b = 1,2,3,4,5$, $f_a \cdot f_b = g_{ab} = \{-\text{diag}\,\eta_{\mu\nu}, +1\} = (4,1)$ can be a corresponding basis for V_s. Here $\{f_a\} = \{i*e_\mu, i*\}$.

For Lorentz signature (3,1) one gets a de Sitter space with signature (2,3). Σ for $g_{ab} = \text{diag}(+ + + + -)$ and Σ for $g_{ab} = \text{diag}(+ + - - -)$ are isomorphic as universal Clifford algebras (cf. Ref. 23). A possible basis for V_s with $g_{ab} = \text{diag}(+ + - - -)$ is $\{f_5, e_\mu\}$. Thus $i \equiv *(4,1)$.

Now let β_s be a principal antiautomorphism in Σ and $\gamma_s \equiv \alpha_s \circ \beta_s$. Acting with γ_s on Σ one gets the following decomposition:

$$\Sigma = \Sigma^c \oplus \Sigma^{ac},$$

where

$$\gamma_s(\Sigma^c) = -\Sigma^c, \gamma_s(\Sigma^{ac}) = \Sigma^{ac}$$

and

$$\Sigma^c = V_s \oplus T_s \oplus P_s,$$
$$\Sigma^{ac} = A_s \oplus \tilde{T}_s.$$

It can be shown that Σ^c/P_s is a 15-dimensional, real Lie algebra of the conformal group $o(2,4) \simeq su(2,2)$. In other words, $\{V \oplus T \oplus iA \oplus iP\}$ and $\{V \longleftrightarrow A\}$ provide two such algebras; the latter occurs due to the duality in Σ: $A_s i = iA_s = V_s$. $V \oplus T$ and $A \oplus T$ are both de Sitter–Lie algebras. For the physical meaning of 15–dimensional, real Dirac–Lie algebra (3.2) see Ref. 18.

Recall that, due to the isomorphism between Σ and $\mathscr{D}^c$, α_s acts on $\mathscr{D}^c$ as complex conjugation while $\beta_s = \beta \circ \alpha$, β and α being principal automorphism and antiautomorphism in $\mathscr{D}^c$.

A conformal group can be introduced by means of exponentiation of its Lie algebra. The connected component of identity of conformal group $\mathscr{K} = \mathscr{C}(1,3) \equiv \{g = \exp(k), k \in \Sigma^c/P_s\}$. Also $\alpha_s(g) = \exp(\alpha_s(k))$, $\beta_s(g) = \exp(\beta_s(k))$, where $g = \exp k$, $k \in \Sigma^c/P_s$. Since $\gamma_s(g) = \exp(-v_s - t_s)$, $v_s \in V_s$, and $t_s \in T_s$, then $\gamma_s(g)g = 1$ for any element g of $\mathscr{K}$. The generators of $\mathscr{K}$ are

(1) *de Sitter translation*: $S_\mu = e_\mu \frac{1}{2}(a + bi*)$, $a = 1 + 1/R^2$, $b = 1 - 1/R^2$. If $a = b = 1$, (one can obtain this by means of the contraction $R \to \infty$) then $S_\mu = \frac{1}{2}e_\mu (1 + i*) \equiv P_\mu$ and one gets a generator of the Poincaré translation. If $-a = b = 1$ then $S_\mu = \frac{1}{2}e_\mu(-1 + i*) \equiv K_\mu$, a generator of the special conformal transformation.

(2) *Lorentz rotation*:
$M_{\mu\nu} = \frac{1}{2}e_\mu \wedge e_\nu = \frac{1}{4}[e_\mu, e_\nu]$.

(3) *Dilatation*:

$$D = -i*.$$

Let $s = \epsilon^\mu S_\mu$, $t = \tau^\mu P_\mu$, $c = \chi^\mu K_\mu$, $d = p'D$ and $\lambda = \lambda^{\mu\nu} M_{\mu\nu}$, $\mu, \nu = 0, 1, 2, 3$ while $\epsilon^\mu, \tau^\mu, \chi^\mu, p', \lambda^{\mu\nu}$ are real numbers. Then after integration one gets the corresponding group elements

(1) *de Sitter transformation*: $g_s = \exp s = \cosh \omega + s \cdot \sinh \omega/\omega$, $\omega \neq 0$, where $\omega^2 = \frac{1}{4}(\epsilon^\mu \epsilon_\mu)(a^2 - b^2)$. When $R \to \infty$ $(\omega \to 0)$, $g_s \to g_p = 1 + s|_{a=b=1} = 1 + t$ (Poincaré translation), and for

(2) *special conformal transformation*: $g_c = \exp c|_{-a=b=1} = 1 + c$.

(3) *dilatation and Lorentz rotation*: $g_d = \exp d = \cosh \zeta + (\sinh \zeta) i*$, $\zeta = -p'/2$ and $g_l = \exp \lambda$, $\lambda = \lambda^{\mu\nu} M_{\mu\nu}$ respectively.

Therefore the enlargement of the Lorentz algebra of tensors of T_pM to the Poincaré algebra naturally results in geometric complexification of the Dirac–Clifford algebra. $\mathscr{D}^c$ for T_pM is isomorphic to the real de Sitter–Clifford algebra Σ for V_s.

IV. PROJECTIONS FROM $\mathscr{D}^c$ ON COMPLEX AFFINE MINKOWSKI SPACE-TIME $(M^a)^c$

We begin this section with some general remarks concerning the subgroups of C^*, the set of all invertible elements of a Clifford algebra $C(Q) \equiv C$ and their different linear representations on C as a natural representation space. We want to emphasize the fact that for a fixed group $G \subset C^*$ one can define its linear representation $\hat{\gamma}$ on C by means of any (cyclic, length 2) antiautomorphism γ of C, invariant under group action in the sense that the following diagram commutes:

$$\begin{array}{ccc} G \times C \ni (g,m) & \xrightarrow{\hat{\gamma}} \hat{\gamma}_g(m) \equiv gm\gamma(g) \in C \\ \downarrow \gamma & \downarrow \gamma \\ G \times C \ni (g,\gamma(m)) & \rightarrow \hat{\gamma}_g(\gamma(m)) \equiv g\gamma(m)\gamma(g) \in C. \end{array} \quad (4.1)$$

In order to enumerate all antiautomorphisms of C, called here γ_i, $i = 1,2,\cdots$, it is sufficient to consider the group of automorphisms, called α_j, $j = 1, 2, \cdots$ of C, Aut C, and one fixed antiautomorphism, say $\bar{\beta}$ of C, for any $\gamma_i = \alpha_i \circ \bar{\beta}$ for some $\alpha_i \in \text{Aut } C$.

In the forthcoming we restrict ourselves to the subgroup $\text{Aut}_1 C$ of Aut C where $\text{Aut}_1 C = \{\alpha_i \in \text{Aut } C \mid \alpha_i^2 = 1\}$. Notice that $\text{Aut}_1 \mathscr{D}^c = \{1, \alpha, \alpha_s, \alpha \circ \alpha_s\}$ where α is a principal automorphism of $\mathscr{D}^c$ while α_s denotes complex conjugation in $\mathscr{D}^c$. Since $\gamma(mn) = \gamma(n)\gamma(m)$ for any $m,n \in C$, it can be easily shown that $\hat{\gamma}_{g_1} \circ \hat{\gamma}_{g_2} = \hat{\gamma}_{g_1 g_2}$ and $\hat{\gamma}: G \xrightarrow[\text{onto}]{} \text{Aut } \gamma$ is a homomorphism where Aut γ denotes a group of linear bijections $\gamma_g: C \to C$ preserving γ.

Diagram (4.1) gives rise, in general, to a so-called twisted adjoint representation (cf. Ref. 1 and 10) $\hat{\gamma}_{\alpha(g)}$ where $\alpha \in \text{Aut}_1 C$. For let $\bar{\beta}(g) = g^{-1}$. Then for a given γ we have $\alpha \in \text{Aut}_1 C$ such that $\gamma(g) = \alpha(g^{-1})$ and

$$\hat{\gamma}_g(m) = gm\alpha(g^{-1}),$$
$$\hat{\gamma}_{\alpha(g)}(m) = \alpha(g)mg^{-1}. \quad (4.2)$$

There are three independent antiautomorphisms in $\mathscr{D}^c$: β, $\alpha \circ \beta$, $\alpha_s \circ \beta$ and their composition $\alpha \circ \alpha_s \circ \beta$. They provide four different realizations (representations) of any group $G \subset C^*$ on C, each of which preserves one antiautomorphism [in the sense of the diagram (4.1)].

We want to point out here the important fact that the condition that a representation $\hat{\gamma}$ of G commute with γ does not lead to any restriction on G itself. However, if one requires that $\hat{\gamma}_{i_g} \circ \gamma_i = \gamma_i \circ \hat{\gamma}_{i_g}$ holds for every $g \in G$ and $i = 1, 2$, then the group elements need to satisfy the relation $\gamma_1(g) = \gamma_2(g)$. In the very same way, $\gamma_1(g) = \gamma_2(g) = \cdots = \gamma_k(g)$ for some k if the antiautomorphisms γ_1 through γ_k are to be preserved by a group representation.

Therefore, for a given group $G \subset C^*$, we have two types of group homomorphisms, ϕ and $\hat{\gamma}$, where

$$C^* \supset G \ni g \xrightarrow[\text{onto}]{\phi} \phi_g \in \text{In } C: \phi_g(m) = gmg^{-1},$$

$$C^* \supset G \ni g \xrightarrow[\text{onto}]{\hat{\gamma}} \hat{\gamma}_g \in \text{Aut } \gamma: \hat{\gamma}_g(m) = gm\gamma(g),$$

and In C denotes a set of inner automorphisms of C. These two homomorphisms are of great importance for all future applications. Here $\hat{\gamma}$ is called the action of $G \subset C^*$ on C.

For a given γ one can define a corresponding norm N_γ which is a map $N_\gamma: C \ni m \to N_\gamma(m) \equiv \gamma(m)m \in Z$, where Z is the center of C. We have the following diagram (for every $m \in C$ and $g \in G$):

$$\begin{array}{ccc} C \ni m & \xrightarrow{N_\gamma} \gamma(m) = m \in C \\ \downarrow g & N_\gamma & \downarrow g \\ C \ni gm\gamma(g) & \longrightarrow g\gamma(m)\gamma(g)gm\gamma(g) = g\gamma(m)m\gamma(g) \end{array} \quad (4.3)$$

If one requires the above diagram to commute then a subgroup G_γ of G for some antiautomorphism γ can be defined as

$$G_\gamma \equiv \{ g \in G, \hat{\gamma}_g(m) = \phi_g(m) \quad \text{for every } m \in C \}. \tag{4.4}$$

Notice that (4.4) is equivalent to $N_\gamma(g) = 1$ and $N_\gamma \circ \hat{\gamma}_g = \hat{\gamma}_g \circ N_\gamma$ for every $g \in G_\gamma$. One concludes that any group G_{γ_i} preserves both γ_i and γ_j, for some i and j, with respect to the homomorphisms $\hat{\gamma}_j$ and ϕ, respectively. According to the definitions of Pin(Q) and Spin(Q) (see Sec. III) one can recognize that for $\mathscr{D}^c$: Pin$(\eta) = G_\beta$ and Spin$(\eta) = G_\beta \cap G_{\alpha\circ\beta}$. Spin$(\eta)$ is a twofold cover of the Lorentz group SO(1,3) which acts on $\mathscr{D}^c$ via ϕ. Therefore, one can also characterize the Lorentz group as $G_\beta \cap G_{\beta_s} \cap G_{\alpha_s\circ\beta}, G_\beta \cap G_{\beta_s} \cap G_{\alpha_s\circ\beta_s}$, or $G_{\beta_s} \cap G_{\alpha_s\circ\beta}$. Linear representations $\hat{\beta}, \hat{\beta}_s, \widehat{\alpha_s\circ\beta}$ *and* $\widehat{\alpha_s\circ\beta_s}$ of the conformal group $\mathscr{K}$ on $\mathscr{D}^c$ are the only ones which become the ϕ action when restricted to the Lorentz subgroup. This is because $g^{-1} = \beta(g) = \beta_s(g) = (\beta\circ\alpha_s)(g) = (\alpha_s \circ \beta_s)(g)$ for any element g of the Lorentz group.

Given now any representation $\hat{\gamma}$ of $\mathscr{K}$ on $\mathscr{D}^c$ one may ask whether or not there exists a projection ρ_γ from $\mathscr{D}^c$ on $(M^a)^c$, defined by the following commutative by assumption diagram:

$$\begin{array}{ccc} \mathscr{K} \ni g : \mathscr{D}^c & \xrightarrow{\hat{\gamma}_g} & \mathscr{D}^c \\ \downarrow \rho_\gamma & & \downarrow \rho_\gamma \\ \mathscr{K} \ni g : (M^a)^c & \xrightarrow[g]{} & (M^a)^c \end{array} \tag{4.5}$$

In what follows, our aim is to construct ρ_β and $\rho_{\alpha_s\circ\beta_s}$.

Let α, α_s, β, and β_s be as before. Then $m \in \mathscr{D}^c$ is said to be

(i) symmetric if $\beta(m) = -m$,

(ii) antisymmetric if $\beta(m) = m$,

(iii) Hermitian if $(\beta_s \circ \alpha_s)(m) = m$,

(iv) antiHermitian if $(\beta_s \circ \alpha_s)(m) = -m$. (4.6)

The definitions (i) and (ii) express the fact that $\beta(f) = -f$, which means that f, the generator of $\mathscr{T}$, is a bivector in $\mathscr{D}^c$. The role played by f in establishing the isomorphism between $\mathscr{D}^c$ and $\mathscr{T} \otimes \mathscr{T}$ will become clear in Sec. V.

A.

Let us consider first the action $\mathscr{K}$ on $\mathscr{D}^c$ preserving decomposition of any $m \in \mathscr{D}^c$ into its symmetric and antisymmetric parts:

$$\mathscr{K} \ni g : \mathscr{D}^c \ni m \to g m \beta(g) \equiv m' \in \mathscr{D}^c. \tag{4.7}$$

If $m = \pm\beta(m)$ then $m' = \pm\beta(m')$. Moreover, if m is symmetric then $m = a + t$, while m antisymmetric means that $m = s + v + p$, where $s \in S$, $v \in V$, $t \in T$, $a \in A$, and $p \in P$.

The following is divided into two steps:

(1) We find the transformation properties of all k-vector parts of $m \in \mathscr{D}^c$ under $\mathscr{K}$, i.e.,

$$\mathscr{D}^c_k \ni \langle m \rangle_k \xrightarrow{g} \langle g m \beta(g) \rangle_k \in \mathscr{D}_k{}^c \tag{4.8}$$

for $k = 0,1,2,3,4$ and $g \in \mathscr{K}$.

(2) Next we look for all combinations of $\langle m \rangle_k$'s which would transform like a vector under $\mathscr{K}$. They will give us all projections from $\mathscr{D}^c$ on $(M^a)^c$.

Let us proceed with step (1). We first list the elements of $\mathscr{K}$ in terms of the Clifford numbers:

(i) *Poincaré translation* g_p:

$$g_p = \exp[\tfrac{1}{2}\tau(1+i*)] = 1 + \tfrac{1}{2}\tau(1+i*), \quad \tau = \tau^\mu e_\mu,$$

(ii) *Special conformal transformation* g_c:

$$g_c = \exp[\tfrac{1}{2}\kappa(1-i*)] = 1 + \tfrac{1}{2}\kappa(1-i*), \tag{4.9}$$

$$\kappa = \kappa^\nu i e_\nu *,$$

(iii) *de Sitter transformation* g_s:

$$g_s = \exp s = \cosh\omega + s(\sinh\omega)/\omega, \quad \omega \neq 0,$$

(iv) *dilatation* g_d:

$$g_d = \exp(p'D) = \cosh\zeta + (i*)\sinh\zeta, \quad \zeta = p'/2,$$

(v) *Lorentz rotation* g_l:

$$g_l = \exp\lambda, \quad \lambda = \lambda^{\mu\nu} M_{\mu\nu}.$$

Combining (4.8) and (4.9) we get the full list of transformations of $\langle m \rangle_k$ (with the help of formulas given in Appendix A):

(i) *Poincaré translation* g_p:

$$A \ni a \xrightarrow{g_p} a + 2\tau \wedge t_1 \in A,$$

$$T_+ \ni t_1 \to t_1 \in T_+,$$

$$T_- \ni t_2 \to t_2 + (\tau\cdot a)(1-i*) + \tau t_1 \tau \in T_-,$$

$$S \ni s \to s + \tau\cdot v + \tfrac{1}{2}s\tau^2 + \tfrac{1}{2}p\tau^2 i* \in S,$$

$$V \ni v \to v + \tau(s + pi*) \in V,$$

$$P \ni p \to p - (\tau\cdot v)i* - \tfrac{1}{2}\tau^2 s i* - \tfrac{1}{2}\tau^2 p \in P, \tag{4.10}$$

where $T = T_+ \oplus T_-, T_\pm = \{ t \in T, i*t = \pm t \}$ for $t = t_1 + t_2$.

The following are the Poincaré invariants:

$$c_1(\infty) \equiv s + ip*, \; t_1 = \tfrac{1}{2}t(1+i*), \; c_2 \equiv I_2/(s+pi*)^2,$$

where

$$I_1 \equiv \left\{ \frac{s - ip*}{s + ip*} - \left(\frac{v}{s + ip*} \right)^2 \right\} (s + ip*)^2 = s^2 - v^2 - p^2$$

is an invariant of the conformal group. We define $(v - ai*)^{-1} \equiv (v - ai*)/(v - ai*)^2$, provided $(v - ai*)^2 \neq 0$, and more generally $(\langle m \rangle_k)^{-1} \equiv \langle m \rangle_k / (\langle m \rangle_k)^2$ provided $(\langle m \rangle_k)^2 \neq 0$, $\langle m \rangle_k \in \mathscr{D}^c{}_k$.

(ii) *Special conformal transformation* g_c:

$$V \ni v \xrightarrow{g_c} v + (pi* - s)\tilde{\kappa} \in V,$$

$$S \ni s \to s - \tilde{\kappa}\cdot v + \tfrac{1}{2}\tilde{\kappa}^2(s - pi*) \in S,$$

$$P \ni p \to p - (\tilde{\kappa}\cdot v)i* + \tfrac{1}{2}\tilde{\kappa}^2(si* - p) \in P,$$

$$T \ni t \to t - (\tilde{\kappa}\cdot a)(1+i*) + \tfrac{1}{2}\tilde{\kappa}t\tilde{\kappa}(1+i*) \in T,$$

$$A \ni a \to a + (\tilde{\kappa}\cdot t)i* - \tilde{\kappa} \wedge t \in A,$$

$$\kappa \equiv \tilde{\kappa}i*, \langle\tilde{\kappa}\rangle_1 = \tilde{\kappa}; \tilde{\kappa}^2 \equiv \tilde{\kappa}^\nu\tilde{\kappa}_\nu. \tag{4.11}$$

It follows that $t_2 \xrightarrow{g_c} t_2$, the antidual part of tensor $t \in T$ is invariant under g_c action. For convenience we put

$\cosh\omega \equiv x$, $a \sinh\omega/\omega \equiv y$, $b \sinh\omega/\omega \equiv z$;

then $1 - x^2 = \frac{1}{4}\epsilon^2(z^2 - y^2)$. Hence

(iii) *de Sitter translation* g_s:

$$
\begin{aligned}
&S \ni s \xrightarrow{g_s} s\left[x^2 + \tfrac{1}{4}(\epsilon)^2(y^2 + z^2)\right] + y\left[x(\epsilon\cdot v) + \tfrac{1}{2}(\epsilon)^2 zpi*\right] \in S,\\
&P \ni p \to p\left[x^2 - \tfrac{1}{4}(\epsilon)^2(y^2 + z^2)\right] - z\left[x(\epsilon\cdot v) + \tfrac{1}{2}(\epsilon)^2 ys\right] i* \in P,\\
&V \ni v \to v + \epsilon\left[x(ys + zpi*) + \tfrac{1}{2}(\epsilon\cdot v)(y^2 - z^2)\right] \in V,\\
&A \ni a \to x^2 a + xz(\epsilon\cdot t)i* + xy(\epsilon\wedge t) + \tfrac{1}{4}(y^2 - z^2)\epsilon a\epsilon \in A,\\
&T_+ \ni t_1 \to x^2 t_1 + x(\epsilon\cdot a)\tfrac{1}{2}(y - z)(1 + i*)\\
&\quad + \tfrac{1}{8}\epsilon t\epsilon(y + zi*)(y - z)(1 + i*) \in T_+,\\
&T_- \ni t_2 \to x^2 t_2 + x(\epsilon\cdot a)\tfrac{1}{2}(y + z)(1 - i*)\\
&\quad + \tfrac{1}{8}\epsilon t\epsilon(y - zi*)(y + z)(1 - i*) \in T_-.
\end{aligned}
\qquad (4.12)
$$

There is the de Sitter invariant $c_1(R) \equiv (s + pi*) - (s - ip*)/R^2$ which becomes the Poincaré invariant $c_1(\infty)$ at the limit $R^2 \to \infty$. One can notice that if $x = y = z = 1$ ($\omega \to 0$ or $R \to \infty$) then (4.12) gives (4.10). Moreover, if $x = -y = z = 1$ then one gets (4.11).

(iv) *dilatation* g_d:

$$
\begin{aligned}
&A \ni a \xrightarrow{g_d} a \in A \text{ is invariant},\\
&T \ni t \to t(\mu^2 + \nu^2 + 2\mu\nu i*) \in T,\\
&S \ni s \to (\mu^2 + \nu^2)s + 2\mu\nu pi* \in S,\\
&V \ni v \to v \in V \text{ is invariant},\\
&P \ni p \to (\mu^2 + \nu^2)p + 2\mu\nu si* \in P,\\
&\cosh\zeta \equiv \mu,\ \sinh\zeta \equiv \nu.
\end{aligned}
\qquad (4.13)
$$

(v) *Lorentz rotation* g_1:

$$
\begin{aligned}
&S \ni s \xrightarrow{g_1} s \in S,\\
&P \ni p \to p \in P,\\
&V \ni v \to \langle(\exp\lambda)(v + a)\exp(-\lambda)\rangle_1 \in V,\\
&A \ni a \to \langle(\exp\lambda)(v + a)\exp(-\lambda)\rangle_3 \in A,\\
&T \ni t \to \langle(\exp\lambda)t(\exp(-\lambda))\rangle_2 \in T.
\end{aligned}
\qquad (4.14)
$$

Let us define two Clifford numbers z_1 and z_2 by

$$z_1 \equiv v(s + ip*)^{-1}, \qquad (4.15a)$$

$$z_2 \equiv [(1 + i*)t^{-1}]\cdot a. \qquad (4.15b)$$

We now seek the transformation laws of z_1 and z_2 under the conformal group $\mathscr{K}$ making extensive use of (4.10)–(4.14). Let us see first how z_1 transforms under de Sitter translation g_s[41]

$$z_1 \xrightarrow{g_s} \frac{(z + y)\{z_1 + [x(ys + zpi*)/(s - pi*)[c_2 + (z_1)^2] + \frac{1}{2}(z_1\cdot\epsilon)(y^2 - z^2)]\epsilon\}}{(z - y)[c_2 + (z_1)^2] + 2x^2(ys + zip*)/(s - pi*)[c_2 + (z_1)^2] + x(y^2 - z^2)(\epsilon\cdot z_1)}, \qquad (4.16)$$

where

$$c_2 = I_1/(s + pi*)^2 = (s - pi*)/(s + pi*) - (z_1)^2$$

and

$$(z_1)^2 \equiv z_1\cdot z_1 = (z_1)^\mu (z_1)_\mu .$$

When $c_2 = 0$ (4.16) gives

$$z_1 \underset{c_2 = 0}{\xrightarrow{g_s}} \frac{(z + y)\{z_1 + [x(ys + zpi*)/(s - pi*)(z_1)^2 + \frac{1}{2}(z_1\cdot\epsilon)(y^2 - z^2)]\epsilon\}}{(z - y)(z_1)^2 + 2x^2[(ys + zpi*)/(s - pi*)](z_1)^2 + x(y^2 - z^2)(z_1\cdot\epsilon)}, \qquad (4.17)$$

the transformation law of Minkowski vector. If $x = y = z = 1$ in (4.17), then $g_s = g_p$ and

$$z_1 \underset{c_2 = 0}{\xrightarrow{g_s}} = z_1 + \tau. \qquad (4.18)$$

From (4.16) one can find also how z_1 transforms under g_c,

$$z_1 \xrightarrow{g_c} \frac{z_1 - \tilde{\kappa}[(z_1)^2 + c_2]}{1 - 2(\tilde{\kappa}\cdot z_1) + (\tilde{\kappa})^2[(z_1)^2 + c_2]}. \qquad (4.19)$$

With respect to the dilatation g_d

$$z_1 \xrightarrow{g_d} \exp(-2\zeta)z_1, \qquad (4.20a)$$

while the Lorentz rotation gives

$$z_1 \xrightarrow{g_1} (\exp\lambda)z_1(\exp(-\lambda)). \qquad (4.20b)$$

Thus from (4.16)–(4.30) one concludes that the Clifford number z_1 transforms under the conformal group $\mathscr{K}$ like a vector from Minkowski complex space-time provided $I_1 = 0$.

Now let us investigate how z_2, defined by (4.15b), transforms. First notice that $t^2 = \langle t^2\rangle_0 + \langle t^2\rangle_4$ for any $t \in T$. Since t^{-1} for $t \neq 0$ can be defined as

$$t^{-1} \equiv t/t^2 = (\langle t^2\rangle_0 + \langle t^2\rangle_4)^{-1}t, \qquad (4.21)$$

$(1 + i*)(t^2)^{-1} = \rho(1 + i*)$, $\rho^{-1} = \langle t^2\rangle_0 + \langle\langle t^2\rangle_4 i*\rangle_0$, and $z_2 = \rho[(1 + i*)t]\cdot a$. We also notice that from t and a one can construct another conformal invariant $I_2 \equiv \frac{1}{2}(t - at^{-1}a)(1 - i*)$. If z_2 undergoes de Sitter translation g_s, then one gets a formula similar to (4.16), provided $I_2 = 0$. Similarly, z_2 satisfies (4.19) under the same condition. Thus if $I_2 = 0$ and $R \to \infty$, z_2 translates by vector τ. It is also easily seen that z_2 satisfies (4.20).

Thus z_1 and z_2 given by (4.15) provide two possible projections from the antisymmetric and symmetric parts of the complex Dirac–Clifford algebra on a complex Minkowski space-time $(M^a)^c$, provided both conformal invariants I_1 and I_2 vanish.

B.

Let us consider now the Hermitian and anti-Hermitian parts of the algebra $\mathscr{D}^c$, preserved under the following action[42] of $\mathscr{K}$ on $\mathscr{D}^c$:

$$\mathscr{K} \ni g: \mathscr{D}^c \ni m \xrightarrow{g} gmg^{-1} \equiv m' \in \mathscr{D}^c. \qquad (4.22)$$

If $(\beta_s \circ \alpha_s)(m) = \pm m$ then $(\beta_s \circ \alpha_s)(m') = \pm m'$. In what follows we deal only with Hermitian Clifford numbers $m \in \mathscr{D}^c$, $m = (\beta_s \circ \alpha_s)(m)$. Again, first we find the transformation laws of all k-vector parts of $m \in \mathscr{D}^c$ under any $g \in \mathscr{K}$, now defined as

$$\mathscr{D}^c_k \ni \langle m\rangle_k \xrightarrow{g} \langle gmg^{-1}\rangle_k \in \mathscr{D}^c_k. \qquad (4.23)$$

Next we look for a certain combination of $\langle m \rangle_k$ which would transform like a vector under $\mathscr{K}$. It will give us another projection from the Hermitian part of $\mathscr{D}^c$ on $(M^a)^c$.

Equations (4.22) and (4.23) show clearly that the scalar part s of m is invariant under $\mathscr{K}$. Moreover, if $g \in \mathscr{K}$ then $\langle g(v+t+a+p)g^{-1} \rangle_0 = 0$.

The elements of $\mathscr{K}$ were listed before [see (4.9)]. Thus we have:

(i) *de Sitter translation* g_s:

$$V \ni v \xrightarrow{g_s} x^2 v + \tfrac{1}{4}\epsilon[2yzai* - (y^2+z^2)v]\epsilon + x(y\epsilon \cdot t + z\epsilon pi*) \in V,$$

$$A \ni a \to x^2 a + \tfrac{1}{4}\epsilon[2yzvi* - (y^2+z^2)a]\epsilon + x(z(\epsilon \cdot t)i* + y\epsilon p) \in A,$$

$$T \ni t \to x^2 t + x\epsilon \wedge (yv - zai*) + \tfrac{1}{4}(z^2 - y^2)\epsilon t \epsilon \in T, \tag{4.24}$$

$$P \ni p \to (2x^2 - 1)p + x\epsilon \wedge (ya - zvi*) \in T.$$

If $x = y = z = 1$ in (4.24) and $\epsilon \equiv \tau$

(ii) *Poincaré translation* g_p:

$$V \ni v \xrightarrow{g_p} v + \tfrac{1}{2}\tau(ai* - v)\tau + \tau \cdot t + \tau pi* \in V,$$

$$A \ni a \to a + \tfrac{1}{2}\tau(vi* - a)\tau + (\tau \cdot t)i* + \tau p \in A,$$

$$T \ni t \to t + \tau \wedge (v - ai*) \in T,$$

$$P \ni p \to p + \tau \wedge (a - vi*) \in P. \tag{4.25}$$

Therefore, there are two Poincaré-invariant Clifford numbers:

$J_3 \equiv t \wedge (v - ai*)^{-1}$ and $J_2 \equiv v - ai*$. If $x = -y = z = 1$ in (4.24) and $\epsilon = \tilde{\kappa}$ then

(iii) *Special conformal transformation* g_c:

$$V \ni v \xrightarrow{g_c} v - \tfrac{1}{2}\tilde{\kappa}(v + ai*)\tilde{\kappa} - \tilde{\kappa} \cdot t + \tilde{\kappa} pi* \in V,$$

$$A \ni a \to a - \tfrac{1}{2}\tilde{\kappa}(a + vi*)\tilde{\kappa} + (\tilde{\kappa} \cdot t)i* - \tilde{\kappa} p \in A,$$

$$T \ni t \to t - \tilde{\kappa} \wedge (v + ai*) \in T,$$

$$P \ni p \to p - \tilde{\kappa} \wedge (a + vi*) \in P. \tag{4.26}$$

It follows that $J_1 \equiv v + ai*$ remains invariant under g_c.

(iv) *Dilatation* g_d:

$$V \ni v \xrightarrow{g_d} (\mu^2 + \nu^2)v - 2\mu\nu ai* \in V,$$

$$A \ni a \to (\mu^2 + \nu^2)a - 2\mu\nu vi* \in A, \tag{4.27}$$

$$P \ni p \to p \in P,$$

$$T \ni t \to t \in T.$$

(v) *Lorentz rotation* g_l:

$$V \ni v \xrightarrow{g_l} \langle g_l(v+a)g_l^{-1} \rangle_1 \in V,$$

$$A \ni a \to \langle g_l(v+a)g_l^{-1} \rangle_3 \in A,$$

$$T \ni t \to \langle g_l t g_l^{-1} \rangle_2 \in T,$$

$$P \ni p \to p + \langle g_l p g_l^{-1} \rangle_4 \in P.$$

Let us now define the Clifford number

$$z := (t - pi*)(v - ai*)^{-1}. \tag{4.28}$$

Notice, that $z = \langle z \rangle_1 + \langle z \rangle_3$, i.e., z contains only vector and axial parts, where

$$\langle z \rangle_1 = t \cdot (v - ai*)^{-1} - pi*(v - ai*)^{-1}, \tag{4.29a}$$

$$\langle z \rangle_3 = t \wedge (v - ai*)^{-1} = J_3. \tag{4.29b}$$

Moreover, $(v - ai*)^{-1} = \pi(v - ai*)$ where π is a scalar and $\pi^{-1} = v^2 - a^2 - 2(v \wedge a)i*$ ($\pi \neq 0$ assumed). It can be shown that the vector part of z transforms like a vector under the conformal group $\mathscr{K}$ (Ref. 43). Let us examine only the Poincaré translation g_p acting on $\langle z \rangle_1$. From (4.25) we find that

$$z \xrightarrow{g_p} z + \tau + J_3$$

or, equivalently,

$$\langle z \rangle_1 \xrightarrow{g_p} \langle z \rangle_1 + \tau \text{ and } J_3 \xrightarrow{g_p} J_3,$$

which simply says that the vector part of z is translated by τ while the axial part remains unchanged.

Let us summarize briefly this section. The following Clifford numbers are obtained by three distinct projections, commuting with $\mathscr{K}$, on a complex Minkowski space-time $(M^a)^c$ from the antisymmetric, symmetric, and Hermitian parts of the complex Dirac–Clifford algebra $\mathscr{D}^c$, provided the corresponding conformal invariants vanish:

$$z_1 = v(s + ip*)^{-1}, \quad I_1 = 0, \tag{4.30a}$$

$$z_2 = [(1 + i*)t^{-1}] \cdot a, \quad I_2 = 0, \tag{4.30b}$$

$$\langle z \rangle_1 = t \cdot (v - ai*)^{-1} - pi*(v - ai*)^{-1}, \tag{4.30c}$$

no condition.

z_1, z_2, and $\langle z \rangle_1$ are all Clifford vectors and their vector

TABLE I. Clifford numbers invariant under conformal group $\mathscr{K} = C(1,3)$.

Antisymmetric $m \in \mathscr{D}^c$ $m = s + v + p$	Symmetric $m \in \mathscr{D}^c$ $m = t + a$	Hermitian $m \in \mathscr{D}^c$ $m = (\alpha \circ \beta)(\bar{m})$	Invariants of
v	a	t,p	dilatation
	$t_2 = \frac{1}{2}t(1 - i*)$	$J_1 = v + ai*$	special conformal transformation
$c_1(\infty) = s + pi*$	$t_1 = \frac{1}{2}t(1 + i*)$	$J_3 = t \wedge (v - ai*)^{-1}$	Poincaré
$c_2 = \dfrac{I_1}{(s + pi*)^2}$		$J_2 = v - ai*$	translation
$c_1(R) = s + pi* - (1/R^2)(s - pi*)$			de Sitter translation
$I_1 = s^2 - v^2 - p^2$	$I_2 = \frac{1}{2}(t - at^{-1}a)(1 - i*)$	s	conformal group $\mathscr{K} = C(1,3)$

coordinates can be written as

$$(z_1)_\mu = v_\mu (s + p')^{-1}, \tag{4.31a}$$

$$(z_2)_\nu = \frac{a^\mu(\overset{*}{t}_{\mu\nu} + t_{\mu\nu})}{t^{\lambda\kappa}(\overset{*}{t}_{\kappa\lambda} + t_{\kappa\lambda})}, \tag{4.31b}$$

$$(\langle z\rangle_1)_\mu = \pi[2t_{\mu\nu}(v^\nu - a^\nu) - p'(v_\mu - a_\mu], \tag{4.31c}$$

where $v = v^\mu e_\mu$, $p' = pi*$, $a^\nu = a^\nu e_\nu i*$, $t = t^{\mu\nu} e_\mu \wedge e_\mu$, $t^{\mu\nu} = -t^{\nu\mu}$, $\overset{*}{t} = \overset{*}{t}^{\mu\nu}(e_\mu \wedge e_\nu)i*$ ($\overset{*}{t}$ denotes a dual tensor to t), $\pi \neq 0$ assumed and given above. Invariants are listed in Table I.

At the end of this section we comment on a possible, geometrical interpretation for I_1 to vanish. Recall from Sec. II, that any vector $w \in V_{4,1}$ can be written as $w = w^a f_a = \tilde{w}i* + w_p$, where $\tilde{w} = w^\mu e_\mu$ and $w_p = w^5 i*$. Thus $w^2 = (w_p)^2 - (\tilde{w})^2$. If we now substitute $w^5 = s$ and $\tilde{w} = v$ then $I_1 = s^2 - v^2 - p^2 = (si*)^2 - v^2 - (pi*)^2 = w^2 - (pi*)^2$. $I_1 = 0$ means that $w^2 = (pi*)^2$, i.e., w lies on a sphere in $V_{4,1}$ with radius $pi*$.

One can also introduce the following basis for $V_{2,4}$ (see Refs. 31 and 44). $\gamma_a = (-if_a, i) = (e_\mu *, *, i)$, $a, b = 1,2,...,6$ where $e_\mu \cdot e_\nu = \eta_{\mu\nu} = (1,3)$. The inner product in $V_{2,4}$ is defined then as $g_{ab} = \gamma_a \cdot \gamma_b = -f_a \cdot f_b + i(-i)$. Therefore any vector $\xi \in V_{2,4}$ can be represented as $\xi = (-iw, (pi*)i) = (v*, s*, (pi*)i)$ for some s, v, and p in $\mathscr{D}^c$. Moreover, $\xi^2 = -s^2 + v^2 + p^2 = -I_1$, and $I_1 = 0$ means that ξ lies on a cone in $V_{2,4}$. If $z = v(s + pi*)^{-1}$ then $z^\mu = \xi^\mu(\xi_5 + \xi_6)^{-1}$, which is a standard projection from $V_{2,4}$ on $V_{1,3}$, provided $\xi^2 = 0$.

V. WITT DECOMPOSITION OF $(T_pM)^C$. TWISTOR SPACE AS AN IDEAL IN CLIFFORD ALGEBRA

For $(T_pM)^c$ one has a Witt decomposition[2,10,23]

$$(T_pM)^c = F \oplus F',$$

where F and F' are two-dimensional, maximal, completely isotropic subspaces. A Witt frame for $(T_pM)^c$ is just the well-known Sachs frame $e_{A\dot{B}} = \sigma^\mu{}_{A\dot{B}} e_\mu$, $A, B, \dot{A}, \dot{B} = 1,2$. Explicitly, $e_{1\dot{1}} = e_0 + e_3$, $e_{1\dot{2}} = e_1 - ie_2 \in F$ and $e_{2\dot{2}} = e_0 - e_3$, $e_{2\dot{1}} = e_1 + ie_2 \in F'$, $\eta|_{F'} = \eta|_F = 0$, and $(e_{A\dot{B}}, e_{C\dot{D}}) = \sigma^\mu{}_{A\dot{B}}\sigma_{\mu C\dot{D}} = 2\epsilon_{AC}\epsilon_{\dot{B}\dot{D}}$ where $\epsilon_{AC} = \epsilon_{\dot{B}\dot{D}} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$. Therefore, in the Sachs frame, η has the matrix form

$$\eta = 2\begin{pmatrix} 0 & \sigma^3 \\ \sigma^3 & 0 \end{pmatrix}.$$

Let f be any tensor of F (or F'), i.e., $f \in \overset{2}{\wedge} F$, $f^2 = 0$. For example, in the Sachs frame one can take $f = f_1$ or f_2 where $f_A := -\frac{i}{8}\epsilon^{\dot{B}\dot{C}}\sigma^\mu{}_{A\dot{B}}\sigma^\nu{}_{A\dot{C}} e_\mu \wedge e_\nu$. Thus $f_2 = \frac{1}{4}(e_0 - e_3) \wedge (ie_1 - e_2)$. f generates a linear, four dimensional twistor space $\mathscr{T}(f)$ over $\mathbb{C}$ as the left ideal in $\mathscr{D}^c$, $\mathscr{T}(f) := \mathscr{D}^c f$. Consider the linear mapping $*$: $T \ni \xi \rightarrow *\xi \in \mathscr{T}$. Since $*^2 + 1 = 0$ (minimal polynomial of $*$) then with respect to $*$ a twistor space is the direct sum of two-dimensional complex eigenspaces of $*$, $\mathscr{T} = S^+ \oplus S^-$, where $*\xi^\pm = \pm i\xi^\pm$ and $\xi^\pm = \frac{1}{2}(1 \pm i*)\xi \in S^\pm$. Moreover, $\mathscr{D}^c_+ S^\pm \subset S^\pm$, $\mathscr{D}^c_- S^\pm \subset S^\mp$, which means that $\mathscr{T}(f)$ is also a Z_2-graded Clifford module and this important gradation, known as the Lorentz spinor decomposition, is solely determined by the orientation $*$ of the tangent space-time T_pM. Change of $*$ will result in $\xi^+ \leftrightarrow \xi^-$. In the Lorentz spinor basis one can write $S^+ \ni \omega = \omega_A \epsilon^A$, $S^- \ni \pi = \pi^{\dot{A}}\epsilon_{\dot{A}}$, $\omega_A, \pi^{\dot{A}} \in \mathbb{C}$. In the Sachs frame a twistor basis $\{\epsilon_\alpha \equiv u_\alpha f\}$ for $\mathscr{T}(f)$ can be chosen as: $u_1 = 1$, $u_2 = \frac{1}{2}e_{1\dot{1}}$, $u_3 = \frac{1}{2}ie_{1\dot{2}}$, $u_4 = \frac{1}{4}ie_{1\dot{1}}e_{1\dot{2}}$. From now on we fix $f = f_2$. We will write $\mathscr{T}(f) \ni \xi = \xi^\alpha \epsilon_\alpha = uf$.

The representation of algebra $\mathscr{D}^c$ in $\mathscr{T}(f)$ is then linear (Refs. 1 and 10).

$$\mathscr{D}^c \ni m : \mathscr{T}(f) \ni \xi \rightarrow m\xi \in \mathscr{T}(f),$$

where $m\epsilon_\alpha = \epsilon_\beta \gamma^\beta{}_\alpha(m)$ ($\mathbb{C} \ni \gamma^\beta{}_\alpha(m) \equiv \epsilon^\beta(m\epsilon_\alpha)$, $\epsilon^\beta \in \mathscr{T}'$-dual twistor space) and determines the matrix representation of $\mathscr{D}^c$. For instance, the Dirac matrices γ_μ can be defined as $e_\mu \epsilon_\alpha = \epsilon_\beta \gamma_\mu{}^\beta{}_\alpha$. One can say, therefore, that the matrix representation of algebra $\mathscr{D}^c$ is determined by the generator f of twistor space $\mathscr{T}$ and orientation $*$ of T_pM up to the choice of basis in Lorentz space.

Let us introduce now the Hermitian correlation $\mathscr{A}$ on $\mathscr{T}(f) \equiv \mathscr{T}$, called also a general Dirac conjugation, and the conformal group U(2,2) as the correlated automorphism of $\mathscr{T}$. Another definition of $\mathscr{A}$ was given in Ref. 10. Correlation $\mathscr{A}$ on $\mathscr{T}$ is a semilinear map such that

$$\mathscr{A} : \mathscr{T} \ni \xi = uf \rightarrow \mathscr{A}(\xi) = -\beta_s(\bar{u}\gamma f) \in \mathscr{T}', \tag{5.1}$$

where $\gamma f = \bar{f}\gamma$; γ is a so-called pure spinor (see Refs. 10 and 22). The essential point for the twistor theory in the Clifford algebra language is to see how the algebra structure alone determines the Hermitian (and symplectic, see below) form on $\mathscr{T}$. This beautiful relation shown by Crumeyrolle is based on the observation that if f and f' are the pseudoscalars of F and F' (Witt decomposition), respectively, then

$$\dim(\mathscr{D}^c f \cap f' \mathscr{D}^c) = 1. \tag{5.2}$$

Hence one can always find $\gamma \in \mathscr{D}^c$ such that $\mathscr{A}$ is well defined by (5.1) and, moreover, the Hermitian from $\mathscr{H}$ on a twistor space $\mathscr{T}$ can be defined[45] in terms of $\mathscr{A}$ as

$$\mathscr{H} : \mathscr{T} \times \mathscr{T} \ni (\xi \equiv uf, \eta \equiv vf) \rightarrow \mathscr{H}(uf, vf)f \equiv \mathscr{A}(uf)vf. \tag{5.3}$$

$\mathscr{H}$ defined here is invariant under the conformal group $\mathscr{K}$. To show this we rewrite (5.3) with a help of (5.1) as

$$\mathscr{H}(uf, vf) = \gamma\beta_s(\bar{u}\bar{f})vf, \tag{5.4}$$

where $uf, vf \in \mathscr{T}$, $\gamma = e_2$, $(e_2)^2 = -1$ and $\bar{u}\bar{f}$ means complex conjugate of uf. Then

$$\overline{\mathscr{H}(uf, vf)f} = \beta_s(\overline{vf})uf\gamma \text{ and } \mathscr{H}(uf, vf) = \overline{\mathscr{H}(vf, uf)}.$$

Since the representation of $\mathscr{D}^c$ in $\mathscr{T}$ defined earlier was linear, $\mathscr{K}$ acts on $\mathscr{T}$ linearly and irreducibly (cf. Ref. 10). Thus the following diagram commutes:

$$\begin{array}{ccc} \mathscr{K} \ni g : \mathscr{T} \ni uf & \longrightarrow & guf \in \mathscr{T} \\ \downarrow \mathscr{A} & & \downarrow \mathscr{A} \\ \mathscr{K} \ni g : \mathscr{A}(uf) & \rightarrow \mathscr{A}(guf) = & \mathscr{A}(uf)g^{-1} \in \mathscr{T}', \end{array} \tag{5.5}$$

for $\mathscr{A}(uf)g^{-1} = -\beta_s(\bar{g}\bar{u}\gamma f) = -\beta_s(\bar{u}\gamma f)(\beta_s \circ \alpha_s(g)$

$= \mathscr{A}(uf)g^{-1}$
and $(\beta_s \circ \alpha_s)(g)g = 1$ for every $g \in \mathscr{K}$. From (5.5) it follows that $\mathscr{H}(guf, gvf)f = \mathscr{A} = (guf)gvf = \mathscr{A}(uf)vf = \mathscr{H}(uf, vcf)f$. Therefore, $\mathscr{H}$ defined by (5.4) is invariant with respect to $\mathscr{K}$ and one can show that the signature of $\mathscr{K}$ is $(+\ +\ -\ -)$.

Notice, that (5.3) provides another possible definition for the conformal group as a group U(2,2) of correlated automorphisms of $\mathscr{T}$ which preserves the Hermitian form $\mathscr{H}$.

$$U(2,2) = \{g \in \mathscr{D}^c \mid \mathscr{A}(g\xi)g\eta = \mathscr{A}(\xi), \forall \xi, \eta \in \mathscr{T}\}. \qquad (5.6)$$

Therefore $(\beta_s \circ \alpha_s)(g)g = 1$ needs to be satisfied for all $g \in U(2.2) \cap \mathscr{D}^c$ or, $(\beta_s \circ \alpha_s)(X) + X = 0$ for all $X \in u(2,2) \cap \mathscr{D}^c$.

Let us remark here that $\mathscr{T}$ can also be endowed with a symplectic, nondegenerate from[10,46] $\theta: \mathscr{T} \times \mathscr{T} \to C$ defined as

$$\theta(uf, vf)f = \beta(uf)vf. \qquad (5.7)$$

θ is symplectic because $-\theta(uf,vf)f = \beta(\theta(uf,vf)f) = -f\beta(v)uf = \theta(vf,uf)f$. In $\mathscr{T}$ one can now consider a symplectic basis (e_α, e_{β^*}), $\alpha,\beta = 1,2$, where $e_1 = \epsilon_4$, $e_2 = \epsilon_3$, $e_{1^*} = \epsilon_1, e_{2^*} = \epsilon_2$, such that $\theta(e_\alpha, e_\beta) = \theta(e_{\alpha^*}, e_{\beta^*}) = 0$, and $\theta(e_\alpha, e_{\beta^*}) = \delta_{\alpha\beta}$. Thus $\mathscr{T} = \mathscr{T}_1 \oplus \mathscr{T}_2$, where both $\mathscr{T}_1$ and $\mathscr{T}_2$ are transverse, Lagrangian subspaces of $\mathscr{T}$, generated by (e_{β^*}) and (e_α), respectively.

It can also be shown that θ, given in Eq. (5.7), is invariant under the left regular action of G_β (see Sec. IV) on $\mathscr{T}$: $G_\beta \ni g: \mathscr{T} \ni uf \to guf \in \mathscr{T}$, for $\theta(guf, gvf)f = \beta(uf)\beta(g)gvf = \theta(uf,vf)f$. Therefore $G_\beta = Sp(4,C)$. We will come back to this point in Sec. VI.

VI. TENSOR PRODUCTS OF TWISTOR SPACES

As we recall from Sec. IV, a twistor space $\mathscr{T}$ is defined as left ideal in $\mathscr{D}^c$, $\mathscr{T}(f) := \mathscr{D}^c f$, f fixed, providing a representation space for linear representation of $\mathscr{D}^c$. More precisely, a spinorial representation of the Clifford group G generated by $\mathscr{D}^c$, acts irreducibly on $\mathscr{T}$. We do not intend to develop this point here but we refer to Refs. 1, 10, and 22 Isomorphisms between $\mathscr{D}^c$ and $\mathscr{T} \otimes \mathscr{T}$ or $\mathscr{D}^c$ and $\mathscr{T} \otimes \mathscr{A}(\mathscr{T})$ will play a crucial role in our present considerations. We give them below[10,22] for the antisymmetric case, in agreement with our definitions of symmetric, antisymmetric and Hermitian Clifford numbers [see (4.6)] considered as symmetric, antisymmetric and Hermitian tensors respectively. They allow any m from algebra $\mathscr{D}^c$ to be treated also as an element of tensor space $\mathscr{T} \otimes \mathscr{T}$ or $\mathscr{T} \otimes \mathscr{A}(\mathscr{T})$. Therefore, one can express the coordinates of k vector parts of any $m \in \mathscr{D}^c$ in terms of the corresponding tensor coordinates of $\mathbf{m}$, $\mathbf{m} \in \mathscr{T} \otimes \mathscr{T}$, or $\mathbf{m} \in \mathscr{T} \otimes \mathscr{A}(\mathscr{T})$.

A. ANTISYMMETRIC CASE

It is natural to introduce $\{u_\alpha f \otimes u_\beta f\}$ as a basis in $\mathscr{T} \otimes \mathscr{T}$ while $\{u_{(\alpha} f \otimes u_{\beta)} f\}$ and $\{u_{[\alpha} f \otimes u_{\beta]} f\}$ provide bases in the space of symmetric $(\mathscr{T} \otimes \mathscr{T})^+$ and antisymmetric $(\mathscr{T} \otimes \mathscr{T})^-$ tensors respectively, $\alpha,\beta = 1,2,3,4$. Then any $\mathbf{m} = \mathbf{m}^+ + \mathbf{m}^-$, $\mathbf{m} \in \mathscr{T} \otimes \mathscr{T}$ and $\mathbf{m}^\pm \in (\mathscr{T})^\pm$. We denote by τ a transposition $\tau(u_\alpha f \otimes u_\beta f) := u_\beta f \otimes u_\alpha f$.

Let β_s and α be as before (see Sec. III), f being fixed. Let $\mathbf{m} \in \mathscr{T} \otimes \mathscr{T}$ and $m \in \mathscr{D}^c$. Then the map i

$$i: \mathscr{T} \otimes \mathscr{T} \ni \mathbf{m} = m^{\alpha\beta} u_\alpha f \otimes u_\beta f \to m^{\alpha\beta} \beta_s(u_\alpha) f \alpha(u_\beta) = m \in \mathscr{D}^c \qquad (6.1)$$

provides an isomorphism between $\mathscr{T} \otimes \mathscr{T}$ and $\mathscr{D}^c$ Ref. 10. Moreover, the following diagram commutes:

$$\begin{array}{ccc} \mathscr{T} \otimes \mathscr{T} \ni u_\alpha f \otimes u_\beta f & \xrightarrow{i} & \beta_s(u_\alpha) f \alpha(u_\beta) \in \mathscr{D}^c \\ \tau \downarrow & & -\beta \downarrow \\ \mathscr{T} \otimes \mathscr{T} \ni u_\beta f \otimes u_\alpha f & \xrightarrow{i} & \beta_s(u_\beta) f \alpha(u_\alpha) \in \mathscr{D}^c. \end{array} \qquad (6.2)$$

From (6.2) we conclude that, if $\mathbf{m}^\pm \in (\mathscr{T} \otimes \mathscr{T})^\pm$ then $\beta(i(\mathbf{m}^\pm)) = \mp i(\mathbf{m}^\pm)$, which agrees with (4.6).

Remark: Due to the fact that $\mathscr{T}$ is endowed with a symplectic form θ, one can construct a so called symplectic Clifford algebra over $\mathscr{T}$ (cf. Ref. 47), $C_s(\theta)$ by means of the quotient $\otimes \mathscr{T}/I(\theta) \simeq C_s(\theta)$ where $I(\theta)$ denotes a two-sided ideal in $\otimes \mathscr{T}$ generated by $\xi \otimes \eta - \eta \otimes \xi - \theta(\xi,\eta)$, $\xi, \eta \in \mathscr{T}$. Therefore, $(\mathscr{T} \otimes \mathscr{T})/I(\theta) \simeq V^2 T \subset C_s^+(\theta)$, the even part of $C_s(\theta)$.

We explore now the relation between the basis $\{u_\alpha f \otimes u_\beta f\}$ of $\mathscr{T} \otimes \mathscr{T}$ and that of $\mathscr{D}^c$.

$$\mathscr{T} \otimes \mathscr{T} \ni u_\alpha f \otimes u_\beta f \equiv \Gamma_{\alpha\beta} + \Gamma^\mu_{\alpha\beta} e_\mu + \tfrac{1}{2} \Gamma^{\mu\nu}_{\alpha\beta} e_\mu \wedge e_\nu + \overset{*}{\Gamma}{}^\mu_{\alpha\beta} e_\mu i* + \overset{*}{\Gamma}_{\alpha\beta} i* \in \mathscr{D}^c. \qquad (6.3)$$

Once the Witt base and f are fixed, the numerical values of Γ coefficients can be uniquely determined. We list them in general form

(i) *scalar*

$$\Gamma_{\alpha\beta} = \Gamma_{[\alpha\beta]} = \langle u_{[\alpha} \otimes u_{\beta]} f \rangle_0 \equiv \langle u_\alpha \wedge u_\beta \rangle_0, \qquad (6.4a)$$

(ii) *vector*

$$\Gamma^\mu_{\alpha\beta} = \Gamma^\mu_{[\alpha\beta]} = e^\mu \cdot (u_{[\alpha} f \otimes u_{\beta]} f) \equiv (e^\mu)_{[\alpha\beta]}, \qquad (6.4b)$$

(iii) *tensor*

$$\Gamma^{\mu\nu}_{\alpha\beta} = \Gamma^{[\mu\nu]}_{(\alpha\beta)} = (e^\nu \wedge e^\mu) \cdot (u_{(\alpha} f \otimes u_{\beta)} f) \equiv (e^\nu \wedge e^\mu)_{(\alpha\beta)}, \qquad (6.4c)$$

(iv) *axial vector*

$$\overset{*}{\Gamma}{}^\mu_{\alpha\beta} = \overset{*}{\Gamma}{}^\mu_{(\alpha\beta)} = -(ie^\mu *) \cdot (u_{(\alpha} f \otimes u_{\beta)} f) \equiv -(ie^\mu *)_{(\alpha\beta)}, \qquad (6.4d)$$

(v) *pseudoscalar*

$$\overset{*}{\Gamma}_{\alpha\beta} = \overset{*}{\Gamma}_{[\alpha\beta]} = \langle i* u_{[\alpha} f \otimes u_{\beta]} f \rangle_0 \equiv \langle i*(u_\alpha \wedge u_\beta) \rangle_0. \qquad (6.4e)$$

In (6.4b), (6.4c), and (6.4d) a dot $\cdot$ denotes the inner product in the Clifford algebra $\mathscr{D}^c$. Now one can express the coordinates of any $m \in \mathscr{D}^c$ in terms of the tensor coordinates of a corresponding (symmetric of antisymmetric) tensor $\mathbf{m}$ from $\mathscr{T} \otimes \mathscr{T}$. Since (4.31a) and (4.31b) give the (vector) coordinates of two Clifford numbers which are of special interest to us, we rewrite them in the following form:

$$(z_1)_\mu = \frac{m^{[\alpha\beta]}(e_\mu)_{[\alpha\beta]}}{m^{[\gamma\delta]} \langle (1 + i*)(u_\gamma \wedge u_\delta) \rangle_0}; \qquad (6.5a)$$

$$(z_2)_\nu = 2\frac{m^{(\alpha\beta)}(ie^\mu *)_{(\alpha\beta)}\,[m^{(\gamma\delta)}(i*(e_\mu\wedge e_\nu)_{(\gamma\delta)} + m^{(\epsilon\theta)}(e_\mu\wedge e_\nu)_{(\epsilon\theta)}]}{m^{(\alpha\beta)}(e^\lambda\wedge e^\kappa)_{(\alpha\beta)}\,[m^{(\gamma\delta)}(i*(e_\kappa\wedge e_\lambda))_{(\lambda\delta)} + m^{(\epsilon\theta)}(e_\kappa\wedge e_\lambda)_{(\epsilon\theta)}]}. \tag{6.5b}$$

Remark. With the help of (6.4) one can express I_1 (see Table I) in terms of $m^{[\alpha\beta]}$: $I_1 = \frac{1}{4}\epsilon_{\alpha\beta\gamma\sigma}m^{[\alpha\beta]}m^{[\gamma\sigma]}$, where $\epsilon_{\alpha\beta\gamma\sigma}$ is totally antisymmetric tensor, $\epsilon_{1234} = 1$. If $\mathbf{m}$ is decomposable, i.e., $\mathbf{m} = \xi\otimes\eta$ for some $\xi,\eta\in\mathscr{T}$, then $I_1 = 0$ and (6.5a) provides the well-known Penrose projection.[7–9]

B. HERMITIAN CASE

In $\mathscr{T}\otimes\mathscr{A}(\mathscr{T})$ we introduce $\{u_\alpha f\otimes\mathscr{A}(u_{\dot\beta} f)\}$ as a basis, where $\mathscr{A}(u_\alpha f)\equiv u_{\dot\alpha} f = -f\gamma\beta_s(\bar u_\alpha)$ ($\bar u_\alpha$ is the complex conjugate of u_α). Then any tensor $\mathbf{m}\in\mathscr{T}\otimes\mathscr{A}(\mathscr{T})$ can be written as $\mathbf{m} = m^{\alpha\dot\beta}u_\alpha f\otimes u_{\dot\beta} f$, and its Hermitian conjugate $\mathbf{m}^h := \bar m^{\alpha\dot\beta}u_\beta f\otimes u_{\dot\alpha} f$.

Let $\mathbf{m}\in\mathscr{T}\otimes\mathscr{A}(\mathscr{T})$ and $m\in\mathscr{D}^c$. Then a linear isomorphism ω between $\mathscr{T}\otimes\mathscr{A}(\mathscr{T})$ and $\mathscr{D}^c$ is given as

$$\mathscr{T}\otimes\mathscr{A}(\mathscr{T})\ni\mathbf{m} = m^{\alpha\dot\beta}u_\alpha f\otimes u_{\dot\beta} f \xrightarrow{\omega} m^{\alpha\dot\beta}u_\alpha f\gamma\beta_s(u_{\dot\beta}) \equiv m\in\mathscr{D}^c, \tag{6.6}$$

$\beta_s(u_{\dot\beta})\equiv(\beta_s\circ\alpha_s)(u_\beta)$, $\gamma f = \bar f\gamma$, $\gamma = e_2$ in our basis. Then, the following diagram is commutative:

$$\begin{array}{ccc}
\mathscr{T}\otimes\mathscr{A}(\mathscr{T})\ni\mathbf{m} = m^{\alpha\dot\beta}u_\alpha f\otimes u_{\dot\beta} f & \xrightarrow{\omega} & m = m^{\alpha\dot\beta}u_\alpha f\gamma\beta_s(u_{\dot\beta})\in\mathscr{D}^c \\
\downarrow h & & \downarrow \beta_s\circ\alpha_s \\
\mathscr{T}\otimes\mathscr{A}(\mathscr{T})\ni\mathbf{m}^h = \bar m^{\alpha\dot\beta}u_\beta f\otimes u_{\dot\alpha} f & \xrightarrow{\omega} & (\beta_s\circ\alpha_s)(m) = \bar m^{\alpha\dot\beta}u_\beta f\gamma\beta_s(u_{\dot\alpha})\in\mathscr{D}^c
\end{array} \tag{6.7}$$

for $(\beta_s\circ\alpha_s)(f) = -\bar f$, $(\beta_s\circ\alpha_s)(\gamma) = -\gamma$. From (6.7) one concludes that if $\mathbf{m}^h = \mathbf{m}$ and $\omega(\mathbf{m}^h) = m$, then $\beta_s(\bar m) = m$, in agreement with (4.6).

Let the basis of $\mathscr{T}\otimes\mathscr{A}(\mathscr{T})$ be related to the basis of $\mathscr{D}^c$ through the Γ coefficients defined by

$$\mathscr{T}\otimes\mathscr{A}(\mathscr{T})\ni u_\alpha f\otimes\mathscr{A}(u_\beta f)\equiv u_\alpha f\otimes u_{\dot\beta} f$$
$$= \Gamma_{\alpha\dot\beta} + \Gamma^\mu_{\alpha\dot\beta}e_\mu + \tfrac{1}{2}\Gamma^{\nu\mu}_{\alpha\dot\beta}e_\nu\wedge e_\mu + \dot\Gamma^\mu_{\alpha\dot\beta}(ie_\mu *) + \dot\Gamma_{\alpha\dot\beta}\,i*\in\mathscr{D}^c. \tag{6.8}$$

Taking inner products of $u_\alpha f\otimes u_{\dot\beta} f$ with each element of the Clifford base in $\mathscr{D}^c$ and then projecting the results on scalars, one can calculate all Γ's and express them in general form.

(i) *scalar*:

$$\Gamma_{\alpha\dot\beta} = \Gamma_{(\alpha\dot\beta)} = \langle u_\alpha f\otimes u_{\dot\beta} f\rangle_0, \tag{6.9a}$$

(ii) *vector*:

$$\Gamma^\nu_{\alpha\dot\beta} = e^\nu\cdot(u_\alpha f\otimes u_{\dot\beta} f)\equiv(e^\nu)_{\alpha\dot\beta}. \tag{6.9b}$$

(iii) *tensor*:

$$\Gamma^{\mu\nu}_{\alpha\dot\beta} = (e^\nu\wedge e^\mu)\cdot(u_\alpha f\otimes u_{\dot\beta} f)\equiv(e^\nu\wedge e^\mu)_{\alpha\dot\beta}. \tag{6.9c}$$

(iv) *axial vector*:

$$\dot\Gamma_{\alpha\dot\beta} = -(e^\nu i*)\cdot(u_\alpha f\otimes u_{\dot\beta} f)\equiv -(e^\nu i*)_{\alpha\dot\beta}. \tag{6.9d}$$

(v) *pseudoscalar*

$$\dot\Gamma_{\alpha\dot\beta} = \dot\Gamma_{(\alpha\dot\beta)} = \langle i*(u_\alpha f\otimes u_{\dot\beta} f)\rangle_0\equiv u_\alpha\circ u_{\dot\beta}. \tag{6.9e}$$

Therefore, the vector coordinates of $\langle z\rangle_1$ [see (4.30c)] can be written in terms of the Hermitian Γ forms as:

$$\begin{aligned}\langle z\rangle_1 = \delta\{&m^{(\alpha\dot\beta)}(e^\nu\wedge e^\mu)_{(\alpha\dot\beta)}\,[m^{(\gamma\dot\delta)}(e_\nu)_{(\gamma\dot\delta)} + m^{(\epsilon\dot\theta)}(e_\nu i*)_{(\epsilon\dot\theta)}]\\
&- m^{(\alpha\dot\beta)}\langle i*(u_\alpha f\otimes u_{\dot\beta} f)\rangle\,[m^{(\beta\dot\gamma)}(e^\mu)_{(\beta\dot\gamma)}\\
&+ m^{(\gamma\dot\delta)}(e^\mu i*)_{(\gamma\dot\delta)}]\},\end{aligned} \tag{6.10}$$

where

$$\begin{aligned}\delta^{-1} = {}&m^{(\alpha\dot\beta)}m^{(\gamma\dot\delta)}(e^\mu)_{(\alpha\dot\beta)}(e^\mu)_{(\gamma\dot\delta)}\\
&+ m^{(\alpha\dot\gamma)}m^{(\beta\dot\delta)}(e^\nu i*)_{(\alpha\dot\gamma)}(e_\nu i*)_{(\beta\dot\delta)}\\
&- 2m^{(\alpha\dot\beta)}m^{(\gamma\dot\delta)}(e^\mu)_{(\alpha\dot\beta)}(e_\mu i*)_{(\gamma\dot\delta)},\end{aligned}$$

$m^{(\alpha\dot\beta)} = \frac{1}{2}(m^{\alpha\dot\beta} + \bar m^{\beta\dot\alpha})$ and $\Gamma_{(\alpha\dot\beta)}$ denotes the Hermitian part of $\Gamma_{\alpha\dot\beta}$.

The interrelation between the three projections derived above will be discussed in the near future. In particular, the new identities between the Dirac bilinears will enable us to show in the case of the decomposable tensors, that all three give the same point in $(M^a)^c$.

Finally, we mention that there is a well-known connection between the supersymmetry theory and the twistor formalism (Ref. 48 and the references therein). This results in introducing the fermionic twistor variables instead of the Penrose bosonic twistors used in the present work. The spinorial charges then are the Jordan roots of the conformal generators. We think that it might be a great deal easier to calculate the Casimir operators of the conformal group and its subgroups with a help of the formalism developed here, in particular, through the Γ coefficients (see Sec. VI).

The subgroups OSp(4,1) and OSp(8,1) have been introduced[48] in the supersymmetry theory. The primary groups, however, to start with are Sp(4) and Sp(8), the real forms of Pin(η) for $\mathscr{D}^c$ (see Secs. IV and V) and Pin ($\tilde g$) for $C^c(\tilde g)$, the complex Clifford algebra over $V_{2,4}$ (see Sec. IV). The Witt decomposition for $V^c_{2,4}$ enables one to introduce an eight-dimensional, complex spinor space S as an ideal (left or right) in $C^c(\bar{g})$, in analogy with $\mathscr{T}$. Therefore, the twistor space $\mathscr{T}$ can be indentified with the even part of S. Moreover, since $S\otimes S \simeq \mathscr{T}\otimes\mathscr{T}\otimes\mathscr{T}\simeq C^c(\bar{g})$ and $\mathscr{D}^c$ is isomorphic with the real, even part of $C^c(\bar{g})$, the components of any $m\in\mathscr{D}^c$ could be expressed in terms of the trilinear forms in twistors. This will make it possible to replace the bilinear forms in (6.5) by the trilinear ones, in the case of the decomposable tensors.

ACKNOWLEDGMENTS

Authors R. A. and J. R. are very grateful to Dr. Bruno Gruber for many helpful discussions and encouragement while they were working on this paper and for his kind hospitality at the Department of Physics and Astronomy, SIU Carbondale, where part of this work was done. R. A. thanks Dr. Alphonse Baartmans for his hospitality at the Department of Mathematics, SIU Carbondale, and Dr. Richard Millman for his criticism.

The authors are indebted to Dr. Garret Sobczyk for instructive discussions on geometric calculus. We benefitted from useful talks with Dr. Jan Mozrzymas. Thanks are due to Mr. Jerzy Kocik for calling our attention to important references and Dr. Krystyna Bugajska for pointing out an error in the first draft of this paper. R. A. thanks Linda Macak for typing this paper.

APPENDIX A

We use here notation from Sec. III. C_k denotes a vector subspace of k-vectors in $C(Q)$, $k = 1,2,\cdots,s = \dim V$. Q is a quadratic, nondegenerate form defined on a vector space V (real or complex). $\cdot$ and $\wedge$ denote wedge and inner products, respectively, in $C(Q)$. If $m,n\in C(Q)$, $m_i\in C_i$, $m_j\in C_j$ and $m_k\in C_k$, then the following formulas hold[12]

$$m_i m_j = \sum_{k=0}^{l} \langle m_i m_j\rangle_{|i-j|+2k};\; l = \tfrac{1}{2}(i+j-|i-j|),$$

$$\langle mn\rangle_r = (-1)^{\frac{1}{2}r(r-1)}\langle\beta(n)\beta(m)\rangle_r,$$

$$m_i\wedge m_j = (-1)^{ij} m_j\wedge m_i,$$

$$m_i\cdot m_j = (-i)^{i(j-1)} m_j\cdot m_i,\quad i\leqslant j,$$

$$m_i\cdot(m_j\cdot m_k) = (m_i\wedge m_j)\cdot m_k,\quad i+j\leqslant k,\quad i,j>0,$$

$$m_i\cdot(m_j\cdot m_k) = (m_i\cdot m_j)\cdot m_k\quad \text{for } i+k\leqslant j.$$

APPENDIX B

Below we give the list of used notations:

M	real, Riemannian space-time,
T_PM	tangent space to M at point $p\in M$ with bilinear form η_p of signature (1,3) and basis $\{e_\mu\}$, $\mu = 0,1,2,3$, $e_\mu\cdot e_\nu = \eta_{\mu\nu}$,
$\mathscr{K}\equiv C(1,3)$	conformal group acting in T_PM,
M^a	real, affine Minkowski space-time homogeneous with respect to the Poincaré group,
$C(Q)$	Clifford algebra over real vector space V, $\dim V = s$, endowed with quadratic, nondegenerate form Q,
B	symmetric, bilinear form defined on $V\times V$, associated with Q,
$\wedge^k V$	k th exterior product of V, $k = 0,1,2,\cdots,s$ with wedge product $\wedge$,
$C_k = \wedge^k V$	linear subspace of k-vectors of $C(Q)$ invariant with respect to Clifford group G,
$\langle m\rangle_k$	k-vector part of $m\in C(Q)$, i.e., $\langle m\rangle_k\in C_k$,
$\cdot$	inner product in $C(Q)$,
Z^*	set of all invertible elements of Z, center of $C(Q)$,
α,β	unique automorphism and antiautomorphism of $C(Q)$,
$\mathscr{D}\equiv C(\eta)$	real Dirac-Clifford algebra over T_pM,
$*$	orientation of T_pM or unit pseudoscalar of $\mathscr{D}$, $*^2 = -1$,
$\mathscr{D}^c$, $(M^a)^c$, $(T_pM)^c$	complexifications,
V_s	de Sitter space endowed with bilinear form g of signature (4,1) and basis $\{f_a\}$, $a = 1,2,3,4,5$, $f_a\cdot f_b = g_{ab}$,
$\Sigma\equiv C(g)$	real de Sitter-Clifford algebra over de Sitter space V_s,
$\mathscr{T}(f)\equiv\mathscr{D}^c f$	twistor space as left ideal of $\mathscr{D}^c$, generated by $f\in\mathscr{D}^c$ and endowed with Hermitian form $\mathscr{H}$,
$\{u_\alpha, f\}$	twistor and basis $\{f_\alpha\}$, $\alpha = 1,2,3,4$,
θ	symplectic form in $\mathscr{T}$,
$\{e_\alpha, e_{\beta^*}\}$	symplectic basis in $\mathscr{T}$, $\alpha,\beta = 1,2$,
$\mathscr{T}^d\equiv\mathscr{A}(\mathscr{T})$	conjugate dual of $\mathscr{T}$ with respect to $\mathscr{H}$ where $\mathscr{A}$ is generalized Dirac conjugation in $\mathscr{T}$,
i	pseudoscalar of Σ or imaginary unit, $i^2 = -1$,
$\bar{f}$	complex conjugation of f.

[1]M. F. Atiyah, R. Bott, and A. Shapiro, Topology 3, 3 (1964).
[2]N. Bourbaki, *Algèbre* (Hermann, Paris, 1959), Chaps. 8 and 9.
[3]A. Crumeyrolle, Ann. Inst., Poincaré XI, 19 (1969).
[4]A. Crumeyrolle, Ann. Inst. Henri Poincaré XVI, 171 (1972).
[5]M. Riesz, *Clifford Numbers and Spinors*, Lecture Series no. 38, The Institute for Fluid Dynamics and Applied Mathematics, University of Maryland, (1958).
[6]A. Crumeyrolle, Periodica Math. Hungarica 6, 2 (1975).
[7]R. Penrose, Rep. Math. Phys. 12, 65 (1977).
[8]R. Penrose, J. Math. Phys. 8, 345 (1967).
[9]R. Penrose and M. A. H.MacCallum, Phys. Rep. C 6, 4, 241 (1972).
[10]A. Crumeyrolle, Cours et Seminaire du Departement de Mathematique de l'Universite de Toulouse III (1974).
[11]D. Hestenes, *Space-Time Algebra* (Gordon and Breach, New York, 1966).
[12]D. Hestenes and G. Sobczyk, *Geometric Calculus* (Benjamin–Cummings, New York, 1980) (Advanced Book Program in press).
[13]J. Rzewuski, Nuovo Cimento 5, 942 (1958).
[14]E. T. Newman and R. Penrose, J. Math. Phys. 3, 566 (1962).
[15]R. Penrose, "Structure of Space-Time" in *Battelle* (Benjamin, New York, 1968).
[16]J. Rzewuski, Wrocław University Preprint No. 473 (1979).
[17]The utilization of Witt decomposition for the construction of spinor representations of Clifford algebras has been well known for a long time (cf. Ref. 2).
[18]A. O. Barut, Phys. Rev. 135, 3B, B839 (1964).
[19]K. Hayashi, Phys. Lett. B 69, 441 (1962).
[20]T. W. B. Kibble, J. Math. Phys. 2, 221 (1961).

[21]P. K. Smrz, J. Austral. Math. Soc. **19**, 376 (1975); Progr. Theor. Phys. **57**, 1771 (1977); J. Austral. Math. Soc. Ser. B **20**, 38 (1977).
[22]K. Bugajska, Int. J. Theor. Phys. **18**, 2, 77 (1979).
[23]D. Husemoller, *Fiber Bundles* (McGraw-Hill, New York, 1966).
[24]A. Sommerfeld, *Atombau und Spektrallinien*, Vol. 2 (Friedr. Vieweg und Sohn, Brauschweig, 1960).
[25]E. A. Lord, Int. J. Theor. Phys. **13**, 89 (1975).
[26]R. Abłamowicz, Z. Oziewicz, and J. Rzewuski, Bull. Acad. Polon. Sci. Série des Sciences Physiques et Astron. **XXVII**, 3 (1969).
[27]R. Abłamowicz, J. Mozrzymas, Z. Oziewicz, and J. Rzewuski, Rep., Math. Phys. **14**, 89 (1978).
[28]L. P. Hughston, "Twistors and Particles," *Lecture Notes in Physics* 97 (1979).
[29]N. M. J. Woodhouse, in *Group Theoretical Methods in Physics* (Springer-Verlag, Berlin, 1976).
[30]W. A. Hepner, Nuovo Cimento **26**, 352 (1962).
[31]A. O. Barut, "Introduction to de Sitter and Conformal Groups and Their Physical Applications" in *Lectures in Theoretical Physics*, Vol. XIII (Univ. of Colorado Press, Denver, 1971).
[32]A. Ramakrishnan *et al.*, J. Math. Anal. Appl. **27**, 563 (1969).
[33]P. A. M. Dirac, Ann. Math. **37**, 429 (1936).
[34]R. K. Smrz, Canad. J. Phys. **46**, 2073 (1968).
[35]S. Araki and S. Okubo, Lett. Nuovo Cimento **3**, 511 (1972).
[36]Thi Hong Nguyen, Progr. Theor. Phys. **56**, 1647 (1976).
[37]Thi Hong Nguyen, Canad. J. Phys. **57**, 2 (1979).
[38]G. Sobczyk, Wrocław University Preprint no. **444** (1978); J. Math. Phys. **2**, 333 (1981).
[39]I. R. Porteous, *Topological Geometry* (Van Nostrand, London, 1969).
[40]S. Helgason, *Differential Geometry and Symmetric Spaces* (Academic, New York, 1962).
[41]We may write $(s + ip*)^{-1} = \delta(s + ip*)$ since $\delta^{-1} = s^2 + 2sip* + p^2$ is a number ($\delta \neq 0$ by assumption).
[42]Recall that for any $g \in \mathscr{K}$, $(\alpha_s \circ \beta_s)(g) = g^{-1}$.
[43]We emphasize that there is no conformal-invariant-vanishing condition like in the other two projections in order to provide this results.
[44]S. B. Berger, Lett. Nuovo Cimento **11**, 2 (1974).
[45]In Ref. 10 $\mathscr{H}$ was given as $\mathscr{H}(uf,vf)\gamma f = -\beta(\overline{uf})\,vf$, however, change $\beta \to \beta_s$ was allowed.
[46]P. Tod, Rep. Math. Phys. **11**, 3 339 (1977).
[47]A. Crumeyrolle, J. Math. Pures Appl. **56**, 205 (1977).
[48]J. Lukierski, Lett. Nuovo Cimento **24**, 9 309 (1979).

Explicit evaluation of certain Gaussian functional integrals arising in problems of statistical physics

Janusz Adamowski,[a] and Bernd Gerlach
Institut für Physik, Universität Dortmund, D-4600 Dortmund 50, West Germany
Hajo Leschke
Institut für Theoretische Physik, Universität Düsseldorf, D-4000 Düsseldorf 1, West Germany

(Received 20 January 1981; accepted for publication 28 August 1981)

An explicit formula is presented for a (conditional) Wiener integral, the integrand of which is an exponential of a general quadratic functional of the path. The functional integrals arising in non-Markovian Gaussian approximations to various problems of statistical physics (e.g., theory of the large polaron, theory of disordered systems) are easily recovered as special cases.

PACS numbers: 02.50. + s, 05.40. + j

I. INTRODUCTION

The purpose of this paper is the explicit evaluation of a certain class of Gaussian functional integrals or path integrals which are frequently encountered in problems of statistical physics.

The class of integrals considered is given by

$$I_{K,\eta}(x,\beta\,|x',0) := \int \delta R\;\delta(R(\beta)-x)\delta(R(0)-x')\mathrm{e}^{-S[R]} \tag{1a}$$

$$S[R] := \int_0^\beta \mathrm{d}\tau[(\gamma/2)\dot R^{2}(\tau) - R(\tau)\eta(\tau)] + \frac{1}{2}\int_0^\beta d\tau\int_0^\beta d\tau'\, K(\tau,\tau')R(\tau)R(\tau'). \tag{1b}$$

Following the convention of most physicists[1] we have used the symbolic notation

$$\int \delta R\delta(R(\beta)-x)\delta(R(0)-x')\exp\left\{-\frac{\gamma}{2}\int_0^\beta d\tau\,\dot R^{2}(\tau)\right\}$$

to indicate (conditional) Wiener integration[2-4] over paths $R(\tau)$ of one-dimensional (pinned) Brownian motion with diffusion constant $1/2\gamma$ starting at time 0 from $x' = R(0)$ and arriving at $x = R(\beta)$ at time $\beta > 0$. The dot denotes (formal) differentiation with respect to τ. The functions $\eta(\tau)$ and $K(\tau,\tau')$ [$= K(\tau',\tau)$, without loss of generality] may be quite general but are, of course, subject to the condition that the integral exists.

Apart from a pure mathematical interest much of the motivation for considering the integral (1) comes from the fact that it naturally arises in Gaussian approximations to various physical problems with a nonadditive "action" functional[5] (being characteristic of non-Markovian behavior). Problems of this type are: the calculation of the energy and mobility of the large polaron,[6-10] the density of electronic states in disordered systems,[11,12] the propagation of waves in random media,[13] etc. In these circumstances $K(\tau,\tau')$ typically is some (approximate) memory kernel and $\eta(\tau)$ serves as a source function to generate the corresponding Gaussian averages via functional differentiation.

The physical interpretations and dimensions of the variables entering the integral via the quadratic "action" functional $S[R]$ depend on the problem one wants to study. For example, in quantum statistical mechanics $\gamma\hbar^2$ is the mass of some particle and $1/\beta k_B$ the absolute temperature ($2\pi\hbar$: Planck's constant, k_B: Boltzmann's constant).

Despite the common dictum that Gaussian functional integrals can be done, we have found no reference where the above integral has been done in sufficient generality and explicitness. Either special cases are considered from the very beginning[6-12] and/or the result is contained in theorems,[4] which look nice but leave the evaluation's hard part to the user if he wants to particularize.

For all these reasons we have found it appropriate to present a closed-form expression for the "value" of the Gaussian functional integral (1) from which previously known explicit results are easily derived as special cases.

In some sense this paper is complementary to a recent work[14] in this Journal on the explicit evaluation of path integrals with a general quadratic but single-time action.

The plan of this paper is as follows. In Sec. II we show that the computation of the integral can be reduced to that of the minimal value of the action functional. In Sec. III we establish the general form of the minimal action and of the resulting expression for the integral. By restricting ourselves to "β-periodic" kernels K in Sec. IV we are able to turn the general expressions into explicit ones. Finally, Sec. V is devoted to an example, which contains as limiting cases some of the explicit results available in the literature.

II. REDUCTION TO THE MINIMAL ACTION

In this section we will reduce the computation of the functional integral (1) to the computation of the minimal action $\bar S(x,x')$, which is the action functional $S[R]$ evaluated at its stationary or the "most probable" path $\bar R(\tau)$ from x' to x.

We start from

$$-\gamma\ddot{\bar R}(\tau) + \int_0^\beta d\tau' K(\tau,\tau')\bar R(\tau') = \eta(\tau), \tag{2a}$$

$$\bar R(0) = x', \quad \bar R(\beta) = x, \tag{2b}$$

$$\bar S(x,x') := S[\bar R]. \tag{3}$$

[a] Permanent and present address: Solid State Physics Department, Academy of Mining and Metallurgy—AGH, 30-059 Kraków, Poland.

Since S is (at most) quadratic in R, the linear substitution

$$R(\tau)\rightarrow R(\tau)-\bar{R}(\tau)$$

transforms the integral (1) into a corresponding one with "vanishing boundary conditions" and vanishing source function, i.e.,

$$I_{K,\eta}(x,\beta\,|x',0)=\mathrm{e}^{-\bar{S}(x,x')}I_{K,0}(0,\beta\,|0,0). \tag{4}$$

Identities of this genre and their derivations along the somewhat formal lines indicated are of course well known.[1,15] Here we want to point out that Eq. (4) and also the subsequent equation can be rigorously justified (under certain technical assumptions), e.g., in the framework of the usual (sequential-limit) definitions[16] of Wiener integrals or within DeWitt–Morette's approach[17] via the Fourier transform of an appropriate prodistribution.

In the remainder of this section we assume $\bar{S}(x,x')$ to be known, in particular as a functional of K and η.

As is suggested from finite dimensional Gaussian integrals, the remaining integral in Eq. (4) is essentially the square root of an infinite dimensional determinant. Here we want to stress that it can be derived from $\bar{S}(0,0)$. To show this, we replace K by λK, λ being a positive parameter. Accordingly, $\bar{S}$ changes to $\bar{S}_\lambda$. Now consider the quantity

$$J_\eta(\lambda):=\ln I_{\lambda K,\eta}(0,\beta\,|0,0), \tag{5}$$

with K and β fixed as a function of λ and a functional of η. According to Eq. (4) we will only need $J_0(1)$ in the final stage. Writing

$$J_0(1)=J_0(0)+\int_0^1 d\lambda\ \frac{\partial J_0(\lambda)}{\partial\lambda} \tag{6}$$

we have to specify the "initial" value $J_0(0)$ and the derivative $\partial J_0(\lambda)/\partial\lambda$. From the normalization factor of the transition-probability density of the Wiener process or, equivalently, from the canonical density matrix of a free particle, it is well known[1-3] that

$$J_0(0)=\ln(\gamma/2\pi\beta)^{1/2}. \tag{7}$$

Concerning the derivative we will use the identity

$$\frac{\partial J_\eta(\lambda)}{\partial\lambda}=-\frac{1}{2}\int_0^\beta d\tau\int_0^\beta d\tau'\,K(\tau,\tau')e^{-J_\eta(\lambda)}\frac{\delta^2}{\delta\eta(\tau)\delta\eta(\tau')}e^{J_\eta(\lambda)}, \tag{8}$$

which is an immediate consequence of the definitions (1) and (5). From Eq. (4) we have (with $K\rightarrow\lambda K$ and $x=x'=0$)

$$e^{J_\eta(\lambda)}=e^{-\bar{S}_\lambda(0,0)}e^{J_0(\lambda)}. \tag{9}$$

Since $J_0(\lambda)$ is independent of η, we finally get

$$\frac{\partial J_0(\lambda)}{\partial\lambda}=\frac{1}{2}\int_0^\beta d\tau\int_0^\beta d\tau'\,K(\tau,\tau')\times\left[\frac{\delta^2\bar{S}_\lambda(0,0)}{\delta\eta(\tau)\delta\eta(\tau')}-\frac{\delta\bar{S}_\lambda(0,0)}{\delta\eta(\tau)}\frac{\delta\bar{S}_\lambda(0,0)}{\delta\eta(\tau')}\right]\Bigg|_{\eta=0}. \tag{10}$$

III. GENERAL FORM OF THE RESULT

According to Sec. II the basic quantity to calculate is the minimal action $\bar{S}(x,x')$ defined by Eqs. (2) and (3). In this section we will first establish the general form of the solution $\bar{R}(\tau)$ of the variational problem (2) and then that of the resulting minimal action. As a consequence we will get a preliminary expression for the value of the integral (1) which gives the general structure and is in "closed form" but is not yet fully explicit.

We start by observing that $\bar{R}(\tau)$ may be written as

$$\bar{R}(\tau)=\mu(\tau)+\int_0^\beta d\tau'\,C(\tau,\tau')\eta(\tau'). \tag{11}$$

Here $\mu(\tau)$ is the solution of Eqs. (2) for the homogeneous case $\eta=0$, and $C(\tau,\tau')$ is the corresponding reciprocal kernel or Green's function with "vanishing boundary conditions":

$$-\gamma\frac{\partial^2}{\partial\tau^2}C(\tau,\tau')+\int_0^\beta d\tau''\,K(\tau,\tau'')C(\tau'',\tau')=\delta(\tau-\tau'), \tag{12a}$$

$$C(0,\tau')=C(\beta,\tau')=0. \tag{12b}$$

For the time being we get from Eqs. (2) and (3), upon integrating by parts,

$$\bar{S}(x,x')=(\gamma/2)[x\dot{\bar{R}}(\beta)-x'\dot{\bar{R}}(0)]-\frac{1}{2}\int_0^\beta d\tau\ \eta(\tau)\bar{R}(\tau). \tag{13}$$

Multiplying Eq. (2a) by $\mu(\tau)$, integrating the resulting equation with respect to τ, integrating the term containing $\ddot{\bar{R}}(\tau)$ two times by parts, and using the symmetry $K(\tau,\tau')=K(\tau',\tau)$ and the fact that $\mu(\tau)$ solves Eqs. (2) for $\eta=0$, we find the identity

$$(\gamma/2)[x\dot{\bar{R}}(\beta)-x'\dot{\bar{R}}(0)]=(\gamma/2)[x\dot{\mu}(\beta)-x'\dot{\mu}(0)]-\tfrac{1}{2}\int_0^\beta d\tau\ \eta(\tau)\mu(\tau). \tag{14}$$

Now Eqs. (13) and (14) combine via Eq. (11) to the general form of the minimal action we are looking for

$$\bar{S}(x,x')=(\gamma/2)[x\dot{\mu}(\beta)-x'\dot{\mu}(0)]-\int_0^\beta d\tau\,\mu(\tau)\eta(\tau)-\tfrac{1}{2}\int_0^\beta d\tau\int_0^\beta d\tau'\,C(\tau,\tau')\eta(\tau)\eta(\tau'). \tag{15}$$

For later purpose it is convenient to have those terms, which are independent of η, rearranged in accord with

$$(\gamma/2)[x\dot{\mu}(\beta)-x'\dot{\mu}(0)]=(\gamma/2\beta)(x-x')^2+\tfrac{1}{2}\int_0^\beta d\tau\int_0^\beta d\tau'\,K(\tau,\tau')\mu_0(\tau)\mu(\tau'), \tag{16}$$

where

$$\mu_0(\tau):=x'+(x-x')(\tau/\beta) \tag{17}$$

is $\mu(\tau)$ for $K=0$. Equation (16) follows from the defining equations for $\mu(\tau)$ and from Eq. (17) after integrating two times by parts.

Having established the form (15) for the minimal action, it is easy to give probabilistic interpretations of the functions $\mu(\tau)$ and $C(\tau,\tau')$. Consider the expectation value

$$\langle A\rangle:=\int\delta R\ \delta(R(\beta)-x)\delta(R(0)-x')P[R]A[R] \tag{18}$$

of a general functional $A[R]$ with respect to the "probability density" functional

$$P[R]: = \left. \frac{\exp(-S[R])}{\int \delta R\, \delta(R(\beta)-x)\delta(R(0)-x')\exp(-S[R])} \right|_{\eta=0}. \tag{19}$$

The associated generating functional

$$Z[\eta] := \langle \exp\left\{\int_0^\beta d\tau\, R(\tau)\eta(\tau)\right\}\rangle \tag{20}$$

is found to be

$$Z[\eta] = \frac{I_{K,\eta}(x,\beta\,|x',0)}{I_{K,0}(x,\beta\,|x',0)} = \exp\left\{\int_0^\beta d\tau\, \mu(\tau)\eta(\tau) + \tfrac{1}{2}\int_0^\beta d\tau \int_0^\beta d\tau'\, C(\tau,\tau')\eta(\tau)\eta(\tau')\right\} \tag{21}$$

according to Eqs. (1), (4), and (15). Hence, P controls a Gaussian stochastic process with mean

$$\langle R(\tau)\rangle = \left.\frac{\delta Z}{\delta\eta(\tau)}\right|_{\eta=0} = \mu(\tau) \tag{22}$$

and covariance

$$\langle R(\tau)R(\tau')\rangle - \langle R(\tau)\rangle\langle R(\tau')\rangle = \left.\frac{\delta^2 Z}{\delta\eta(\tau)\delta\eta(\tau')}\right|_{\eta=0} - \mu(\tau)\mu(\tau') = C(\tau,\tau'). \tag{23}$$

The boundary condition (12b) is a reflection of the fact that the fluctuations around the average path vanish at both endpoints and are independent of their respective values.

For many purposes (mainly in quantum statistics) one considers instead of the above average $\langle\cdot\rangle$ over paths with both endpoints fixed a "trace-like" average $\langle\cdot\rangle_{\sim}$ over closed paths. It is characterized by the following generating functional:

$$\tilde{Z}[\eta] := \frac{\int dx\, I_{K,\eta}(x,\beta\,|x,0)}{\int dx\, I_{K,0}(x,\beta\,|x,0)}. \tag{24}$$

In order to perform the integations over x in Eq. (24), we exploit the fact that $\mu(\tau)$ for $x' = x$ varies linearly with x, i.e.,

$$\mu(\tau) = x\,\rho(\tau), \text{ for } \quad x' = x, \tag{25}$$

where $\rho(\tau)$ is the homogeneous solution of Eq. (2a) subject to the boundary conditions

$$\rho(0) = \rho(\beta) = 1. \tag{26}$$

From Eqs. (4), (15), and (25) we see that $I_{K,\eta}(x,\beta\,|x,0)$ is a Gaussian of x. Under the assumption

$$\dot\rho(\beta) - \dot\rho(0) > 0 \tag{27}$$

we can perform the corresponding integrations and end up with

$$\tilde{Z}[\eta] = \exp\left\{\frac{1}{2}\int_0^\beta d\tau \int_0^\beta d\tau'\, \tilde{C}(\tau,\tau')\eta(\tau)\eta(\tau')\right\}, \tag{28}$$

where

$$\tilde{C}(\tau,\tau') := C(\tau,\tau') + \frac{1}{\gamma}\frac{\rho(\tau)\rho(\tau')}{\dot\rho(\beta)-\dot\rho(0)} = \langle R(\tau)R(\tau')\rangle_{\sim}. \tag{29}$$

In the two types of averages considered above the actual value of the "remaining integral" $I_{K,0}(0,\beta\,|0,0)$ in Eq. (4) played no role. It dropped out as a sort of normalization factor. Nevertheless its computation according to the recipe given in Sec. II can be carried one step further. In fact, making use of Eqs. (15) and (25) the functional derivatives of the minimal action in Eq. (10) can be taken to yield

$$\frac{\partial J_0(\lambda)}{\partial\lambda} = -\frac{1}{2}\int_0^\beta d\tau \int_0^\beta d\tau'\, K(\tau,\tau')C_\lambda(\tau',\tau). \tag{30}$$

Here C_λ is C for K replaced by λK.

It is instructive to exhibit the general structure of the right-hand side of Eq. (30) at a formal, algebraic level. Let us interpret $\delta(\tau-\tau')$, $K(\tau,\tau')$, and $C_\lambda(\tau,\tau')$ as kernels of (symmetric) integral operators ∂^2, $\hat K$, and $\hat C_\lambda$, respectively. Then we can formally write

$$\hat C_\lambda = (-\gamma\partial^2 + \lambda\hat K)^{-1} \tag{31}$$

and

$$\frac{\partial J_0(\lambda)}{\partial\lambda} = -\frac{1}{2}\operatorname{tr}\hat K\hat C_\lambda = -\frac{1}{2}\frac{\partial}{\partial\lambda}\operatorname{tr}\ln(-\gamma\partial^2+\lambda\hat K) = -\frac{1}{2}\frac{\partial}{\partial\lambda}\ln\det(-\gamma\partial^2+\lambda\hat K) \tag{32}$$

so that

$$J_0(1) = J_0(0) - \frac{1}{2}\ln\det(1+\hat K\hat C_0). \tag{33}$$

Employing the definition (5) and Eq. (7) we eventually arrive at

$$I_{K,0}(0,\beta\,|0,0) = (\gamma/2\pi\beta)^{1/2}[\det(1+\hat K\hat C_0)]^{-1/2}. \tag{34}$$

Here $\hat C_0$, according to Eqs. (31) and (12), has the kernel

$$C_0(\tau,\tau') = (1/\gamma)(\min\{\tau,\tau'\} - \tau\tau'/\beta). \tag{35}$$

While C_0 is the covariance, μ_0 [see Eq. (17)] is the mean of (pinned) Brownian motion.

Taken together, Eqs. (4), (15), (16), and (34) give the general form of the value of the functional integral (1):

$$I_{K,\eta}(x,\beta\,|x',0) = \left(\frac{\gamma}{2\pi\beta}\right)^{1/2}[\det(1+\hat K\hat C_0)]^{-1/2}\exp\Big\{-\frac{\gamma}{2\beta}(x-x')^2 - \frac{1}{2}\int_0^\beta d\tau\int_0^\beta d\tau'\, K(\tau,\tau')\mu_0(\tau)\mu(\tau') + \int_0^\beta d\tau\,\mu(\tau)\eta(\tau) + \frac{1}{2}\int_0^\beta d\tau\int_0^\beta d\tau'\, C(\tau,\tau')\eta(\tau)\eta(\tau')\Big\}. \tag{36}$$

Clearly, this is the result one can guess at from formal analogy to finite dimensional Gaussian integrals.

Concerning the infinite dimensional determinant in Eqs. (34) and (36) we want to remark the following. All that we actually have done is to use the left-hand side of the equation

$$\det(1+\hat K\hat C_0) = \exp\left\{\int_0^1 d\lambda \int_0^\beta d\tau \int_0^\beta d\tau'\, K(\tau,\tau')C_\lambda(\tau',\tau)\right\} \tag{37}$$

as a suggestive abbreviation for its right-hand side. This can be seen by comparing the expression (34) for $I_{K,0}(0,\beta\,|0,0)$ with the expression directly obtained from Eqs. (5), (6), (7), and (30). Although it is possible to give an "intrinsic" definition of the determinant,[18] the problem often remains to find a more explicit expression for it. We will accomplish this in the next section for a large class of kernels K by computing the right-hand side of Eq. (37).

IV. EXPLICIT RESULT FOR β-PERIODIC KERNELS

The closed-form expression of Sec. III becomes fully explicit, when the most probable path $\bar{R}(\tau)$ or, equivalently, the mean $\mu(\tau)$ and the covariance $C(\tau,\tau')$ are explicitly known. Clearly, for general kernels $K(\tau,\tau')$ this is not possible, because nobody can explicitly solve the most general, though linear integral equation. However, we can proceed for a restricted class of kernels which is large enough to cover many applications in statistical physics and, in particular, those mentioned in the introduction. This class is formed by kernels of the type

$$K(\tau,\tau')=f(\tau-\tau'), \tag{38}$$

where f is a real valued, even function in the interval $[-\beta,\beta]$ supposed to fulfill

$$f(\tau-\beta)=f(\tau), \quad \text{for} \quad \tau\in[o,\beta]. \tag{39}$$

The last equation represents a necessary condition for the action functional S (with $\int_0^\beta d\tau\,\eta(\tau)=0$) to be invariant with respect to constant translations

$$R(\tau)\to R(\tau)+\text{const},$$

which implies that the functional integral (1) depends on x and x' only through the difference $x-x'$. Conversely, translation invariance follows, if Eq. (39) holds and additionally

$$\int_0^\beta d\tau\, f(\tau)=0. \tag{40}$$

In the remainder of this paper we will derive explicit results exclusively for kernels of the type (38). While Eq. (39) is imposed throughout, Eq. (40) may hold or not. We will refer to these kernels as β periodic, because Eq. (39) renders possible an extension of f to the real line, which has period β.

For β-periodic kernels it is possible to solve Eqs. (2) by Fourier analysis. We have found it convenient to write $\bar{R}(\tau)$ in the form

$$\bar{R}(\tau)=\mu_0(\tau)-\frac{B_0}{2\gamma}\tau(\tau-\beta)+\frac{1}{\gamma}\sum_{n\neq 0}\frac{B_n}{\nu_n^2}(e^{i\nu_n\tau}-1), \tag{41}$$

where the sum goes over all integers n except zero, $\mu_0(\tau)$ is defined in Eq. (17), and

$$\nu_n:=2\pi n/\beta \tag{42}$$

denote the nth Fourier frequency. Obviously, the ansatz (41) obeys the boundary condition (2b). For the unknown coefficients B_n Eq. (2a) requires

$$B_n=\eta_n-f_n\int_0^\beta d\tau\, e^{-i\nu_n\tau}\bar{R}(\tau), \tag{43}$$

where

$$\eta_n:=\frac{1}{\beta}\int_0^\beta d\tau\, e^{-i\nu_n\tau}\eta(\tau),$$
$$f_n:=\frac{1}{\beta}\int_0^\beta e^{-i\nu_n\tau}f(\tau)=f_{-n} \tag{44}$$

are the Fourier coefficients of η and f, respectively. We can determine the coefficients B_n explicitly by inserting Eq. (41) into Eq. (43) and performing the integration over τ term by term. The result is

$$B_0\left[1+\sum_{n\neq 0}\frac{\beta f_0}{\gamma\nu_n^2+\beta f_n}\right]=\eta_0-\beta f_0\left[\frac{1}{2}(x+x')-\sum_{n\neq 0}\frac{\eta_n}{\gamma\nu_n^2+\beta f_n}\right] \tag{45}$$

and for $n\neq 0$

$$B_n\left[1+\frac{\beta f_n}{\gamma\nu_n^2}\right]=\eta_n+\frac{x-x'}{i\nu_n}f_n+\frac{\beta f_n}{\gamma\nu_n^2}B_0. \tag{46}$$

From Eqs. (41), (45), and (46), for $\bar{R}(\tau)$, we can infer explicit expressions for the mean $\mu(\tau)$ and the covariance $C(\tau,\tau')$. According to Eq. (11), $\mu(\tau)$ is simply $\bar{R}(\tau)$ for $\eta=0$ and $C(\tau,\tau')$ equals the functional derivative $\delta\bar{R}(\tau)/\delta\eta(\tau')$. Since the resulting expressions are somewhat lengthy, we leave it to the reader to write them down. However, at the cost of having *a priori* no longer pointwise (let alone uniform) convergence for $\tau=0$ and $\tau=\beta$, we can turn these expressions into fairly compact ones by using the Fourier expansions of $\mu_0(\tau)$ and $\tau(\tau-\beta)$ and rearranging the infinite series after one or two partial fraction decompositions. In this way we get

$$\mu(\tau)=\frac{1}{2}(x+x')\frac{1+\beta^2 f_0 D(\tau)}{1+\beta^2 f_0 D(0)}+(x-x')\gamma\dot{D}(\tau) \tag{47}$$

$$C(\tau,\tau')=D(\tau-\tau')+D(0)-D(\tau)-D(\tau')-\frac{\beta^2 f_0}{1+\beta^2 f_0 D(0)}[D(\tau)-D(0)][D(\tau')-D(0)], \tag{48}$$

where we have introduced the function

$$D(\tau-\tau'):=\frac{1}{\beta}\sum_{n\neq 0}\frac{e^{i\nu_n(\tau-\tau')}}{\gamma\nu_n^2+\beta f_n}, \tag{49}$$

which is basic for β-periodic kernels.

While in Eq. (48) the boundary conditions (12b) are manifest, the boundary conditions (2b) are obeyed by Eq. (47) only *a posteriori*, because $D(\tau)=D(-\tau)$ is not differentiable at $\tau=0$ (and $\tau=\beta$) due to a cusp. But, of course, the interpretations

$$\dot{D}(\beta-0)=-\dot{D}(+0)=1/2\gamma \tag{50}$$

are natural and sufficient.

If we distinguish explicitly between $f_0=0$ (translation invariance) and $f_0\neq 0$, we can further simplify Eqs. (47) and (48). The simplifications occurring for $f_0=0$ are obvious. For $f_0\neq 0$ we can write

$$\mu(\tau)=\tfrac{1}{2}(x+x')[\tilde{D}(\tau)/\tilde{D}(0)]+(x-x')\gamma\dot{\tilde{D}}(\tau), \tag{51}$$

$$C(\tau,\tau')=\tilde{D}(\tau-\tau')-\tilde{D}(\tau)\tilde{D}(\tau')/\tilde{D}(0), \tag{52}$$

where

$$\tilde{D}(\tau-\tau') := D(\tau-\tau') + 1/\beta^2 f_0 = \frac{1}{\beta}\sum_n \frac{e^{i\nu_n(\tau-\tau')}}{\gamma\nu_n^2+\beta f_n} \tag{53}$$

is the sum in the definition (49) with the term corresponding to $n=0$ included.

The expressions found for μ and C are explicit in the sense that they are reduced to the single function D. The more explicit our knowledge of D is, the more explicit is our knowledge of μ and C.

In order to complete the particularization of the general form (36) for β-periodic kernels, it remains to specify the η-independent terms in the exponent and the determinant. The former can be found either directly from Eqs. (41), (45), and (46) or from Eqs. (16) and (47). In any case one gets

$$\begin{aligned}\frac{\gamma}{2}[x\dot{\mu}(\beta) - x'\dot{\mu}(0)] &= \frac{\gamma}{2\beta}(x-x')^2 + \frac{1}{2}\int_0^\beta d\tau\int_0^\beta d\tau' f(\tau-\tau')\mu_0(\tau)\mu(\tau') \\ &= \frac{\gamma}{2\beta}(x-x')^2\left[1+\sum_{n\neq 0}\frac{\beta f_n}{\gamma\nu_n^2+\beta f_n}\right] \\ &\quad + \frac{1}{8}(x+x')^2\frac{\beta^2 f_0}{1+\beta^2 f_0 D(0)}.\end{aligned} \tag{54}$$

For the determinant one obtains from Eqs. (37) and (48) by a straightforward calculation

$$\det(1+\hat{K}\hat{C}_0) = [1+\beta^2 f_0 D(0)]\left[\prod_{n=1}^{\infty}\left(1+\frac{\beta f_n}{\gamma\nu_n^2}\right)\right]^2. \tag{55}$$

Equations (47)–(50), (54), and (55) in combination with the general formula (36) constitute the explicit result for β-periodic kernels which we want to present.

From the above expressions one can read off the following necessary and sufficient conditions for the existence of the functional integral (1) in the case of β-periodic kernels (and appropriate source functions η):

$$1+\beta f_n/\gamma\nu_n^2 > 0, \quad n\neq 0 \tag{56a}$$

$$1+\beta^2 f_0 D(0) > 0, \tag{56b}$$

$$\sum_{n\neq 0}\frac{\beta|f_n|}{\gamma\nu_n^2} < \infty. \tag{56c}$$

These conditions are in agreement with those given in Ref. 17.

We close this section by specifying the covariance $\tilde{C}(\tau,\tau')$ for β-periodic kernels. From Eq. (54) we see that the condition (27) becomes equivalent to the requirement $f_0 > 0$. Specializing Eqs. (51) and (54) to $x' = x$ we find from Eqs. (29) and (52)

$$\tilde{C}(\tau,\tau') = \tilde{D}(\tau-\tau'). \tag{57}$$

Hence, although we have introduced D (or $\tilde{D}$) mainly as a convenient abbreviation, it has a direct probabilistic meaning.

V. EXAMPLE

For illustrative purposes and in order to make contact with some of the explicit results available in the literature, we use this section to particularize the expressions of Secs. III and IV for a special class of β-periodic kernels. This class is characterized by

$$f(\tau-\tau') = \gamma(\epsilon^2+E^2)\delta(\tau-\tau') - \gamma E^2 M_W(\tau-\tau'). \tag{58}$$

Here ϵ, E, and W are three positive parameters. The overall constant γ is present for dimensional reasons. The function

$$M_W(\tau-\tau') := \frac{W\cosh(\beta/2 - |\tau-\tau'|)W}{2\sinh(\beta W/2)} \tag{59}$$

serves to model "memory effects" on a "time scale" of the order $1/W$. Note the properties

$$\int_0^\beta d\tau' M_W(\tau-\tau') = 1, \tag{60a}$$

$$\lim_{W\downarrow 0} M_W(\tau-\tau') = 1/\beta, \tag{60b}$$

$$\lim_{W\to\infty} M_W(\tau-\tau') = \delta(\tau-\tau'). \tag{60c}$$

The kernel (58) leads to the following quadratic terms in the action functional (1b):

$$\begin{aligned}\tfrac{1}{2}\int_0^\beta d\tau\int_0^\beta d\tau' f(\tau-\tau')R(\tau)R(\tau') &= \tfrac{1}{2}\gamma\epsilon^2\int_0^\beta d\tau R^2(\tau) + \tfrac{1}{4}\gamma E^2\int_0^\beta d\tau\int_0^\beta d\tau' \\ &\quad \times M_W(\tau-\tau')|R(\tau)-R(\tau')|^2.\end{aligned} \tag{61}$$

The first (translation noninvariant) term induces the (conditional) Uhlenbeck–Ornstein or harmonic-oscillator process. The second (translation invariant) term has been extensively used in Gaussian theories of an electron, which is either coupled to optical lattice vibrations[6–10] ("polaron") or moves in a random potential[11,12] ("disordered system"). In the latter case mainly the extreme non-Markovian limit $W\to 0$ has been considered.[12]

Let us now come to the particularization of the expression in Sec. IV for the kernel (58). We start from the Fourier expansion

$$M_W(\tau-\tau') = \frac{W^2}{\beta}\sum_n \frac{e^{i\nu_n(\tau-\tau')}}{\nu_n^2+W^2}, \tag{62}$$

which is in fact a generalization of the Mittag–Leffler expansion of the hyperbolic function $\coth(\beta W/2)$. From this relation we can read off the Fourier coefficients of f:

$$\beta f_n = \gamma\epsilon^2 + \frac{\gamma E^2\nu_n^2}{\nu_n^2+W^2}. \tag{63}$$

By a partial fraction decomposition we find

$$\frac{\gamma}{\gamma\nu_n^2+\beta f_n} = \frac{1}{W_+^2-W_-^2}\left(\frac{W_+^2-W^2}{\nu_n^2+W_+^2}+\frac{W^2-W_-^2}{\nu_n^2+W_-^2}\right), \tag{64}$$

where the positive numbers W_+ and W_- are defined through

$$2W_\pm^2 := \epsilon^2+E^2+W^2 \pm [(\epsilon^2+E^2+W^2)^2 - 4\epsilon^2W^2]^{1/2}. \tag{65}$$

As an immediate consequence of Eqs. (62) and (64) we get the $\tilde{D}$ function associated with the kernel (58):

$$\tilde{D}(\tau-\tau') = \frac{1}{\gamma(W_+^2 - W_-^2)} \times\left[\left(1-\frac{W^2}{W_+^2}\right)M_{W_+}(\tau-\tau') + \left(\frac{W^2}{W_-^2}-1\right)M_{W_-}(\tau-\tau')\right]. \quad (66)$$

According to Eqs. (53) and (63) it is related to the D function via

$$D(\tau-\tau') = \tilde{D}(\tau-\tau') - 1/\gamma\beta\epsilon^2. \quad (67)$$

While $\tilde{D}$ becomes singular in the limit $\epsilon\to 0$, D does not. Noting that $W_+W_- = \epsilon W$, one finds

$$\lim_{\epsilon\downarrow 0} D(\tau-\tau') = \frac{1}{\gamma V^2}\left(1-\frac{W^2}{V^2}\right)\left[M_V(\tau-\tau') - \frac{1}{\beta}\right] + \frac{W^2}{\gamma V^2}\left[\frac{1}{2\beta}|\tau-\tau'|(|\tau-\tau'|-\beta) + \frac{\beta}{12}\right], \quad (68)$$

where

$$V := (E^2 + W^2)^{1/2}. \quad (69)$$

Inserting Eq. (66) [resp. (67)] into Eqs. (51) and (52) [resp. (47) and (48)] we get the mean μ and the covariance C associated with the kernel (58). The covariance $\tilde{C}$ is directly given by Eq. (66) according to Eq. (57).

It remains to specialize Eqs. (54) and (55) to the kernel (58). Because of $\beta^2 f_0 = \gamma\beta\epsilon^2$ Eqs. (66), (67), and (59) give for $\tau=\tau'$ and $\epsilon>0$:

$$1+\beta^2 f_0 D(0) = \gamma\beta\epsilon^2\tilde{D}(0) = \frac{1}{W_+^2 - W_-^2}\left[(\epsilon^2 - W_-^2)\frac{\beta W_+}{2}\coth\frac{\beta W_+}{2} + (W_+^2 - \epsilon^2)\frac{\beta W_-}{2}\coth\frac{\beta W_-}{2}\right]. \quad (70)$$

Similarly, from a partial fraction decomposition analogous to Eq. (64), we find

$$1+\sum_{n\neq 0}\frac{\beta f_n}{\gamma\nu_n^2 + \beta f_n} = \frac{1}{W_+^2 - W_-^2}\left[(W_+^2 - W^2)\frac{\beta W_+}{2}\coth\frac{\beta W_+}{2} + (W^2 - W_-^2)\frac{\beta W_-}{2}\coth\frac{\beta W_-}{2}\right]. \quad (71)$$

Finally we observe

$$1+\frac{\beta f_n}{\gamma\nu_n^2} = \frac{(1+W_+^2/\nu_n^2)(1+W_-^2/\nu_n^2)}{1+W^2/\nu_n^2}, \qquad n\neq 0, \quad (72)$$

so that by the Weierstrass–Hadamard factorization of the hyperbolic sine we get

$$\prod_{n=1}^{\infty}\left(1+\frac{\beta f_n}{\gamma\nu_n^2}\right) = \frac{2}{\beta\epsilon}\,\frac{\sinh(\beta W_+/2)\sinh(\beta W_-/2)}{\sinh(\beta W/2)}. \quad (73)$$

Equations (70), (71), and (73) inserted into Eqs. (54) and (55) complete the particularization of the general expressions for the kernel (58). The resulting value for the corresponding functional integral, although already considerably specialized in comparison with Sec. IV, still includes the limiting cases $E\to 0$ ("harmonic oscillator") and $\epsilon\to 0$ ("polaron", "disordered system") considered in the literature. In particular, for the latter case we find by collecting the above results:

$$\int \delta R\,\delta(R(\beta)-x)\delta(R(0)-x')\exp\left\{-\frac{\gamma}{2}\int_0^\beta d\tau\,\dot{R}^2(\tau) - \tfrac{1}{4}\gamma E^2\int_0^\beta d\tau\int_0^\beta d\tau'\,M_W(\tau-\tau')|R(\tau)-R(\tau')|^2 + \int_0^\beta d\tau\,R(\tau)\eta(\tau)\right\}$$
$$= \left(\frac{\gamma}{2\pi\beta}\right)^{1/2}\frac{V}{W}\frac{\sinh(\beta W/2)}{\sinh(\beta V/2)}\exp\left\{-\frac{\gamma}{2\beta}(x-x')^2 \times\left[\frac{W^2}{V^2} + \left(1-\frac{W^2}{V^2}\right)\frac{\beta V}{2}\coth\frac{\beta V}{2}\right] + \int_0^\beta d\tau\,\mu(\tau)\eta(\tau) + \tfrac{1}{2}\int_0^\beta d\tau\int_0^\beta d\tau'\,C(\tau,\tau')\eta(\tau)\eta(\tau')\right\}, \quad (74)$$

with

$$\mu(\tau) = \tfrac{1}{2}(x+x') - \tfrac{1}{2}(x-x') \times\left[\left(1-\frac{W^2}{V^2}\right)\frac{\sinh(\beta/2-\tau)V}{\sinh(\beta V/2)} + \frac{W^2}{V^2}\left(1-\frac{2\tau}{\beta}\right)\right] \quad (75)$$

and

$$C(\tau,\tau') = \frac{2}{\gamma V}\left(1-\frac{W^2}{V^2}\right)\frac{\cosh\frac{1}{2}(\tau-\tau')V}{\sinh(\beta V/2)}\sinh\frac{\tau_- V}{2} \times\sinh\tfrac{1}{2}(\beta-\tau_+)V + \frac{W^2}{\gamma V^2}\left(\tau_- - \frac{\tau\tau'}{\beta}\right). \quad (76)$$

Here V is defined in Eq. (69) and

$$\tau_+ := \max\{\tau,\tau'\}, \quad \tau_- := \min\{\tau,\tau'\} \quad (77)$$

denotes the larger and the smaller one of the two times τ and τ', respectively.

For results in the literature corresponding to Eqs. (74)–(77) see Refs. 9 and 11. The limiting case $W\to 0$ (with $\eta = 0$) has also been computed in Refs. 5, 12, and 19. The subsequent limit $V\to 0$ (i.e., $E\to 0$) gives the "free particle" case, which is determined by Eqs. (7), (17), and (35). As it should be, this case can be obtained alternatively and more directly by letting $V\to W$.

VI. ADDITIONAL REMARKS

Most of our results have an obvious generalization to the case of Wiener integrals over paths $R(\tau)$ in multi-dimensional space. Moreover, they can be formally turned into results for Feynman path integrals by going over to imaginary τ.

ACKNOWLEDGMENT

This work was supported in part by the Deutsche Forschungsgemeinschaft.

[1]R. P. Feynman and A. Hibbs, *Quantum Mechanics and Path Integrals* (McGraw-Hill, New York, 1965).
[2]I. M. Gel'fand and A. M. Yaglom, Usp. Mat. Nauk. **11**, 77 (1956) [J. Math. Phys. **1**, 48 (1960)].
[3]L. Streit, Acta Phys. Austriaca Suppl. **2**, 2 (1966).
[4]B. Simon, *Functional Integration and Quantum Physics* (Academic, New York, 1979).
[5]V. M. Chetverikov, Teor. Mat. Fiz. **24**, 211 (1975) [Theor. Math. Phys. **24**, 780 (1975)].
[6]R. P. Feynman, Phys. Rev. **97**, 660 (1955).
[7]M. A. Krivoglaz and S. I. Pekar, Izv. Akad. Nauk. SSSR, Ser. Fiz. **21**, 16 (1957) [Bull. Acad. Sci. USSR, Phys. Ser. **21**, 13 (1957)]; Y. Ōsaka, Prog. Theor. Phys. **22**, 437 (1959).
[8]R. W. Hellwarth and P. M. Platzman, Phys. Rev. **128**, 1599 (1962); K. K. Thornber, Phys. Rev. B **3**, 1929 (1971).
[9]V. Sa-yakanit, Phys. Rev. B **19**, 2377 (1979).
[10]J. Adamowski, B. Gerlach, and H. Leschke, in *Functional Integration, Theory and Aplications*, edited by J.-P. Antoine and E. Tirapegui (Plenum, New York, 1980), pp. 291–301; M. Saitoh, J. Phys. Soc. Jpn. **49**, 878 (1980).
[11]V. Samathiyakanit, J. Phys. C **7**, 2849 (1974).
[12]E. P. Gross, J. Stat. Phys. **17**, 265 (1977); V. Sa-yakanit, Phys. Rev. B **19**, 2266 (1979).
[13]P.-L. Chow, J. Math. Phys. **13**, 1224 (1972); R. Dashen, J. Math. Phys. **20**, 894 (1979).
[14]J. T. Marshall and J. L. Pell, J. Math. Phys. **20**, 1297 (1979).
[15]R. P. Feynman, Phys. Rev. **84**, 108 (1951), Appendix C.
[16]For these definitions see, e.g., Refs. 2–4.
[17]C. DeWitt-Morette, A. Maheshwari, and B. Nelson, Phys. Reports **50**, 255 (1979).
[18]See, e.g., Theorem 3.11 in Ref. 4 and also B. Simon, Adv. Math. **24**, 244 (1977).
[19]G. J. Papadopoulos, J. Phys. A **7**, 183 (1974); A. Maheshwari, *ibid.* **8**, 1019 (1975).

Random walks on lattices. The problem of visits to a set of points revisited

Robert J. Rubin
National Bureau of Standards, Washington, D.C. 20234
George H. Weiss
National Institutes of Health, Bethesda, Maryland 20205

(Received 8 May 1981; accepted for publication 28 August 1981)

A general method is outlined for calculating the statistical properties of the number of visits to a set of points in a random walk. In illustrative examples, known results and new results are easily derived.

PACS numbers: 02.50. + s, 05.90. + m

I. INTRODUCTION

Montroll and Weiss[1] analyzed the problem of finding the distribution function of the number of visits to a given lattice point in a discrete n-step random walk. At about the same time, Rubin,[2] in a study of the adsorption of a chain polymer at a solution surface, derived statistical properties of a one-dimensional random walk in which the number of visits to one or a pair of lattice points in the presence of an absorbing point was of interest. In this paper we present a formalism with which to calculate statistical properties of the number of visits to a set of points $S = \{\mathbf{R}_i\}$, $i = 1, 2, \ldots, m$. The formalism can be derived in two ways, the first related to the analysis of Montroll[3,4] on random walks in the presence of traps, and the second related to the analysis by Rubin. As Montroll[3] has noted, the method had been used earlier in studies of properties of harmonic crystals containing defects.[5,6] Spitzer[7] has also discussed closely related problems in a much more complicated (but also more detailed) way.

II. GENERAL FORMALISM

We will assume that the random walker is always initially at $\mathbf{r} = \mathbf{0}$. Let the single-step transition probabilities be $p(\mathbf{r})$ and let $P_n(\mathbf{r}|l_1, l_2, \ldots, l_m)$ be the probability that the random walker is at $\mathbf{r}$ at step n having visited point $\mathbf{R}_j$ l_j times, $j = 1, 2, \ldots, m$. One can write an obvious set of recursion relations for the $P_n(\mathbf{r}|l_1, l_2, \ldots, l_m)$

$$P_{n+1}(\mathbf{r}|\mathbf{l}) = \sum_{\boldsymbol{\rho}} p(\mathbf{r}-\boldsymbol{\rho})P_n(\boldsymbol{\rho}|\mathbf{l}), \quad \mathbf{r}\notin S, \tag{1a}$$

$$P_{n+1}(\mathbf{R}_j|\mathbf{l}) = \sum_{\boldsymbol{\rho}} p(\mathbf{R}_j-\boldsymbol{\rho})P_n(\boldsymbol{\rho}|l_1, l_2, \ldots, l_j - 1, \ldots, l_m), \quad j = 1, 2, \ldots m. \tag{1b}$$

These equations are valid when $\mathbf{0}$ does not belong to S. To resolve these equations it is expedient to introduce the generating functions

$$Q_n(\mathbf{r}|\mathbf{x}) = \sum_{l_1=0}^{\infty} \cdots \sum_{l_m=0}^{\infty} P_n(\mathbf{r}|l_1, l_2, \ldots, l_m) x_1^{l_1} x_2^{l_2} \cdots x_m^{l_m}, \tag{2}$$

in terms of which Eq. (1) can be represented as

$$Q_{n+1}(\mathbf{r}|\mathbf{x}) = \sum_{\boldsymbol{\rho}} p(\mathbf{r}-\boldsymbol{\rho})Q_n(\boldsymbol{\rho}|\mathbf{x}), \quad \mathbf{r}\notin S, \tag{3a}$$

$$Q_{n+1}(\mathbf{R}_j|\mathbf{x}) = x_j \sum_{\boldsymbol{\rho}} p(\mathbf{R}_j-\boldsymbol{\rho})Q_n(\boldsymbol{\rho}|\mathbf{x}), \quad j = 1, 2, \ldots, m. \tag{3b}$$

The reader will notice that this last relation can be written rather simply in terms of matrices if we define $\mathbf{Q}_n(\mathbf{x})$ to be a D dimensional matrix whose elements are the $Q_n(\mathbf{r}|\mathbf{x})$, $\mathbf{p}$ to be the matrix of single step transition probabilities, and $\mathbf{w}$ to be the diagonal matrix

$$\mathbf{w} = \begin{bmatrix} \ddots & & & & & & & & & & \\ & 1 & & & & & & & & & \\ & & \ddots & & & & & & & & \\ & & & 1 & & & & & & & \\ & & & & x_1 & & & & & & \\ & & & & & 1 & & & & & \\ & & & & & & \ddots & & & & \\ & & & & & & & x_2 & & & \\ & & & & & & & & \ddots & & \\ & & & & & & & & & x_m & \\ & & & & & & & & & & \ddots \end{bmatrix} \tag{4}$$

with x_j the entry in the position corresponding to $\mathbf{R}_j$. The matrix equivalent to Eqs. (3a) and (3b) is

$$\mathbf{Q}_{n+1} = \mathbf{w}\mathbf{p}\mathbf{Q}_n \tag{5}$$

or $\mathbf{Q}_n = (\mathbf{w}\mathbf{p})^n\mathbf{Q}_0$. In this form the resulting expression constitutes a generalization of the formalism of Rubin. To make contact with Montroll's formalism we can write Eqs. (3a) and (3b) as

$$Q_{n+1}(\mathbf{r}) = \sum_{\boldsymbol{\rho}} \left[p(\mathbf{r}-\boldsymbol{\rho})Q_n(\boldsymbol{\rho}) + \sum_{j=1}^{m} \delta_{\mathbf{r},\mathbf{R}_j}(x_j - 1)p(\mathbf{r}-\boldsymbol{\rho})Q_n(\boldsymbol{\rho}) \right] \tag{6}$$

where, for simplicity we have dropped the argument $\mathbf{x}$. This equation suggests the use of a Green's function technique as used by Montroll for random walks in the presence of traps and by Rubin[8] for the vibrational motion of particles in a harmonic crystal containing m isotopic defects. Equation (6) is valid when none of the $\mathbf{R}_j$ is equal to the starting point $\mathbf{0}$. If, however, $\mathbf{R}_1 = \mathbf{0}$, we can take the change into account by writing $Q_0(\mathbf{r}|\mathbf{x}) = x_1\delta_{\mathbf{r},\mathbf{0}}$ rather than $Q_0(\mathbf{r}|\mathbf{x}) = \delta_{\mathbf{r},\mathbf{0}}$.

To solve Eq. (3) or (5) we eliminate the space variables by going to a Fourier representation. For this purpose introduce the generating functions

$$G_n(\boldsymbol{\phi},\mathbf{x}) = \sum_{\mathbf{r}} Q_n(\mathbf{r}|\mathbf{x})\exp(i\mathbf{r}\cdot\boldsymbol{\phi}), \tag{7}$$

$$\Gamma(\boldsymbol{\phi},\mathbf{x},z) = \sum_{n=0}^{\infty} G_n(\boldsymbol{\phi},\mathbf{x})z^n. \tag{8}$$

By hypothesis $G_0(\boldsymbol{\phi},\mathbf{x}) = 1$ (except, as noted above, when $\mathbf{R}_1 = \mathbf{0}$, in which case $G_0(\boldsymbol{\phi},\mathbf{x}) = x_1$). It is also convenient to have the generating function

$$\xi(\mathbf{r},\mathbf{x},z) = \sum_{n=0}^{\infty} Q_n(\mathbf{r}|\mathbf{x})z^n. \tag{9}$$

We note that $\Gamma(\boldsymbol{\phi},\mathbf{x},z)$ can be written in terms of $\xi(\mathbf{r},z)$ as

$$\Gamma(\boldsymbol{\phi},\mathbf{x},z) = \sum_{\mathbf{r}} \xi(\mathbf{r},z)\exp(i\mathbf{r}\cdot\boldsymbol{\phi}) \tag{10}$$

or

$$\xi(\mathbf{r},z) = (2\pi)^{-D}\int_{-\pi}^{\pi}\cdots\int \Gamma(\boldsymbol{\phi},\mathbf{x},z)\exp(-i\mathbf{r}\cdot\boldsymbol{\phi})\,d^D\boldsymbol{\phi}. \tag{11}$$

In addition to these we also define the structure factor

$$\lambda(\boldsymbol{\phi}) = \sum_{\mathbf{r}} p(\mathbf{r})\exp(i\mathbf{r}\cdot\boldsymbol{\phi}). \tag{12}$$

Equations (3a) and (3b) are then equivalent to

$$\Gamma(\boldsymbol{\phi},\mathbf{x},z) = [1 - z\lambda(\boldsymbol{\phi})]^{-1} \times\left\{1 + \sum_{j=1}^{m}(1 - x_j^{-1})\xi(\mathbf{R}_j,\mathbf{x},z)\exp(i\mathbf{R}_j\cdot\boldsymbol{\phi})\right\}. \tag{13}$$

If we take Eq. (10) into account, we find that the $\xi(\mathbf{R}_j,z)$ satisfy a set of simultaneous linear equations which can be solved formally.

The solution to these equations can be written in terms of the quantities

$$P(\mathbf{r},z) = (2\pi)^{-D}\int_{-\pi}^{\pi}\cdots\int [1 - z\lambda(\boldsymbol{\phi})]^{-1}\exp(-i\mathbf{r}\cdot\boldsymbol{\phi})\,d^D\boldsymbol{\phi}. \tag{14}$$

which are generating functions with respect to step number for the state probabilities. The solution can be written

$$\Gamma(\boldsymbol{\phi},\mathbf{x},z) = [1 - z\lambda(\boldsymbol{\phi})]^{-1}\left\{1 + \sum_{j=1}^{m}(x_j - 1)(D_j/D)\exp(i\mathbf{R}_j\cdot\boldsymbol{\phi})\right\}, \tag{15}$$

where

$$D(\mathbf{x}) = \begin{vmatrix} x_1 + (1-x_1)P(\mathbf{0},z) & (1-x_2)P(\mathbf{R}_1 - \mathbf{R}_2,z) & \cdots & (1-x_m)P(\mathbf{R}_1 - \mathbf{R}_m,z) \\ (1-x_1)P(\mathbf{R}_2 - \mathbf{R}_1,z) & x_2 + (1-x_2)P(\mathbf{0},z) & \cdots & (1-x_m)P(\mathbf{R}_2 - \mathbf{R}_m,z) \\ \vdots & \vdots & & \vdots \\ (1-x_1)P(\mathbf{R}_m - \mathbf{R}_1,z) & (1-x_2)P(\mathbf{R}_m - \mathbf{R}_2,z) & \cdots & x_m + (1-x_m)P(\mathbf{0},z) \end{vmatrix} \tag{16}$$

and D_j is obtained from D by replacing its jth column by

$$\begin{pmatrix} P(\mathbf{R}_1,z) \\ P(\mathbf{R}_2,z) \\ \vdots \\ P(\mathbf{R}_m,z) \end{pmatrix}.$$

It follows from Eqs. (11) and (15) that

$$\xi(\mathbf{r},\mathbf{x},z) = P(\mathbf{r},z) - \sum_{j=1}^{m}(1 - x_j)\frac{D_j(\mathbf{x})}{D(\mathbf{x})}P(\mathbf{r} - \mathbf{R}_j,z), \quad \mathbf{r} \neq R_j, \tag{17a}$$

$$= x_j D_j(\mathbf{x})/D(\mathbf{x}), \quad \mathbf{r} = \mathbf{R}_j. \tag{17b}$$

If one is interested in statistical properties of the occupation time of the set S, then Eq. (15) can be used directly to generate moments. For example the mean occupancy of S during an n-step walk has the generating function

$$\mu_s(z) = \left.\frac{\partial}{\partial x}\Gamma(\mathbf{0},x,x,\ldots x,z)\right|_{x=1} = (1-z)^{-1}\sum_{j=1}^{m} P(\mathbf{R}_j,z), \tag{18}$$

as is otherwise obvious. Somewhat less obvious is the expression for the generating function for the second moments

$$\nu_s(z) = \left(\frac{\partial^2}{\partial x^2} + \frac{\partial}{\partial x}\right)\Gamma(\mathbf{0},x,\ldots x,z)|_{x=1}$$
$$= (1-z)^{-1}\left\{[2P(0,z) - 1]\sum_{j=1}^{m} P(\mathbf{R}_j,z) + 2\sum_{j=1}^{m} P(\mathbf{R}_j,z)\sum_{\substack{k=1 \\ k\neq j}}^{m} P(\mathbf{R}_k - \mathbf{R}_j,z)\right\}. \tag{19}$$

These results can also be obtained from the theory developed by Darling and Kac.[9] It is also possible to derive generating functions for moments and correlations of the occupation of multisite sets starting from the expression for $\Gamma(\boldsymbol{\phi},\mathbf{x},z)$. For example the second-order correlation function for two sites $\mathbf{R}_1$ and $\mathbf{R}_2$ has the generating function

$$\frac{\partial^2}{\partial x_1 \partial x_2}\Gamma(0,x_1,x_2,z)|_{x_1 = x_2 = 1} = -2(1-z)^{-1}[1 - P(0,z)][P(\mathbf{R}_1,z) + P(\mathbf{R}_2,z)] \tag{20}$$

and higher-order correlations can be dealt with in a similar manner.

III. APPLICATIONS

We now consider applications of the foregoing formalism in several special cases.

(i) $m = 1$. If the set S consists of one point, namely $\mathbf{R}$, and if $\mathbf{R}$ is not the starting point $\mathbf{0}$, then

$$\xi(\mathbf{R},z) = \frac{xP(\mathbf{R},z)}{x + (1-x)P(\mathbf{0},z)} \tag{21}$$

and

$$\xi(\mathbf{r},z) = P(\mathbf{r},z) + \frac{(x-1)P(\mathbf{r}-\mathbf{R},z)P(\mathbf{R},z)}{x + (1-x)P(\mathbf{0},z)}. \tag{22}$$

If $\mathbf{R}=\mathbf{0}$ is the starting point and the initial weighted probability is

$$Q_0(\mathbf{r}|\mathbf{x}) = x\delta_{\mathbf{r},\mathbf{0}},$$

then Eqs. (21) and (22) remain valid

$$\xi(\mathbf{0},z) = \frac{xP(\mathbf{0},z)}{x+(1-x)P(\mathbf{0},z)} \tag{23}$$

and

$$\xi(\mathbf{r},z) = \frac{xP(\mathbf{r},z)}{x+(1-x)P(\mathbf{0},z)} \,. \tag{24}$$

The generating function for the probability of returning to the starting point for the first time at the nth step is the coefficient of x^2 in the expansion of (23) in powers of x, namely

$$1 - 1/P(\mathbf{0},z). \tag{25}$$

The generating function of the probability of reaching $\mathbf{R}\neq\mathbf{0}$ for the first time at the nth step is the coefficient of x in the expansion of (21) in powers of x,

$$P(\mathbf{R},z)/P(\mathbf{0},z). \tag{26}$$

Finally, the generating function for the probability of reaching r at the nth step having visited $\mathbf{R}\neq\mathbf{r}$ (or $\mathbf{0}$) exactly s times is the coefficient of x^s in the expansion of (22) in powers of x,

$$\frac{P(\mathbf{r}-\mathbf{R},z)P(\mathbf{R},z)}{[P(\mathbf{0},z)]^2}\left[1-\frac{1}{P(\mathbf{0},z)}\right]^{s-1}, \quad s\geqslant 1 \tag{27a}$$

$$P(\mathbf{r},z) - \frac{P(\mathbf{r}-\mathbf{R},z)P(\mathbf{R},z)}{P(\mathbf{0},z)}, \quad s=0. \tag{27b}$$

These generating functions have been obtained by Montroll and Weiss.[1]

The generating function for the probability-distribution function of a random walk with an excluded origin[10] can be obtained from Eq. (24). It is the coefficient of x in Eq. (24)

$$P(\mathbf{r},z)/P(\mathbf{0},z). \tag{28}$$

If we compare (28) with (26) we recognize the obvious, namely the probability-distribution function for first arrival at $\mathbf{R}$ from $\mathbf{0}$ is identical with the probability-distribution function for going from $\mathbf{0}$ to $\mathbf{r}$ without revisiting $\mathbf{0}$.

(ii) $m=2$. If the set S consists of two points, $\mathbf{R}_1=\mathbf{0}$ and $\mathbf{R}_2$, and the random walk starts at $\mathbf{0}$, then

$$\xi(\mathbf{R}_2,z) = x_1x_2P(\mathbf{R}_2,z)/D, \tag{29}$$

where

$$D = \begin{vmatrix} x_1+(1-x_1)P(\mathbf{0},z) & (1-x_2)P(-\mathbf{R}_2,z) \\ (1-x_1)P(\mathbf{R}_2,z) & x_2+(1-x_2)P(\mathbf{0},z)\end{vmatrix}$$

and from Eqs. (15) and (17),

$$\Gamma(\mathbf{0},z) = (1-z)^{-1}x_1\{x_2+(1-x_2)[P(\mathbf{0},z)-P(\mathbf{R}_2,z)]\}/D. \tag{30}$$

The coefficient of $x_1\,x_2$ in Eq. (29) is the generating function of the probability-distribution function for a random walk to go from $\mathbf{0}$ to $\mathbf{R}_2$ with no visits of intermediate steps to either $\mathbf{0}$ or $\mathbf{R}_2$ (a generalization of the excluded-origin random walk)

$$\Pi_{\mathbf{0},\mathbf{R}_2}(z) = \frac{P(\mathbf{R}_2,z)}{[P(\mathbf{0},z)]^2 - P(\mathbf{R}_2,z)P(-\mathbf{R}_2,z)} \,. \tag{31}$$

The probability of starting at $\mathbf{0}$ and ultimately arriving at $\mathbf{R}_2$ with no intermediate visits to $\mathbf{0}$ or $\mathbf{R}_2$ is obtained from the expression for $\Pi_{\mathbf{0},\mathbf{R}_2}(1)$. In the case of a symmetric random walk,

$$\Pi_{\mathbf{0},\mathbf{R}_2}(1) = \Pi_{|\mathbf{R}_2|}(1) = \frac{1/2}{P(\mathbf{0},1)-P(\mathbf{R}_2,1)} - \frac{1/2}{P(\mathbf{0},1)+P(\mathbf{R}_2,1)} \,. \tag{32}$$

The probability of ultimately arriving at $\mathbf{R}_2$ in case there are no restrictions on the number of visits to $\mathbf{0}$ is well known to be equal to one in the case of one- and two-dimensional random walks. This will no longer be the case if restrictions are placed on the number of visits to $\mathbf{0}$ and/or $\mathbf{R}_2$. For example, in the case of the nearest-neighbor symmetric random walk in one and two dimensions, the structure factors are

$$\lambda(\boldsymbol{\phi}) = \cos\phi$$

and

$$\lambda(\boldsymbol{\phi}) = \tfrac{1}{2}(\cos\phi_1 + \cos\phi_2),$$

respectively. The corresponding perfect-lattice random-walk generating functions follow from Eq. (14). In the one-dimensional case, there is the explicit result

$$P(\mathbf{R},z) = (1-z^2)^{-1/2}\left(\frac{1-(1-z^2)^{1/2}}{z}\right)^{|R|}, \tag{33}$$

so that the probability $\Pi_{|R_2|}(1)$ has the value

$$\Pi_{|R_2|}(1) = [2|R_2|]^{-1}. \tag{34}$$

The form of $P(\mathbf{R},z)$ is more complicated for the plane square lattice. However, it is known[1] that

$$P(\mathbf{R},z) = P(\mathbf{0},z) + g(\mathbf{R},z), \tag{35}$$

where $g(\mathbf{R},z)$ is not singular at $z=1$. Therefore, it follows that

$$\Pi_{|R_2|}(1) = [2g(|R_2|,1)]^{-1}. \tag{36}$$

Van der Pol[11] has given an explicit formula for the nonsingular part of $P(\mathbf{R},z)$ in case $\mathbf{R}$ is a diagonal point such as (m,m). In our notation, van der Pol's result is [for $\mathbf{R}_2=(m,m)$]

$$g(|R_2|,1) = \frac{4}{\pi}\left(1+\frac{1}{3}+\frac{1}{5}+\cdots+\frac{1}{2|m|+1}\right) = \frac{2}{\pi}\,[\,\gamma + 2\ln z + \psi(|m|+\tfrac{3}{2})]\,, \tag{37}$$

where $\psi(x) = (d/dx)\ln\Gamma(x)$ is the logarithmic derivative of the gamma function and where $\gamma = 0.5772...$ is the Euler–Mascheroni constant. Therefore, in the limit of large $|m|$,

$$\Pi_{|R_2|}(1) \cong \pi[4\ln(|m|)]^{-1}. \tag{38}$$

Thus, the probability of ultimately reaching $\mathbf{R}_2$ from $\mathbf{0}$ without returning to $\mathbf{0}$ decreases to zero logarithmically with the distance $|\mathbf{R}_2|$.

It is a simple matter to generalize the result given in Eq. (32) for $\Pi_{\mathbf{0},\mathbf{R}_2}(1)$. For example, one can calculate the probability of starting at $\mathbf{0}$ and ultimately arriving at $\mathbf{R}_2$ with s intermediate visits to the set of points $\mathbf{0}$ and $\mathbf{R}_2$. This probability, $\Pi^{(s)}_{\mathbf{0},\mathbf{R}_2}(1)$, can be obtained from Eq. (29) by first setting $x_1=x_2=x$ and then determining the coefficient of x^{2+s} in the expansion in powers of x. The first few probabilities are, first from Eq. (32),

$$\Pi^{(0)}_{\mathbf{0},\mathbf{R}_2}(1) = \Pi_{\mathbf{0},\mathbf{R}_2}(1),$$

then

$$\Pi^{(1)}_{\mathbf{0},\mathbf{R}_2}(1) = 2\left(1 - \frac{P(\mathbf{0},1)}{[P(\mathbf{0},1)]^2 - [P(\mathbf{R}_2,1)]^2}\right)\Pi^{(0)}_{\mathbf{0},\mathbf{R}_2}(1), \quad (39)$$

and

$$\Pi^{(2)}_{\mathbf{0},\mathbf{R}_2}(1) = \left\{4\left(1 - \frac{P(\mathbf{0},1)}{[P(\mathbf{0},1)]^2 - [P(\mathbf{R}_2,1)]^2}\right)^2 - 1 + \frac{2P(\mathbf{0},1) - 1}{[P(\mathbf{0},1)]^2 - [P(\mathbf{R}_2,1)]^2}\right\}\Pi^{(0)}_{\mathbf{0},\mathbf{R}_2}(1). \quad (40)$$

In contrast to the applications considered thus far, where the x_i have all been treated strictly as counting variables, we consider a final application of Eq. (30), where one of the x_i is given a thermodynamic significance. In this last application, we treat a random walk model of polymer chain adsorption at a plane solution surface. Each random walk configuration of n steps is weighted by a Boltzmann factor $\exp(n\theta)$ where n is the number of visits of that configuration to the surface layer and where the reduced energy, θ, equals ϵ/kT, where the energy, $\epsilon > 0$, is the energy gained for each step in the surface layer. In the simplest version of the model, configurations are regarded as nearest-neighbor one-dimensional random walks between lattice planes parallel to the solution surface. For random walks which start in the surface layer (labeled $\kappa = 0$), we only wish to consider configurations which avoid the lattice plane $\kappa = -1$ (outside the solution). Thus, in this model, the special set of points is $R_2 = -1$ with associated weight x_2 and the other point is $R_1 = 0$ with associated weight $x_1 = \exp\theta$. Then, the generating function for the weighted probability of random walks which start at $R_1 = 0$ and never visit $R_2 = -1$ is given by the coefficient of x_2^0 in Eq. (30). This coefficient is simply determined by setting $x_2 = 0$ in that expression

$$\Gamma(\mathbf{0},z)|_{x_2=0} = \frac{(1-z)^{-1}e^{\theta}[P(0,z) - P(1,z)]}{\begin{vmatrix} e^{\theta} + (1-e^{\theta})P(0,z) & P(1,z) \\ (1-e^{\theta})P(1,z) & P(0,z) \end{vmatrix}}. \quad (41)$$

If the nearest-neighbor single-step transition probabilities for steps between layers are

$$p(\pm 1) = \tfrac{1}{2}a$$

and $p(0) = 1 - a$ with $p(\pm|j|) = 0$ for $|j| \geqslant 2$, then the structure factor is

$$\lambda(\phi) = 1 - a + a\cos\phi,$$

and the random walk propagator in the perfect lattice is

$$P(\mathbf{r},z) = \{(1-z)[1-(1-2a)z]\}^{-1/2} \times \left(\frac{1-(1-a)z - \{(1-z)[1-(1-2a)z]\}^{1/2}}{az}\right)^{|r|}. \quad (42)$$

If Eq. (42) is used in (41) for $\Gamma(\mathbf{0},z)|_{x_2=0}$, one can obtain the result

$$\Gamma(\mathbf{0},z)|_{x_2=0} = \frac{1 + [1-(1-2a)z]^{1/2}(1-z)^{-1/2}}{\{(1-z)[1-(1-2a)z]\}^{1/2} - 1 - (1-2a)z + 2e^{-\theta}}, \quad (43)$$

which is identical with the one obtained by Rubin [Eqs. (22), (24), and (26)],[2] who used a similar method to count visits to layer 0 when layer -1 was treated as an absorbing layer.

[1]E. W. Montroll and G. H. Weiss, J. Math. Phys. **6**, 167 (1965).
[2]R. J. Rubin, J. Chem. Phys. **43**, 2392 (1965).
[3]E. W. Montroll, Proc. Symp. Appl. Math., Am. Math. Soc. **16**, 193 (1964).
[4]E. W. Montroll, J. Phys. Soc. Japan, **26**, Suppl. 6 (1969).
[5]E. W. Montroll and R. Potts, Phys. Rev. **100**, 525 (1955).
[6]R. J. Rubin, *Proc. Intl. Symp. on Transport Proc. in Stat. Mech.*, Aug. 1956, edited by I. Prigogine (Interscience, New York, 1958), p. 155.
[7]F. Spitzer, *Principles of Random Walk*, 2nd ed. (Springer, New York, 1976).
[8]R. J. Rubin, J. Math. Phys. **9**, 2252 (1968).
[9]D. A. Darling and M. Kac, Trans. Am. Math. Soc. **84**, 444 (1957).
[10]R. J. Rubin, J. Math. Phys. **8**, 576 (1967).
[11]B. van der Pol, *Probability and Related Topics in Physical Sciences*, edited by M. Kac (Interscience, New York, 1959), Appendix IV.

Hypervirial calculation of integrals involving Bessel functions

F. M. Fernández, A. Mesón, and E. A. Castro
INIFTA, Sección Química Teórica, Sucursal 4—Casilla de Correo 16, La Plata 1900, Argentina

(Received 5 August 1981; accepted for publication 4 September 1981)

A general and simple procedure is presented for evaluating matrix elements that involve Bessel functions. The method is based upon hypervirial relationships for systems subjected to Dirichlet boundary conditions.

PACS numbers: 02.70. + d

The quantum mechanical study of physical systems enclosed within spherical or cylindrical surfaces[1,2] requires quanties of the following kind:

$$I_{ik}^{N}(c)=\frac{\int_0^1 x^N J_c(j_{c,i}x)J_c(j_{c,k}x)\,dx}{\{\int_0^1 xJ_c^2(j_{c,i}x)\,dx\int_0^1 xJ_c^2(j_{c,k}x)\,dx\}^{1/2}},\qquad(1)$$

where $J_c(x)$ $(c\geqslant 0)$ is a Bessel function, and $j_{c,i}$ the corresponding ith zero. Suprisingly enough, the current literature on this subject [3–7] does not record a single simple formula that enables one to calculate the integrals (1).

The purpose of this communication is to present a general and easy procedure for evaluating the matrix elements (1) for N odd, through the use of a recursion formula. The method is based upon hypervirial relationships for systems subjected to Dirichlet boundary conditions, and which were recently deduced.[8–11]

Let us start from the stationary unidimensional Schrödinger equation

$$H\phi_i=E\phi_i,\quad H=-D^2/2+V(x),\quad D\equiv d/dx,\qquad(2)$$

with the following boundary conditions

$$\phi_i(0)=\phi_i(b)=0.\qquad(3)$$

From the hypervirial relations, Eq. (4) can be deduced without any difficulty[9]:

$$\tfrac14 N(N-1)(N-2)\langle i|x^{N-3}|j\rangle+N(E_i+E_j)\langle i|x^{N-1}|j\rangle-2N\langle i|x^{N-1}V|j\rangle-\langle i|x^NV'|j\rangle+\frac{(E_i-E_j)^2}{N+1}\langle i|x^{N+1}|j\rangle=b^N\left\{\frac{\partial E_i}{\partial b}\frac{\partial E_j}{\partial b}\right\}^{1/2}.\qquad(4)$$

When

$$V(x)=t/2x^2,\quad t\geqslant-\tfrac14\qquad(5)$$

Eq. (4) is transformed in a recursion relation for the matrix elements of the x powers:

$$A_{ij}^{N+1}(t,b)=\frac{N+1}{(E_i-E_j)^2}\left[b^N\left(\frac{\partial E_i}{\partial b}\frac{\partial E_j}{\partial b}\right)^{1/2}+(N-1)\left(t-\frac{N(N-2)}{4}\right)A_{ij}^{N-3}(t,b)-N(E_i+E_j)A_{ij}^{N-1}(t,b)\right],\qquad(6)$$

where

$$A_{ij}^{N}(t,b)\equiv\langle i|x^N|j\rangle.\qquad(7)$$

Equation (2) for the potential function (5) adopts the form

$$-\frac12\phi''(x)+\frac{t}{2x^2}\phi_i(x)=E_i\phi_i(x).\qquad(8)$$

The change of variables

$$x=p_iy,\quad p_i=(2E_i)^{-1/2}\qquad(9)$$

transforms Eq. (8) into

$$\phi_i''(p_iy)+[1-(t/y^2)\phi_i(p_iy)]=0.\qquad(10)$$

Obviously, the solutions of this last differential equation are related with the Bessel function $J_c(y)$ in the following way:

$$\phi_i(p_iy)=y^{1/2}J_c(y),\quad c=(t+\tfrac14)^{1/2}.\qquad(11)$$

The boundary condition (3) associates eigenvalues E_i with the zeros of J_c through the formula

$$E_i=j_{c,i}^2/2b^2.\qquad(12)$$

The substitution of Eq. (12) in the recursion relationship (6) for $b=1$ allowed us to obtain

$$A_{ik}^{N+1}(t)=\frac{4(N+1)}{(j_{c,i}^2-j_{c,k}^2)^2}\left[j_{c,i}j_{c,k}+(N-1)\left(t-\frac{N(N-2)}{4}\right)\times A_{ik}^{N-3}(t)-\frac{N}{2}(j_{c,i}^2+j_{c,k}^2)A_{i,k}^{N-1}(t)\right].\qquad(13)$$

From Eq. (11) the following equality is deduced at once:

$$A_{ik}^{N}(t)=I_{ik}^{N+1}(c).\qquad(14)$$

The starting point for the recursion relationship (13) is the orthonormalization condition

$$A_{ij}^{0}(t)=\delta_{ij}.\qquad(15)$$

The two first matrix elements are

$$A_{ik}^{2}(t)=8j_{c,i}j_{c,k}/(j_{c,i}^2-j_{c,k}^2)^2=I_{ik}^{3}(c),\qquad(16)$$

$$A_{ik}^{4}(t)=16j_{c,i}j_{c,k}\left(1-\frac{12(j_{c,i}^2+j_{c,k}^2)}{(j_{c,i}^2-j_{c,k}^2)}\right)\times(j_{c,i}^2-j_{c,k}^2)^{-2}=I_{ik}^{5}(c).\qquad(17)$$

When $i=k$, Eq. (13) cannot directly be used. However, by a simple rearrangement we get

$$A_{ii}^{N-1}(t)=\frac{j_{c,i}^2+(N-1)[t-N(N-2)/4]A_{ii}^{N-3}(t)}{Nj_{c,i}^2}\qquad(18)$$

and consequently we obtain

$$A_{ii}^{2}(t)=\frac{j_{c,i}^2+2(t-\frac34)}{3j_{c,i}^2}=\frac13\left(1+\frac{2(c^2-1)}{j_{c,i}^2}\right)=I_{ii}^{3}(c),\qquad(19)$$

and so forth.

The finite induction principle allows us to prove that recursion formulae (13) and (18) permit the calculation of any matrix element (1) for N odd and $c\geqslant 0$.

While studying the magnetic properties of small quantum systems, Dingle[1,2] used the integrals (16) and (19), which he obtained from Straubel's work[4] and Schafheithins formula,[3] respectively. Both results can be deduced as particular cases from our earlier more general equations.

[1]R. B. Dingle, Proc. R. Soc. London Ser. A **212**, 47 (1952).
[2]R. B. Dingle, Proc. Camb. Philos. Soc. **49**, 103 (1953).
[3]G. N. Watson, *Theory of Bessel Functions* (Cambridge U.P., Cambridge, 1922).
[4]R. Straubel, Ingen.-Arch. **13**, 14 (1942).
[5]A. Gray and T. McRobert, *A Treatise of Bessel Functions* (McMillan, New York, 1952).
[6]E. T. Whittaker and G. N. Watson, *A Course of Modern Analysis* (Cambridge U.P., Cambridge, 1963).
[7]G. N. Watson, *A Treatise on the Theory of Bessel Functions* (Cambridge U.P., Cambridge, 1966).
[8]F. M. Fernández and E. A. Castro, Int. J. Quantum Chem. **19**, 521 (1981).
[9]F. M. Fernández and E. A. Castro, Int. J. Quantum Chem. **19**, 533 (1981).
[10]F. M. Fernández and E. A. Castro, Int. J. Quantum Chem. **20**, 623 (1981).
[11]F. M. Fernández and E. A. Castro, J. Math. Phys. **22**, 1669 (1981).

Poincaré–Cartan integral invariant and canonical transformations for singular Lagrangians: An addendum

D. Dominici
Istituto Nazionale di Fisica Nucleare—Sezione di Firenze, Istituto di Fisica Teorica dell'Università di Firenze, 50125 Firenze, Italy

J. Gomis
Departament de Física Teòrica de la Universitat de Barcelona, Barcelona, Spain

(Received 21 May 1981; accepted for publication 2 September 1981)

The results of a previous work, concerning a method for performing the canonical formalism for constrained systems, are extended when the canonical transformation proposed in that paper is explicitly time dependent.

PACS numbers: 03.20. + i

In a previous paper[1] we discussed in the framework of the Poincaré–Cartan integral invariant, a method for performing the canonical formalism for constrained systems. The basic idea consists of considering a canonical transformation which brings the constraints into a subset of the canonical variables. Thus the physical variables can be easily obtained by means of a reduction of the phase space. Our method is different from the path-integral approach of Faddeev[2] (see also Ref. 3), which use in addition a set of gauge-fixing conditions, one for each first-class constraint. Two applications of our procedure concerning action-at-a-distance relativistic models have been recently studied.[4]

In this note we extend the method by considering a time-dependent general canonical transformation, such that all the constraints acquire an explicit time dependence.

Let us consider a dynamical system described in terms of $2n$ degrees of freedom in the phase space q_s, p_s $(s = 1,...,n)$ and constrained to the hypersurface S defined by

$$\Omega_\alpha(q,p) = 0 \quad (\alpha = 1,...,T - W), \tag{1}$$

$$\Omega_\beta(q,p) = 0 \quad (\beta = T - W + 1,...,T), \tag{2}$$

where Ω_α are $T - W$ first-class[5] and Ω_β W second-class constraints. In order to guarantee the stability of S during the evolution, the Ω_α are required to satisfy

$$\{\Omega_\alpha, H_c\} \approx 0, \tag{3}$$

where H_c is the canonical Hamiltonian. The notation "$\approx$" means equality on the hypersurface S ("weak" equality).

Now, given the set (2), according to some theorems on function groups[6] and involutory systems[7] it is possible, at least locally, to find a canonical transformation

$$\{q_s, p_s, \quad s = 1,...,n\} \to \{Q'_s, P'_s, \quad s = 1,...,n\}, \tag{4}$$

such that the equations

$$Q'_f = P'_f = 0 \quad (f = n_2 + 1,...,n), (n_2 = n - W/2), \tag{5}$$

define the same surface as Eqs. (2) and the following equations,

$$\{Q'_s, P'_{s'}\} = \delta_{ss'},$$
$$\{Q'_s, Q'_{s'}\} = \{P'_s, P'_{s'}\} = 0, \tag{6}$$

are identically (and not only "weakly") satisfied.

If we denote the generating function by F, defined as

$$p_s\delta q_s - H_c\delta t = P'_s\delta Q'_s - K_c\delta t - \delta F, \tag{7}$$

the Hamilton equations for the new variables are given by

$$\dot{Q}'_s \approx \{Q'_s, K(Q'_s, P'_s, t)\}, \quad \dot{P}'_s \approx \{P'_s, K(Q'_s, P'_s, t)\} \tag{8}$$

where K,

$$K = K_c + l_\alpha \tilde{\Omega}_\alpha - \tilde{\Omega}_\beta \tilde{C}_{\beta\beta'}\left[\{\tilde{\Omega}_{\beta'}, K_c\} + \frac{\partial \tilde{\Omega}_{\beta'}}{\partial t}\right], \tag{9}$$

is the extended Hamiltonian with l_α arbitrary functions. $\tilde{\Omega}_{\alpha,\beta}$ are obtained from Eqs. (1) and (2) by substitution of variables, and $\tilde{C}_{\beta\beta'}$ is defined by

$$\tilde{C}_{\beta\beta'}\{\tilde{\Omega}_{\beta'}, \tilde{\Omega}_{\beta''}\} \approx \delta_{\beta\beta''}. \tag{10}$$

In I we have shown that it is possible to write the equations of motion for the reduced set of variables $R' = \{Q'_j, P'_j, j = 1,...,n_2\}$ which are free with respect to the second-class constraints (5) in a simple form

$$\dot{Q}'_j \approx \{Q'_j, \bar{K}\}_{R'}, \quad \dot{P}'_j \approx \{P'_j, \bar{K}\}_{R'}, \tag{11}$$

$$\bar{K} = \bar{K}(Q'_j, P'_j, t) = \bar{K}_c(Q'_j, P'_j, t) + l_\alpha \bar{\Omega}_\alpha(Q'_j, P'_j, t) \tag{12}$$

where $\{\ ,\ \}_{R'}$ denote the Poisson brackets defined on the space R' and $\bar{K}_c$ and $\bar{\Omega}_\alpha$ are obtained by setting equal to zero the variables Q'_f and P'_f, corresponding to the second-class constraints, in K_c and $\tilde{\Omega}_\alpha$ of Eq. (9). As shown in I the $\bar{\Omega}_\alpha$ so obtained are first class, i.e.,

$$\{\bar{\Omega}_\alpha, \bar{\Omega}_{\alpha'}\}_{R'} \approx 0 \tag{13}$$

and, as a consequence of $(d/dt)\Omega_\alpha(q,p) \approx 0$, satisfy the stability condition

$$\frac{d}{dt}\bar{\Omega}_\alpha = \frac{\partial \bar{\Omega}_\alpha}{\partial t} + \{\bar{\Omega}_\alpha, \bar{K}_c\}_{R'} \approx 0. \tag{14}$$

In Eq. (14) we have now supposed the $\bar{\Omega}_\alpha$ explicitly time dependent, unlike what we did for the sake of simplicity in I.

A similar procedure of reduction of the phase space can be performed also for the first-class constraints. In fact, a theorem on involutory systems[7] guarantees that it is possible, at least locally, to replace the $\bar{\Omega}_\alpha$ by an equivalent set of equations

$$P_e(Q'_j, P'_j, t) = 0 \quad (e = n_1 + 1,...,n_2), \tag{15}$$

$(n_1 = n - T + W/2)$, which are in involution. For instance, the set (15) can be obtained by solving the equations

$$\bar{\Omega}_\alpha(Q'_j,P'_j,t)=0 \quad (\alpha=1,...,n_2-n_1) \tag{16}$$

with respect to an equal number n_2-n_1 of momenta. Without loss of generality we suppose Eq. (16) be solved with respect to $P'_e (e=n_1+1,...,n_2)$, or

$$|\partial\bar{\Omega}_\alpha/\partial P'_e|\neq 0. \tag{17}$$

Let

$$P_e=P'_e-f_e(Q'_e,Q'_k,P'_k,t) \quad (k=1,...,n_1) \tag{18}$$

be the expression of the equations in involution. The stability of the hypersurface (18) can be easily proved. In fact, from

$$\bar{\Omega}_\alpha(Q'_k,Q'_e,P'_k,P'_e=f_e(Q'_k,Q'_e,P'_k,t),t)=0 \tag{19}$$

we get

$$\frac{\partial\bar{\Omega}_\alpha}{\partial t}\approx-\frac{\partial\bar{\Omega}_\alpha}{\partial P'_e}\frac{\partial f_e}{\partial t}=\frac{\partial\bar{\Omega}_\alpha}{\partial P'_e}\frac{\partial P_e}{\partial t}, \tag{20}$$

$$\begin{cases} -\{P'_j,\bar{\Omega}_\alpha\}_{R'}=\dfrac{\partial\bar{\Omega}_\alpha}{\partial Q'_j}\approx-\dfrac{\partial\bar{\Omega}_\alpha}{\partial P'_e}\dfrac{\partial f_e}{\partial Q'_j}=-\dfrac{\partial\bar{\Omega}_\alpha}{\partial P'_e}\dfrac{\partial P_e}{\partial Q'_j}\\ \{Q'_j\bar{\Omega}_\alpha\}_{R'}=\dfrac{\partial\bar{\Omega}_\alpha}{\partial P'_j}\approx-\dfrac{\partial\bar{\Omega}_\alpha}{\partial P'_e}\dfrac{\partial f_e}{\partial P'_j}=-\dfrac{\partial\bar{\Omega}_\alpha}{\partial P'_e}\dfrac{\partial P_e}{\partial P'_j}\end{cases} \tag{21}$$

Therefore, from Eq. (14) we get

$$\frac{\partial\bar{\Omega}_\alpha}{\partial P'_e}\left[\frac{\partial P_e}{\partial t}+\{P_e,\bar{K}_c\}_{R'}\right]\approx 0, \tag{22}$$

and using Eq. (17)

$$\frac{\partial P_e}{\partial t}+\{P_e,\bar{K}_c\}_{R'}\approx 0. \tag{23}$$

As a final step we make a transformation

$$\{Q'_j,P'_j,j=1,...,n_2\}\rightarrow\{Q_k,P_k,Q_e,P_e,k=1,...,n_1,\ e=n_1+1,...,n_2\} \tag{24}$$

with

$$\{Q_k,P_{k'}\}=\delta_{kk'},\quad \{Q_e,P_{e'}\}=\delta_{ee'}, \tag{25}$$

where part of the momenta are the set of functions in the involution (18) which are equivalent to the first-class constraints.

If we denote the new canonical Hamiltonian by K'_c and the new expression for the constraints by

$$\hat{\Omega}_\alpha(Q_k,P_k,Q_e,P_e,t)=\bar{\Omega}_\alpha(Q'_j(Q_k,P_k,Q_e,P_e,t),P'_j(Q_k,P_k,Q_e,P_e,t),t), \tag{26}$$

the Hamiltonian equations are given by

$$\begin{cases}\dot{Q}_k\approx\{Q_k,K'_c+l_\alpha\hat{\Omega}_\alpha\}_R\\ \dot{P}_k\approx\{P_k,K'_c+l_\alpha\hat{\Omega}_\alpha\}_R\end{cases} \tag{27}$$

$$\begin{cases}\dot{Q}_e\approx\{Q_e,K'_c+l_\alpha\hat{\Omega}_\alpha\}_R\\ \dot{P}_e\approx\{P_e,K'_c+l_\alpha\hat{\Omega}_\alpha\}_R\end{cases} \tag{28}$$

where now $\{\ ,\ \}_R$ denote the Poisson brackets with respect to the set

$$R=\{Q_k,P_k,Q_e,P_e,k=1,...,n_1,e=n_1+1,...,n_2\}.$$

With respect to the stability of the hypersurface $\hat{\Omega}_\alpha=0$, after the canonical transformations (24) we have

$$\frac{\partial}{\partial t}\hat{\Omega}_\alpha+\{\hat{\Omega}_\alpha,K'_c\}_R\approx 0. \tag{29}$$

On the other hand, due to the equivalence between $\hat{\Omega}_\alpha$ and P_e we may write

$$\hat{\Omega}_\alpha(Q_k,P_k,Q_e,P_e,t)=g_{\alpha e'}(Q_k,P_k,Q_e,P_e,t)P_{e'},\quad \det|g|\neq 0, \tag{30}$$

where we introduced the strong equality notation "≡" following Sudarshan and Mukunda.[8]

Thus from Eq. (30) we have

$$\frac{\partial\hat{\Omega}_\alpha}{\partial t}\approx 0, \tag{31}$$

and using Eqs. (30) and (29) in Eq. (28), we get

$$\dot{P}_e=\{P_e,K'_c\}=\frac{\partial K'_c}{\partial Q_e}\approx 0. \tag{32}$$

In other words the variables Q_e are ignorable variables.

Finally, the remaining equations (27) and (28) become

$$\begin{cases}\dot{Q}_k\approx\{Q_k,K'_c\}_R\\ \dot{P}_k\approx\{P_k,K'_c\}_R\end{cases} \tag{33}$$

and

$$\dot{Q}_e=\{Q_e,K'_c\}_R+\lambda_e, \tag{34}$$

where $\lambda_e=g_{e\alpha}l_\alpha$ are arbitrary functions.

We can now consider the reduced space $[Q_k,P_k,Q_e]$, where Q_k and P_k satisfy

$$\dot{Q}_k=\frac{\partial\mathscr{K}_c}{\partial P_k},\quad \dot{P}_k=-\frac{\partial\mathscr{K}_c}{\partial Q_k}\quad (k=1,...,n_1), \tag{35}$$

with

$$\mathscr{K}_c=\mathscr{K}_c(Q_k,P_k,t)=K'_c(Q_k,P_k,Q_e,P_e,t)|_{P_e=0}. \tag{36}$$

where the Q_e dependence disappears due to Eq. (32) and the Q_e's are gauge-dependent variables

$$\dot{Q}_e=\left.\frac{\partial K'_c}{\partial P_e}\right|_{P_e=0}+\lambda_e\quad (e=n_1+1,...,n_2). \tag{37}$$

In conclusion, we have isolated the set of the gauge-dependent variables Q_e from a set of physical (gauge-independent) variables Q_k,P_k.

[1]D. Dominici and J. Gomis, J. Math. Phys. 21, 2124 (1980); from now on we will call it I.

[2]L. D. Faddeev, Theor. Math. Phys. 1, 1 (1970).

[3]P. Senjanovic, Ann. Phys. (NY) 100, 227 (1976); T. Maskawa and H. Nakajima, Prog. Theor. Phys. 56, 1295 (1976).

[4]D. Dominici, J. Gomis, and G. Longhi, "Two applications of a General Method for Studying Constrained Dynamical Systems" to be published in Nuovo Cimento A.

[5]P. A. M. Dirac, Can. J. Math. 2, 129 (1950); see also "Lectures on Quantum Mechanics", Belfer Graduate School of Science, Yeshiva University, New York, 1954.

[6]L. P. Eisenhart, *Continuous Groups of Transformations* (Dover, New York, 1961), Chap. VI.

[7]J. A. Schouten and W. der Kulk, *Pfaff's Problem and its Generalizations* (Clarendon, Oxford, 1949); C. R. Chester, *Techniques in Partial Differential Equations* (McGraw-Hill, Kogakusha, Tokyo, 1971), Chap. XV.

[8]E. C. G. Sudarshan and N. Mukunda, *Classical Dynamics: a Modern Perspective* (Wiley, New York, 1974).

Solution of a Schrödinger inverse scattering problem with a polynomial spectral dependence in the potential [a)]

M. Jaulent and C. Jean [b)]
Department de Physique Mathématique, Université des Sciences et Techniques du Languedoc, 34060-Montpellier Cedex, France

(Received 11 March 1981; accepted for publication 5 June 1981)

The inverse scattering problem for the scalar Schrödinger equation $y'' + \left[E - \sum_{p=0}^{n} (E^{1/2n})^p u_p(x)\right] y = 0, x\in\mathbb{R}$, is considered. It is solved by reduction to the inverse scattering problem for a matrix Schrödinger equation: $Y'' + [EI - (U(x) + E^{1/2}Q(x))]\, Y = 0$, $x\in\mathbb{R}$.

PACS numbers: 03.65.Nk, 02.30. – f

I. INTRODUCTION

The inverse scattering problem (ISP) associated with the scalar Schrödinger equation

$$y'' + [E - (u(x) + E^{1/2}q(x))]\; y = 0, \quad x\in\mathbb{R}, \tag{1.1}$$

where the complex parameter E is the "energy" and u, q are the "potentials", has been extensively studied.[1] Indeed the square root $E^{1/2}$ being an analytic function on a Riemann's two sheet surface, it is convenient to set $E = k^2 (k\in\mathbb{C})$ and to represent $E^{1/2}$ by "$+k$" or "$-k$". (1.1) is then replaced by a pair of equations

$$y^{\pm''} + [\,k^2 - (u(x) \pm k\, q(x))\,]y^{\pm} = 0, \quad x\in\mathbb{R} \tag{1.2}$$

in which the indices $\pm$ correspond to each other. There are other ISP in physics, especially in absorbing media which can be reduced to this ISP.[2] Furthermore a family of nonlinear evolution equations has been exhibited[3] which can be solved by the method of the Inverse Scattering Transform (IST) for (1.2), and a Hamiltonian formulation can be given.[4] There is a one-to-one correspondence[5] between these equations and those derived from the IST for the Zakharov–Shabat system.[6–8] This transformation is canonical.[9]

In this paper we are interested in the following generalization of (1.1):

$$y'' + \left[E - \sum_{p=0}^{n} (E^{1/2n})^p u_p(x)\right]y = 0, \quad x\in\mathbb{R}, \tag{1.3}$$

where the complex parameter E is the "energy" and $u_0,u_1,...,u_n$ are the $(n+1)$ "potentials", supposed to be sufficiently regular complex functions decreasing fast enough as $|x|\to\infty$. In Ref. 10 the Gel'fand–Dikii method has been applied to an equation more general than (1.3)

$$y'' + \left[E - \sum_{p=0}^{n-1} (E^{1/n})^p u_p(x)\right]y = 0, \quad x\in\mathbb{R} \tag{1.4}$$

and a family of nonlinear Hamiltonian equations has been derived[11] which can be solved using the IST for (1.4) provided that the ISP for (1.4) can be solved. It is the aim of this paper to solve the ISP for (1.3) and thus to continue the work undertaken in Ref. 11. This paper contains proofs of results announced in Ref. 12.

II. EQUIVALENT REPRESENTATIONS OF EQUATION (1.3)

In Eq. (1.3) the $(2n)^{\text{th}}$ root $E^{1/2n}$ is an analytic function on a Riemann's $2n$ sheet surface. A simple way to take this into account is to set $E = \lambda^{2n} (\lambda\in C)$ and to represent $E^{1/2n}$ by $\lambda e^{il\pi/n}$, where l can take the values $l = 0,1,...,2n-1$. Equation (1.3) is then represented by the $2n$ scalar Schrödinger equations:

$$y_l'' + \left[\lambda^{2n} - \sum_{p=0}^{n} \lambda^p e^{ipl\pi/n} u_p(x)\right]y_l = 0, \quad x\in\mathbb{R}, \quad l = 0,1,..,2n-1. \tag{2.1}$$

Let us remark that Eq. (1.2) corresponds to the case $n = 1$ and $\lambda = k$. It is worthwhile to note that we are led to consider the whole system of $2n$ equations and not just only the single equation corresponding to $l = 0$ in order to get a well posed inverse problem.

Clearly if $u_p = 0$ for $p = 1,2,...,n-1$, $(2.1)_l$ reduces to $(2.1)_+$ for even l and to $(2.1)_-$ for odd l, with $\lambda^n = k, u_0 = u$, and $u_n = q$. This leads us to the conjecture that it is also possible in the general case to put (2.1) in a form which is "analogous" to (1.2) in some way to be specified. Once this conjecture will be verified, the ISP will then be solved in analogy with Ref. 1.

To prove this conjecture we first separate Eqs. (2.1) for even l from those for odd l and group each block of equations. Explicitly we obtain a pair of matrix Schrödinger equations, which can be viewed as another "representation" of (1.3),

$$Y^{\pm''} + [\lambda^{2n}I - V^{\pm}(\lambda,x)]\,Y^{\pm} = 0, \quad x\in\mathbb{R} \tag{2.2}$$

$$V^{\pm}(\lambda,x) = \sum_{p=0}^{n} u_p(x)[\sigma^{\pm}(\lambda)]^p = V^{\mp}(\lambda e^{\mp i\pi/n},x), \tag{2.3}$$

where I is the $n\times n$ identity matrix,

[a)] This work has been done as part of the program "Recherche Coopérative sur Programme No. 264: Etude interdisciplinaire des problèmes inverses".

[b)] Physique Mathématique et Théorique, Equipe de recherche associée au CNRS, No. 154.

$$Y^+ = \begin{pmatrix} y_0 \\ y_2 \\ \vdots \\ y_{2l} \\ \vdots \\ y_{2(n-1)} \end{pmatrix}, \quad Y^- = \begin{pmatrix} y_1 \\ y_3 \\ \vdots \\ y_{2l+1} \\ \vdots \\ y_{2n-1} \end{pmatrix},$$

$$\sigma^+(\lambda) = \lambda \begin{pmatrix} 1 & & & & & 0 \\ & \alpha & & & & \\ & & \ddots & & & \\ & & & \alpha^l & & \\ & & & & \ddots & \\ 0 & & & & & \alpha^{n-1} \end{pmatrix}, \tag{2.4}$$

$$\alpha = e^{2i\pi/n}, \quad \sigma^-(\lambda) = \sigma^+(\lambda e^{i\pi/n}). \tag{2.5}$$

It is important to remark that $\sigma^+(\lambda)$ and $\sigma^-(\lambda)$ obey the identity

$$[\sigma^+(\lambda)]^n = \lambda^n I, \quad [\sigma^-(\lambda)]^n = -\lambda^n I. \tag{2.6}$$

At this step we note that there are other matrices $\tilde{\sigma}^+(\lambda^n)$ and $\tilde{\sigma}^-(\lambda^n)$, whose dependence in λ is only expressed in term of λ^n and which are also, respectively, the n^{th} root of $\lambda^n I$ and $-\lambda^n I$:

$$\tilde{\sigma}^\pm(\lambda^n) = \begin{pmatrix} 0 & 0 & \cdots & 0 & \pm\lambda^n \\ 1 & 0 & & & 0 \\ 0 & 1 & 0 & & \vdots \\ \vdots & 0 & \ddots & \ddots & \\ & & \ddots & & \\ 0 & 0 & \cdots 0 & 1 & 0 \end{pmatrix},$$

$$[\tilde{\sigma}^\pm(\lambda^n)]^n = \pm\lambda^n I. \tag{2.7}$$

Clearly, $\sigma^\pm(\lambda)$ and $\tilde{\sigma}^\pm(\lambda^n)$ are equivalent matrices, i.e., there exists a matrix $P^\pm(\lambda)$ such that

$$\tilde{\sigma}^\pm(\lambda^n) = [P^\pm(\lambda)]^{-1}\sigma^\pm(\lambda)P^\pm(\lambda). \tag{2.8}$$

We first calculate $P^+(\lambda)$. To this end, we introduce two bases in $\mathbb{R}^n$: $(e_1,...,e_n)$ and $(e'_1,...,e'_n)$ and we note the linear application f such that its matricial representation in the basis $(e_1,...,e_n)$ is $\sigma^+(\lambda)$ and its matricial representation in the basis $(e'_1,...,e'_n)$ is $\tilde{\sigma}^+(\lambda^n)$, i.e.,

$$f(e_i) = \lambda\alpha^{i-1}e_i, \quad i = 1,...,n \tag{2.9}$$

$$f(e'_i) = e'_{i+1}, \quad i = 1,...,n-1, f(e'_n) = \lambda^n e'_1. \tag{2.10}$$

We recall that if x_i^j $(j = 1,...,n)$ are the components of e'_i in the basis $(e_1,...,e_n)$ then x_i^j $(j = 1,...,n)$ are the elements of the ith column of $P^+(\lambda)$. Substituting e'_i by $\Sigma_{j=1}^n x_i^j e_j$ in (2.10) and using (2.9) we obtain, for all i

$$\begin{cases} x_2^i = \lambda\alpha^{i-1}x_1^i \\ x_p^i = \lambda\alpha^{i-1}x_{p-1}^i, \quad p = 2,...,n. \\ x_n^i = (\lambda^n/\alpha^{i-1})x_1^i \end{cases} \tag{2.11}$$

If we choose $x_1^i = 1$, $i = 1,...,n$, we find

$$P^+(\lambda) = M(\alpha)D(\lambda), \tag{2.12}$$

where

$$M(\alpha) = \begin{pmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & \alpha & \alpha^2 & \cdots & \alpha^{n-1} \\ 1 & \alpha^2 & \alpha^4 & \cdots & \alpha^{2(n-1)} \\ \cdot & \cdot & \cdot & & \cdot \\ \cdot & \cdot & \cdot & & \cdot \\ \cdot & \cdot & \cdot & & \cdot \\ 1 & \alpha^{n-1} & \alpha^{2(n-1)} & & \alpha^{(n-1)^2} \end{pmatrix},$$

$$D(\lambda) = \begin{pmatrix} 1 & & & & 0 \\ & \lambda & & & \\ & & \lambda^2 & & \\ & & & \ddots & \\ 0 & & & & \lambda^{n-1} \end{pmatrix}. \tag{2.13}$$

A glance at (2.3) allows us to write

$$P^-(\lambda) = P^+(\lambda e^{i\pi/n}). \tag{2.14}$$

Using properties of α, for example, $1 + \alpha + \alpha^2 + ... + \alpha^{n-1} = 0$, it is not difficult to calculate $[P^+(\lambda)]^{-1}$:

$$[P^+(\lambda)]^{-1} = (1/n)D(\lambda^{-1})\, M(\alpha^{-1}). \tag{2.15}$$

Setting now

$$\tilde{Y}^\pm = [P^\pm(\lambda)]^{-1}Y^\pm, \tag{2.16}$$

we deduce from Eq. (2.2) another representation of Eq. (2.1),

$$\tilde{Y}^{\pm\prime\prime} + [k^2 I - \tilde{V}^\pm(k,x)]\tilde{Y}^\pm = 0, \quad x\in\mathbb{R} \tag{2.17}$$

$$\tilde{V}^\pm(k,x) = [P^\pm(\lambda)]^{-1}V^\pm(\lambda,x)P^\pm(\lambda), \quad k = \lambda^n. \tag{2.18}$$

Substituting $V^\pm(\lambda,x)$ by (2.3) in (2.18) and using (2.8), (2.7), and cyclical properties of $\tilde{\sigma}^\pm(k)$, we finally obtain

$$V^\pm(k,x) = \sum_{p=0}^{n} u_p(x)[\tilde{\sigma}^\pm(k)]^p = U(x) \pm k\,Q(x), \tag{2.19}$$

where

$$U = \begin{pmatrix} u_0 & & & & \\ u_1 & u_0 & & 0 & \\ u_2 & u_1 & u_0 & & \\ & & \ddots & \ddots & \\ u_{n-1} & \cdots & u_2 & u_1 & u_0 \end{pmatrix},$$

$$Q = \begin{pmatrix} u_n & u_{n-1} & \cdots & u_1 \\ & u_n & \ddots & \vdots \\ & & \ddots & \vdots \\ & 0 & \ddots & u_{n-1} \\ & & & u_n \end{pmatrix} \tag{2.20}$$

To sum up, we have obtained three equivalent "representations" (2.1), (2.2), (2.17) for Eq. (1.3). We now have to verify that these equivalences are "canonical", i.e., they "preserve" the scattering data information, so that we can go easily from one formulation of the ISP in a representation to another. The "natural" representation is the first one, (2.1). The family of nonlinear equations obtained in Ref. 11 originates from it. The "good" representation to solve the ISP is the third one, (2.17), because of the analogy with (1.2). Note that going from (2.2) to (2.17) we have lost the diagonality of the potential $V(\lambda,x)$ but we have won an easier dependence in λ.

III. COMPARISON BETWEEN THE SCATTERING DATA FOR (2.1), (2.2), AND (2.17)

First we remark that in order to define the scattering data for a matrix Schrödinger equation we have in general to make use of matrix solutions (for the energy-independent case and for the solution of the corresponding ISP, see Ref. 13). Nevertheless, because of the special features of the matrix potentials in (2.2) and (2.17), it is sufficient here to make use of vector solutions. This will greatly simplify the solution of the ISP.

A. The right and left Jost solutions

We suppose that the potentials u_i $(i = 0...n)$ satisfy the following conditions D_1 and D_2:

D_1: For $i = 0,...,(n-1)$, $u_i(x)$ $(x\in\mathbb{R})$ is continuously differentiable, and $x\,u_i(x)$, $u'_i(x)$ are integrable on $\mathbb{R}$.

D_2: $u_n(x)$ $(x\in\mathbb{R})$ is twice continuously differentiable, and $u_n(x)$, $u'_n(x)$, $u''_n(x)$ are integrable on $\mathbb{R}$.

The right and left Jost solutions $f_l(\lambda,x)$ and $\overleftarrow{f}_l(\lambda,x)$ of $(2.1)_l$, respectively, $F^{\pm}(\lambda,x)$ and $\overleftarrow{F}^{\pm}(\lambda,x)$ of $(2.2)_{\pm}$, respectively, $\tilde F^{\pm}(k,x)$ and $\overleftarrow{\tilde F}^{\pm}(k,x)$ of $(2.17)_{\pm}$, are defined as follows:

$$f_l(\lambda,x) \underset{x\to\infty}{\sim} e^{i\lambda^n x}, \quad \overleftarrow{f}_l(\lambda,x) \underset{x\to-\infty}{\sim} e^{-i\lambda^n x}, \tag{3.1}$$

$$F^{\pm}(\lambda,x) \underset{x\to\infty}{\sim} e^{i\lambda^n x}(1,..,1)^T, \quad \overleftarrow{F}^{\pm}(\lambda,x) \underset{x\to-\infty}{\sim} e^{-i\lambda^n x}(1,..,1)^T, \tag{3.2}$$

$$\tilde F^{\pm}(k,x) \underset{x\to\infty}{\sim} e^{ikx}V, \quad \overleftarrow{\tilde F}^{\pm}(k,x) \underset{x\to-\infty}{\sim} e^{-ikx}V, \; k=\lambda^n, \tag{3.3}$$

where T means "transposed" and $V = (1,0,..,0)^T$.

$f_l(\lambda,x)$ and $\overleftarrow{f}_l(\lambda,x)$ are defined equivalently as the solution in the class of continuous functions for real x of the following integral equations:

$$f_l(\lambda,x) = e^{i\lambda^n x} + \int_x^{\infty} \frac{\sin\lambda^n(y-x)}{\lambda^n} \times\Big[\sum_{p=0}^{n}\lambda^p e^{ipl\pi/n}u_p(y)\Big] f_l(\lambda,y)dy, \tag{3.4}$$

$$\overleftarrow{f}_l(\lambda,x) = e^{-i\lambda^n x} + \int_{-\infty}^{x} \frac{\sin\lambda^n(x-y)}{\lambda^n} \times\Big[\sum_{p=0}^{n}\lambda^p e^{ipl\pi/n}u_p(y)\Big] \overleftarrow{f}_l(\lambda,x)dy. \tag{3.5}$$

$f_l(\lambda,x)$ and $\overleftarrow{f}_l(\lambda,x)$ are (for fixed x) defined and continuous for $0\leqslant\arg\lambda\leqslant\pi/n$, analytic for $0<\arg\lambda<\pi/n$ and obey the bounds

$$|f_l(\lambda,x)|\leqslant e^{-bx}e^{d(x)}, \quad 0\leqslant\arg\lambda\leqslant\pi/n, \quad b=\operatorname{Im}\lambda^n\geqslant 0, \tag{3.6}$$

$$|\overleftarrow{f}_l^{\pm}(\lambda,x)|\leqslant e^{bx}e^{\overleftarrow{d}(x)}, \quad 0\leqslant\arg\lambda\leqslant\pi/n, \quad b\geqslant 0, \tag{3.7}$$

where

$$d(x) = 2\int_x^{\infty}(y-x+1)\sum_{p=0}^{n}|u_p(y)|dy, \tag{3.8}$$

$$\overleftarrow{d}(x) = 2\int_{-\infty}^{x}(x-y+1)\sum_{p=0}^{n}|u_p(y)|dy. \tag{3.9}$$

It is clear that $F^{\pm}(\lambda,x)$ and $\overleftarrow{F}^{\pm}(\lambda,x)$ are also defined and continuous for $0\leqslant\arg\lambda\leqslant\pi/n$, analytic for $0<\arg\lambda<\pi/n$ and verify

$$F^{\pm}(\lambda,x) = e^{i\lambda^n x}(1,...,1)^T + \int_x^{\infty}\frac{\sin\lambda^n(y-x)}{\lambda^n}V^{\pm}(\lambda,x) \times F^{\pm}(\lambda,y)dy, \tag{3.10}$$

$$\overleftarrow{F}^{\pm}(\lambda,x) = e^{-i\lambda^n x}(1,...,1)^T + \int_{-\infty}^{x}\frac{\sin\lambda^n(x-y)}{\lambda^n}V^{\pm}(\lambda,x) \times \overleftarrow{F}^{\pm}(\lambda,y)dy. \tag{3.11}$$

Using the formulas (2.15) and (2.17), we have

$$\tilde F^{\pm}(k,x) = e^{ikx}V + \int_x^{\infty}\frac{\sin k(y-x)}{k}\tilde V^{\pm}(k,y) \times \tilde F^{\pm}(k,y)dy, \tag{3.12}$$

$$\overleftarrow{\tilde F}^{\pm}(k,x) = e^{-ikx}V + \int_{-\infty}^{x}\frac{\sin k(x-y)}{k}\tilde V^{\pm}(k,y) \times \overleftarrow{\tilde F}^{\pm}(k,y)dy. \tag{3.13}$$

It is not difficult to prove that $\tilde F^{\pm}(k,x)$ and $\overleftarrow{\tilde F}^{\pm}(k,x)$ are defined and continuous for $\operatorname{Im} k\geqslant 0$, analytic for $\operatorname{Im} k>0$, and admit the following bounds:

$$\|\tilde F^{\pm}(k,x)\|\leqslant e^{-bx}e^{h(x)}, \quad x\in\mathbb{R}, \quad b=\operatorname{Im}k\geqslant 0, \tag{3.14}$$

$$\|\overleftarrow{\tilde F}^{\pm}(k,x)\|\leqslant e^{bx}e^{\overleftarrow{h}(x)}, \quad x\in\mathbb{R}, \quad b\geqslant 0, \tag{3.15}$$

where

$$h(x) = 2\int_x^{\infty}\Big[(y-x)\sum_{p=0}^{n-1}|u_p(y)| + \sum_{p=1}^{n}|u_p(y)|\Big]dy, \tag{3.16}$$

$$\overleftarrow{h}(x) = 2\int_{-\infty}^{x}\Big[(x-y)\sum_{p=0}^{n-1}|u_p(y)| + \sum_{p=1}^{n}|u_p(y)|\Big]dy, \tag{3.17}$$

$$\|(v_1,...,v_n)\| = \max_{i=1..n}|v_i|, (v_1,...,v_n)^T\in\mathbb{R}^n. \tag{3.18}$$

B. Reflection coefficients

For $\lambda>0$, $f_l(\lambda,x)$ and $f_{l-1}(\lambda e^{i\pi/n},x)$ form a fundamental system of solutions of $(2.1)_l$. So, for all l and with the convention $f_{-1} = f_{2n-1}$, we have the relation

$$\overleftarrow{f}_l(\lambda,x) = b_l(\lambda)f_l(\lambda,x) + a_l(\lambda)f_{l-1}(\lambda e^{i\pi/n},x), \quad \lambda>0 \tag{3.19}$$

where

$$a_l(\lambda) = (1/2i\lambda^n)W[\overleftarrow{f}_l(\lambda,x),f_l(\lambda,x)], \tag{3.20}$$

$$b_l(\lambda) = -(1/2i\lambda^n)W[\overleftarrow{f}_l(\lambda,x),f_{l-1}(\lambda e^{i\pi/n},x)]; \tag{3.21}$$

$W[f,g]$ is the Wronskian of f and g.

We see from formula (3.20) that the function $a_l(\lambda)$ admits a unique continuous extension $a_l(\lambda)$ $(0\leqslant\arg\lambda\leqslant\pi/n)$ which is analytic for $0<\arg\lambda<\pi/n$. [For $\lambda=0$, by using supplementary conditions on potentials it is possible to get over the difficulty. For the greatest accuracy, see Ref. 1. (case $x\in\mathbb{R}$)]. Because of the convention $f_{-1} = f_{2n-1}$, we have

$$[f_{2n-1}(\lambda,x), f_1(\lambda,x),\dots,f_{2n-3}(\lambda,x)]^T = \tilde{\sigma}^+(1)[f_1(\lambda,x),\dots,f_{2n-1}(\lambda,x)]^T, \tag{3.22}$$

where $\tilde{\sigma}^+(1)$ has been defined in (2.7).
It follows from (3.19) and (3.22) that diagonal matrices $A^{\pm}(\lambda)$ and $B^{\pm}(\lambda)$ exist and verify

$$\overleftarrow{F}^+(\lambda,x) = B^+(\lambda)F^+(\lambda,x) + A^+(\lambda)\tilde{\sigma}^+(1)F^-(\lambda e^{i\pi/n},x), \quad \lambda>0 \tag{3.23}$$

$$\overleftarrow{F}^-(\lambda,x) = B^-(\lambda)F^-(\lambda,x) + A^-(\lambda)F^+(\lambda e^{i\pi/n},x), \quad \lambda>0 \tag{3.24}$$

and have the form

$$A^+(\lambda) = \begin{pmatrix} a_0(\lambda) & & & & 0 \\ & \ddots & & & \\ & & a_{2l}(\lambda) & & \\ & & & \ddots & \\ 0 & & & & a_{2(n-1)}(\lambda) \end{pmatrix},$$

$$A^-(\lambda) = \begin{pmatrix} a_1(\lambda) & & & & 0 \\ & \ddots & & & \\ & & a_{2l+1}(\lambda) & & \\ & & & \ddots & \\ 0 & & & & a_{2n-1}(\lambda) \end{pmatrix}, \tag{3.25}$$

$$B^+(\lambda) = \begin{pmatrix} b_0(\lambda) & & & & 0 \\ & \ddots & & & \\ & & b_{2l}(\lambda) & & \\ & & & \ddots & \\ 0 & & & & b_{2(n-1)}(\lambda) \end{pmatrix},$$

$$B^-(\lambda) = \begin{pmatrix} b_1(\lambda) & & & & 0 \\ & \ddots & & & \\ & & b_{2l+1}(\lambda) & & \\ & & & \ddots & \\ 0 & & & & b_{2n-1}(\lambda) \end{pmatrix}. \tag{3.26}$$

Clearly, $A^{\pm}(\lambda)$ can be defined and continuous for $0\leqslant\arg\lambda\leqslant\pi/n$ and analytic for $0<\arg\lambda<\pi/n$.
Starting from relations (3.23) and (3.24), taking into account the formula (2.16) and the equality

$$[P^+(\lambda)]^{-1}\tilde{\sigma}^+(1) = [P^-(\lambda e^{i\pi/n})]^{-1}, \tag{3.27}$$

we obtain

$$\overleftarrow{\tilde{F}}^{\pm}(k,x) = \tilde{B}^{\pm}(k)\tilde{F}^{\pm}(k,x) + \tilde{A}^{\pm}(k)\tilde{F}^{\pm}(-k,x), \quad k=\lambda^n, \quad k\in\mathbb{R} \tag{3.28}$$

where

$$\tilde{A}^{\pm}(k) = [P^{\pm}(\lambda)]^{-1}A^{\pm}(\lambda)P^{\pm}(\lambda),$$
$$\tilde{B}^{\pm}(k) = [P^{\pm}(\lambda)]^{-1}B^{\pm}(\lambda)P^{\pm}(\lambda). \tag{3.29}$$

The function $\lambda = k^{1/n}$ being continuous for $0\leqslant\arg k\leqslant\pi$, analytic for $0<\arg k<\pi$, $\tilde{A}^{\pm}(k)$ is continuous for $\text{Im}\, k\geqslant 0$ and analytic for $\text{Im}\, k>0$. The reflection coefficients (to the right) $r_l(\lambda)$ for $(2.1)_l$, respectively, $R^{\pm}(\lambda)$ for $(2.2)_{\pm}$, respectively, $\tilde{R}^{\pm}(k)$ for $(2.17)_{\pm}$, are defined and connected as follows:

$$r_l(\lambda) = b_l(\lambda)/a_l(\lambda), \quad \lambda>0, \tag{3.30}$$

$$R^{\pm}(\lambda) = [A^{\pm}(\lambda)]^{-1}B^{\pm}(\lambda), \quad \lambda>0, \tag{3.31}$$

$$\tilde{R}^{\pm}(k) = [\tilde{A}^{\pm}(k)]^{-1}\tilde{B}^{\pm}(k) = [P^{\pm}(\lambda)]^{-1}R^{\pm}(\lambda)P^{\pm}(\lambda), k=\lambda^n, \quad k\in\mathbb{R}. \tag{3.32}$$

$r_l(\lambda)$ and $R^{\pm}(\lambda)$ are continuous functions for $\lambda>0$. And so, $\tilde{R}^{\pm}(k)$ is continuous for $k\in\mathbb{R}$ and $\tilde{B}^{\pm}(k)$ too.

C. Bound states

The "bound states" of $(2.1)_l$, i.e., the square integrable solutions, correspond to the zeros $\lambda_{l,j}$ $(j=1,2,..,J_l)$ of $a_l(\lambda)$. We impose the condition D_3:

$$\begin{cases} \text{The zeros } \lambda_{l,j} \text{ of } a_l(\lambda) \text{ are simple, in finite,} \\ \text{number } J_l;\ 0<\arg\lambda_{l,j}<\pi/n \text{ and } \lambda_{l,j}\neq\lambda_{l',j}, \text{ if } l\neq l', \\ \text{have the same parity.} \end{cases}$$

The "bound states" of $(2.2)_{\pm}$ [respectively, of $(2.17)_{\pm}$], i.e., the square integrable vector solutions, correspond to the zeros $\lambda_{m\pm}$ $(m^{\pm}=1,2,\dots,M^{\pm})$ of $\det A^{\pm}(\lambda)$ [respectively, to the zeros $k_{m\pm} = (\lambda_{m\pm})^n$ $(m^{\pm}=1,2,..,M^{\pm})$ of $\det\tilde{A}^{\pm}(k)$.
It is clear that

$$\{\lambda_{m^+}, m^+=1,2,..,M^+\} = \{\lambda_{2l,j}; l=0,1,..,n-1; \; j=1,2,..,J_l\}, \tag{3.33}$$

$$\{\lambda_{m^-}, m^-=1,2,..,M^-\} = \{\lambda_{2l+1,j}; l=0,1,..,n-1; \; j=1,2,..,J_l\}. \tag{3.34}$$

To each zero $\lambda_{l,j}$ of $a_l(\lambda)$, respectively, $\lambda_{m\pm}$ of $\det A^{\pm}(\lambda)$, respectively, $k_{m\pm}$ of $\tilde{A}^{\pm}(k)$, we associate a constant scalar $c_{l,j}$, respectively, matrix $C_{m\pm}$:

$$c_{l,j} = i\lim_{\lambda\to\lambda_{l,j}}(\lambda-\lambda_{l,j})\frac{b_l(\lambda)}{a_l(\lambda)}, \tag{3.35}$$

$$C_{m^+} = n(\lambda_{m^+})^{n-1}\begin{pmatrix} 0 & & & 0 \\ & 0 & & \\ & & c_{2l,j} & \\ & & & 0 \\ 0 & & & 0 \end{pmatrix} \quad \text{if } \lambda_{m^+} = \lambda_{2l,j},$$

$$C_{m^-} = n(\lambda_{m^-})^{n-1}\begin{pmatrix} 0 & & & 0 \\ & 0 & & \\ & & c_{2l+1,j} & \\ & & & 0 \\ 0 & & & 0 \end{pmatrix} \quad \text{if } \lambda_{m^-} = \lambda_{2l+1,j} \tag{3.36}$$

$$\tilde{C}_{m\pm} = [P^{\pm}(\lambda_{m\pm})]^{-1}C_{m\pm}P^{\pm}(\lambda_{m\pm}), \quad k_{m\pm} = (\lambda_{m\pm})^n. \tag{3.37}$$

We define the scattering data s for (2.1), respectively, $\mathscr{S}$ for (2.2), respectively, $\tilde{\mathscr{S}}$ for (2.17) by

$$s = \{r_l(\lambda),(\lambda>0);\lambda_{l,j};c_{l,j}(j=1,2,..,J_l),(l=0,1,..,2n-1)\}, \tag{3.38}$$

$$\mathscr{S} = \{R^{\pm}(\lambda),(\lambda>0);\lambda_{m^{\pm}};C_{m^{\pm}}(m^{\pm}=1,2,..,M^{\pm})\}, \tag{3.39}$$

$$\tilde{\mathscr{S}} = \{\tilde{R}^{\pm}(k),(k\in\mathbb{R});k_{m^{\pm}};\tilde{C}_{m^{\pm}}(m^{\pm}=1,2,..,M^{\pm})\}. \tag{3.40}$$

Clearly, the scattering data s, $\mathscr{S}$, and $\tilde{\mathscr{S}}$ are equivalent. So are the corresponding ISP for (2.1), (2.2), and (2.17). Therefore, later on we just consider the ISP for (2.17).

IV. STUDY OF THE ISP FOR (2.17)

First, we start from Eq. (2.17) with the potentials satisfying the conditions D_1 and D_2. In Sec. 4.1, we show that $\tilde{F}^{\pm}(k,x)$ can be determined by two functions $f^{\pm}(x)$ (scalar) and $A^{\pm}(x,t)$ ($\mathbb{R}^n$ vector), through an integral representation; $f^{\pm}(x)$ and $A^{\pm}(x,t)$ are solutions of a partial differential equation system. In Sec. 4.2 we deduce some properties of $\tilde{A}^{\pm}(k)$, $\tilde{B}^{\pm}(k)$, and $\tilde{R}^{\pm}(k)$. In Sec. 4.3, we establish the "inversion integral equations" with a coupling condition. Finally, in Sec. 4.4, we show how to construct the potentials from the scattering data $\tilde{\mathscr{S}}$.

A. The Jost solutions

We recall the $\tilde{F}^{\pm}(k,x)$ respectively, $\overleftarrow{\tilde{F}}^{\pm}(k,x)$, is defined equivalently as the solution in the class of continuous functions for real x of Eq. (3.12), respectively, (3.13) and $\tilde{F}^{\pm}(k,x)$ and $\overleftarrow{\tilde{F}}^{\pm}(k,x)$ are (for fixed x) continuous for Im $k\geqslant 0$ and analytic for Im $k>0$. By applying the successive approximation method to Eq. (3.12) and (3.13), we find the behavior for large values of $|k|$ of these functions. If $\tilde{F}_0^{\pm}(k,x),\ldots,\tilde{F}_{n-1}^{\pm}(k,x)$, respectively, $\overleftarrow{\tilde{F}}_0^{\pm}(k,x),\ldots,\overleftarrow{\tilde{F}}_{n-1}^{\pm}(k,x)$ are the components of $\tilde{F}^{\pm}(k,x)$, respectively, $\overleftarrow{\tilde{F}}^{\pm}(k,x)$, we can write the results whose proof is given in the Appendix:

$$\tilde{F}^{\pm}(k,x) = e^{ikx}f^{\pm}(x)V + \frac{e^{ikx}}{k}W(x) + O\left(\frac{1}{k^2}\right), \quad \text{Im } k\geqslant 0,\ |k|\to\infty, \tag{4.1}$$

$$\overleftarrow{\tilde{F}}^{\pm}(k,x) = e^{-ikx}\overleftarrow{f}^{\pm}(x)V + \frac{e^{-ikx}}{k}\overleftarrow{W}(x) + O\left(\frac{1}{k^2}\right), \quad \text{Im } k\geqslant 0,\ |k|\to\infty, \tag{4.2}$$

where

$$f^{\pm}(x) = \exp\left[\pm\frac{i}{2}\int_x^{\infty}u_n(t)dt\right], \tag{4.3}$$

$$\overleftarrow{f}^{\pm}(x) = \exp\left[\pm\frac{i}{2}\int_{-\infty}^{x}u_n(t)dt\right], \tag{4.4}$$

$W(x)$ and $\overleftarrow{W}(x)$ are $\mathbb{R}^n$ vectors.

Consequently, $F^{\pm}(k,x) - e^{ikx}f^{\pm}(x)V$, for fixed x, belongs to $L_2(\mathbb{R})$, and admits a Fourier transform. In fact, similarly to the case $n=1$ (cf. Ref. 1), $\tilde{F}^{\pm}(k,x)$ has the following representation:

$$\tilde{F}^{\pm}(k,x) = e^{ikx}f^{\pm}(x)V + \int_x^{\infty}A^{\pm}(x,t)e^{ikt}dt, \quad \text{Im } k\geqslant 0,\ x\in\mathbb{R}, \tag{4.5}$$

where $f^{\pm}(x)$ has been defined by the formula (4.3) and $A^{\pm}(x,t) = (A_0^{\pm}(x,t),\ldots,A_{n-1}^{\pm}(x,t))$ is the $\mathbb{R}^n$-valued function solution of the partial differential equation system:

$$\left[\frac{\partial^2}{\partial x^2} - \frac{\partial^2}{\partial t^2} - U(x) \mp iQ(x)\frac{\partial}{\partial t}\right]A^{\pm}(x,t) = 0, \quad t>x, \tag{4.6}$$

$$f^{\pm\prime\prime}(x)V - 2\frac{d}{dx}A^{\pm}(x,x) \mp iQ(x)A^{\pm}(x,x) - U(x)f^{\pm}(x)V = 0, \tag{4.7}$$

and the condition $A^{\pm}(x,\infty) = 0$. (4.8)

Indeed, we start from Eq. (2.17) in which we substitute $Y^{\pm}$ by $\tilde{F}^{\pm}(k,x)$ given by (4.5). After different integration by parts, we obtain

$$[2if^{\pm\prime}(x)V \mp f^{\pm}(x)Q(x)V]ke^{ikx} + \left[f^{\pm\prime\prime}(x)V - 2\frac{d}{dx}A^{\pm}(x,x) \mp iQ(x)A^{\pm}(x,x) - U(x)f^{\pm}(x)V\right]e^{ikx} + \int_x^{\infty}\left[\frac{\partial^2}{\partial x^2} - \frac{\partial^2}{\partial t^2} - U(x) \mp iQ(x)\frac{\partial}{\partial t}\right]A^{\pm}(x,t)e^{ikt}dt = 0; \tag{4.9}$$

and then we deduce the formula (4.3) and the relations (4.6) and (4.7). It is important to remark that, if we seek U and Q in the form given by (2.20), we can construct them from f^+, f^-, A^+, and A^-. Using the formula (2.20) in (4.6) and taking into account the relation (4.3) we obtain the triangular system with $(n+1)$ equations and $(n+1)$ unknown values $u_0,u_1,\ldots,u_n$:

$$\mp i\sum_{i=0}^{n-1}u_{n-i}(x)A_i^{\pm}(x,x) - u_0(x)f^{\pm}(x) = 2\frac{d}{dx}A_0^{\pm}(x,x) + f^{\pm\prime\prime}(x)$$

$$\mp i\sum_{i=0}^{n-p-1}u_{n-i}(x)A_{i+p}^{\pm}(x,x) - u_p(x)f^{\pm}(x) = 2\frac{d}{dx}A_p^{\pm}(x,x); \quad p=1,2,..,(n-1), \tag{4.10}$$

$$u_n(x) = \pm 2i\frac{f^{\pm\prime}(x)}{f^{\pm}(x)}.$$

Clearly, $u_0(x),\ldots,u_n(x)$ are uniquely determined by the system (4.10).

B. Some properties of $\tilde{A}^{\pm}(k)$, $\tilde{B}^{\pm}(k)$, $\tilde{A}^{\pm}(k)^{-1}$, and $\tilde{R}^{\pm}(k)$

First, we precise the form of the matrices $\tilde{A}^{\pm}(k)$ and $\tilde{B}^{\pm}(k)$. We remark that the matrices $[\sigma^{+}(\lambda)]^{p}$, respectively, $[\sigma^{-}(\lambda)]^{p}$, $p=0,\dots,n-1$, being respectively, the nth of $\lambda^{n}I$ and $(-\lambda^{n})I$, form a basic system in the space of the diagonal matrices. So, we can find unique scalar $\alpha_p^{\pm}(\lambda)$ and $\beta_p^{\pm}(\lambda)$ such that

$$A^{\pm}(\lambda)=\sum_{p=0}^{n-1}\alpha_p^{\pm}(\lambda)[\sigma^{\pm}(\lambda)]^{p},\quad \lambda>0 \tag{4.11}$$

$$B^{\pm}(\lambda)=\sum_{p=0}^{n-1}\beta_p^{\pm}(\lambda)[\sigma^{\pm}(\lambda)]^{p},\quad \lambda>0. \tag{4.12}$$

Applying the formula (3.29) and by analogy with the computation of $\tilde{V}^{\pm}(k,x)$ from $V^{\pm}(\lambda,x)$, we obtain the matrices

$$\tilde{A}^{\pm}(k)=\begin{pmatrix}\tilde{\alpha}_0^{\pm}(k) & (\pm k)\tilde{\alpha}_{n-1}^{\pm}(k) & \cdots & (\pm k)\tilde{\alpha}_1^{\pm}(k)\\ \tilde{\alpha}_1^{\pm}(k) & \ddots & \ddots & \vdots\\ \vdots & \ddots & \tilde{\alpha}_0^{\pm}(k) & (\pm k)\tilde{\alpha}_{n-1}(k)\\ \tilde{\alpha}_{n-1}^{\pm}(k) & \cdots & \tilde{\alpha}_1^{\pm}(k) & \tilde{\alpha}_0^{\pm}(k)\end{pmatrix}$$

$$\tilde{B}^{\pm}(k)=\begin{pmatrix}\tilde{\beta}_0^{\pm}(k) & (\pm k)\tilde{\beta}_{n-1}^{\pm}(k) & \cdots & (\pm k)\tilde{\beta}_1^{\pm}(k)\\ \tilde{\beta}_1^{\pm}(k) & \ddots & \tilde{\beta}_0^{\pm}(k) & \vdots\\ \vdots & \ddots & \ddots & (\pm k)\tilde{\beta}_{n-1}^{\pm}(k)\\ \tilde{\beta}_{n-1}^{\pm}(k) & \cdots & \tilde{\beta}_1^{\pm}(k) & \tilde{\beta}_0^{\pm}(k)\end{pmatrix}, \tag{4.13}$$

where $\tilde{\alpha}_p^{\pm}(k)=\alpha_p^{\pm}(\lambda)$, $\tilde{\beta}_p^{\pm}(k)=\beta_p^{\pm}(\lambda)$, $p=0,\dots,(n-1)$. Let us note that $[\tilde{A}^{\pm}(k)]^{-1}$ and $\tilde{R}^{\pm}(k)$ can have a matricial representation as $\tilde{A}^{\pm}(k)$.

We also need to know the estimate for $|k|\to\infty$ of $\tilde{A}^{\pm}(k)$, $\tilde{B}^{\pm}(k)$, $[\tilde{A}^{\pm}(k)]^{-1}$, and $\tilde{R}^{\pm}(k)$. To this end, it is convenient to rewrite Eq. (3.13) thus:

$$\overleftarrow{\tilde{F}}^{\pm}(k,x)=e^{ikx}\Big[V-\int_{-\infty}^{x}\frac{e^{iky}}{2ik}\tilde{V}^{\pm}(k,y)\overleftarrow{\mathbf{F}}^{\pm}(k,y)dy\Big] + e^{ikx}\Big[\int_{-\infty}^{x}\frac{e^{-iky}}{2ik}\tilde{V}^{\pm}(k,y)\tilde{F}^{\pm}(k,y)dy\Big]. \tag{4.14}$$

Looking at the formulas (3.28) and (4.14) when $x\to\infty$, we obtain

$$\tilde{A}^{\pm}(k)V=V-\int_{-\infty}^{\infty}\frac{e^{iky}}{2ik}\tilde{V}^{\pm}(k,y)\overleftarrow{\tilde{F}}^{\pm}(k,y)\,dy,\quad \operatorname{Im}k\geqslant 0 \tag{4.15}$$

$$\tilde{B}^{\pm}(k)V=\int_{-\infty}^{\infty}\frac{e^{-iky}}{2ik}\tilde{V}^{\pm}(k,y)\overleftarrow{\tilde{F}}^{\pm}(k,y)\,dy,\quad k\in\mathbf{R}. \tag{4.16}$$

Using different integrations by parts and thanks to the bound (4.2), we find

$$\tilde{A}^{\pm}(k)V=(\tilde{\alpha}_0^{\pm}(k),\dots,\tilde{\alpha}_{n-1}^{\pm}(k))^{T} =\overleftarrow{f}^{\pm}(\infty)+\frac{W}{k}+O\Big(\frac{1}{k^2}\Big),\quad \operatorname{Im}k\geqslant 0,\ |k|\to\infty, \tag{4.17}$$

$$\tilde{B}^{\pm}(k)V=(\tilde{\beta}_0^{\pm}(k),\dots,\tilde{\beta}_{n-1}^{\pm}(k))^{T} =\frac{1}{k^2}O\Big(\frac{1}{k}\Big),\quad k\in\mathbf{R},\ |k|\to\infty, \tag{4.18}$$

where W is a constant $\mathbf{R}^n$ vector.

Thanks to the matricial representation of $\tilde{A}^{\pm}(k)$ and $\tilde{B}^{\pm}(k)$, we derive the following results for $|k|\to\infty$:

$$\tilde{A}^{\pm}(k)=\overleftarrow{f}^{\pm}(\infty)I+T+O\Big(\frac{1}{k}\Big),\quad \operatorname{Im}k\geqslant 0 \tag{4.19}$$

$$[\tilde{A}^{\pm}(k)]^{-1}=\overleftarrow{f}^{\mp}(\infty)+T'+O\Big(\frac{1}{k}\Big),\quad \operatorname{Im}k\geqslant 0,\ k\neq k_{m^{\pm}}, \tag{4.20}$$

where T and T' are constant superior triangular matrices with zeros on the diagonal, and

$$\tilde{B}^{\pm}(k)=\frac{1}{k}O(1),\quad k\in\mathbf{R}, \tag{4.21}$$

$$\tilde{R}^{\pm}(k)=\frac{1}{k}O(1),\quad k\in\mathbf{R}, \tag{4.22}$$

$$\det\tilde{A}^{\pm}(k)=[\overleftarrow{f}^{\pm}(\infty)]^{n}+O\Big(\frac{1}{k}\Big),\quad \operatorname{Im}k\geqslant 0 \tag{4.23}$$

$$\det[\tilde{A}^{\pm}(k)]^{-1}=[\overleftarrow{\mathbf{f}}^{\mp}(\infty)]^{n}+O\Big(\frac{1}{k}\Big),\quad \operatorname{Im}k\geqslant 0,\ k\neq k_{m^{\pm}}. \tag{4.24}$$

Let us remark that $\tilde{R}^{\pm}(k)$ has a Fourier transform in $L_2(\mathbf{R})$.

C. Inversion equations and coupling condition

In order to establish these equations, we start from the formula (3.28) written in the form

$$[\tilde{A}^{\pm}(k)]^{-1}\overleftarrow{\tilde{\mathbf{F}}}^{\pm}(k,x)-[\tilde{A}^{\pm}(k)]^{-1}\tilde{B}^{\pm}(k)\tilde{F}^{\pm}(k,x) =\tilde{F}^{\mp}(-k,x),\quad k\in\mathbf{R} \tag{4.25}$$

and in the equivalent form, for fixed x,

$$G_x^{\pm}(k)-H_x^{\pm}(k)=\tilde{F}^{\mp}(-k,x)-e^{-ikx}f^{\mp}(x)V =\int_{x}^{\infty}A^{\mp}(x,t)e^{-ikt}dt, \tag{4.26}$$

where

$$G_x^{\pm}(k)=[\tilde{A}^{\pm}(k)]^{-1}\overleftarrow{\tilde{\mathbf{F}}}^{\pm}(k,x)-e^{-ikx}f^{\mp}(x)V, \tag{4.27}$$

$$H_x^{\pm}(k)=[\tilde{A}^{\pm}(k)]^{-1}\tilde{B}^{\pm}(k)\tilde{F}^{\pm}(k,x) =\tilde{R}^{\pm}(k)\tilde{F}^{\pm}(k,x). \tag{4.28}$$

Let us evaluate the Fourier transform of these two functions. The function $G_x^{\pm}(k)$ is continuous for $\operatorname{Im}k\geqslant 0$, $k\neq k_{m^{\pm}}$, and analytic for $\operatorname{Im}k>0$, $k\neq k_{m^{\pm}}$. It is obvious, from the formulas (4.3) and (4.4) that

$$\overleftarrow{\mathbf{f}}^{\pm}(x)\overleftarrow{\mathbf{f}}^{\mp}(\infty)=f^{\mp}(x), \tag{4.29}$$

and then $G_x^{\pm}(k)$ can be expressed as

$$G_x^{\pm}(k)=[[(\tilde{A}^{\pm}(k)]^{-1}-\overleftarrow{\mathbf{f}}^{\mp}(\infty)I] \times[\overleftarrow{\tilde{\mathbf{F}}}^{\pm}(k,x)-\overleftarrow{f}^{\pm}(x)e^{-ikx}V] + [[\tilde{A}^{\pm}(k)]^{-1}\overleftarrow{f}^{\mp}(\infty)I]\overleftarrow{f}^{\pm}(x)e^{-ikx}V +\overleftarrow{f}^{\mp}(\infty)[\tilde{F}^{\pm}(k,x)-\overleftarrow{f}^{\pm}(x)e^{-ikx}V]. \tag{4.30}$$

Using the bounds (4.20) and (4.2) we obtain

$$G_x^{\pm}(k)=e^{-ikx}O\Big(\frac{1}{k}\Big),\quad \operatorname{Im}k\geqslant 0,\ k\neq k_{m^{\pm}}. \tag{4.31}$$

We now consider the integral $\int_\Gamma G_x^{\pm}(k)e^{ikt}dk$ $(t>x)$ where Γ is the closed path in the upper half of the complex k plane and consisting of the segment $[-\rho,\rho]$ and of the half-circle $|k|=\rho$. Thanks to (4.31) we can apply a Jordan lemma to prove that the integral along the half-circle vanishes for $t>x$ and $\rho\to\infty$. So, we have

$$\lim_{\rho\to\infty}\int_{-\rho}^{\rho} G_x^{\pm}(k)e^{ikt}\,dk = 2i\pi\sum_{m^\pm=1}^{M^\pm}\operatorname{Res}([\tilde A^{\pm}(k)]^{-1}\overleftarrow{F}(k,x)e^{ikt},k_{m^\pm}). \quad (4.32)$$

It is clear that

$$\begin{aligned}\operatorname{Res}([\tilde A^{\pm}(k)]^{-1}\overleftarrow{\tilde F}^{\pm}(k,x)e^{ikt},k_{m^\pm}) &= \operatorname{Res}([\tilde A^{\pm}(k)]^{-1}\tilde B^{\pm}(k)\tilde F^{\pm}(k,x)e^{ikt},k_{m^\pm})\\ &= \lim_{k\to k_{m^\pm}}(k-k_{m^\pm})[\tilde A^{\pm}(k)]^{-1}\tilde B^{\pm}(k)\tilde F^{\pm}(k,x)e^{ikt}.\end{aligned} \quad (4.33)$$

Resorting to the formulas (3.29), (3.36), and (3.37), we finally find

$$\lim_{R\to\infty}\int_{-R}^{+R} G_x^{\pm}(k)e^{ikt}dk = 2\pi A^{\mp}(x,t) = \sum_{m^\pm=1}^{M^\pm}\tilde C_{m^\pm}\tilde F^{\pm}(k_{m^\pm},x)e^{i(k_{m^\pm})t}. \quad (4.34)$$

To obtain the Fourier transform of $H_x^{\pm}(k)$, we write thus

$$H_x^{\pm}(k) = \tilde R^{\pm}(k)[\tilde F^{\pm}(k,x) - f^{\pm}(x)e^{ikx}V] + \tilde R^{\pm}(k)f^{\pm}(x)e^{ikx}V. \quad (4.35)$$

Recalling the formulas (4.22) and (4.5) and taking into account the result (4.34), we obtain the "inversion equations"

$$A^{\pm}(x,t) = f^{\mp}(x)S^{\mp}(x+t)V + \int_x^\infty S^{\mp}(t+y)A^{\mp}(x,y)dy,\quad t>x, \quad (4.36)$$

where

$$S^{\pm}(x) = -\frac{1}{2\pi}\operatorname*{l.i.m}_{\rho\to\infty}\int_{-\rho}^{\rho}\tilde R^{\pm}(k)e^{ikx}dk + \sum_{m^\pm=1}^{M^\pm}\tilde C_{m^\pm}e^{i(k_{m^\pm})x}, \quad (4.37)$$

and l.i.m. stands for "limit in mean".

To the system of coupled Fredholm integral equations (4.36) and (4.37) we add a coupling condition. We start from the last component of $(4.7)_+$ and $(4.7)_-$. We find

$$\begin{aligned}u_{n-1}(x) &= [f^+(x)]^{-1}\left[-2\frac{d}{dx}A^+_{n-1}(x,x) - iu_n(x)A^+_{n-1}(x,x)\right]\\ &= [f^-(x)]^{-1}\left[-2\frac{d}{dx}A^-_{n-1}(x,x) + iu_n(x)A^-_{n-1}(x,x)\right];\end{aligned} \quad (4.38)$$

recalling that $u_n(x) = \pm 2i\dfrac{f^{\pm\prime}(x)}{f^{\pm}(x)}$, we obtain

$$f^-(x)\frac{d}{dx}A^+_{n-1}(x,x) + f^{-\prime}(x)A^+_{n-1}(x,x) = f^+(x)\frac{d}{dx}A^-_{n-1}(x) + f^{+\prime}(x)A^+_{n-1}(x,x) \quad (4.39)$$

and using the condition (4.8), we have the coupling condition

$$f^-(x)A^+_{n-1}(x,x) = f^+(x)A^-_{n-1}(x,x),\quad n>1, x\in\mathbb{R}. \quad (4.40)$$

D. Construction of potentials from scattering data $\mathscr{S}$

In the ISP for (2.17), the scattering data $\tilde{\mathscr{S}}$ [formula (3.40)] is given and we seek the matrix potentials $U(x)$ and $Q(x)$ written with $n+1$ scalar potentials u_p $(p=0,\dots,n)$ as in the formula (2.20), $U(x)$ and $Q(x)$ admitting $\tilde{\mathscr{S}}$ as scattering data. To this end, we construct $S^{\pm}(x)$ from $\tilde{\mathscr{S}}$ through the formula (4.37). If we suppose that (4.36) has a unique solution $(A^+(x,t),A^-(x,t))$ for given $f^+(x)$ and $f^-(x)$, we seek to make the dependence of $A^{\pm}(x,t)$ on $f^{\pm}(x)$ explicit. Let $C^{\pm}(x,t)$ be the solution of Eq. (4.36) corresponding to $f^{\pm}(x)=1$ and $D^{\pm}(x,t)$ be the one corresponding to $f^{\pm}(x)=1$ and $D^{\pm}(x,t)$ be the one corresponding to $f^{\pm}(x)=\mp i$. Let $\hat C^{\pm}(x,t)$ and $\hat D^{\pm}(x,t)$ be the functions defined for $(t\geqslant x, x\in\mathbb{R})$ by

$$\hat C^{\pm}(x,t) = \frac{C^{\pm}(x,t)\mp iD^{\pm}(x,t)}{2},\qquad \hat D^{\pm}(x,t) = \frac{C^{\pm}(x,t)\pm iD^{\pm}(x,t)}{2}. \quad (4.41)$$

It is easy to find the relation

$$A^{\pm}(x,t) = f^{\pm}(x)\hat C^{\pm}(x,t) + f^{\mp}(x)\hat D^{\pm}(x,t),\quad t\geqslant x, x\in\mathbb{R}, \quad (4.42)$$

which we can also write

$$A_p^{\pm}(x,t) = f^{\pm}(x)\hat C_p^{\pm}(x,t) + f^{\mp}(x)\hat D_p^{\pm}(x,t),\quad p=0,\dots,n-1 \quad (4.43)$$

where $\hat C_p^{\pm}(x,t)$, respectively, $\hat D_p^{\pm}(x,t)$, is the p^{th} component of $\hat C^{\pm}(x,t)$, respectively $\hat D^{\pm}(x,t)$. Using (4.43) in the case $p=n-1$ in (4.40), we have the equation

$$y\hat D^-_{n-1}(x,x) - (1/y)D^+_{n-1}(x,x) = \hat C^+_{n-1}(x,x) - \hat C^-_{n-1}(x,x),\quad y=[f^+(x)]^2, \quad (4.44)$$

whose solution gives y [note that for $n=1$, instead of (4.44), we obtain a Riccati type equation—see Ref. 1]. Hence $f^+(x)$, $f^-(x)$, $A^+(x,t)$ and $A^-(x,t)$, U, and Q in the form (2.20) are then obtained (the proof has been given in Sec. 4.1).

ACKNOWLEDGMENTS

We would like to acknowledge many fruitful discussions with Prof. F. Guil Guerrero and Prof. L. Martinez Alonso. One of us (M. J.) also thanks Prof. A. F. Rañada for his hospitality at the University Complutense of Madrid, where part of this work was done.

APPENDIX

We want to estimate the behavior of $\tilde F^{\pm}(k,x)$ for $|k|\to\infty$. We start from Eq. (3.12) which we write

$$\tilde F^{\pm}(k,x) = e^{ikx}V + \int_x^\infty\frac{\sin k(y-x)}{k}U(y)\tilde F(k,y)dy \pm\int_x^\infty \sin k(y-x)Q(y)\tilde F^{\pm}(k,y)dy, \quad (A1)$$

for $b=\operatorname{Im}k\geqslant 0$, $x\in\mathbb{R}$.

We consider the Neumann series of $\tilde{F}^{\pm}(k,x)$, more especially the Neumann series of each component $\tilde{F}_s^{\pm}(k,x)$ of $\tilde{F}^{\pm}(k,x)$, which we define

$$\tilde{F}_s^{\pm}(k,x) = \sum_{p=0}^{\infty} \tilde{F}_s^{\pm}(k,x)_p, \quad s = 0,\dots,n-1 \tag{A2}$$

where

$$\tilde{F}_0^{\pm}(k,x)_0 = e^{ikx}, \quad \tilde{F}_s^{\pm}(k,x)_0 = 0, \quad s = 1,\dots,n-1, \tag{A3}$$

$$\tilde{F}_s^{\pm}(k,x)_{p+1} = \int_x^{\infty} \frac{\sin k(y-x)}{k} U(y)\tilde{F}_s^{\pm}(k,y)_p dy \pm \int_x^{\infty} \sin k(y-x) Q(y)\tilde{F}_s^{\pm}(k,y)_p dy. \tag{A4}$$

Throughout the proof, we use the following results which are easily established:
if $u(x)$ satisfies the condition D_1, so

$$\int_x^{\infty} \frac{\sin k(y-x)}{k} u(y)e^{iky}dy = \frac{e^{ikx}}{k}\alpha(x);$$

$$|\alpha(x)| \leqslant \int_x^{\infty} |u(y)|\,dy, \tag{A5}$$

$$\int_x^{\infty} \sin k(y-x)u(y)e^{iky}dy = e^{ikx}\left[\frac{i}{2}\int_x^{\infty} u(y)dy\right] + \frac{e^{ikx}}{k}\left[\frac{u(x)}{4} + \beta(x)\right], \tag{A6}$$

where $|\beta(x)| \leqslant \int_x^{\infty} |u'(y)|\,dy$.

We give explicitly the computations for $\tilde{F}_s^{\pm}(k,x)_1$ and $\tilde{F}_s^{\pm}(k,x)_2$, for all, s, $s = 0,\dots,n-1$. We first start from the relation (A4) for $p = 0$ in which we substitute $\tilde{F}_s^{\pm}(k,y)$ by the formula (A3) and U and Q by their representation (2.20):

$$\tilde{F}_0^{\pm}(k,x)_1 = \int_x^{\infty} \frac{\sin k(y-x)}{k} u_0(y)e^{iky}dy \pm \int_x^{\infty} \sin k(y-x)u_n(y)e^{iky}dy, \tag{A7}$$

$$\tilde{F}_s^{\pm}(k,x)_1 = \int_x^{\infty} \frac{\sin k(y-x)}{k} u_s(y)e^{iky}dy, \quad s = 1,\dots,n-1. \tag{A8}$$

Applying the results (A5) and (A6), we have

$$\tilde{F}_0^{\pm}(k,x)_1 = e^{ikx}\left[\pm\frac{i}{2}\int_x^{\infty} u_n(y)dy\right] + \frac{e^{ikx}}{k}\left[\pm\frac{u_n(x)}{4} + \hat{v}_{01}(x) + v_{01}(x)\right], \tag{A9}$$

$$\tilde{F}_s^{\pm}(k,x)_1 = \frac{e^{ikx}}{k} v_{s1}(x) \quad s = 1,\dots,n-1 \tag{A10}$$

where

$$|\hat{v}_{01}(x)| \leqslant \int_x^{\infty} |u'_n(t)|\,dt, \tag{A11}$$

$$|v_{s1}(x)| \leqslant \int_x^{\infty} \sum_{p=0}^{n} |u_p(t)|\,dt = \gamma(x). \tag{A12}$$

Let us now consider the relation (A4) for $p = 1$. Thanks to (2.20), we obtain, for $s = 1,\dots,n-1$,

$$\tilde{F}_0^{\pm}(k,x)_2 = \int_x^{\infty} \frac{\sin k(y-x)}{k} u_0(y)\tilde{F}_0^{\pm}(k,y)_1 dy \pm \int_x^{\infty} \sin k(y-x) \sum_{p=1}^{n} u_p(y)\tilde{F}_{n-p}^{\pm}(k,y)_1 dy, \tag{A13}$$

$$\tilde{F}_s^{\pm}(k,x)_2 = \int_x^{\infty} \frac{\sin k(y-x)}{k} \sum_{p=0}^{s} u_p(y)\tilde{F}_{s-p}^{\pm}(k,y)_1 dy \pm \int_x^{\infty} \sin k(y-x) \sum_{p=1}^{n-s} u_{s+p}(y)\tilde{F}_{n-p}^{\pm}(k,y)_1 dy. \tag{A14}$$

We remark that, in every component $\tilde{F}_s^{\pm}(k,x)_2$, we find all the values of $u_i(y)$ and $\tilde{F}_i^{\pm}(k,y)_1$ once and only once. But, we have $\tilde{F}_0^{\pm}(k,y)_1$ in the second term only in the expression (A13) which we rewrite thus:

$$\tilde{F}_0^{\pm}(k,x)_2 = \int_x^{\infty} \frac{\sin k(y-x)}{k} u_0(y)\tilde{F}_0^{\pm}(k,y)_1 dy \pm \int_x^{\infty} \sin k(y-x)u_n(y)\tilde{F}_0^{\pm}(k,y)_1 dy \pm \int_x^{\infty} \sin k(y-x) \sum_{p=1}^{n-1} u_p(y)\tilde{F}_{n-p}^{\pm}(k,y)_1 dy. \tag{A15}$$

Thanks to (A5), (A6) and (A9), (A10), we can write

$$\int_x^{\infty} \frac{\sin k(y-x)}{k} u_0(y)\tilde{F}_0^{\pm}(k,y)_1 dy = \frac{e^{ikx}}{k} w_1(x) + O\left(\frac{1}{k^2}\right), \tag{A16}$$

$$\text{where } |w_1(x)| \leqslant \int_x^{\infty} |u_0(y)|\,dy \int_y^{\infty} |u_n(t)|\,dt; \tag{A17}$$

$$\pm\int_x^{\infty} \sin k(y-x)u_n(y)\tilde{F}_0^{\pm}(k,y)_1 dy = \pm\int_x^{\infty} \sin k(y-x)u_n(y)\left[\pm\frac{i}{2}\int_y^{\infty} u_n(t)dt\right]e^{iky}dy \pm \int_x^{\infty} \frac{\sin k(y-x)}{k}u_n(y)\left[\pm\frac{u_n(y)}{4} + \hat{v}_{01}(y) + v_{01}(y)\right]e^{iky}dy, \tag{A18}$$

where

$$\pm\int_x^{\infty} \sin k(y-x)u_n(y)\left[\pm\frac{i}{2}\int_y^{\infty} u_n(t)dt\right]e^{iky}dy = e^{ikx}\left[\pm\frac{i}{2}\int_x^{\infty} u_n(y)dy\right]^2\Big/2! + \frac{e^{ikx}}{k}\left[\pm\frac{u_n(x)}{4}\left(\pm\frac{i}{2}\int_x^{\infty} u_n(y)dy\right) + w_2(x)\right], \tag{A19}$$

$$\pm\int_x^{\infty} \frac{\sin k(y-x)}{k} u_n(y)\left[\pm\frac{u_n(y)}{4}\right]e^{iky}dy = \frac{e^{ikx}}{k} w_3(x), \tag{A20}$$

$$\pm\int_x^{\infty} \frac{\sin k(y-x)}{k} u_n(y)\hat{v}_{01}(y)e^{iky}dy = \frac{e^{ikx}}{k} w_4(x); \tag{A21}$$

$$\text{where } |w_i(x)| \leqslant \tfrac{1}{4}\int_x^{\infty} |u'_n(y)|\,dy \int_x^{\infty} |u_n(y)|\,dy, \quad i = 2,3,4; \tag{A22}$$

$$\pm \int_x^\infty \frac{\sin k(y-x)}{k} u_n(y) v_{01}(y) e^{iky} dy = \frac{e^{ikx}}{k} w_5(x), \quad \text{(A23)}$$

$$\text{where } |w_5(x)| \leqslant \int_x^\infty |u_n(y)| \gamma(y) dy; \quad \text{(A24)}$$

$$\pm \int_x^\infty \sin k(y-x) \sum_{p=1}^{n-1} u_p(y) \tilde{F}^{\pm}_{n-p}(k,y)_1 dy$$

$$= \pm \int_x^\infty \sin k(y-x) \sum_{p=1}^{n-1} u_p(y) v_{(n-p)1}(y) \frac{e^{iky}}{k} dy$$

$$= \frac{e^{ikx}}{k} w_6(x), \quad \text{(A25)}$$

$$\text{where } |w_6(x)| \leqslant \int_x^\infty \left(\sum_{p=1}^{n-1} |u_p(y)| \right) \gamma(y) dy. \quad \text{(A26)}$$

Adding (A17), (A24), and (A26), we obtain

$$|v_{02}(x)| = |w_1(x) + w_5(x) + w_6(x)| \leqslant \gamma^2(x)/2!. \quad \text{(A27)}$$

We deduce from (A22) that

$$|\hat{v}_{02}(x)| = |w_2(x) + w_3(x) + w_4(x)| \leqslant \sigma(x)\gamma(x), \quad \text{(A28)}$$

$$\text{where } \sigma(x) = \int_x^\infty |u'_n(t)| dt. \quad \text{(A29)}$$

Collecting all these results, we can write that

$$\tilde{F}_0^{\pm}(k,x)_2 = e^{ikx} \left[\pm \frac{i}{2} \int_x^\infty u_n(y) dy \right]^2 \Big/ 2!$$

$$+ \frac{e^{ikx}}{k} \left[\pm \frac{u_n(x)}{4} \left(\pm \frac{i}{2} \int_x^\infty u_n(y) dy \right) + \hat{v}_{02}(x) + v_{02}(x) \right]$$

$$+ O\left(\frac{1}{k^2}\right).$$

It is easier to evaluate the behavior of $\tilde{F}_s^{\pm}(k,x)_2$, $s = 1,...,n-1$. Let us start from (A14) and consider each term of these relations:

$$\int_x^\infty \frac{\sin k(y-x)}{k} \sum_{p=0}^{s} u_p(y) \tilde{F}^{\pm}_{s-p}(k,y)_1 dy$$

$$= \int_x^\infty \frac{\sin k(y-x)}{k} u_s(y) \tilde{F}_0^{\pm}(k,y)_1 dy$$

$$+ O\left(\frac{1}{k^2}\right)$$

$$= \int_x^\infty \frac{\sin k(y-x)}{k} u_s(y) \left[\pm \frac{i}{2} \int_y^\infty u_n(t) dt \right] e^{iky} dy$$

$$+ O\left(\frac{1}{k^2}\right) \quad \text{(A30)}$$

$$\pm \frac{e^{ikx}}{k} w_7(x) + O\left(\frac{1}{k^2}\right), \quad \text{(A31)}$$

$$\text{where } |w_7(x)| \leqslant \int_x^\infty |u_s(y)| \gamma(y) dy \quad \text{(A32)}$$

$$\pm \int_x^\infty \sin k(y-x) \sum_{p=1}^{n-s} u_{s+p}(y) \tilde{F}^{\pm}_{n-p}(k,y)_1 dy$$

$$= \pm \int_x^\infty \frac{\sin k(y-x)}{k} \sum_{p=1}^{n-s} u_{s+p}(y) v_{s1}(y) e^{iky} dy$$

$$= \frac{e^{ikx}}{k} w_8(x), \quad \text{(A33)}$$

$$\text{where } |w_8(x)| \leqslant \int_x^\infty \left(\sum_{p=1}^{n-s} |u_{s+p}(y)| \right) \gamma(y) dy. \quad \text{(A34)}$$

If we add (A32) and (A33), we obtain

$$\tilde{F}_s^{\pm}(k,x)_2 = \frac{e^{ikx}}{k} v_{s2}(x) + O\left(\frac{1}{k^2}\right), \quad s = 1,...,n-1 \quad \text{(A35)}$$

$$\text{where } |v_{s2}(x)| \leqslant \gamma^2(x)/2!. \quad \text{(A36)}$$

One can prove similarly, by recurrence, that we have, for $p \geqslant 2$,

$$\tilde{F}_0^{\pm}(k,x)_p = e^{ikx} \left[\pm \frac{i}{2} \int_x^\infty u_n(t) dt \right]^p \Big/ p! + \frac{e^{ikx}}{k}$$

$$\times \left\{ \pm \frac{u_n(x)}{4} \left[\pm \frac{i}{2} \int_x^\infty u_n(t) dt \right]^{p-1} \Big/ (p-1)! \right.$$

$$\left. + \hat{v}_{0p}(x) + v_{0p}(x) \right\}$$

$$+ O\left(\frac{1}{k^2}\right), \quad \text{(A37)}$$

$$\tilde{F}_s^{\pm}(k,x)_p = \frac{e^{ikx}}{k} v_{sp}(x) + O\left(\frac{1}{k^2}\right), \quad s = 1,...,n-1 \quad \text{(A38)}$$

$$\text{where } |\hat{v}_{0p}(x)| \leqslant \sigma(x)\gamma(x) \frac{(\gamma(x))^{p-2}}{(p-2)!}, \quad \text{(A39)}$$

$$|v_{sp}(x)| \leqslant \frac{(\gamma(x))^p}{p!}, \quad s = 0,...,n-1 \quad \text{(A40)}$$

$\sigma(x)$ and $\gamma(x)$ being defined by (A29) and (A12), respectively. For that we are obliged, in particular, to use the inequality

$$\frac{p-2}{p-1} + \frac{1}{(p-1)2^p} + \frac{1}{2^p} < 1, \quad p \geqslant 2.$$

It is clear that, similarly, we can estimate the behavior of $\overleftarrow{\mathbf{F}}^{\pm}(k,x)$ for $|k| \to \infty$.

[1]M. Jaulent and C. Jean, Commun. Math. Phys. **28**, 177 (1972) (radial case $x \geqslant 0$); see also Ann. Inst. Henri Poincaré **25**, 105, 119 (1976) (case $x \in \mathbf{R}$).

[2]M. Jaulent, J. Math. Phys. **17**, 1351 (1976).

[3]M. Jaulent and I. Miodek, Lett. Math. Phys. **1**, 243 (1976).

[4]F. Guil Guerrero and L. Martinez Alonso, Lett. Nuovo Cimento **27**, 28 (1980).

[5]M. Jaulent and I. Miodek, Lett. Nuovo Cimento **20**, 655 (1977).

[6]V. E. Zakharov and A. B. Shabat, Sov. Phys.—JETP **34**, 62 (1972).

[7]M. J. Ablowitz, D. J. Kaup, A. C. Newell, and H. Segur, Stud. Appl. Math. **53**, 294 (1974).

[8]F. Calogero and A. Degasperis, Nuovo Cimento **32B**, 201 (1976).

[9]F. Guil Guerrero and L. Martinez Alonso, Lett. Nuovo Cimento **27**, 85 (1980).

[10]L. Martinez Alonso, Lett. Math. Phys. **4**, 215 (1980).

[11]L. Martinez Alonso, J. Math. Phys. **21**, 2342 (1980).

[12]M. Jaulent and C. Jean, Preprint Physique Mathematique Montpellier PM/80/12.

[13]M. Wadati and T. Kamijo, Prog. Theor. Phys. **52**, 397 (1974).

Generalized second-order Coulomb phase shift functions

S. Rosendorff
Department of Physics, Technion-Israel Institute of Technology, Haifa, Israel

(Received 29 May 1980; accepted for publication 19 December 1980)

Some specific properties and the evaluation of the generalized second-order Coulomb phase shift functions (two-dimensional integrals of four spherical cylinder functions) are discussed. The dependence on the three momenta $k_1,\bar{k},k_2$, corresponding to the final, intermediate, and initial states is illustrated.

PACS numbers: 03.80. + r, 34.40. + n

I. INTRODUCTION

The expression of the unitary scattering amplitude for particle–atom collisions derived recently[1] includes second-order phase shift functions. The theory is valid both for elastic and inelastic collisions. In the derivation of this amplitude the summation over the intermediate states is performed by introducing a finite, mean excitation energy of the atom. Consequently, both the amplitude and the phase shift functions depend on three momenta, $k_1,\bar{k},k_2$, corresponding to the final, intermediate, and initial state, respectively. We name these phase shifts the generalized phase shift functions. There are four contributions to the generalized phase shift functions: The two particle–nucleus and particle–electron double scattering processes, the particle–electron–nucleus, and particle–nucleus–electron processes. The last two processes are simply related to each other but difficult to deal with analytically. On the other hand, the phase shift functions of the two double processes are proportional to the generalized second-order Coulomb phase shift functions. The discussion of the latter and its evaluation is the subject of the present communication.

The generalized second-order Coulomb phase shift functions are defined by

$$\delta_l^{(2)}(k_1,\bar{k},k_2) = I_l^{(1)}(k_1,\bar{k},k_2) + I_l^{(2)}(k_1,\bar{k},k_2), \tag{1}$$

where

$$I_l^{(1)} = -\bar{k}^2\int_0^\infty r\,dr\,j_l(k_1r)n_l(\bar{k}r)\int_0^r r'dr'j_l(\bar{k}r')j_l(k_2r'),$$

$$I_l^{(2)} = -\bar{k}^2\int_0^\infty r\,dr\,j_l(k_1r)j_l(\bar{k}r)\int_r^\infty r'dr'n_l(\bar{k}r')j_l(k_2r'). \tag{1'}$$

Here j_l and n_l are spherical Bessel and Neumann functions, respectively. Note that $I_l^{(1)}(k_1,\bar{k},k_2) = I_l^{(2)}(k_2,\bar{k},k_1)$. In Sec. II the two integrals $I_l^{(1)}$ and $I_l^{(2)}$ are calculated in a straightforward way, for $l = 0$. However for higher angular momenta this method becomes extremely unmanageable. A different, very simple, but indirect method based on the partial wave expansion of the second-order amplitude is discussed in Sec. III. This method gives the phase shift functions of Eq. (1), but it does not give the two integrals of Eq. (1′) separately. The phase shifts are calculated for $k_1 = \bar{k} = k_2$, as well as for the general case when all three or two of the momenta are different from each other.

II. *S*-WAVE PHASE SHIFTS

We shall calculate the s-wave phase shifts directly. We distinguish between six different possibilities.

(i) $k_1 = \bar{k} = k_2 \equiv k$. In this case the two integrals of Eq. (1′) are equal. We have for $k \neq 0$,

$$I_0 = \frac{1}{2k^2}\int_0^\infty \frac{\sin 2kx}{x}dx\int_0^x \frac{\sin^2 kx'}{x'}dx'. \tag{2}$$

Introducing the new variable k' by $kx' = k'x$, this becomes

$$I_0 = \frac{1}{4k^2}\int_0^\infty \frac{dx}{x}\int_0^k \frac{dk'}{k'} \times[\sin 2kx - \tfrac{1}{2}\sin 2(k+k')x - \tfrac{1}{2}\sin 2(k-k')x]. \tag{3}$$

Next, introducing an upper, finite limit T, it is permissible to reverse the order of integration

$$I_0 = \frac{1}{4k^2}\lim_{T\to\infty}\int_0^k \frac{dk'}{k'}\int_0^T \frac{dx}{x} \times[\sin 2kx - \tfrac{1}{2}\sin 2(k+k')x - \tfrac{1}{2}\sin 2(k-k')x]. \tag{4}$$

Integration by parts of the last two integrals yields

$$I_0 = \frac{1}{4k^2}\lim_{T\to\infty}\left(\frac{\pi}{4}\ln k + \int_0^k \ln k'dk'\frac{\sin 2(k+k')T}{2(k+k')} - \int_0^k \ln k'dk'\frac{\sin 2(k-k')T}{2(k-k')}\right) \tag{5}$$

where we have made use of the step function

$$\eta(\alpha) = \frac{2}{\pi}\int_0^\infty \frac{\sin\alpha x}{x}dx$$

equal[2] to

$$= \begin{cases} 1, & \alpha > 0 \\ 0, & \alpha = 0. \\ -1, & \alpha < 0 \end{cases} \tag{6}$$

Now, let us take advantage of Dirichlet's limit formula[3]

$$\frac{\pi}{2}f(k) = \lim_{T\to\infty}\int_0^a f(x+k)\frac{\sin xT}{x}dx, \tag{7}$$

which holds for arbitrary positive values of a; the function f is supposed to be sectionally smooth. The same result is obtained if the integral is taken from $-a$ to 0. Applying Eq. (7) to Eq. (5) shows immediately that the second term is zero, and the third term is equal to $\frac{1}{4}\pi\ln k$. It follows that

$$I_0 \equiv 0 \tag{8}$$

i.e., the s-wave phase shift function is identically zero for every $k \neq 0$.

We have also proved by the same kind of analysis that the p-wave phase shift function is zero.

(ii) $k_1 < \bar{k} < k_2$. The first integral of Eq. (1) becomes, after putting $kx' = k'x$,

$$I_0^{(1)}(k_1,\bar{k},k_2) = \frac{1}{k_1 k_2}\int_0^\infty \frac{\sin k_1 x \cos \bar{k} x}{x} dx \times \int_0^{\bar{k}} \frac{\sin k' x \sin\left(\frac{k_2}{\bar{k}} k' x\right)}{k'} dk', \quad (9)$$

which becomes, as is easily verified,

$$I_0^{(1)} = \frac{1}{8k_1k_2} \lim_{T\to\infty} \lim_{\epsilon\to 0} \int_\epsilon^{\bar{k}} \frac{dk'}{k'} \int_0^T \frac{dx}{x} \times \sum_{i=1}^{4} [\sin(f_i(k_1,\bar{k},k_2;k')x) - \sin(f_i(k_1,\bar{k},-k_2;k')x)], \quad (10)$$

where

$$f_i(k_1,\bar{k},k_2;k') = k_1 + (-1)^{i(i+1)/2}\bar{k} + (-1)^i(1 - k_2/\bar{k})k'. \quad (10')$$

Integration by parts of each of the above eight double integrals gives

$$\int_\epsilon^{\bar{k}} \frac{dk'}{k'} \int_0^T \frac{dx}{x} \sin(f_i(k')x) = \ln\bar{k} \int_0^T \frac{\sin f_i(\bar{k})x}{x} dx - \ln\epsilon \int_0^T \frac{\sin f_i(0)x}{x} dx - \int_0^{\bar{k}} \ln k' dk' f_i'(k') \frac{\sin f_i(k')T}{f_i(k')}, \quad (11)$$

where we have put $f_i(k_1,\bar{k},k_2;k') \equiv f_i(k')$ for short. We then find that among the first four integrals of Eq. (10), the second and the third cancel each other, and among the last four integrals, the first and the fourth cancel each other. As to the remaining four integrals, they are easily worked out, making use of the definition of η and the limit formula, Eqs. (6) and (7), respectively. We then find for $I_0^{(1)}$ from Eq. (10),

$$I_0^{(1)}(k_1,\bar{k},k_2) = \frac{\pi}{16k_1k_2}\left\{[1 - \eta(k_1 + k_2 - 2\bar{k})]\ln\frac{\bar{k} - k_1}{k_2 - \bar{k}} - [1 - \eta(k_2 - 2\bar{k} - k_1)]\ln\frac{k_1 + \bar{k}}{k_2 - \bar{k}}\right\}. \quad (12)$$

The second integral $I_0^{(2)}(k_1,\bar{k},k_2)$ of Eq. (1) is related to $I_0^{(1)}$ by

$$I_0^{(2)}(k_1,\bar{k},k_2) = I_0^{(1)}(k_2,\bar{k},k_1). \quad (13)$$

Going through the same analysis as above, we find

$$I_0^{(1)}(k_2,\bar{k},k_1) = -I_0^{(1)}(k_1,\bar{k},k_2). \quad (14)$$

In other words, we have proved that the generalized second-order s-wave Coulomb phase shift functions, $\delta_0^{(2)}(k_1,\bar{k},k_2)$ are equal to zero for all values of $k_1,\bar{k}$ and k_2 which satisfy the inequality $k_1 < \bar{k} < k_2$.

(iii) $(k_1 \leqslant k_2 < \bar{k})$. We find

$$\delta_0^{(2)}(k_1,\bar{k},k_2) = \frac{\pi}{8k_1k_2}\left\{[\eta(k_2 - k_1) - 1]\ln\left|\frac{\bar{k} + k_2}{\bar{k} - k_2}\right| + [\eta(k_1 - k_2) - 1]\ln\left|\frac{\bar{k} + k_1}{\bar{k} - k_1}\right|\right\}. \quad (15)$$

For the particular case $k_1 = k_2 \equiv k$, this becomes

$$\delta_0^{(2)}(k_1,\bar{k},k_2) = -\frac{\pi}{4k^2}\ln\frac{\bar{k} + k}{\bar{k} - k}. \quad (15')$$

It thus has a logarithmic divergence when $\bar{k}\to k$. However, for $\bar{k} = k$, we have $\delta_0^{(2)} = 0$ according to (i). For $k_1 < k_2$ and $\bar{k}\to k_2 + 0$, we get according to Eq. (15),

$$\delta_0^{(2)}(k_1,k_2 + 0,k_2) = -\frac{\pi}{4k_1k_2}\ln\frac{k_2 + k_1}{k_2 - k_1}. \quad (15'')$$

Approaching k_2 from the left, $\bar{k}\to k_2 - 0$, we have $\delta_0^{(2)}(k_1,k_2 - 0,k_2) = 0$ according to (ii). The function therefore has a discontinuity at the point $\bar{k} = k_2$.

(iv) $\bar{k} < k_1 \leqslant k_2$. In this case the phase shift function is obtained from Eq. (15) by multiplying it by -1 and interchanging k_1 with k_2 in the logarithmic functions.

(v) $k_1 < k_2 = \bar{k}$. We find

$$\delta_0^{(2)}(k_1,\bar{k},k_2) = -\frac{\pi}{8k_1k_2}\ln\frac{k_2 + k_1}{k_2 - k_1}. \quad (16)$$

Hence the value of $\delta_0^{(2)}$ at $\bar{k} = k_2$ is half the value one obtains by approaching k_2 from the right.

(vi) $\bar{k} = k_1 < k_2$. The phase shift function is equal to Eq. (16) multiplied by -1.

III. PHASE SHIFTS FOR ANY ANGULAR MOMENTUM

The calculation of the generalized second-order Coulomb phase shift functions for any angular momentum l by the method of Sec. II is obviously not feasible. Even the calculation of the p-wave phase shift by this method is a considerable task. On the other hand, to obtain the generalized phase shifts through the second-order amplitude is straightforward. The only disadvantage is that this method yields only the phase shift functions $\delta_0^{(2)}(k_1,\bar{k},k_2)$ of Eq. (1), but not the two integrals $I_l^{(1)}$ and $I_l^{(2)}$ separately.

Let us start with the second-order scattering amplitude

$$f^{(2)}(\mathbf{k}_1,\bar{k},\mathbf{k}_2) = \frac{1}{2\pi^2}\int f_1^{(1)}(\mathbf{k}_1 - \mathbf{q})\frac{d\mathbf{q}}{q^2 - \bar{k}^2 - i\epsilon} f_2^{(1)}(\mathbf{q} - \mathbf{k}_2). \quad (17)$$

Generally, the magnitude of the three momenta $k_1,\bar{k},k_2$ are not equal to each other. We assume that at the two vertices, act two different, spherical symmetric potentials functions $V_1(r)$ and $V_2(r)$. Here $f_1^{(1)}$ and $f_2^{(1)}$ are the corresponding first-order scattering amplitudes. They are given by $i = 1,2$,

$$f_i^{(1)}(\mathbf{k}_1 - \mathbf{k}_2) = -\frac{1}{4\pi}\int e^{-i\mathbf{k}_1\cdot\mathbf{r}} V_i(\mathbf{r}) e^{i\mathbf{k}_2\cdot\mathbf{r}} d\mathbf{r}. \quad (18)$$

The partial wave expansion of these functions is easily derived. We have

$$f_i^{(1)}(\mathbf{k}_1 - \mathbf{k}_2) = \frac{1}{\sqrt{k_1k_2}}\sum_l (2l + 1)P_l(\cos\vartheta_{k_1,k_2})\delta_{i,l}^{(1)}(k_1,k_2), \quad (19)$$

where the first-order generalized phase shift functions are given by

$$\delta_{i,l}^{(1)}(k_1,k_2) = -\sqrt{k_1k_2}\int_0^\infty r^2 dr j_l(k_1r) V_i(r) j_l(k_2r). \quad (20)$$

Putting these expressions of the first-order amplitudes back into Eq. (17), and making use of the integral representation of the radial Green's function

$$g_l(r,r')=\frac{2}{\pi}\int_0^\infty q^2 j_l(qr)\frac{dq}{q^2-\bar{k}^2-i\epsilon}j_l(qr'),$$

which is equal to $i\bar{k}j_l(\bar{k}r_<)h_l^{(1)}(\bar{k}r_>)$, we obtain, for the real part of the second-order amplitude, the expression

$$\mathrm{Re}f^{(2)}(\mathbf{k}_1,\bar{k},\mathbf{k}_2)=\frac{1}{\bar{k}}\sum_l(2l+1)P_l(\cos\vartheta_{k_1k_2})\delta_l^{(2)}(k_1,\bar{k},k_2), \tag{21}$$

where the second-order generalized phase shift functions $\delta_l^{(2)}$ are given by

$$\delta_l^{(2)}(k_1,\bar{k},k_2)=-\bar{k}^2\Big[\int_0^\infty r^2dr j_l(k_1r)V_1(r)n_l(\bar{k}r)\int_0^r r'^2dr' j_l(\bar{k}r')V_2(r')j_l(k_2r') + \int_0^\infty r^2dr j_l(k_1r)V_1(r)j_l(\bar{k}r)\int_r^\infty r'^2dr' n_l(\bar{k}r')V_2(r')j_l(k_2r')\Big]. \tag{22}$$

We shall now specify the form of the functions V_i by choosing them to be the Yukawa potentials $e^{-\lambda_i r}/r$. The corresponding first-order amplitudes [Eq. (18)] are then given by

$$f_i^{(1)}(\mathbf{k}_1,\mathbf{k}_2)=-1/(\lambda_i^2+Q^2),$$

where $\mathbf{Q}=\mathbf{k}_1-\mathbf{k}_2$ is the momentum transfer. Substitution of this expression into Eq. (17) gives

$$f^{(2)}(\mathbf{k}_1,\bar{k},\mathbf{k}_2)=\frac{1}{2\pi^2}\int\frac{d\mathbf{q}}{[(\mathbf{k}_1-\mathbf{q})^2+\lambda_1^2](q^2-\bar{k}^2-i\epsilon)[(\mathbf{q}-\mathbf{k}_2)^2+\lambda_2^2]}. \tag{23}$$

Now making use of the well-known Feynman technique,[4] Eq. (23) is easily transformed into an one-dimensional integral, the real part of which is given by

$$\mathrm{Re}f^{(2)}(\mathbf{k}_1,\bar{k},\mathbf{k}_2) = -\tfrac{1}{4}\int_{-1}^1\frac{\bar{k}^2-P^2-\Lambda^2}{\Lambda[(\bar{k}^2-P^2-\Lambda^2)^2+4\bar{k}^2\Lambda^2]}dz, \tag{24}$$

where

$$\Lambda^2=\tfrac{1}{4}Q^2(1-z^2)+\tfrac{1}{2}(\lambda_1^2+\lambda_2^2)+\tfrac{1}{2}(\lambda_2^2-\lambda_1^2)z, \tag{24'}$$

and

$$\Lambda^2+P^2=\tfrac{1}{2}(k_1^2+k_2^2+\lambda_1^2+\lambda_2^2)+\tfrac{1}{2}(k_2^2-k_1^2+\lambda_2^2-\lambda_1^2)z. \tag{24''}$$

The expression [Eq. (24)] can be converted by a simple transformation into the integral

$$\mathrm{Re}f^{(2)}(\mathbf{k}_1,\bar{k},\mathbf{k}_2)=-\frac{1}{Q}\int_{t_1}^{t_2}\frac{(\alpha+\beta t^2)dt}{(\alpha+\beta t^2)^2+\left(\frac{2\bar{k}}{Q}\right)^2[(\lambda_2^2-\lambda_1^2)^2+Q^4+2Q^2(\lambda_1^2+\lambda_2^2)]t^2}. \tag{25}$$

Here the coefficients α, β, t_1, and t_2 are known algebraic functions of λ_1 and λ_2, the three momenta $k_1,\bar{k},k_2$, and the momentum transfer Q.

As in Sec. II, here also we shall discuss the six different possibilities (i)–(vi) one by one.

(i) $k_1=\bar{k}=k_2\equiv k$. Here we put $\lambda_1\neq 0$ and $\lambda_2=0$. We then find

$$\alpha=-(\lambda_1^2/Q^2)(\lambda_1^2+Q^2),\quad \beta=0.$$

and

$$t_1=\lambda_1/Q,\quad t_2=\infty. \tag{26}$$

The real part of the second-order amplitude thus becomes, according to Eq. (25),

$$\mathrm{Re}f^{(2)}(\mathbf{k}_1,\bar{k},\mathbf{k}_2)=[2k(Q^2+\lambda_1^2)]^{-1}\tan^{-1}(\lambda_1/2k). \tag{27}$$

Applying now Eq. (21) in order to calculate the phase shifts, we get in conjunction with Eq. (22),

$$\int_0^\infty r\,dr j_l^2(kr)e^{-\lambda_1 r}\int_r^\infty r'dr' n_l(kr')j_l(kr') + \int_0^\infty r\,dr j_l(kr)e^{-\lambda_1 r}n_l(kr)\int_0^r r'dr' j_l^2(kr') = -\frac{1}{4k^2}\tan^{-1}\frac{\lambda_1}{2k}\int_{-1}^1\frac{P_l(\mu)d\mu}{2k^2(1-\mu)+\lambda_1^2}. \tag{28}$$

The integral on the rhs is known to be equal to $k^{-2}Q_l(1+\lambda_1^2/2k^2)$, where Q_l is the Legendre function of the second kind. Now in the limit $\lambda_1\to 0$,

$$Q_l(1+\lambda_1^2/2k^2)\to -\ln(\lambda_1/2k);$$

thus the rhs of Eq. (28) vanishes when $\lambda_1\to 0$. Now, the two integrals on the lhs of Eq. (28) are both continuous functions of λ_1. Therefore, the processes of integration and of taking the limit $\lambda_1\to 0$ can be interchanged. To show that the integrals are continuous in λ_1 we have to prove that they converge uniformly (in λ_1) in an interval which includes $\lambda_1=0$. To prove this for the first integral of Eq. (28), consider the expression of the remainder for $Ak\gg 1$,

$$R_A=\int_A^\infty r\,dr j_l^2(kr)e^{-\lambda_1 r}\int_r^\infty r'dr' n_l(kr')j_l(kr'). \tag{29}$$

Making use of the asymptotic expansions for j_l and n_l, this becomes

$$k^4R_A=\tfrac{1}{2}(-1)^l\int_A^\infty\frac{dr}{r}\sin^2\!\left(kr-\frac{\pi}{2}l\right)e^{-\lambda_1 r}\mathrm{si}(2kr), \tag{29'}$$

where $\mathrm{si}(x)=-\int_x^\infty \sin t\,dt/t$ is the sine integral. The dominant term of si for $x \gg 1$ is $-\cos x/x$; hence,

$$k^4|R_A| = \frac{1}{4k}\left|\int_A^\infty \frac{dr}{r^2}\sin^2\!\left(kr - \frac{\pi}{2}l\right)e^{-\lambda_1 r}\cos 2kr\right| < \frac{1}{4kA}. \tag{29''}$$

From the fact that the rhs does not depend on λ_1 follows the uniform convergence of the integral. The uniform convergence of the second integral in Eq. (28) can be proved along similar lines, as

$$\int_0^\infty r\,dr\,j_l n_l e^{-\lambda_1 r}\int_0^r r'dr' j_l^2 = \int_0^\infty r dr j_l^2 \int_r^\infty r'dr' j_l n_l e^{-\lambda_1 r'}. \tag{30}$$

To summarize, we have proved that for all $k \neq 0$, and *all* angular momenta $l = 0,1,2,\cdots$,

$$\int_0^\infty r\,dr\,j_l(kr)n_l(kr)\int_0^r r'dr' j_l^2(kr') \equiv 0, \tag{31}$$

i.e., all the second-order Coulomb phase shift functions are identically zero.

For all the other possibilities we take $\lambda_1 = \lambda_2 = 0$, i.e., $V_1 = V_2 = 1/r$.

(ii) $k_1 < \bar{k} < k_2$. We find

$$\alpha = \bar{k}^2 - k_2^2, \quad \beta = \bar{k}^2 - k_1^2,$$
$$t_1 = 0, \quad t_2 = \infty. \tag{32}$$

Hence, the real part of the generalized second-order amplitude of Eq. (25) is determined by the expression

$$\mathrm{Re} f^{(2)}(\mathbf{k}_1,\bar{k},\mathbf{k}_2) = -\frac{1}{Q}\int_0^\infty \frac{[\bar{k}^2 - k_2^2 + (\bar{k}^2 - k_1^2)t^2]\,dt}{[\bar{k}^2 - k_2^2 + (\bar{k}^2 - k_1^2)t^2]^2 + 4\bar{k}^2Q^2t^2}. \tag{33}$$

This integral has been dealt with extensively in the literature.[5] We find that subject to the above condition on the k's, the above integral is identically zero. We therefore conclude by Eqs. (21) and (22) that the generalized second-order Coulomb phase shifts are identically zero for every angular momentum l, and all values of $k_1,\bar{k},k_2$, provided $k_1 < \bar{k} < k_2$. Thus, according to Eq. (1), we have

$$\delta_0^{(2)}(k_1,\bar{k},k_2) = I_l^{(1)}(k_1,\bar{k},k_2) + I_l^{(1)}(k_2,\bar{k},k_1) = 0. \tag{34}$$

(iii) $k_1 \leqslant k_2 < \bar{k}$. The coefficients α and β are the same as in (ii). According to Ref. 5 we get

$$\mathrm{Re} f^{(2)}(\mathbf{k}_1,\bar{k},\mathbf{k}_2) = -(\pi/2Q) \times [\bar{k}^2Q^2 + (\bar{k}^2 - k_1^2)(\bar{k}^2 - k_2^2)]^{-1/2}. \tag{35}$$

Hence the phase shift functions are determined by

$$\delta_0^{(2)}(k_1,\bar{k},k_2) = I_l^{(1)}(k_1,\bar{k},k_2) + I_l^{(1)}(k_2,\bar{k},k_1) = -\frac{\pi\bar{k}}{4}\int_{-1}^1 \frac{P_l(\mu)d\mu}{Q\,[\bar{k}^2Q^2 + (\bar{k}^2 - k_1^2)(\bar{k}^2 - k_2^2)]^{1/2}}, \tag{36}$$

where $Q^2 = (k_1 - k_2)^2 + 2k_1k_2(1-\mu)$. It does not seem to be possible to express the above integral in terms of known functions. However, for given l the evaluation is simple. In particular, for $l = 0$ calculation of the integral yields the result of Eq. (15). For $l \gg 1$, the above expression is generally rather cumbersome; yet, for the particular case when $k_1 = k_2 \equiv k$, it is easily estimated by making use of the eikonal approximation $P_l(\mu) \approx J_0(bQ)$, b being the impact parameter $(l+\frac{1}{2})/k$. Thus for $l \gg 1$ and $\Delta k/k \ll 1$ with $\Delta k = \bar{k} - k$,

$$\delta_l^{(2)}(k_1,\bar{k},k_2) \approx -\frac{\pi\bar{k}}{4k^2}\int_0^\infty \frac{J_0(bQ)}{[\bar{k}^2Q^2 + (\bar{k}^2 - k^2)^2]^{1/2}}dQ \approx -\frac{\pi}{4k^2}I_0(b\Delta k)K_0(b\Delta k), \tag{37}$$

where I_0 and K_0 are the modified Bessel and Hankel functions of zero order, respectively. Comparison with Eq. (15') shows that even for s-states the above result is very satisfactory.

(iv) $\bar{k} < k_1 \leqslant k_2$. As the coefficients α and β are the same as above, but negative, the phase shifts are given by the same expression as in (iii) multiplied by -1. Hence,

$$\delta_l^{(2)}(k_1,\bar{k},k_2) = \frac{\pi\bar{k}}{4}\int_{-1}^1 \frac{P_l(\mu)d\mu}{Q\,[\bar{k}^2Q^2 + (\bar{k}^2 - k_1^2)(\bar{k}^2 - k_2^2)]^{1/2}}. \tag{38}$$

Suppose now that $\bar{k}_1$ and $\bar{k}_2$ are the two roots of the quadratic equation $(\bar{k}^2 - k_1^2)(\bar{k}^2 - k_2^2) = \bar{k}^2k_0^2$, where k_0 is an arbitrary, but nonvanishing momentum; then the two expressions of Eqs. (36) and (38) are equal and of opposite sign. Hence,

$$\delta_l^{(2)}(k_1,\bar{k}_1,k_2) + \delta_l^{(2)}(k_1,\bar{k}_2,k_2) = 0. \tag{39}$$

or in terms of the I-integrals of Eq. (1),

$$I_l^{(1)}(k_1,\bar{k}_1,k_2) + I_l^{(1)}(k_2,\bar{k}_1,k_1) + I_l^{(1)}(k_1,\bar{k}_2,k_2) + I_l^{(1)}(k_2,\bar{k}_2,k_1) = 0. \tag{39'}$$

In other words, to every $\bar{k}_2 > k_2$ at which the phase shift is given by the expression of Eq. (36), corresponds a $\bar{k}_1 < k_1$ at which the phase shift is equal but of opposite sign. In particular, to the point $\bar{k}_2 \to k_2 + 0$ corresponds the point $\bar{k}_1 \to k_1 - 0$, and to $\bar{k}_2 \to \infty$ corresponds $\bar{k}_1 \to 0$.

(v) $k_1 < k_2 = \bar{k}$. In this case $\alpha = 0$ and $\beta = k_2^2 - k_1^2$. Thus, by Eq. (25) we get

$$\mathrm{Re} f^{(2)}(\mathbf{k}_1,\bar{k},\mathbf{k}_2) = -\pi/4\bar{k}Q^2,$$

and the phase shifts become

$$\delta_l^{(2)}(k_1,\bar{k},k_2) = -(\pi/8k_1k_2)Q_l(1 + (k_2 - k_1)^2/2k_1k_2). \tag{40}$$

Comparison with Eq. (36) shows that this is half the value one obtains when $\bar{k} \to k_2 + 0$.

(vi) $\bar{k} = k_1 < k_2$. The phase shift function is equal to minus the value of Eq. (40).

IV. SUMMARY

We have calculated the generalized Coulomb phase shift functions defined by Eq. (1). We have to distinguish essentially between two possibilities: (i) all the three momenta involved are equal to some k. Then for all values of k different from zero and *all* angular momenta l the phase shift functions, proportional to the two-dimensional integral $I_l(kkk)$ of Eq. (1'), are identically zero. (ii) The momenta k_1 and k_2 are different from each other, say $k_1 < k_2$. Then considering the phase shifts as functions of $\bar{k}$, we find that for $k_1 < \bar{k} < k_2$ they are equal to zero, again for all angular momenta l. Thus in terms of the integrals of Eq. (1') this is equivalent to the interesting relation $I^{(1)}(k_1,\bar{k},k_2) = -I^{(1)}(k_2,\bar{k},k_1)$. For $\bar{k} > k_2$ and $\bar{k} < k_1$, the phase shift functions are determined by Eqs. (36) and (38). It is relatively easy to evaluate these expressions for any given angular momentum l. For $l \gg 1$ and the particular case $k_1 = k_2 \equiv k$ with $|\bar{k} - k|/k \ll 1$, a good approximation is provided by the explicit expression Eq. (37). When $\bar{k}$ approaches k from the right or from the left the phase shift functions (for $k_1 = k_2$) diverge logarithmically like $\pm \ln|\bar{k} - k|$. The value of the phase shift for $\bar{k}$ close to k is essentially independent of l. When $k_1 < k_2$, the phase shifts are finite everywhere, however, they have discontinuities at the points $k_1 = \bar{k}$ and $k_2 = \bar{k}$.

ACKNOWLEDGMENT

This work was supported by a grant from the fund for the encouragement of research at the Technion.

[1]S. Rosendorff, Phys. Rev. A **23**, 683 (1981).
[2]The value of $\eta(x)$, including $x = 0$, is easily evaluated by making use of Dirichlet's discontinuous factor.
[3]See, e.g., R. Courant, *Differential and Integral Calculus* (Blackie, London, 1945), Vol. II, p. 321.
[4]R. Dalitz, Proc. R. Soc. London, Ser. A **206**, 509 (1951).
[5]A. Birman and S. Rosendorff, Nuovo Cimento B **63**, 89 (1969).

Pre-atlas manifolds and coordinate-free general relativity

S. Kloster and A. Das [a]
Department of Mathematics, Simon Fraser University, British Columbia, V5A 1S6, Canada

(Received 21 November 1980; accepted for publication 24 April 1981)

Einstein's equations can be expressed in the tetrad form so that coordinates do not appear explicitly. Tetrads, however, are usually defined on a manifold, which means that coordinates have been introduced. The notion of a manifold without coordinates (a pre-atlas manifold) is described here and it is shown that Einstein's equations can be expressed in this setting without introducing coordinates at any stage. Conditions on a pre-atlas manifold are given which ensure that a C^0-atlas can be generated. The motivation for this formulation is the desire to incorporate the philosophy of relativity, which asserts that the mathematical laws of nature are essentially independent of observers or coordinates. "The introduction of numbers as coordinates...is an act of violence."—H. Weyl

PACS numbers: 04.20. − q, 02.40. + m

I. INTRODUCTION

An n-dimensional differentiable manifold is usually defined[1] to be a pair $(M,\mathscr{A})$ where M is a topologized set (the manifold topology is one possibility) and $\mathscr{A}$ is an atlas providing the differential structure. The atlas consists of n-coordinate pairs (ϕ,U), where U is a subset of M and ϕ is a 1–1 map of U onto an open set in R^n. It is possible for different atlases to give the same differential structure, in which case they may be called equivalent. The manifold may be thought of as $(M,[\mathscr{A}])$, where $[\mathscr{A}]$ is an equivalence class of atlases. It should be noted that a given topological space can have inequivalent atlases.[2]

A semi-Riemannian manifold is a quadruple $(M,M_p,g,\mathscr{A})$, where for each $p\in M$, M_p is the tangent space at p, g_p is the metric defined on $M_p\times M_p$, and $\mathscr{A}$ is the atlas. A second metric $\hat{g}$ gives a second semi-Riemannian manifold $(M,M_p,\hat{g},\mathscr{A})$, which is defined to be isometric to the first one if g_p and $\hat{g}_p$ are related by the usual change of coordinate formulas. A given manifold $(M,M_p,\mathscr{A})$ can generally have many nonisometric metrics (e.g., g and $\hat{g}$ may be conformally related).

For the application of differentiable manifolds to the theory of general relativity, Einstein's equation have to be incorporated. In the usual treatment, these are written as a set of partial differential equations in coordinate patch of an atlas, thus presupposing the existence of an atlas. These are tensor equations and thus covariant under a *change* of coordinates. But it would conform more closely to the objective philosophy of relativity if these could be expressed without even *using* coordinates. In this paper we find such formulations of Einstein's equations by defining the notion of a pre-atlas differentiable manifold $(M,M_p,g,\mathscr{F})$.

We also consider the question of how to define the atlas $\mathscr{A}$ operationally if $(M,M_p,g,\mathscr{F})$ is known. Synge[3] has shown one way to do this using geodescis and clocks. The concept of a pre-atlas differentiable manifold is also useful here to give a rigorous foundation for these attempts. In this paper, we find conditions on such structures which guarantee that a C^0 atlas can be defined operationally. It is an open problem whether these conditions guarantee a C^∞ atlas. However, the *major* open problem that remains is how to integrate these equations in a coordinate-free manner.

[a] Both authors belong to the Theoretical Science Institute, S.F.U.

2. PRE-ATLAS MANIFOLDS AND EINSTEIN'S EQUATIONS

An atlas serves to determine which functions on the manifold are differentiable and which are not. The approach here will be to turn this around by starting with a family $\mathscr{F}$ of real valued functions, all of which are to be smooth, and deriving the atlas from them. (This might be called a basis for a smoothness structure in the terminology of Milnor and Stasheff[4]). Furthermore, these functions should all be physical quantities that can be determined by experiment. Some examples would be the eigenvalues of the energy–momentum tensor, radar coordinates,[5] etc. A collection of such fields, which are smooth functions of the proper time of the observers, could be used for the family $\mathscr{F}$.

Definition 1. A family $\mathscr{F}$ of real-valued continuous functions defined on open subsets of a topological space M is called an *algebra of functions* if the following properties hold:

(1) If A is an open subset of the domain of f, $f\in\mathscr{F}$, then $f|_A$ is in $\mathscr{F}$.

(2) If $A=UA_\alpha$ and f is defined on A with $f|_{A_\alpha}\in\mathscr{F}$, then f is in $\mathscr{F}$.

(3) If f and h belong to $\mathscr{F}$ and are both defined on an open subset $A\neq\phi$, then

(a) $af+bh\in\mathscr{F}$, for all $a,b\in R$,

(b) $f\cdot h\in\mathscr{F}$, and

(c) $f\div h\in\mathscr{F}$, if $h\neq 0$ on A.

These operations all preserve smooth functions. $\mathscr{F}$ cannot be closed under the extraction of roots because this does not preserve smoothness.

Definition 2. A *tangent vector* at p is a real valued function X_p on $\mathscr{F}_p=\{f\in\mathscr{F}: p\in\text{dom} f\}$ such that

$$X_p h(f_1,f_2)=h_{,1}X_p f_1+h_{,2}X_p f_2,$$

whenever $f_1,f_2\in\mathscr{F}_p$, $h(f_1(p),f_2(p))\in\mathscr{F}_p$, and h is a differentiable mapping from R^2 into R.

 0022-2488/82/020272-04$02.50

The following properties of tangent vectors can be deduced:

$$X_p(f+h) = X_p(f) + X_p(h), \quad X_p cf = cX_p f,$$
$$X_p(f\cdot h) = (X_p f)h(p) + f(p)X_p h.$$

Definition 3. The *tangent space* to M at p, denoted by M_p, is the set of all tangent vectors at p.

It can be considered as a vector space over the reals if we set $(X_p + Y_p)f = X_p f + Y_p f$ and $(bX_p)f = b(X_p f)$.

Definition 4. A *pre-atlas differentiable manifold* (PDM) is a triple $(M, M_p, \mathscr{F})$ with $\mathscr{F}$ and M_p as defined above.

Definition 5. A *vector field* X on a set $A \subseteq M$ is a mapping that assigns to each point p in A a vector X_p in M_p.

A vector field X is *smooth* on A if A is open and for each $f \in \mathscr{F}$ the function f_X defined by $f_X(p) = X_p f$ belongs to $\mathscr{F}_p$. In this case we see that Xf_X (which may be denoted f_{XX}), Xf_{XX}, etc., will all be in $\mathscr{F}$. In this way the functions of $\mathscr{F}$ are infinitely differentiable with respect to a smooth vector field.

Note that the existence of a smooth vector field X guarantees a certain form of consistency of $\mathscr{F}$. For example, suppose there is a function f in $\mathscr{F}_p$ with $f(p) = 0$ and $X_p f = c \neq 0$. Then the function $h \equiv f^{1/3}$ cannot belong to $\mathscr{F}_p$, for it if it did, then by Definition 2, when q is near p, but not equal to p,

$$X_q h = \tfrac{1}{3} f^{-2/3} X_q f.$$

As q approaches p, $X_q f \to c$ and $f^{2/3} \to 0$, so $X_q h$ cannot be defined continuously at p, and hence does not belong to $\mathscr{F}_p$. Thus h does not belong to $\mathscr{F}_p$.

In the following, we suppose that $(M, M_p, \mathscr{F})$ is a PDM and A is a subset of M.

Definition 6. A *metric field* g on A specifies a mapping $g_p: M_p \times M_p \to R$ for each $p \in A$ such that g_p is bilinear, symmetric, and nondegenerate. It is *smooth* if for any two smooth vector fields X, Y, the function taking p to $g_p(X_p, Y_p)$ belongs to $\mathscr{F}$. It will be assumed that g is smooth in the following.

Definition 7. A *covariant differentiation operator* on $(M, M_p, g, \mathscr{F})$ is an operator D that assigns to each pair of smooth vector fields X and Y with domain A, a smooth vector field $D_X Y$, with the same domain; and if Z is a smooth vector field on A and $f \in \mathscr{F}$, then D satisfies the following six axioms:

(1) $D_X(Y+Z) = D_X Y + D_X Z$,

(2) $D_{(X+Y)} Z = D_X Z + D_Y Z$,

(3) $D_{(fX)} Y = f D_X Y$,

(4) $D_X(fY) = (Xf)Y + f D_X Y$,

(5) $D_X Y - D_Y X = [X, Y]$,

(6) $Zg(X,Y) = g(D_Z X, Y) + g(X, D_Z Y)$.

Definition 8. The *curvature operator* R of a covariant differentiation operator D is defined by $R(X,Y)Z \equiv D_X D_Y Z - D_Y D_X Z - D_{[X,Y]} Z$, where X, Y, and Z are smooth vector fields.

Definition 9. The *Riemann–Christoffel curvature tensor* field is defined by $K(\omega, X, Y, Z) \equiv \omega(R(Y,Z)X)$, where ω is a dual vector field and X, Y, Z are smooth vector fields.

An open set A is said to be *framed* [1] if the space of smooth vector fields on A has a basis of smooth vector fields $X_{(i)}$, $i = 1,\ldots,n$. (This *implies* that $\dim M_p = n$ for all $p \in A$). If the tensors defined above are evaluated at a point p of a framed set, the values obtained will depend only on the values of the vectors fields at p. In the following, it is assumed that we are working on a framed set. Then we may write $D_{X_{(i)}} X_{(r)} = \Gamma^{(k)}{}_{(ri)} X_{(k)}$, where the Γ's are real valued functions on A. One can check that these all belong to $\mathscr{F}$. Note that the order of i and r gets reversed, and that (ri) *does not* denote symmetrization.

Axiom (5) implies that $[X_{(i)}, X_{(j)}]$ must be a smooth vector field, so we may write $[X_{(i)}, X_{(j)}] = \gamma_{(ij)}{}^{(k)} X_{(k)}$. These are related to the Γ's by the formula $\hat{\gamma}_{(ij)}{}^{(k)} = \Gamma^{(k)}{}_{(ji)} - \Gamma^{(k)}{}_{(ij)}$.

If g is a smooth metric, then the real functions $g_{(ij)} \equiv g(X_{(i)}, X_{(j)})$ belong to $\mathscr{F}$ and $\det[g_{(ij)} \neq 0]$. Thus one can define $g^{(ij)} = (g^{-1})_{ij}$, and one can prove from the above axioms that

$$\Gamma^{(k)}{}_{(ij)} = \tfrac{1}{2} g^{(kr)} [X_{(i)} g_{(rj)} + X_{(j)} g_{(ri)} - X_{(r)} g_{(ij)}] + \tfrac{1}{2} [\hat{\gamma}^{(k)}{}_{(ij)} + \hat{\gamma}^{(k)}{}_{(ji)} - \hat{\gamma}_{(ij)}{}^{(k)}]. \tag{1}$$

Indices are raised with $g^{(ij)}$.

Let us define

$$\begin{Bmatrix} (k) \\ (ij) \end{Bmatrix} \equiv \tfrac{1}{2} g^{(kr)} [X_{(i)} g_{(rj)} + X_{(j)} g_{(ri)} - X_{(r)} g_{(ij)}]$$

and

$$\gamma^{(k)}{}_{(ij)} \equiv \tfrac{-1}{2} [\hat{\gamma}^{(k)}{}_{(ij)} + \hat{\gamma}^{(k)}{}_{(ji)} - \hat{\gamma}_{(ij)}{}^{(k)}].$$

Thus Eq. (1) becomes

$$\Gamma^{(k)}{}_{(ij)} = \begin{Bmatrix} (k) \\ (ij) \end{Bmatrix} - \gamma^{(k)}{}_{(ij)}. \tag{2}$$

One also finds

$$\hat{\gamma}_{(ij)}{}^{(k)} = \gamma^{(k)}{}_{(ij)} - \gamma^{(k)}{}_{(ji)}.$$

Using a dual basis $\omega^{(i)}$, $i = 1,\ldots,n$, for M_p^*, $p \in A$, one has[6]

$$\begin{aligned} R^{(m)}{}_{(ijk)} &\equiv K(\omega^{(m)}, X_{(i)}, X_{(j)}, X_{(k)}) \\ &= X_{(j)} \begin{Bmatrix} (m) \\ (ik) \end{Bmatrix} - X_{(k)} \begin{Bmatrix} (m) \\ (ij) \end{Bmatrix} + \begin{Bmatrix} (m) \\ (nj) \end{Bmatrix} \begin{Bmatrix} (n) \\ (ik) \end{Bmatrix} - \begin{Bmatrix} (m) \\ (nk) \end{Bmatrix} \begin{Bmatrix} (n) \\ (ij) \end{Bmatrix} \\ &\quad - X_{(j)} \gamma^{(m)}{}_{(ik)} + X_{(k)} \gamma^{(m)}{}_{(ij)} + \gamma^{(m)}{}_{(nj)} \gamma^{(n)}{}_{(ik)} - \gamma^{(m)}{}_{(nk)} \gamma^{(n)}{}_{(ij)} \\ &\quad - \begin{Bmatrix} (m) \\ (nj) \end{Bmatrix} \gamma^{(n)}{}_{(ik)} - \gamma^{(m)}{}_{(nj)} \begin{Bmatrix} (n) \\ (ik) \end{Bmatrix} + \begin{Bmatrix} (m) \\ (nk) \end{Bmatrix} \gamma^{(n)}{}_{(ij)} + \gamma^{(m)}{}_{(nk)} \begin{Bmatrix} (n) \\ (ij) \end{Bmatrix} \\ &\quad - \begin{Bmatrix} (m) \\ (in) \end{Bmatrix} (\gamma^{(n)}{}_{(jk)} - \gamma^{(n)}{}_{(kj)}) + \gamma^{(m)}{}_{(in)} (\gamma^{(n)}{}_{(jk)} - \gamma^{(n)}{}_{(kj)}). \end{aligned}$$

Definition 10. The Ricci tensor is defined by[1]

$$\mathrm{Ric}(X,Y) \equiv (\mathrm{tr}^{1,3} K)(X,Y) \equiv \sum_{k=1}^{n} K(\omega^{(k)}, X, Y, X_{(k)}).$$

One can check that this definition is independent of the particular basis used.

In terms of a basis we may write $R_{(ij)} = \mathrm{Ric}(X_{(i)}, X_{(j)}) =$

$$
\begin{aligned}
R^{(k)}{}_{(ijk)} = {} & X_{(j)}\begin{Bmatrix}(k)\\(ik)\end{Bmatrix} - X_{(k)}\begin{Bmatrix}(k)\\(ij)\end{Bmatrix} + \begin{Bmatrix}(k)\\(nj)\end{Bmatrix}\begin{Bmatrix}(n)\\(ik)\end{Bmatrix} - \begin{Bmatrix}(k)\\(nk)\end{Bmatrix}\begin{Bmatrix}(n)\\(ij)\end{Bmatrix} \\
& - X_{(j)}\gamma^{(k)}{}_{(ik)} + X_{(k)}\gamma^{(k)}{}_{(ij)} + \gamma^{(k)}{}_{(nj)}\gamma^{(n)}{}_{(ik)} - \gamma^{(k)}{}_{(nk)}\gamma^{(n)}{}_{(ij)} \\
& - \begin{Bmatrix}(k)\\(nj)\end{Bmatrix}\gamma^{(n)}{}_{(ik)} - \gamma^{(k)}{}_{(nj)}\begin{Bmatrix}(n)\\(ik)\end{Bmatrix} + \begin{Bmatrix}(k)\\(nk)\end{Bmatrix}\gamma^{(n)}{}_{(ij)} + \gamma^{(k)}{}_{(nk)}\begin{Bmatrix}(n)\\(ij)\end{Bmatrix} \\
& - \begin{Bmatrix}(k)\\(in)\end{Bmatrix}(\gamma^{(n)}{}_{(jk)} - \gamma^{(n)}{}_{(kj)}) + \gamma^{(k)}{}_{(in)}(\gamma^{(n)}{}_{(jk)} - \gamma^{(n)}{}_{(kj)}).
\end{aligned}
$$

Definition 11. Let $(M, M_p, g, \mathscr{F})$ be a PDM with smooth metric g. An open set $A \subset M$ is called *special Einstein* provided that for all X, Y on A

$$\mathrm{Ric}(X,Y) = 0. \tag{3}$$

Condition (3) can be expressed as

$$R_{(ij)} = 0. \tag{4}$$

A third equivalent form of these equations is

$$R^{(m)}{}_{(ijk)} = C^{(m)}{}_{(ijk)}, \tag{5}$$

where $C^{(m)}{}_{(ijk)}$ is the Weyl conformal curvature tensor.[6]

In case $\dim M_p = 4$ and the signature of g_p is -2 (or else $+2$) any of the above three equations represent Einstein's vacuum equations. In case an orthonormal basis set is chosen,

$\begin{Bmatrix}(k)\\(ij)\end{Bmatrix} = 0$, and the $\gamma^{(k)}{}_{(ij)}$ are the *usual* Ricci rotation coefficients. The complex linear combinations of (5) give predecessors of the Newman–Penrose equations. In case a coordinate basis is chosen, $\gamma^{(k)}{}_{(ij)} = 0$, and the

$\begin{Bmatrix}(k)\\(ij)\end{Bmatrix}$ are the *usual* Christoffel symbols and Eqs. (4) are the standard vacuum equations.

3. CONSTRUCTION OF COORDINATES IN A PDM

To construct coordinate patches on a PDM $(M, M_p, \mathscr{F})$, the following assumptions are used.

(a) Each point has a neighborhood on which there exists a basis of n smooth vector fields $X^1, X^2, \ldots, X^n$ (locally framed).
(b) Local path connectivity hypothesis: For each point r and each neighborhood $\mathscr{N}$ of r there is a neighborhood $\mathscr{U} \subseteq \mathscr{N}$ of r such that for any two points p, q, $p \neq q$, in $\mathscr{U}$ there exists in $\mathscr{U}$ a continuous curve $\sigma(t)$, $0 \leqslant t \leqslant 1$, joining p to q. Furthermore, for all f in $\mathscr{F}_{\sigma(t)}$,

$$\frac{d}{dt} f \cdot \sigma(t) = \sigma_i X^i_{\sigma(t)}(f), \tag{6}$$

for some constant vector $\boldsymbol{\sigma} = (\sigma_i)$ of Euclidean length 1.

Theorem. If a PDM $(M, M_p, \mathscr{F})$ satisfies (a) and (b), then C^0 atlas exists.

Proof: Let $r \in M$. It will be shown that there exist n functions f_j in $\mathscr{F}_r$ such that the matrix $J^{ij}_r \equiv X^i_r f_j$ is nonsingular. By way of contradiction, suppose that for every n functions this matrix were singular. Expanding the determinant along the last column, we find $a_k X^k f_n = 0$, where the a_k's are cofactors. Since f_n is arbitrary and the X^k's are independent, all the a_k's must vanish. In particular,

$0 = a_n = \det[X^i f_j]$, $i, j \in \{1, 2, \ldots, n-1\}$, so we can expand again along the last column. Continuing in this way we find $X^1 f_1 = 0$ for all f_1. This contradicts the linear independence of the X^k's.

Let us write $\mathbf{J}(q, \boldsymbol{\sigma})$ for the vector with components $\sigma_i J^{ij}_q$. Since this depends continuously on q, one can show that there is a neighborhood $\mathscr{U}$ of r such that

$$|\mathbf{J}(q,\boldsymbol{\sigma}) - \mathbf{J}(r,\boldsymbol{\sigma})| \leqslant \tfrac{1}{2}|\mathbf{J}(r,\boldsymbol{\sigma})|.$$

for all $p, q, \in \mathscr{U}$ and all $\boldsymbol{\sigma}$. Let the neighborhood $\mathscr{U}$ be chosen small enough[7] so that in addition (a) and (b) hold. Let $m = \min\{|\mathbf{J}(r,\boldsymbol{\sigma})| : |\boldsymbol{\sigma}| = 1\}$.

If $\phi: \mathscr{U} \to R^n$ is defined by

$$\phi(q) = (f_1(q), f_2(q), \ldots, f_n(q)),$$

then $(\phi, \mathscr{U})$ will be a coordinate patch on M. It suffices to show that ϕ is 1–1. Let p and q be distinct points in $\mathscr{U}$, and let σ be the path joining them, given by (b). Then

$$
\begin{aligned}
\phi(q) - \phi(p) = \phi \cdot \sigma(1) - \phi \cdot \sigma(0) &= \int_0^1 \frac{d}{dt}\phi \cdot \sigma(t)\, dt \\
&= \int_0^1 \mathbf{J}(\sigma(t), \boldsymbol{\sigma})\, dt
\end{aligned}
$$

and

$$
\begin{aligned}
|\phi(q) - \phi(p)| &= \left| \int_0^1 \mathbf{J}(r,\sigma)\, dt + \int_0^1 [\mathbf{J}(\sigma(t),\boldsymbol{\sigma}) - \mathbf{J}(r,\boldsymbol{\sigma})]\, dt \right| \\
&\geqslant \left| \int_0^1 \mathbf{J}(r,\boldsymbol{\sigma})\, dt \right| - \left| \int_0^1 [\mathbf{J}(\sigma(t),\boldsymbol{\sigma}) - \mathbf{J}(r,\boldsymbol{\sigma})]\, dt \right| \\
&\geqslant |\mathbf{J}(r,\boldsymbol{\sigma})| - \tfrac{1}{2}|\mathbf{J}(r,\boldsymbol{\sigma})| \geqslant m/2.
\end{aligned}
$$

Thus ϕ is 1–1.

4. PHYSICAL INTERPRETATION OF THE COORDINATES

Let M be a topological space and let θ be the set of all curves $\sigma: (a,b) \to M$ representing the world lines of observers with C^∞ acceleration, where the parameter t of $\sigma(t)$ corresponds to the proper time (or any C^∞ transformation thereof). Let $\mathscr{F}$ be the collection of all scalar functions f on M such that f can be experimentally determined and $f \cdot \sigma(t)$ is C^∞. Then $\mathscr{F}$ is an algebra of functions. Let θ_p the set of these curves which pass through p; i.e., $\theta_p = \{\sigma \in \theta : p = \sigma(t)$ for some $t \in (a,b)\}$.

The tangent space M_p can now be defined. If $f \in \mathscr{F}_p$, then $f \cdot \sigma$ is C^∞, so one can define a function $\sigma^*: \mathscr{F}_p \to R$ by the rule $\sigma^*(f) = (d/dt) f \cdot \sigma(t)|_p$. It can be shown that σ^* is a tangent vector, and we define M_p to be the vector space generated by the σ^*'s; i.e.,

$$M_p \equiv \left\{ \sum_{i=1}^{n} a_i \sigma_i^* : \sigma_i \in \theta_p \text{ and } n \in N \right\}.$$

It is assumed that $\dim M_p$ is four for a pre-atlas semi-Riemannian manifold of space–time events.

ACKNOWLEDGMENTS

The authors are indebted to NSERC Grant No. 3993. They also had informal discussions with graduate students T. Biech, J. Gegenberg, and S. Toporowski.

[1] N. J. Hicks, *Notes on Differential Geometry* (Van Nostrand Reinhold, New York, 1965).
[2] J. W. Milnor, Ann. Math. **64**, 399 (1956).
[3] J. L. Synge, *Relativity: The General Theory* (North-Holland, Amsterdam, 1960).
[4] J. Milnor and J. Stasheff, *Characteristic Classes* (Princeton University, Princeton, 1974), p. 10; S. Kobayashi and K. Nomizu, *Foundations of Differential Geometry* (Interscience, New York, 1981), p. 5.
[5] J. Ehlers, F. A. E. Pirani, and A. Schild, *General Relativity: Papers in honour of J. L. Synge* (Clarendon, Oxford, 1972), p. 63; P. Carlson, J. Math. Phys. **21**, 1149 (1980).
[6] C. W. Misner, K. S. Thorn, and J. A. Wheeler, *Gravitation* (W. H. Freeman, San Francisco, 1973), pp. 223, 227; T. J. Willmore, *An Introduction to Differential Geometry* (Clarendon, Oxford, 1966), p. 273.
[7] This can be done by intersecting a sufficient number of open sets.

Structure of asymptotic twistor space

Gabriel Lugo
Department of Mathematics, University of Kansas, [a]*Lawrence, Kansas 66045*

(Received 12 May 1980; accepted for publication 27 February 1981)

We show that asymptotic projective twistor space $\mathscr{P}\mathscr{T}^+$ is an Einstein–Kähler manifold of positive curvature. We then use the Chern–Moser theory of hypersurfaces in complex manifolds to show that the Kähler curvature of $\mathscr{P}\mathscr{T}^+$ is closely related to the CR curvature of its boundary. We also give a proof that the Kähler potential function defining the boundary satisfies the complex Monge–Ampere equations.

PACS numbers: 04.20.Cv

INTRODUCTION

During the past few years a great deal of work has been done in the applications of complex manifolds to the theory of general relativity. In particular, much attention has been paid to half-flat space–times; that is, four-dimensional complex manifolds with Ricci flat, self-dual curvatures.[1–4] It has been shown by Newman and his coworkers[2] that given an asymptotically flat space–time $(\mathscr{M}, g)$, there exists a naturally associated half-flat manifold called an $\tilde{\mathscr{H}}$ space.

It is also possible to construct from $(\mathscr{M}, g)$ an asymptotic twistor space $\mathscr{T}(\mathscr{M})$ and a corresponding projective twistor space $\mathscr{P}\mathscr{T}(\mathscr{M})$.[5] In the presence of gravitational radiation the space $\mathscr{T}$ is a curved four-dimensional Kähler manifold of signature $(+\ +\ -\ -)$. The Kähler potential $L(z^\alpha, \bar{z}^\alpha)$ is a real valued function built out of a solution to Newman's good cut equation.

The equation $L = 0$ defines in $\mathscr{P}\mathscr{T}$ a five-dimensional real hypersurface $\mathscr{P}\mathscr{N}$ with a nondegenerate Levi form of signature $(+\ -)$. The hypersurface bounds a region $\mathscr{P}\mathscr{T}^+ = \{L > 0\}$ which has been suggested as representing a nonlinear graviton of positive helicity.[3–5]

It has been known for some time that the Ricci tensor of the Kähler metric of $\mathscr{T}$ vanishes, while the full curvature tensor contains information about the radiation field of the original space–time.[5]

Aside from the facts just mentioned very little is known about this Kähler structure. There have been suggestions that this structure is well suited to the implementation of a scattering theory of nonlinear gravitons,[4] but this goal is far from complete. It also seems likely that there is enough information coded in the CR structure of $\mathscr{P}\mathscr{N}$ to extract an intrinsic definition of the elusive concept of asymptotic flatness of $\tilde{\mathscr{H}}$ spaces, but how this is to be done is still an unanswered question.

With this motivation in mind we study in this paper the Kähler structure induced on the projective twistor space $\mathscr{T}$. We show that $\mathscr{P}\mathscr{T}^+$ is an Einstein–Kähler manifold of positive curvature. We then apply the Chern–Moser theory of pseudoconformal geometry of real hypersurfaces in complex manifolds[7] to conclude that the Kähler curvature of $\mathscr{P}\mathscr{T}^+$ is closely related to the Chern–Moser invariants of the boundary. We also give a short proof that the function defining $\mathscr{P}\mathscr{N}$ satisfies the complex Monge–Ampere equations.

The reader is assumed to have some familiarity with the spinor formalism and with the ð operator of Newman and Penrose.[2–4] The paper will be arranged as follows. In Sec. 1 we briefly review the construction of $\tilde{\mathscr{H}}$ spaces and asymptotic twistor spaces. In Sec. 2 we discuss the Kähler structure of $\mathscr{P}\mathscr{T}^+$ and in Sec. 3 we relate this structure to the CR structure of the boundary.

I. HALF FLAT SPACES

A. $\tilde{\mathscr{H}}$-space

Let $(\mathscr{M}, g)$ be an asymptotically flat space–time with complexified null infinity $\mathbb{C}\mathscr{I}^+$. Let ζ and $\tilde{\zeta}$ denote the stereographic coordinates on the complexified two-sphere and introduce the quantity

$$P_0 = \tfrac{1}{2}(1 + \zeta\tilde{\zeta}). \qquad (1.1)$$

By a "*good cut*" one means a cross section $u = X(\zeta, \tilde{\zeta})$ of $\mathbb{C}\mathscr{I}^+$ satisfying the equation

$$\tilde{ð}^2 X = \tilde{\sigma}^0(X, \zeta, \tilde{\zeta}). \qquad (1.2)$$

Here, the operators $\tilde{ð}$ and its dual ð (edth) acting on a function η of spin weight s are defined by

$$ð\eta = 2P_0^{1-s}\frac{\partial}{\partial\zeta}(P_0^s\eta), \qquad (1.3)$$

$$\tilde{ð}\eta = 2P_0^{1+s}\frac{\partial}{\partial\tilde{\zeta}}(P_0^{-s}\eta),$$

and $\tilde{\sigma}^0$ is the asymptotic shear of a Bondi family of null surfaces.[2] In a Bondi coordinate system the commutator of ð and $\tilde{ð}$ is given by

$$(ð\tilde{ð} - \tilde{ð}ð)\eta = 2s\eta. \qquad (1.4)$$

The nonlinear differential equation (1.2) has a four-parameter family of solutions for $\tilde{\sigma}^0$ sufficiently close to zero. The manifold of such solutions is the $\tilde{\mathscr{H}}$-space associated with $(\mathscr{M}, g)$.

If $\tilde{\sigma}^0 = 0$, the solutions to (1.2) are of the form

$$X = (\omega^0 + \tilde{\zeta}\omega^1)/2P_0, \qquad (1.5)$$

with

$$\omega^0 = u + x\zeta, \qquad \omega^1 = y + v\zeta. \qquad (1.6)$$

The four quantities (u, v, x, y) parametrize the $\tilde{\mathscr{H}}$ space

[a]Part of this work was done at the University of California at Berkeley.

 0022-2488/82/020276-07$02.50

which in this case is isomorphic to complexified Minkowski space $\mathbb{C}M_{1,3}$.

B. Asymptotic twistor space

Denote by ι^A the spinor field pointing along the generators of $\mathbb{C}\mathscr{I}^+$. By a *hypersurface twistor* with respect to $\mathbb{C}\mathscr{I}^+$ one means a pair $(\pi_{\dot{A}}, \Gamma)$, where Γ is a complex curve in $\mathbb{C}\mathscr{I}^+$ and $\pi_{\dot{A}}$ is a spinor field on Γ satisfying

$$\begin{aligned}&\text{(a)}\ t^a = \sigma^a_{A\dot{B}}\iota^A\pi^{\dot{B}} \text{ is tangent to } \Gamma,\\&\text{(b)}\ \iota^A\pi^{\dot{B}}\nabla_{A\dot{B}}\pi_{\dot{C}} = 0 \text{ on } \Gamma,\\&\text{(c)}\ \pi_{\dot{1}} \neq 0 \text{ on } \Gamma.\end{aligned} \tag{1.7}$$

Here $\sigma^a A\dot{B}$ are the usual Pauli spin matrices. Condition (b) states that $\pi_{\dot{A}}$ is parallelly propagated along the curve Γ which is referred to as a *twistor curve*. Condition (c) is added to avoid the possibility of having Γ coincide with one of the generators of $\mathbb{C}\mathscr{I}^+$.

The collection of all such pairs $(\pi_{\dot{A}}, \Gamma)$ is a four-dimensional complex manifold associated with $\mathbb{C}\mathscr{I}^+$ known as the *asymptotic twistor space* $\mathscr{T}$ of $\mathscr{M}$. The *projective asymptotic twistor space* $\mathscr{T}$ is obtained by taking the quotient of $\mathscr{T}$ with the equivalence relation $(\pi_{\dot{A}}, \Gamma_1) \sim (\rho_{\dot{A}}, \Gamma_2)$ iff $\Gamma_1 = \Gamma_2$ and $\pi_{\dot{A}} = c\rho_{\dot{A}}$ for some number $c \in \mathbb{C}^*$.

In a similar fashion we may define the space $\mathscr{T}^*$ of dual hypersurface twistors to be the space of pairs $(\eta_A, \tilde{\Gamma})$ satisfying

$$\begin{aligned}&\text{(a)}\ \tilde{t}^a \equiv \sigma^a{}_{A\dot{B}}\eta^A\tilde{\iota}^{\dot{B}} \text{ is tangent to } \tilde{\Gamma},\\&\text{(b)}\ \lambda^A\tilde{\iota}^{\dot{B}}\nabla_{A\dot{B}}\eta_c = 0 \text{ on } \tilde{\Gamma},\\&\text{(c)}\ \eta_1 \neq 0 \text{ on } \tilde{\Gamma}.\end{aligned} \tag{1.8}$$

The space $\mathscr{T}^*$ also has a projective structure defined exactly as above. We will denote the corresponding *dual asymptotic projective twistor space* by $\mathscr{P}\mathscr{T}^*$. The curves $\tilde{\Gamma}$ are called *dual twistor curves*.

It turns out that if $\mathscr{M}$ is taken to the Minkowski space, then the definition of the asymptotic twistor space given here coincides with the usual flat twistor space $\mathbb{T}$. Furthermore, it has been shown[2] that the space $\mathscr{P}\mathscr{T}$ arises as a deformation of the complex structure of a region on the flat projective twistor space $\mathbb{PT} \simeq \mathbb{CP}_3$.

The relation between the asymptotic twistor spaces and the $\mathscr{H}$ spaces associated with a space–time $\mathscr{M}$ is contained in the following theorems.

Theorem 1.1: (Penrose[2,3]). If $\mathscr{P}\mathscr{T}$ is a sufficiently small deformation of $\mathbb{PT}$, then there exists in $\mathscr{P}\mathscr{T}$ a four-complex-parameter family of compact holomorphic curves with the same homology class as a sphere S^2. Furthermore, the manifold parametrizing these curves is isomorphic to $\mathscr{H}$.

Theorem 1.2: Each point of $\mathscr{P}\mathscr{T}$ corresponds to a twistor curve in $\mathbb{C}\mathscr{I}^+$ and the good cuts of $\mathbb{C}\mathscr{I}^+$ are ruled by a $\mathbb{CP}_1$'s worth of twistor curves.

Define two points in $\mathscr{H}$ to be null separated if the corresponding holomorphic curves in $\mathscr{P}\mathscr{T}$ intersect. This endows $\mathscr{H}$ with a conformal structure and we have:

Theorem 1.3. The conformal structure of $\mathscr{H}$ is half-flat.

For the convenience of the reader we breifly explain the correspondence between $\mathscr{P}\mathscr{T}$ and $\mathscr{H}$. Consider a particular surface $u = X(\zeta, \tilde{\zeta})$ in $\mathbb{C}\mathscr{I}^+$ satisfying Eq. (1.2). For a fixed value of $\zeta = \zeta_1$, the good cut equation becomes an ordinary differential equation and the solution is a twistor curve Γ lying in $X(\zeta, \tilde{\zeta})$. In fact, it is not hard to see that the twistor curve is a null geodesic in $\mathbb{C}\mathscr{I}^+$. The curve may be parametrized by three numbers $(\omega^0, \omega^1, \zeta_1)$, where we may think of $\omega^A = (\omega^0, \omega^1)$ as the constants of integration of the differential equation. The quantities $(\omega^0, \omega^1, \zeta_1)$ define the twistor line up to proportionality, and thus they may be regarded as the local coordinates of a point in $\mathscr{P}\mathscr{T}$. As the value of ζ varies along the good cut we get a curve in $\mathscr{P}\mathscr{T}$ which is holomorphic since $X(\zeta, \tilde{\zeta})$ is assumed to vary holomorphically with respect to both ζ and $\tilde{\zeta}$. The curve so obtained is an element of the four-parameter family of curves whose existence is guaranteed by Theorem 1.1.

II. KÄHLERIAN STRUCTURES

A. Kähler structure of $\mathscr{T}$

In the space of asymptotic twistors it is possible to define a scalar product and a Kähler structure using the ideas of local twistors.[4,5] Consider two arbitrary hypersurface twistors $(\Gamma, \pi_{\dot{A}}) \in \mathscr{T}$ and $(\tilde{\Gamma}, \eta_A) \in \mathscr{T}^*$ respectively. In general, there exists at most one generator γ of $C\mathscr{I}^+$ intersecting both of the twistor curves Γ and $\tilde{\Gamma}$. Supposing that such a generator exists, we represent the hypersurface twistor $(\Gamma, \pi_{\dot{A}})$ by a local twistor $(\omega^A, \pi_{\dot{A}}) = (0, \pi_{\dot{A}})$ at P and the dual hypersurface twistor $(\tilde{\Gamma}, \eta_A)$ by a local twistor $(\eta_A, 0)$ at Q, where P and Q are the corresponding points of intersection of Γ and $\tilde{\Gamma}$ with the generator γ.

The scalar product between $(\Gamma, \pi_{\dot{A}})$ and $(\tilde{\Gamma}, \eta_A)$ is defined by propagating the local twistor $Z^\alpha = (\omega^A, \pi_{\dot{A}})$ up the generator γ using local twistor transport

$$\iota^A\tilde{\iota}^{\dot{B}}\nabla_{A\dot{B}}\omega^C(x) = -i\pi_{\dot{1}}(x)\iota^C, \tag{2.1}$$

$$\iota^A\tilde{\iota}^{\dot{B}}\nabla_{A\dot{B}}\pi_{\dot{C}}(x) = -iP_{1B\dot{1}\dot{C}}\omega^B, \tag{2.2}$$

and then taking the local scalar product with $\tilde{Z}_\alpha = (\eta_A, \xi^{\dot{A}})$ at Q. Thus, the scalar product is given by

$$L(Z^\alpha, \tilde{Z}_\alpha) = (\omega^A\eta_A + \pi_{\dot{A}}\xi^{\dot{A}})(Q) = \omega^A\eta_A(Q). \tag{2.3}$$

We could, of course, propagate the dual twistor Z_α from Q to P and take the local inner product at P, but this will clearly yield the same answer. It may happen that there exists no generator γ intersecting both of the twistor curves. In this case the twistor scalar product is not defined.

The Kähler structure of asymptotic twistor space is obtained by viewing the scalar product as a potential for a Kähler form on $\mathscr{T}$ defined by

$$\phi = [\partial^2 L(Z^\alpha, \overline{Z}_\alpha)/\partial z^\alpha \partial z^{\bar{\beta}}]dz^\alpha \wedge dz^{\bar{\beta}}, \tag{2.4}$$

where $\overline{Z}_\alpha$ is the dual asymptotic twistor associated with the complex conjugate of the curve defining the twistor Z^α, and z^α $(\alpha = 0, 1, 2, 3)$ are local coordinates of $\mathscr{T}$. The curvature properties of the Kähler metric of $\mathscr{T}$ have been computed by Penrose, Newman, and Ko[5] using local coordinates $z^\alpha = (\omega^0, \omega^1, \zeta, \lambda)$, where λ is essentially the component $\pi_{\dot{1}}$ taken with respect to a conformal rescaling K^A of the spinor ι^A, the conformal factor being taken so as to make the induced metric $\mathbb{C}\mathscr{I}^+$ become flat. In the paper just mentioned, it was found that in local coordinates the scalar product can be expressed in terms of the good cut function $X(\omega^0, \omega^1, \zeta, \bar{\zeta})$ as

$$L(Z^\alpha, \bar{Z}_\alpha) = i\lambda\bar{\lambda}(1+\zeta\bar{\zeta})(X-\bar{X}), \quad (2.5)$$

and the curvature tensor is expressible in terms of the Ψ_4^0, Ψ_3^0 and $\mathrm{Im}\Psi_2^0$ components of the Weyl tensor of the space-time $\mathscr{M}$. In this sense the Kähler structure codes into the asymptotic twistor space the information about the gravitational radiation field of the space-time.

Remark: The scalar product (2.5) agrees with the usual flat twistor scalar product when we consider the space of asymptotic twistors of Minkowski space. To see this write the solution to the good cut equation as in (1.5). Choosing coordinates

$$\begin{aligned} z^0 &= \tfrac{1}{2}\sqrt{2\lambda}\,(i\omega^1 + \zeta), \\ z^1 &= \tfrac{1}{2}\sqrt{2\lambda}\,(i\omega^1 - \zeta), \\ w &= 2\lambda\omega^0, \\ s &= \lambda. \end{aligned} \quad (2.6)$$

We get the inner product in the form

$$L(Z, \bar{Z}) = |z^0|^2 - |z^1|^2 + \tfrac{1}{2}i(w\bar{s} - s\bar{w}). \quad (2.7)$$

Note that these coordinates are homogeneous in λ.

B. Kähler structure of $\mathscr{PT}$

Since the quantities $t^\alpha = (\omega^A, \zeta)$ are good inhomogeneous coordinates in $\mathscr{PT}$, we can replace L by

$$K = 2iP_0(X - \bar{X}), \quad (2.8)$$

and we get a Kähler metric in the region $\mathscr{PT}^+$ defined by $K > 0$ by taking

$$ds^2 = 2g_{\alpha\bar{\beta}}\,dt^\alpha\,dt^{\bar{\beta}} = 4(\partial^2 \ln K/\partial t^\alpha \partial t^{\bar{\beta}})\,dt^\alpha\,dt^{\bar{\beta}}. \quad (2.9)$$

Using subscripts to denote the derivatives of K (i.e., $K_\alpha = \partial K/\partial t^\alpha$, $K_{\bar{\beta}} = \partial K/\partial t^{\bar{\beta}}$, $K_A = \partial K/\partial \omega^A$, etc.) we can write the metric tensor as

$$g_{\alpha\bar{\beta}} = K^{-2}(KK_{\alpha\bar{\beta}} - K_\alpha K_{\bar{\beta}}). \quad (2.10)$$

Setting $V = X - \bar{X}$ and observing that $V_A = X_A$, $V_{\bar{B}} = X_{\bar{B}}$, and $V_{A\bar{B}} = 0$, we find a more explicit form of the metric

$$g_{\alpha\bar{\beta}} = \left\| \begin{array}{c|c} -2V_A V_{\bar{B}}/V^2 & 2(VV_{A\bar{\zeta}} - V_A V_{\bar{\zeta}})/V^2 \\ \hline 2(VV_{\zeta\bar{B}} - V_\zeta V_{\bar{B}})/V^2 & 2(VV_{\zeta\bar{\zeta}} - V_\zeta V_{\bar{\zeta}})/V^2 + \tfrac{1}{2}P_0^{\,2} \end{array} \right\|. \quad (2.11)$$

Following the conventions of,[6] we find that the only nonvanishing components of the connection are

$$\Gamma^\mu_{\alpha\gamma} = g_{\alpha\bar{\nu},\gamma}\,g^{\mu\bar{\nu}}, \quad \Gamma^{\bar{\mu}}_{\bar{\beta}\bar{\delta}} = \overline{\Gamma^{\mu}_{\beta\delta}} = g_{\mu\bar{\beta},\bar{\delta}}\,g^{\mu\bar{\nu}}, \quad (2.12)$$

where $g^{\mu\bar{\nu}}$ is the matrix inverse of $g_{\alpha\bar{\beta}}$ (i.e, $g_{\alpha\bar{\beta}}g^{\gamma\bar{\beta}} = \delta^\gamma_\alpha$).

The Riemann curvature is given locally by the expression

$$R^{\ \mu}_{\alpha\gamma\bar{\delta}} = \Gamma^{\mu}_{\alpha\gamma,\bar{\delta}}, \quad (2.13)$$

and the Ricci tensor defined as

$$R_{\alpha\bar{\beta}} = R^{\ \mu}_{\alpha\mu\bar{\beta}} \quad (2.14)$$

can be obtained from the determinant g of the metric

$$R_{\alpha\bar{\beta}} = -\frac{\partial^2}{\partial t^\alpha\,\partial t^{\bar{\beta}}}\ln g. \quad (2.15)$$

A long but straightforward computation using (2.10) and (2.13) gives the following formula for the Riemann tensor.

$$\begin{aligned} R_{\alpha\bar{\beta}\gamma\bar{\delta}} &= -\tfrac{1}{2}(g_{\alpha\bar{\beta}}g_{\gamma\bar{\delta}} - g_{\alpha\bar{\delta}}g_{\gamma\bar{\beta}}) + 2K^{-2}(KK_{\alpha\bar{\beta}\gamma\bar{\delta}} - K_{\alpha\gamma}K_{\bar{\beta}\bar{\delta}}) \\ &\quad + 2K^{-4}g^{\mu\bar{\nu}}(KK_{\alpha\bar{\nu}\gamma} - K_{\alpha\gamma}K_{\bar{\nu}})(KK_{\nu\bar{\beta}\bar{\delta}} - K_{\bar{\beta}\bar{\delta}}K_\mu). \end{aligned} \quad (2.16)$$

As a consequence of this we have

Proposition 2.1: The space $\mathbb{PT}^+$ is an Einstein–Kähler manifold of constant holomorphic sectional curvature equal to 1.

Proof: In this case the solution to the good cut equation is

$$X = (2P_0)^{-1}(\omega^0 + \bar{\zeta}\omega^1),$$

so that

$$K = i(\omega^0 + \bar{\zeta}\omega^1 - \bar{\omega}^0 - \zeta\bar{\omega}^1).$$

It follows immediately from (2.16) that

$$R_{\alpha\bar{\beta}\gamma\bar{\delta}} = \tfrac{1}{2}(g_{\alpha\bar{\beta}}g_{\gamma\bar{\delta}} - g_{\alpha\bar{\delta}}g_{\gamma\bar{\beta}}), \quad (2.17)$$

which is the form of a Riemann tensor of constant holomorphic curvature equal to 1. That the space is Einstein then follows trivially.

Remark: The metric on $\mathbb{PT}^+$ has signature $(+\ -\ -)$ and is the semidefinite analog of the Fubini–Study metric of $\mathbb{CP}^3$.

Proposition 2.2: The Kähler structure of $\mathbb{PT}^+$ is invariant under the action of SU(2,2), and, up to a scalar multiple, it is the only one with this property.

The proof follows from the fact that the action of SU(2, 2) preserves the Hermitian form $(+\ +\ -\ -)$ as well as the complex structure of $\mathbb{C}^4$.

The analog of the Weyl tensor for a Kähler manifold of complex dimension n is given by the *Bochner tensor*

$$\begin{aligned} C_{\alpha\bar{\beta}\gamma\bar{\delta}} &= R_{\alpha\bar{\beta}\gamma\bar{\delta}} + (n+2)^{-1} \\ &\quad \times (R_{\alpha\bar{\beta}}g_{\gamma\bar{\delta}} + R_{\gamma\bar{\beta}}g_{\alpha\bar{\delta}} + g_{\alpha\bar{\beta}}R_{\gamma\bar{\delta}} + g_{\gamma\bar{\beta}}R_{\alpha\bar{\delta}}) \\ &\quad - R(n+1)^{-1}(n+2)^{-1}(g_{\alpha\bar{\beta}}g_{\gamma\bar{\delta}} + g_{\gamma\bar{\beta}}g_{\alpha\bar{\delta}}), \end{aligned} \quad (2.18)$$

where R is the scalar curvature $R = g^{\alpha\bar{\beta}}R_{\alpha\bar{\beta}}$.

Proposition 2.3: The Bochner tensor of $\mathbb{PT}^+$ vanishes.

Proof: Contracting the expression for the Riemann curvature with the metric tensor we get

$$R_{\alpha\bar{\beta}} = 2g_{\alpha\bar{\beta}},$$
$$R = 6.$$

Substituting into (2.18) and making further use of (2.17) we find

$$C_{\alpha\bar{\beta}\gamma\bar{\delta}} = R_{\alpha\bar{\beta}\gamma\bar{\delta}} - \tfrac{8}{5}R_{\alpha\bar{\beta}\gamma\bar{\delta}} + \tfrac{3}{5}R_{\alpha\bar{\beta}\gamma\bar{\delta}} = 0.$$

Theorem 2.1: The space $\mathscr{PT}^+$ is an Einstein–Kähler manifold of constant scalar curvature equal to 6.

Proof: First, we notice that the upper left corner block of the metric (2.11) has vanishing determinant. Thus, expanding the full determinant g of the metric by minors along the bottom row, we find after some cancellations that g can be expressed in the rather simple form

$$\begin{aligned} g &= V^{-4}\rho\bar{\rho}, \\ \rho &= \epsilon^{CD}X_C X_{D\bar{\zeta}}. \end{aligned} \quad (2.19)$$

The proof then follows by a long computation using (2.15) and the good cut equation. The details are found in the Appendix. Theorem (2.1) states that the nontrivial part of the curvature is contained in the Bochner tensor. This tensor may be calculated using (2.16), (2.18), and the lemma in the Appendix. We will spare the reader the long details of the computation and we just remark that the tensor is expressible in terms of space–time Weyl components ψ_4^0, ψ_3^0, and Im ψ_2^0. In other words, the gravitational radiation data of $\mathscr{M}$ is coded into the bochner tensor of $\mathscr{P}\mathscr{T}^+$.

III. CR STRUCTURE

A real hypersurface on $\mathbb{C}^{n+1}$ inherits from the ambient space an intrinsic structure called a pseudoconformal or CR structure. That is, there exists a $2n$-dimensional subbundle of the holomorphic tangent bundle having a complex vector space structure on each fiber. The theory of pseudoconformal structures has been used by Chern–Moser[7] to study the invariants of strongly pseudoconvex hypersurfaces (i.e., hypersurfaces with positive definite Levi form) under biholomorphic mappings. The known invariants at present consist of a curvature tensor and certain real curves called chains.

Pseudoconformal geometry is important in complex analysis because one often has to deal with domains in $\mathbb{C}^{n+1}$ and the boundaries of domains are real hypersurfaces. In this context, biholomorphic invariants have also been independently studied by Fefferman[8] by constructing a defining function for the hypersurface which is an approximate solution to the complex Monge–Ampere equations. The invariants found by Fefferman have been related to those of Chern–Moser by work of Burns and Shnider[9] and Webster.[10]

Most of the known results on the subject until now make use of the assumption of strong pseudoconvexity and they do not apply to situations where the hypersurface does not have positive definite Levi form. Although the case of $\mathscr{P}\mathscr{N}$ is not a favorable one in the sense that its Levi form has signature $(+\ -)$, it is still possible to extend some of the known results to our situation.

The CR structure of $\mathscr{P}\mathscr{N}$ is important on several grounds. First, we need to understand the boundary to clarify the notion of positive frequency of nonlinear gravitons. Secondly, the CR structure contains important information about the space–time. In fact, it seems plausible that all of $\mathscr{H}$ space may be recovered by studying the chains in $\mathscr{P}\mathscr{N}$. Finally, the results of this paper (in particular Theorem 3.2) together with the observation that the good cut equation is the Dolbeault version of twistor deformations[11] brings into play powerful machinery of complex theory which hopefully can be used to gain further insight into the structure of the half-flat Einstein equations.

The theory of pseudoconformal structures is perhaps not well known to mathematical physicists and unfortunately it is impossible to present a detailed description of the geometry in these few pages. Thus, we will content ourselves with drawing a quick sketch of the ideas involved by evoking our knowledge of the theory of surfaces in Euclidean three space. For a more complete account, the reader is referred to the extensive work in the literature.[7–10]

The simplest and most important surfaces in $\mathbb{R}^3$ are the planes. They are the prototypes of flat surfaces in Riemannian geometry. The theory of surfaces in Euclidean space is a generalization of the geometry of planes. To be specific, let us consider a smooth surface which near the origin is defined by an equation of the form

$$z = f(x^i), \quad i = 1, 2. \tag{3.1}$$

The local properties of the surface may be understood classically by expanding the function f in a Taylor series around the origin:

$$z = z_0 + b_i x^i + (1/2!) b_{ij} x^i x^j + \cdots, \tag{3.2}$$

where the coefficients $b_{ij}\cdots$, denote the value of the partial derivatives of f at the origin. If we neglect all but the linear terms of the series we get the equation of a plane. This is the osculating plane which best approximates the surface at the origin. The quadratic coefficients b_{ij} represent the components of the second fundamental form. By applying a linear transformation, if necessary, we can rotate the surface so that $b_i = 0$; in other words, we can choose our frames such that the osculating plane becomes horizontal. With this choice, the determinant

$$G = |b_{ij}| \tag{3.3}$$

is the Gaussian curvature of the surface at the point in question. The fundamental theorem of geometry states that although the second fundamental form depends on the embedding, the Gaussian curvature is an intrinsic property of the surface and it is a bending invariant.

The geometric structure of the surface may also be viewed in terms of a principal fiber bundle with a connection. Classically, the fiber of the bundle at a given point on the surface is the tangent plane at that point. The unique Levi–Civita connection defines a law of parallel transport which allows one to construct the tangent space of a point in terms of that of a neighboring point. The curvature of the connection, defined by the second structure equations, has only one independent component which is the Gaussian curvature of the surface.

The situation for real hypersurfaces in $\mathbb{C}^{n+1}$ is, of course, more complicated but we can get some feeling for the geometry by trying to emulate the discussion above. It may be helpful for the reader to keep in mind the following lexicon of corresponding concepts.

Surfaces on $\mathbb{R}^3$	Real hypersurfaces on $\mathbb{C}^{n+1}$
Planes	Real hyperquadrics
Riemannian structure	Pseudoconformal structure
Fundamental form	Levi form
Group of motions	Biholomorphic transformations
Bundle of frames	Pseudoconformal bundle
Geodesics	Chains.

Suppose that near the origin the hypersurface is gven locally by a real valued function

$$r(z^\alpha, \bar{z}^\alpha) = 0. \tag{3.4}$$

We assume that not all of the partial derivatives of r

vanish at the origin. In particular, if we define

$$w = z^{n+1} = u + iv \tag{3.5}$$

we may assume that

$$r_w \neq 0. \tag{3.6}$$

By a linear change of coordinates, Eq. (3.4) may be put into the form

$$v = F(z, \bar{z}, u). \tag{3.7}$$

What Chern and Moser do is to take the last equation and write it in normal form; that is, they express F in terms of a power series in z and $\bar{z}$ with coefficients depending on u. Then, by applying appropriate coordinate transformations, the power series is reduced to the simplest possible form.

$$v = \langle z, \bar{z}\rangle + N_{22} + \sum_{k+l\geqslant 5} N_{kl}, \tag{3.8}$$

where

$$\langle z, \bar{z}\rangle = h_{\alpha_1\bar{\beta}_1} z^{\bar{\alpha}_1} z^{\bar{\beta}_1}, \tag{3.9}$$

$$N_{22} = b_{\alpha_1\alpha_2\bar{\beta}_1\bar{\beta}_2} z^{\alpha_1} z^{\alpha_2} z^{\bar{\beta}_1} z^{\bar{\beta}_2}, \tag{3.10}$$

$$N_{kl} = b_{\alpha_1\cdots\alpha_k\bar{\beta}_1\cdots\bar{\beta}_l} z^{\alpha_1}\cdots z^{\alpha_k} z^{\bar{\beta}_1}\cdots z^{\bar{\beta}_l}. \tag{3.11}$$

Here, the quantities N_{22} and N_{kl} are symmetric on the α's and β's, and they satisfy some trace conditions that we will not discuss.

The equation

$$v = \langle z, z\rangle \tag{3.12}$$

represents a real hyperquadric of signature $(p+1, q+1)$, where (p, q) is the signature of the quadratic form $h_{\alpha_1\beta_1}$, also called the Levi form of the surface. It will be assumed that the Levi form is nondegenerate. Equation (3.8) states that any nondegenerate real hypersurface in $\mathbb{C}^{n+1}$ may be osculated to high order by a quadric, whose Levi form is of the same signature of that of the given surface. For this reason real hyperquadrics are of fundamental importance in the theory. In fact, transformations of a surface into normal form are unique only up to the group of matrices which preserves the quadric, as well as the origin. The group H of such transformations is the isotropy subgroup of $SU(p+1, q+1)$ which leaves the origin fixed. The group H plays a role in pseudoconformal geometry analogous to that of the orthogonal group in Riemannian geometry.

The quartic coefficients $b_{\alpha_1\alpha_2\bar{\beta}_1\bar{\beta}_2}$ depend on the first four derivatives of the defining function r. They may be regarded (up to a constant factor) as the components of the so-called fourth-order Chern–Moser tensor. The tensor is an intrinsic quantity and it is invariant under the pseudogroup of biholomorphic transformations. There are, of course, higher order invariants associated with the fifth- and sixth-order coefficients but we will not be concerned with those here.

As in the Riemannian case, there is a coordinate-free formulation of pseudoconformal structures in terms of a bundle with connection. For a given point on the hypersurface, the fiber is a homogeneous space, namely, the tangent hyperquadric. The fundamental theorem of Chern–Moser is that on the bundle there exists a torsion-free Cartan connection π defining a notion of "parallel transport" among tangent hyperquadrics of infinitesimally close points. The pseudoconformal invariants are defined in terms of the connection via Cartan's structure equations

$$d\pi + \pi\wedge\pi = \Pi. \tag{3.13}$$

CR structure of $\hat{\mathscr{N}}$

As before, we let $t = \{t^\alpha\}$ be the natural coordinates in an open set U_1 of $\mathscr{P}\mathscr{T}^+$. The forms $\{\theta^\alpha, \theta^{\bar{\alpha}}\}$, where $\theta^\alpha = dt^\alpha$, then define a basis for the cotangent space over U_1. In terms of this coframe the Kähler metric (2.9) and the corresponding Kähler two-form Σ can be written as

$$ds^2 = 2g_{\alpha\bar{\beta}}\theta^\alpha\theta^{\bar{\beta}}, \tag{3.14}$$

$$\sum = -2ig_{\alpha\bar{\beta}}\theta^\alpha\wedge\theta^{\bar{\beta}}. \tag{3.15}$$

From Eq. (2.9) it follows that we can also write the Kähler form as

$$\sum = -4i\partial\bar{\partial}\ln K(t^\alpha, t^{\bar{\alpha}}). \tag{3.16}$$

Consider now the real valued function r on $\mathscr{T}$ defined by the equation

$$r = s\bar{s}K(t^\alpha, t^{\bar{\alpha}}) - 1 \tag{3.17}$$

$$= L - 1. \tag{3.18}$$

The real hypersurface $\hat{\mathscr{N}}$ given by

$$r = 0 \tag{3.19}$$

is then a circle bundle over $\mathscr{P}\mathscr{T}^+$. The hypersurface is a CR manifold whose Levi form is simply the lift of the Kähler form (3.15) to the bundle. This kind of hypersurface is exactly of the kind studied by Webster[10a] except that in his paper he was only concerned with the the positive definite case.

To get an idea of what kind of bundle we are dealing with, we consider the situation when the base manifold is taken to be the "flat" projective twistor space $\mathbb{PT}^+$. This manifold admits a homogeneous space representation of the form

$$\mathbb{PT}^+ = U(2,2)/(U(1,2)\times U(1,0)). \tag{3.20}$$

It follows from this observation that the corresponding circle bundle $\hat{\mathscr{N}}$ is just the Steifel bundle

$$U(2,2)/U(2,1) \xrightarrow{U(1,0)} \mathbb{PT}^+ \tag{3.21}$$

of frames of type (1,2) in T, with fiber $U(1,0) = S^1$.

To understand the nature of the circle bundle in (3.21) it is helpful to compare (3.17) with (2.5). It follows at a glance with the surface $\hat{N}$ associated with $\mathbb{PT}^+$ is nothing else but the hyperboloid of twistors of unit helicity in $\mathbb{T}$. The topology of this hyperboloid is $S^2\times R^4$ as can be seen from the equation

$$L = |z^0|^2 - |z^1|^2 + \tfrac{1}{2}i(w\bar{s} - s\bar{w}) = 1. \tag{3.22}$$

In view of this it becomes evident that the nontrivial part of the bundle (3.21) $S^3\times R^4 \xrightarrow{S^1} S^2\times R^4$ is just the usual Hopf fibration $S^3 \xrightarrow{S^1} S^2$.

Introduce a new variable $z^4 \in \mathbb{C}$, and "homogenize" (3.17) by letting $s \rightarrow s/z^4$. Then, $z^\alpha \mapsto z^\alpha/z^4$ and $t^\alpha \mapsto t^\alpha$. In terms of these quantities we get a "projective" version $\mathscr{P}\hat{\mathscr{N}}$ of the surface $\mathscr{N}$ defined by the equation

$$s\bar{s}K(t^\alpha, t^{\bar{\alpha}}) - |z^4| = 0. \tag{3.23}$$

The manifold of null projective twistors $\mathscr{P}\mathscr{N}$ then sits as the submanifold of $\mathscr{P}\hat{\mathscr{N}}$ defined by setting $z^4 = 0$.

To gain more geometrical insight into what is going on, we will again look at the flat $\hat{N}$. In this case, the coordinates z^i give an embedding of $\mathbb{C}^4$ as an open set in $\mathbb{CP}^4$ and we get the following diagram

$$\begin{array}{ccc} \mathbb{P}\hat{N} & \longrightarrow & \mathbb{CP}^4 \\ \downarrow & & \downarrow \\ \mathbb{P}N & \longrightarrow & \mathbb{CP}^3. \end{array}$$

The expression (3.23) then becomes a Hermitian form of signature $(+\ +\ -\ -\ -)$, and it defines a quadric $Q(2, 3)$ in $\mathbb{CP}^4$. The quadric has topology $S^3 \times S^4$ and it is, of course, the compactification of $N = S^3 \times R^4$. The map $(s, \omega^A, \zeta, z^4) \mapsto (s, \omega^A, \zeta, 0)$ gives a natural embedding of $\mathbb{CP}^3$ into $\mathbb{CP}^4$. The induced embedding gives $\mathbb{P}N$ as the submanifold defined by the intersection of $\mathbb{P}\hat{N}$ with the plane at infinity $\{z^4 = 0\}$.

The interesting fact is that there exists a theorem of Webster [10a] which can easily be extended to Kähler manifolds with indefinite signatures to give the following

Proposition 3.1: The Bochner tensor of $\mathscr{P}\mathscr{T}^+$ is equal to the fourth-order Chern–Moser tensor of $\hat{\mathscr{N}}$.

On the manifold $\hat{\mathscr{N}}$ one can find certain distinguished real curves called chains defined by the differential system $\pi_{-1} = \pi_1 = 0$.[12] Chains are the analogs of geodesics in CR manifolds. Very little is known about these objects, but they seem to carry a lot of information about the intrinsic geometry of the hypersurface. In the case of $\mathbb{P}N$, the chains are obtained by the intersection of complex secant lines with the quadric. The chains have topology S^1 and there is an eight-real-parameter family of them. The space of chains has the structure of a complex manifold which can be identified with $\mathbb{C}M_{1,3}$.

In a general CR manifold the space of chains is not so nice. In fact, it seems that generally there will exist a large number of chains which will spiral into a point.[8] In the situation here the base manifold of the bundle $\hat{\mathscr{N}} \xrightarrow{S^1} \mathscr{P}\mathscr{T}^+$ has constant scalar curvature. This is necessary and sufficient condition for vertical curves to be chains. Furthermore, on $\hat{\mathscr{N}}$ we have a free S^1 action and no spiralling occurs.

The final result in this paper is the following.

Theorem 3.2: The function K defining the null twistors satisfies the complex Monge–Ampere equations.

Proof: Define a new function R on $U \times \mathbb{C}$, where U is an open set in $\mathscr{P}\mathscr{T}^+$ with $U\text{n}\ \mathscr{P}\mathscr{N} \neq \varnothing$, by the equation

$$R = (z^0 \bar{z}^0)^p K.$$

Here $z^0 \in \mathbb{C}$ and p is a positive constant. One can then define a Kähler metric on a circle bundle over $\mathscr{P}\mathscr{N}$ using R as the Kähler potential. By a computation exactly analogous to that of Theorem (2.1) we find that the Ricci tensor of this metric vanishes. By the results of Ref. (10b) the assertion of the theorem follows.

CONCLUSIONS

The results in this paper show that asymptotic twistor space possesses a rich geometrical structure which up to now has been neglected. We expect to explore this structure further by applying the ideas to some specific self-dual metrics recently found by the author, but clearly much work remains to be done.

APPENDIX

Lemma: If $V(\zeta, \bar{\zeta})$ is a scalar function of spin weight $s = 0$ on the sphere then the first fourth-order derivatives of V with respect to ζ and $\bar{\zeta}$ can be written in terms of the edth operator according to the following equations:

$$\begin{aligned}
V_\zeta &= (2P_0)^{-1} \eth V, \\
V_{\zeta\zeta} &= (2P_0)^{-2}(\eth^2 V - 2\bar{\zeta}\eth V), \\
V_{\zeta\zeta\zeta} &= (2P_0)^{-3}(\eth^3 V - 6\bar{\zeta}\eth^2 V + 6\zeta\bar{\zeta}\eth V), \\
V_{\zeta\zeta\zeta\zeta} &= (2P_0)^{-4}(\eth^4 V - 12\bar{\zeta}\eth^3 V + 36\zeta\bar{\zeta}\eth^2 V - 24\bar{\zeta}\bar{\zeta}\bar{\zeta}\eth V), \\
V_{\bar{\zeta}} &= (2P_0)^{-1} \bar{\eth} V, \\
V_{\zeta\bar{\zeta}} &= (2P_0)^{-2} \eth\bar{\eth} V, \\
V_{\zeta\zeta\bar{\zeta}} &= (2P_0)^{-3}(\eth^2\bar{\eth} V - 2\bar{\zeta}\eth\bar{\eth} V), \\
V_{\zeta\zeta\zeta\bar{\zeta}} &= (2P_0)^{-4}(\eth^3\bar{\eth} V - 4\bar{\zeta}\eth^2\bar{\eth} V + 3\bar{\zeta}\bar{\zeta}\eth\bar{\eth} V), \\
V_{\bar{\zeta}\bar{\zeta}} &= (2P_0)^{-2}(\bar{\eth}^2 V - 2\zeta\bar{\eth} V), \\
V_{\zeta\zeta\zeta} &= (2P_0)^{-3}(\eth\bar{\eth}^2 V - 2\zeta\eth\bar{\eth} V - 2\bar{\eth} V), \\
V_{\zeta\zeta\zeta\zeta} &= (2P_0)^{-4}(\eth^2\bar{\eth}^2 V - 2\zeta\eth^2\bar{\eth} V - 4\eth\bar{\eth} V \\
&\quad + 4\bar{\zeta}\,\bar{\eth} V - 2\bar{\zeta}\eth\bar{\eth}^2 V + 4\zeta\bar{\zeta}\eth\bar{\eth} V).
\end{aligned}$$

Remark: The formulas above follow by recursive use of the definition of the edth operator. The operator $\bar{\eth}$ is defined exactly as in (1.2) but replacing $\bar{\zeta}$ by $\bar{\zeta}$. We note that the rest of the formulas for derivatives of order $\leqslant 4$ are obtained from the above by complex conjugation. It is in fact possible to write a general formula for derivatives of any order but it is not simple and we do not need it here.

Proof of Theorem 2.1: Using Eq. (2.15) we compute the Ricci tensor. For the $R_{A\bar{B}}$ components we have

$$\begin{aligned}
R_{A\bar{B}} &= 4\partial_A \partial_{\bar{B}} \ln V - \partial_A \partial_{\bar{B}} \ln \rho - \partial_A \partial_{\bar{B}} \ln \rho \\
&= -4 V_A V_{\bar{B}} / V^2 \\
&= 2 g_{A\bar{B}}.
\end{aligned} \tag{A1}$$

Next we compute the $R_{A\bar{2}}$ component which is given by

$$R_{A\bar{2}} = -4\partial_A \partial_{\bar{\zeta}} \ln V - \partial_A \partial_{\bar{\zeta}} \ln \rho - \partial_A \partial_{\zeta} \ln \bar{\rho}. \tag{A2}$$

Using the previous lemma and the facts that $\bar{\eth}^2 X = \bar{\sigma}^0$ and $\eth^2 \bar{X} = \sigma^0$, we find that the second term in the right-hand side of the last equation becomes

$$\begin{aligned}
&\rho^{-2}\epsilon^{CD}\epsilon^{DF}\{X_E X_{F\bar{\zeta}}(X_{CA} X_{D\bar{\zeta}\bar{\zeta}} + X_C X_{AD\bar{\zeta}\bar{\zeta}}) \\
&\quad - X_E X_{F\zeta\bar{\zeta}}(X_{CA} X_{D\bar{\zeta}} + X_C X_{DA\bar{\zeta}})\} \\
&= \rho^{-2}\epsilon^{CD}\epsilon^{EF}(2P_0)^{-4}\{X_E \bar{\eth} X_F(\bar{\eth} X_D \bar{\eth} X_{CA} \\
&\quad + \bar{\eth} X_C \bar{\eth} X_{DA} + X_{CA}(\bar{\sigma}^0 X_D - 2\zeta\bar{\eth} X_D) \\
&\quad + X_C[\bar{\sigma}^0 X_{DA} - 2\zeta\bar{\eth} X_{DA})] - X_E(\bar{\sigma}^0 X_F - 2\zeta\bar{\eth} X_F) \\
&\quad \times (X_{CA}\bar{\eth} X_D + X_C \bar{\eth} X_{DA})\} = 0.
\end{aligned}$$

In the computation above we have used the fact that the contraction of ϵ^{CD} with a symmetric spinor F_{CD} is identically equal to zero. The last term of (A2) also vanishes by virtue of the remarks following Eq. (2.10). Hence, the only surviving term is

$$R_{A\bar{2}} = 2g_{A\bar{2}}.$$

Computation of the $R_{2\bar{2}}$ component is a bit nastier since it involves third-order derivatives with respect to ζ and $\bar{\zeta}$,

$$R_{2\bar{2}} = 4\partial_\zeta \partial_{\bar{\zeta}} \log V - \partial_\zeta \partial_{\bar{\zeta}} \log \rho - \partial_\zeta \partial_{\bar{\zeta}} \log \rho. \qquad \text{(A3)}$$

By further use of the lemma, we find some nice cancellations take place and the second term in the right hand side of the equation above reduces to

$$\partial_\zeta \partial_{\bar{\zeta}} \ln \rho = \rho^{-2} \epsilon^{CD} \epsilon^{EF} (2P_0)^{-4} X_E \bar{\eth} X_F (\sigma^0 X_D \eth X_C + X_C \bar{\eth}^2 \eth X_D).$$

Applying formula (1.4) to the quantity $\bar{\eth} X_D$ we get

$$\bar{\eth}^2 \eth X_D = \eth \bar{\eth}^2 X_D - 2\bar{\eth} Z_D.$$

Substituting this into the last equation and using (1.2) yields

$$\partial_\zeta \partial_{\bar{\zeta}} \ln \rho = -2\rho^{-2} (2P_0)^{-4} (\epsilon^{CD} X_C \bar{\eth} X_D)(\epsilon^{EF} X_K \bar{\eth} X_L) = -1/2P_0.$$

Since the last term in (A3) is the complex conjugate of the preceding one, we see that

$$R_{2\bar{2}} = 4\partial_\zeta \partial_{\bar{\zeta}} \ln V + 1/P_0^2 = 2g_{22}.$$

Thus we have shown that

$$R_{\alpha\bar{\beta}} = 2g_{\alpha\bar{\beta}} \therefore R = 6,$$

concluding the proof of the theorem.

ACKNOWLEDGMENTS

I would like to express my gratitude to R. O. Hansen for his unfailing guidance and support, to D. Lerner for many helpful remarks and to Judy LaFollette for typing the manuscript.

[1]C. P. Boyer, J. D. Finley, and J. F. Plebański, "Complex general relativity, $\mathscr{H}$ and $\mathscr{H}\mathscr{H}$-spaces," *Einstein Memorial Volume* (Plenum, New York, 1979).

[2]R. O. Hansen, E. T. Newman, R. Penrose, and K. P. Tod, "The metric and curvature properties of $\mathscr{H}$-space," Proc. R. Soc. London Ser. A **363**, 445–68 (1978).

[3]R. Penrose, "Non-linear gravitons and curved twistor theory," Gen. Relativ. Gravit. 7, 35–52 (1976).

[4]R. S. Ward, Ph. D. Thesis, Oxford, 1977.

[5]M. Ko, E. T. Newman, and R. Penrose, "The Kähler structure of asymptotic twistor space," J. Math. Phys. **18**, 58–64 (1977).

[6]S. Kobayashi and K. Nomitzu, *Differential Geometry, Vol. II* (Interscience, New York, 1963).

[7]S. S. Chern and J. K. Moser, "Real hypersurfaces in complex manifolds," Acta Mat. **133**, 219–70 (1974).

[8]C. L. Fefferman, "Mange–Ampere equations, the Bergmann kernal and geometry of pseudoconvex domains," Ann. Math. **103**, 395–416 (1976).

[9]D. Burns and S. Shnider, "Real hypersurfaces in complex manifolds," Proc. Symp. pure. Math. **30**, 141–167 (1977).

[10](a)S. Webster, "On the pseudoconformal geometry of Kähler manifolds," Math. Z. **157**, 265–70 (1977). (b) S. Webster, "Kahler metrics associated with real hypersurfaces." Comm. Math. Helv. **52**, 235–50 (1977).

[11]P. Tod, Twistor Newletter 9, Oxford preprint, 1979.

[12]See. Ref. 7. The Lie algebra $\mathfrak{g}$ of the group SU(2,2) admits a five-step grading $\mathfrak{g} = \mathfrak{g}_{-2} + \mathfrak{g}_{-1} + \mathfrak{g}_0 + \mathfrak{g}_{-1} + \mathfrak{g}_2$. What we mean by π_{-1} and π_1 here are the parts of the connection π valued in $\mathfrak{g}_{-1}$ and $\mathfrak{g}_1$, respectively.

Timelike infinity

S. Persides
Astronomy Department, University of Thessaloniki, Thessaloniki, Greece

(Received 18 June 1980; accepted for publication 12 December 1980)

A new formulation is presented for analyzing the structure of a space–time at timelike infinity. An asymptotically simple space–time is defined as a space–time $(\mathscr{M}, \mathbf{g})$ which can be imbedded in a space $(\hat{\mathscr{M}}, \hat{\mathbf{g}})$ with boundary $\mathscr{T}$, a C^∞ metric $\hat{\mathbf{g}}$ and a C^∞ scalar field Ω, such that $\Omega = 0$ on $\mathscr{T}$, $\Omega > 0$ on $\hat{\mathscr{M}} - \mathscr{T}$ and $\hat{g}^{\mu\nu} - \hat{g}^{\mu\lambda}\hat{g}^{\nu\rho}\Omega_{|\lambda}\Omega_{|\rho} = \Omega^{-2}g^{\mu\nu} - \Omega^{-4}g^{\mu\lambda}g^{\nu\rho}\Omega_{;\lambda}\Omega_{;\rho}$ in a neighborhood of $\mathscr{T}$. Demanding that $\mathscr{T} = \mathscr{T}^- \cup \mathscr{T}^+$, where each one of $\mathscr{T}^-$ and $\mathscr{T}^+$ is isometric to the unit spacelike hyperboloid, and $\hat{g}^{\mu\nu}\Omega_{|\mu}\Omega_{|\nu} = \Omega^{-4}g^{\mu\nu}\Omega_{;\mu}\Omega_{;\nu} = 1$ on $\mathscr{T}$, we have an almost asymptotically flat (at timelike infinity) space–time. The group of asymptotic symmetries of $(\mathscr{M}, \mathbf{g})$ at timelike infinity is found to be isomorphic to the Lorentz group. Some properties of the space–time near $\mathscr{T}$ are shown.

PACS numbers: 04.20.Cv

1. INTRODUCTION

In the framework of general relativity and other similar metric theories of gravity it is believed that a bounded source, e.g., a binary neutron star or a dust cloud, generates a curved space–time which at large distances from the source becomes in some sense more and more Minkowskian. Although a mathematically rigorous study of the near zone is extremely difficult, substantial progress has been made in the study of the asymptotic region.

To study null infinity (i.e., the asymptotic region reached when traveling infinite affine distances along null geodesics) Penrose[1] has introduced the idea of conformal completion of the space–time. Using Penrose's technique we can attach a three-dimensional null boundary $\mathscr{I}$ to the space–time, define a C^∞ four-metric in a neighborhood of $\mathscr{I}$ and use ordinary local differential geometry as in any other regular region of the space–time. Conformal mapping, however, does not give satisfactory results at spatial and timelike infinities, that is, for the asymptotic regions at infinite spacelike and timelike distances respectively. In fact conformal mapping "shrinks" spatial and timelike infinites too much, so that they become the single points i^0 (spatial infinity), i^- (past timelike infinity), and i^+ (future timelike infinity). Several studies[2,3] along these lines have shown some awkward and undesired features for the space–time at i^0, e.g. the unphysical metric is only $C^{>0}$ at i^0. For timelike infinity a new complication arises, since the source itself reaches timelike infinity (after infinite time), while it does not reach null or spatial infinity. Thus the structure attributed to i^- and i^+ by the conformal mapping technique, although not investigated yet, is expected to be much more complicated than that of i^0.

An alternative approach has been presented for spatial infinity[4] and for timelike infinity.[5,6] The central idea is to attach three-dimensional boundaries to the space–time (one at spatial infinity and two, past and future, at timelike infinity) and define only projective structure near each boundary. Thus in the projective completion approach we avoid the definition of a four-metric smooth on the boundary. However, projective structure seems to be inadequate when we are dealing with physical questions. Thus, e.g., the conformal completion, although resulting in an awkward structure at i^0, is more effective in the study of the physical fields than the elegant projective completion.

To overcome these difficulties a new approach has been proposed[7,8] to define asymptotic flatness at spatial infinity. In this formulation a three-dimensional boundary $\mathscr{S}$ (a unit timelike hyperboloid) is attached to the physical space–time $(\mathscr{M}, \mathbf{g})$ and a C^∞ four-metric $\hat{\mathbf{g}}$ is defined on the extended manifold $\hat{\mathscr{M}} = \mathscr{M} \cup \mathscr{S}$. Working along the same lines we propose in this paper a similar formulation for past and future timelike infinity. The basic requirement is that the space–time admit a natural timelike boundary $\mathscr{T}$ consisting of two separate pieces $\mathscr{T}^-$ (past timelike infinity) and $\mathscr{T}^+$ (future timelike infinity). The term "natural boundary" associates with $\mathscr{T}$ the following three properties: (a) $\mathscr{T}$ is three-dimensional; (b) the unphysical metric $\hat{\mathbf{g}}$ is C^∞ on a neighborhood $\hat{U}$ of $\mathscr{T}$; (c) on $\hat{U}$ the unphysical metric $\hat{\mathbf{g}}$ is determined uniquely from the physical metric $\mathbf{g}$ (and a scalar field Ω) and vice versa. In building up the structure near $\mathscr{T}$ we follow a step-by-step process, imposing at each step only the requirements which are necessary. Thus we define first asymptotic simplicity at timelike infinity. In the second step we specify the intrinsic structure of $\mathscr{T}$. In the third step we describe how $\mathscr{T}$ is attached to the space–time. Finally, in the fourth step we examine the physical fields and determine other conditions which will give a rich and physically interesting space–time.

In this paper we examine only the geometrical fields which provide the background geometry and some physical fields which do not affect this geometry. Since a world tube, which includes the bounded source, reaches $\mathscr{T}$ eventually, it is expected that for the most interesting space–times (e.g., a binary star, a Schwarzschild or a Kerr black hole) the physical fields will affect (at least on some part of $\mathscr{T}$) the background geometry. This is an important difference from the corresponding case of spatial infinity. A future detailed investigation of the structure at the points where the source touches $\mathscr{T}$ is expected to give a classification of asymptotically flat space–times and a better understanding of the Cauchy problem.

In the Eardely–Sachs formulation[5] a space–time $(\mathscr{M},g)$ has a C^k *regular future projective infinity* τ iff there exists a C^k Hausdorff manifold $\overline{\mathscr{M}} = \mathscr{M} \cup \tau$ with boundary τ and a C^{k-2} symmetric connection $\overline{\Gamma}$ such that $(\mathscr{M},g)$ and $(\overline{\mathscr{M}},\overline{\Gamma})$ have the same geodesics and each timelike geodesic of $\overline{\mathscr{M}}$ can be extended to intersect τ. It is obvious that such a space–time may or may not be asymptotically flat, whatever "asymptotic flatness" means. We can add some conditions which will ensure similarity of τ with the future timelike boundary of Minkowski's space–time. Such an approach, however, will not give a metric structure to $\overline{\mathscr{M}}$. On the other hand an almost asymptotically flat (at timelike infinity) space–time, as it will be defined in Sec. 3, does not have, in general, a regular future projective infinity, unless an additional condition is fulfilled (Sec. 5).

In Sec. 2 we define the concept of asymptotic simplicity at timelike infinity. In Sec. 3 we give the conditions which define an almost asymptotically flat space–time. In Sec. 4 we examine the group of asymptotic symmetries. Some properties of almost asymptotically flat space–times are presented in Sec. 5. In our notation Greek indices λ,μ,ν, etc., take values 0,1,2,3 while Latin indices i,j,k, etc., take values 0,2,3. Covariant derivatives with respect to the physical metric are denoted by ∇_μ, with respect to the conformal metric by $\tilde{\nabla}_\mu$ or a semicolon, and with respect to the unphysical metric by $\hat{\nabla}_\mu$ or a vertical rule. If $\Omega^{-n}\Psi$ admits a smooth extension to $\mathscr{T}$, we write $\Psi = O_n$. Finally, the symbol $\hat{=}$ denotes a relation which holds on $\mathscr{T}$ only.

2. ASYMPTOTIC SIMPLICITY AT TIMELIKE INFINITY

To determine the relation connecting g and $\hat{g}$ we start from the Minkowski metric in coordinates t,r,θ,φ which is $\mathrm{diag}[1, -1, -r^2, -r^2\sin^2\theta]$ (we take $c = 1$). To study timelike infinity we set $\gamma = (t^2 - r^2)^{1/2}$ with $r < |t|$ and $t = \epsilon\gamma\cosh\chi$ $r = \gamma\sinh\chi$ ($\epsilon = 1$ and $\epsilon = -1$ respectively for future and past timelike infinity). Thus we have a new set of coordinates $\chi,\gamma,\theta,\varphi$. Since timelike infinity is reached when $\gamma \to +\infty$ with χ,θ,φ constant, we set $\omega = \gamma^{-1}$. In coordinates $\chi,\omega,\theta,\varphi$ the *physical metric* of Minkowski's space–time is

$$h_{\mu\nu} = \mathrm{diag}[-\omega^{-2}, \omega^{-4}, -\sinh^2\chi\omega^{-2}, -\sinh^2\chi\sin^2\theta\omega^{-2}]. \tag{1}$$

Let now Ω be a scalar field such that in coordinates $\chi,\omega,\theta,\varphi$ we have $\Omega = \omega$. We define the conformal metric

$$\tilde{h}_{\mu\nu} = \Omega^2 h_{\mu\nu} = \mathrm{diag}[-1, \omega^{-2}, -\sinh^2\chi, -\sinh^2\chi\sin^2\theta], \tag{2}$$

with contravariant form

$$\tilde{h}^{\mu\nu} = \mathrm{diag}[-1, \omega^2, -\sinh^{-2}\chi, -\sinh^{-2}\chi\sin^{-2}\theta]. \tag{3}$$

Obviously $\tilde{h}_{\mu\nu}$ is singular at $\omega = 0$ because $\tilde{h}_{11} = \omega^{-2}$ or because $\tilde{h}^{11} = \omega^2$. To eliminate this singular behavior we define a new metric $\hat{h}_{\mu\nu}$ whose contravariant components satisfy an equation of the form

$$\hat{h}^{\mu\nu} - \hat{h}^{\mu\lambda}\hat{h}^{\nu\rho}\Omega_{|\lambda}\Omega_{|\rho} = \tilde{h}^{\mu\nu} - \tilde{h}^{\mu\lambda}\tilde{h}^{\nu\rho}\Omega_{;\lambda}\Omega_{;\rho}. \tag{4}$$

Note that in our coordinate system $\Omega_{;\lambda} = \Omega_{|\lambda} = \delta^1_\lambda$ and (4) changes $\hat{h}^{11}$ only (it makes it different than zero) without affecting the other components near the hypersurface $\Omega = 0$.

We state now the definition of asymptotic symplicity.

Definition: A space–time $(\mathscr{M},g)$ is *asymptotically simple at timelike infinity* iff there exist (a) a space $(\hat{\mathscr{M}},\hat{g})$ with a nonempty boundary $\mathscr{B}$ $(\mathscr{B} \subset \hat{\mathscr{M}})$ and a C^∞ metric $\hat{g}$ on some open neighborhood $\hat{U}$ *of* $\mathscr{B}$ $(\mathscr{B} \subset \hat{U})$, (b) a diffeomorphism $f: U \to \hat{U} - \mathscr{B}$ from an open subset U of $\mathscr{M}$ to $\hat{U} - \mathscr{B}$, and (c) a C^∞ scalar field Ω on $\hat{U}$, positive on $\hat{U} - \mathscr{B}$ and zero on $\mathscr{B}$, such that on $\hat{U}$

$$\hat{g}^{\mu\nu} - \hat{g}^{\mu\lambda}\hat{g}^{\nu\rho}\Omega_{|\lambda}\Omega_{|\rho} = \Omega^{-2}g^{\mu\nu} - \Omega^{-4}g^{\mu\lambda}g^{\nu\rho}\Omega_{;\lambda}\Omega_{;\rho}. \tag{5}$$

Equation (5) is the basic relation which essentially determines the unphysical metric tensor $\hat{g}$ from the physical metric tensor g and the scalar field Ω. It should be emphasized that the quadratic terms in both sides of (5) have a minus sign in front of them, contrary to the case of spatial infintiy.

A consequence of the above definition is that an asymptoticaly simple space–time accepts a boundary which satisfies condition (b) of Sec. 1 for being a natural boundary. Furthermore we can prove now the following theorem.

Theorem 1: For an asymptotically simple (at timelike infinity) space–time (a) if on some part $\mathscr{N}$ of $\mathscr{B}$

$$\Omega^{-2}\tilde{g}^{\mu\nu}\Omega_{;\mu}\Omega_{;\nu} \hat{=} -1 \quad \text{and} \quad \hat{g}^{\mu\nu}\Omega_{|\mu}\Omega_{|\nu} \hat{=} 0, \tag{6}$$

then on some open neighborhood on $\mathscr{N}$ we have

$$\hat{g}^{\mu\nu} = \tilde{g}^{\mu\nu}, \tag{7}$$

while (b) if on some part $\mathscr{T}$ of $\mathscr{B}$

$$\Omega^{-2}\tilde{g}^{\mu\nu}\Omega_{;\mu}\Omega_{;\nu} \hat{=} 1 \quad \text{and} \quad \hat{g}^{\mu\nu}\Omega_{|\mu}\Omega_{|\nu} \hat{=} 1, \tag{8}$$

then on some open neighborhood of $\mathscr{T}$ we have in a coordinate system x^μ with $\Omega = x^1$ ($i,j = 0,2,3$)

$$\hat{g}^{11} = 1 - \tilde{g}^{11}, \tag{9}$$

$$\hat{g}^{i1} = \tilde{g}^{i1}(-1 + 1/\hat{g}^{11}), \tag{10}$$

$$\hat{g}^{ij} = \tilde{g}^{ij} + \tilde{g}^{i1}\tilde{g}^{j1}[-1 + (1 - 1/\tilde{g}^{11})^2]. \tag{11}$$

Proof: It is enough to show Eqs. (7) and (9)–(11) in a coordinate system x^μ with $\Omega = x^1$. In such a system (5) gives

$$\hat{g}^{\mu\nu} - \hat{g}^{1\mu}\hat{g}^{1\nu} = \tilde{g}^{\mu\nu} - \tilde{g}^{1\mu}\tilde{g}^{1\nu}. \tag{12}$$

If (6) hold, then $\tilde{g}^{11} \hat{=} 0$, $\hat{g}^{11} \hat{=} 0$, and (12) gives Eq. (7). If (8) hold, then $\tilde{g}^{11} \hat{=} 0$, $\hat{g}^{11} \hat{=} 1$, and (12) gives Eqs. (9)–(11).

This theorem indicates an unexpected relation of asymptotic simplicity at null infinity and at timelike infinity. This property and the corresponding in the case of spatial infinity[8] will be used in a future paper to unify asymptotic simplicity and flatness at timelike, null, and spatial infinity. Also from the previous theorem we conclude that if we impose conditions (8) then our asymptotically simple space–time admits a boundary which fulfills conditions (b) and (c) of Sec. 1. For Minkowski's space–time the unphysical metric can be obtained easily from (3) and (9)–(11). We have in coordinates $\chi,\omega,\theta,\varphi$

$$\hat{h}^{\mu\nu} = \mathrm{diag}[-1, 1 - \omega^2, -\sinh^{-2}\chi, -\sinh^{-2}\chi\sin^{-2}\theta], \tag{13}$$

$$\hat{h}_{\mu\nu} = \mathrm{diag}[-1, (1 - \omega^2)^{-1}, -\sinh^2\chi, -\sinh^2\chi\sin^2\theta]. \tag{14}$$

Thus $\hat{h}_{\mu\nu}$ and $\hat{h}^{\mu\nu}$ are C^∞ on a neighborhood of the hypersurface $\omega = 0$, which is the timelike boundary of $\hat{\mathscr{M}}$. This hypersurface has two separate identical parts $\mathscr{T}^-$ and $\mathscr{T}^+$ ($\mathscr{T} = \mathscr{T}^- \cup \mathscr{T}^+$) corresponding, respectively, to $t \to -\infty$ and $t \to +\infty$ (a separate coordinate system $\chi,\omega,\theta,\varphi$ is needed to cover a neighborhood of each $\mathscr{T}^-$ and $\mathscr{T}^+$, while $\omega > 0$ on $\mathscr{M}$). The induced metric on each part of $\mathscr{T}$, i.e., on $\mathscr{T}^-$ or $\mathscr{T}^+$ is

$$\gamma_{ij} = \mathrm{diag}[-1, -\sinh^2\chi, -\sin^2\chi^2\sin^2\theta], \tag{15}$$

namely the metric of a three-dimensional unit spacelike hyperboloid. The coordinates χ,θ,φ take always values in the intervals $-\infty < \chi < +\infty, 0 \leqslant \theta \leqslant \pi, 0 \leqslant \varphi \leqslant 2\pi$ ($\varphi = 0$ is identified with $\varphi = 2\pi$).

3. ASYMPTOTIC FLATNESS AT TIMELIKE INFINITY

The space–time generated by a bounded source can be regarded as a space–time for which the energy-momentum tensor vanishes outside a world tube which extends from $\mathscr{T}^-$ to $\mathscr{T}^+$. Consequently, asymptotic flatness at timelike infinity is expected to be a much more delicate and complicated concept than it is at null and spatial infinty. There is, however, an encouraging element in the definitions of asymptotic flatness at null[9] and spatial[8] infinity. These definitions do not assume that the energy-momentum tensor is zero near the null or the spatial boundary. Thus we can proceed and define the concept of almost asymptotically flat space–time (AAFS) at timelike infinity as a space–time in which the energy-momentum tensor and the curvature are *not* strong enough to destroy the background geometry at $\mathscr{T}^-$ or $\mathscr{T}^+$.

Working along these lines we require for the space–time to be asymptotically simple and a part of its boundary identical to the timelike boundary of Minkowski's space–time. Some additional conditions are needed to specify how the boundary is fastened to the interior of the space–time. These conditions, and a way to define asymptotic flatness in terms of tensor relations as well as in terms of the existence of a special class of coordinate systems, are indicated by the following theorem.

Theorem 2: For an asymptotically simple space–time with boundary $\mathscr{B}$, $\mathscr{T}'$ a part of $\mathscr{B}$ and $\hat{U}$ an open neighborhood of $\mathscr{T}'$ the following conditions are equivalent.

(a) $\mathscr{T}'$ is isometric to the unit spacelike hyperboloid and on $\mathscr{T}'$ the conditions (8) hold.

(b) There exists a coordinate system $(\chi,\omega,\theta,\varphi)$ on $\hat{U}$ in which we have $\Omega = \omega$, $\tilde{g}^{11} = \omega^2 + O_3$, and

$$\hat{g}_{\mu\nu} = \begin{bmatrix} -1+O_1 & \alpha+O_1 & O_1 & O_1 \\ \alpha+O_1 & \beta+O_1 & \gamma+O_1 & \delta+O_1 \\ O_1 & \gamma+O_1 & -\sinh^2\chi+O_1 & O_1 \\ O_1 & \delta+O_1 & O_1 & -\sinh^2\chi\sin^2\theta+O_1 \end{bmatrix}. \tag{16}$$

with α,γ,δ arbitrary functions of χ,θ,φ and

$$\beta = 1 - \alpha^2 - \gamma^2\sinh^{-2}\chi - \delta^2\sinh^{-2}\chi\sin^{-2}\theta. \tag{17}$$

(c) There exists a coordinate system $(\chi,\omega,\theta,\varphi)$ on $\hat{U}$ in which we have $\Omega = \omega$, $\tilde{g}^{11} = \omega^2 + O_3$, and ($\alpha,\beta,\gamma,\delta$ as before)

$$\hat{g}^{\mu\nu} = \begin{bmatrix} -1+\alpha^2+O_1 & \alpha+O_1 & \alpha\gamma\sinh^{-2}\chi+O_1 & \alpha\delta\sinh^{-2}\chi\sin^{-2}\theta+O_1 \\ \alpha+O_1 & 1+O_1 & \gamma\sinh^{-2}\chi+O_1 & \delta\sinh^{-2}\chi\sin^{-2}\theta+O_1 \\ \alpha\gamma\sinh^{-2}\chi+O_1 & \gamma\sinh^{-2}\chi+O_1 & (-1+\gamma^2\sinh^{-2}\chi)\sinh^{-2}\chi+O_1 & \gamma\delta\sinh^{-4}\chi\sin^{-2}\theta+O_1 \\ \alpha\delta\sinh^{-2}\chi\sin^{-2}\theta+O_1 & \delta\sinh^{-2}\chi\sin^{-2}\theta+O_1 & \gamma\delta\sinh^{-4}\sin^{-2}\theta+O_1 & (-1+\delta^2\sinh^{-2}\chi\sin^{-2}\theta)\sinh^{-2}\chi\sin^{-2}\theta+O_1 \end{bmatrix}. \tag{18}$$

(d) There exists a coordinate system $(\chi,\omega,\theta,\varphi)$ on $\hat{U}$ in which we have $\Omega = \omega$, $\hat{g}^{11} = 1 + O_1$ and

$$\tilde{g}_{\mu\nu} = \begin{bmatrix} -1+O_1 & \alpha+O_1 & O_1 & O_1 \\ \alpha+O_1 & \omega^{-2}+O_1 & \gamma+O_1 & \delta+O_1 \\ O_1 & \gamma+O_1 & -\sinh^2\chi+O_1 & O_1 \\ O_1 & \delta+O_1 & O_1 & -\sinh^2\chi\sin^2\theta+O_1 \end{bmatrix}. \tag{19}$$

(e) There exists a coordinate system $(\chi,\omega,\theta,\varphi)$ on $\hat{U}$ in which we have $\Omega = \omega$, $\hat{g}^{11} = 1 + O_1$, and

$$\tilde{g}^{\mu\nu} = \begin{bmatrix} -1+O_1 & \alpha\omega^2+O_3 & O_1 & O_1 \\ \alpha\omega^2+O_3 & \omega^2+O_3 & \gamma\sinh^{-2}\chi\omega^2+O_3 & \delta\sinh^{-2}\chi\sin^{-2}\theta\omega^2+O_3 \\ O_1 & \gamma\sinh^{-2}\chi\omega^2+O_3 & -\sinh^{-2}\chi+O_1 & O_1 \\ O_1 & \delta\sinh^{-2}\chi\sin^{-2}\theta\omega^2+O_3 & O_1 & -\sinh^{-2}\chi\sin^{-2}\theta+O_1 \end{bmatrix}. \tag{20}$$

We will omit the proof since it is similar to the corresponding case at spatial infinity.[8] However, the formulas (16)–(20) are different from the corresponding cases at spatial infinity. Another difference is that a space–time which is asymptotically similar to Minkowski's space–time is expected to have two separate pieces of its boundary, each one similar to $\mathscr{T}'$. This is incorporated in the following definition.

Definition: A space–time is *almost asymptotically flat* (or *Minkowskian*) *at timelike infinity* iff it is asymptotically simple and for each one of two separate parts $\mathscr{T}^-$ and $\mathscr{T}^+$ of its boundary $\mathscr{B}$ one of the conditions (a)–(e) of Theorem 2 is satisfied.

Obviously an AAFS admits a natural boundary $\mathscr{T} = \mathscr{T}^- \cup \mathscr{T}^+$ in the sense of Sec. 1. Near $\mathscr{T}^-$ and $\mathscr{T}^+$ we can give a general expression of the physical metric using the relation $g_{\mu\nu} = \Omega^{-2}\tilde{g}_{\mu\nu}$. From (19) we have

$$g_{\mu\nu} = \begin{bmatrix} -\omega^{-2} + O_{-1} & O_{-2} & O_{-1} & O_{-1} \\ O_{-2} & \omega^{-4} + O_{-3} & O_{-2} & O_{-2} \\ O_{-1} & O_{-2} & -\sinh^2\chi\omega^{-2} + O_{-1} & O_{-1} \\ O_{-1} & O_{-2} & O_{-1} & -\sinh^2\chi\sin^2\theta\omega^{-2} + O_{-1} \end{bmatrix}. \quad (21)$$

This expression is very useful for practical calculations near the boundary, something which is not possible in general in the conformal or projective completion. The explicitly given terms represent the "geometrical" part of $g_{\mu\nu}$, while the "physical" part is hidden in the terms O_n.

4. ASYMPTOTIC SYMMETRIES

Since $\mathscr{T}$ is a real boundary of the space–time manifold with a C^∞ four-metric on it, we can define asymptotic symmetries in three equivalent ways. The most natural definition is based on the form of the physical metric tensor near the boundary and defines the group of asymptotic symmetries as the group of coordinate transformations $(\chi,\omega,\theta,\varphi) \rightarrow (\chi',\omega',\theta',\varphi')$ which preserve the form (21). The most general transformation which maps $\omega = 0$ to $\omega' = 0$ and the region $\omega > 0$ to the region $\omega' > 0$ is

$$\chi = \chi_0 + O_1, \quad (22)$$

$$\omega = \omega_1\omega' + O_2, \quad (23)$$

$$\theta = \theta_0 + O_1, \quad (24)$$

$$\varphi = \varphi_0 + O_1, \quad (25)$$

with $\omega_1 > 0$ and $\chi_0,\theta_0,\varphi_0$ functions of χ',θ',φ'. Calculating $g'_{\mu\nu}$ and demanding that it be of the same form as (21), we have

$$\left(\frac{\partial\chi_0}{\partial\chi'}\right)^2 + \left(\frac{\partial\theta_0}{\partial\chi'}\right)^2\sinh^2\chi_0 + \left(\frac{\partial\varphi_0}{\partial\chi'}\right)^2\sinh^2\chi_0\sin^2\theta_0 = 1, \quad (26)$$

$$\frac{\partial\chi_0}{\partial\chi'}\frac{\partial\chi_0}{\partial\theta'} + \frac{\partial\theta_0}{\partial\chi'}\frac{\partial\theta_0}{\partial\theta'}\sinh^2\chi_0 + \frac{\partial\varphi_0}{\partial\chi'}\frac{\partial\varphi_0}{\partial\theta'}\sinh^2\chi_0\sin^2\theta_0 = 0, \quad (27)$$

$$\frac{\partial\chi_0}{\partial\chi'}\frac{\partial\chi_0}{\partial\varphi'} + \frac{\partial\theta_0}{\partial\chi'}\frac{\partial\theta_0}{\partial\theta'}\sinh^2\chi_0 + \frac{\partial\varphi_0}{\partial\chi'}\frac{\partial\varphi_0}{\partial\varphi'}\sinh^2\chi_0\sin^2\theta_0 = 0, \quad (28)$$

$$\left(\frac{\partial\chi_0}{\partial\theta'}\right)^2 + \left(\frac{\partial\theta_0}{\partial\theta'}\right)^2\sinh^2\chi_0 + \left(\frac{\partial\varphi_0}{\partial\theta'}\right)^2\sinh^2\chi_0\sin^2\theta_0 = \sinh^2\chi', \quad (29)$$

$$\left(\frac{\partial\chi_0}{\partial\varphi'}\right)^2 + \left(\frac{\partial\theta_0}{\partial\varphi'}\right)^2\sinh^2\chi_0 + \left(\frac{\partial\varphi_0}{\partial\varphi'}\right)^2\sinh^2\chi_0\sin^2\theta_0 = \sinh^2\chi'\sin^2\theta', \quad (30)$$

$$\frac{\partial\chi_0}{\partial\theta'}\frac{\partial\chi_0}{\partial\varphi'} + \frac{\partial\theta_0}{\partial\theta'}\frac{\partial\theta_0}{\partial\varphi'}\sinh^2\chi_0 + \frac{\partial\varphi_0}{\partial\theta'}\frac{\partial\varphi_0}{\partial\varphi'}\sinh^2\chi_0\sin^2\theta_0 = 0. \quad (31)$$

After some calculations we obtain also the useful relation

$$\left(\frac{\partial\chi_0}{\partial\chi'}\right)^2 + \left(\frac{\partial\chi_0}{\partial\theta'}\right)^2\sinh^{-2}\chi' + \left(\frac{\partial\chi_0}{\partial\varphi'}\right)^2\sinh^{-2}\chi'\sin^{-2}\theta = 1, \quad (32)$$

and for the Jacobian of the transformation $(\theta_0, \varphi_0) \rightarrow (\theta',\varphi')$,

$$J(\theta_0, \varphi_0; \theta', \varphi') = \pm\frac{\partial\chi_0}{\partial\chi'}\frac{\sinh^2\chi'\sin\theta'}{\sinh^2\chi_0\sin\theta_0}. \quad (33)$$

We can also define the group of asymptotic symmetries as the group of the transformations $(\chi,\theta,\varphi) \rightarrow (\chi',\theta',\varphi')$ which preserve the intrinsic geometry of $\mathscr{T}^+$ (or $\mathscr{T}^-$), i.e. preserve the three-metric (15). This condition gives again equations (26)–(31) with $\chi_0,\theta_0,\varphi_0$ replaced by χ,θ,φ. Thus the two definitions are equivalent. Finally we can consider the group of asymptotic symmetries as the group generated by the asymptotic Killing vector fields ξ^μ of the unphysical space–time $(\hat{\mathscr{M}},\hat{\mathbf{g}})$. As in the case of spatial infinity, this definition implies that the restriction of ξ^μ to $\mathscr{T}^-$ (or $\mathscr{T}^+$) is a Killing tensor field of the intrinsic geometry of $\mathscr{T}^-$ (or $\mathscr{T}^+$) and consequently generates the same symmetry group. It should be noted also that linearization of Eqs. (26)–(31) gives the Killing equations for the six linearly independent Killing vector fields of (15) which can be solved directly.

To identify the group of asymptotic symmetries we consider the three-dimensional hypersurface $\eta_{\mu\nu}x^\mu x^\nu = 1$ of a four-dimensional Minkowskian space-time with $\eta_{\mu\nu} = \mathrm{diag}(1, -1, -1, -1)$. With

$$\begin{aligned} x^0 &= \cosh\chi, \\ x^1 &= \sinh\chi\sin\theta\cos\varphi, \\ x^2 &= \sinh\chi\sin\theta\sin\varphi, \\ x^3 &= \sinh\chi\cos\theta, \end{aligned} \quad (34)$$

the induced metric on the hypersurface is given by (15) and the hypersurface is the unit spacelike hyperboloid. Any transformation $(\chi,\theta,\varphi)\rightarrow(\chi',\theta',\varphi')$ which preserves (15) can be written as $H^{-1}LH$ where H is the fixed (with no arbitrary parameters) transformation (34), H^{-1} its inverse, and L an arbitrary Lorentz transformation $x^\mu\rightarrow x'^\mu$. The group of asymptotic transformations $H^{-1}LH$ is the group of asymptotic symmetries and is obviously isomorphic to the Lorentz group.

5. PROPERTIES OF AAFS NEAR $\mathscr{T}$

In this section we derive some properties of almost asymptotically flat space–times at timelike infinity assuming that the conditions of the definition (Sec. 3) are satisfied at every point of $\mathscr{T}$. Besides describing AAFS, these properties indicate some of the quantities which should be examined in detail for specific bounded sources near $\mathscr{T}$.

The first property refers to the way $\mathscr{T}^+$ (and $\mathscr{T}^-$) is fastened to the space–time. Let n_μ be the unit normal on a hypersurface $\Omega=\text{const}$ (which is close to $\Omega=0$), $q_{\mu\nu}=g_{\mu\nu}-n_\mu n_\nu$ the intrinsic metric induced on the hypersurface, and $p_{\mu\nu}=q^\lambda_\mu q^\rho_\nu\nabla_\lambda n_\rho$ its extrinsic curvature. Using the form (21) of the physical metric with a few higher order terms as defined in Theorem 2, we obtain, after some straightforward calculations,

$$\Omega^2 q_{\mu\nu}=-\Omega p_{\mu\nu}+O_1=\begin{bmatrix}-1+O_1 & \alpha+O_1 & O_1 & O_1\\ \alpha+O_1 & -1+\beta+O_1 & \gamma+O_1 & \delta+O_1\\ O_1 & \gamma+O_1 & -\sinh^2\chi+O_1 & O_1\\ O_1 & \delta+O_1 & O_1 & -\sinh^2\chi\sin^2\theta+O_1\end{bmatrix}. \tag{35}$$

Hence on $\mathscr{T}$ we have

$$\Omega^2 q_{\mu\nu}\triangleq-\Omega p_{\mu\nu}. \tag{36}$$

Straightforward calculations of the Riemann, Ricci, and Weyl tensors give the same order relations as at spatial infinity, namely

$$R_{\lambda\mu\rho\nu}=O_{-1-n},\quad R^{\lambda\mu\rho\nu}=O_{7+m},\quad R_{\mu\nu}=O_{1-n}, \tag{37}$$

$$R_\mu{}^\nu=O_{3+m-n},\quad R^{\mu\nu}=O_{5+m},\quad R=O_3, \tag{38}$$

$$C_{\lambda\mu\rho\nu}=O_{-1-n},\quad C_{\lambda\mu}{}^{\rho\nu}=O_{3+m-n},\quad C^{\lambda\mu\rho\nu}=O_{7+m},\quad C^*_{\lambda\mu\rho\nu}=O_{-1-n}, \tag{39}$$

where m (n) is the number of upper (lower) indices which are equal to 1. From these relations we have for the Weyl tensor of the conformal metric $\tilde g_{\mu\nu}=\Omega^2 g_{\mu\nu}$

$$\tilde C^{\lambda\mu\rho\nu}\triangleq0,\ \Omega^{-1}C^{\lambda\mu\rho\nu}\Omega_{;\mu}\triangleq0,\ \Omega^{-2}C^{\lambda\mu\rho\nu}\Omega_{;\mu}\Omega_{;\nu}\triangleq0. \tag{40}$$

Further calculations give for the electric $E_{\mu\nu}=C_{\lambda\mu\rho\nu}n^\lambda n^\rho$ and the magnetic $B_{\mu\nu}=C^*_{\lambda\mu\rho\nu}n^\lambda n^\rho$ part of the Weyl tensor $E_{\mu\nu}=O_1$ and $B_{\mu\nu}=O_1$ near $\mathscr{T}$ or

$$E_{\mu\nu}\triangleq0\quad\text{and}\quad B_{\mu\nu}\triangleq0. \tag{41}$$

Consequently $\Omega^{-1}E_{\mu\nu}$ and $\Omega^{-1}B_{\mu\nu}$ induce on $\mathscr{T}^+$ (and $\mathscr{T}^-$) two (symmetric and trace-free) smooth tensor fields expected to have physical content.

Finally we examine the relation between these two concepts: a space–time with regular future projective infinity (PI) and an almost asymptotically flat (at future timelike infinity) space–time. It is obvious that the boundary of a PI space–time is not necessarily a spacelike hyperboloid, since some well-known cosmological models are PI space–times.[5] On the other hand, the conditions for a space–time to be almost asymptotically flat at timelike infinity are not enough to imply a regular future projective infinity for the space–time. From the Eardley–Sachs definition given in Sec. 1 the key question is whether a connection $\bar\Gamma$ can be defined on $\bar{\mathscr{M}}=\mathscr{M}\cup\mathscr{T}^+$ so that $(\bar{\mathscr{M}},\bar\Gamma)$ be projectively equivalent to $(\mathscr{M},\Gamma)$, where Γ is the metric connection on the physical space–time as determined by $\mathbf{g}$ ($\bar\Gamma$ is not necessarily a metric connection). For this to be possible the two connections should satisfy on $\mathscr{M}$ the relations

$$\begin{aligned}&\bar\Gamma^\lambda_{\mu\nu}-\tfrac{1}{5}(\bar\Gamma^\rho_{\mu\rho}\delta^\lambda_\nu+\bar\Gamma^\rho_{\nu\rho}\delta^\lambda_\mu)\\&=\Gamma^\lambda_{\mu\nu}-\tfrac{1}{5}(\Gamma^\rho_{\mu\rho}\delta^\lambda_\nu+\Gamma^\rho_{\nu\rho}\delta^\lambda_\mu).\end{aligned} \tag{42}$$

Since $\bar\Gamma$ can be smoothly extended to $\mathscr{T}^+$, so can both sides of (42). But a straightforward calculation shows that the right-hand side of (42) is smooth on $\mathscr{T}^+$, except for (λ,μ,ν) equal to (0,1,1), (2,1,1), and (3,1,1) which cases give $\frac{1}{2}\alpha_{11,2}\omega^{-1}+O_0$, $\frac{1}{2}\alpha_{11,2}\sinh^{-2}\chi\omega^{-1}+O_0$, and $\frac{1}{2}\alpha_{11,3}\times\sinh^{-2}\chi\times\sin^{-2}\theta\omega^{-1}+O_0$ respectively, where we have set $g_{11}=\omega^{-4}+\alpha_{11}\omega^{-3}+O_{-2}$. Thus we must have $\alpha_{11,0}=\alpha_{11,2}=\alpha_{11,3}=0$. This condition can be put in tensor form if we consider the scalar

$$\Phi=\Omega^{-5}g^{\lambda\rho}\Omega_{;\lambda}\Omega_{;\rho}-\Omega^{-1}. \tag{43}$$

In our coordinates $(\chi,\omega,\theta,\varphi)$ we have

$$\Phi=\omega^{-5}g^{11}-\omega^{-1}=-\alpha_{11}+O_1. \tag{44}$$

Hence an almost asymptotically flat space–time has a regular future projective infinity iff Φ is constant on $\mathscr{T}^+$ or equivalently

$$\Phi_{;\mu}-(\Phi_{;\lambda}n^\lambda)n_\mu\triangleq0, \tag{45}$$

where n_μ is the unit normal on $\mathscr{T}^+$. This condition is expected to imply some restrictions on the physical fields and at this stage it seems unwise to impose it on the space–time in a definition of asymptotic flatness.

6. GENERAL REMARKS

The definition given in Sec. 3, the theorems of Secs. 2 and 3, and the properties of almost asymptotically flat space–times presented in Secs. 4 and 5 establish a new formulation for the study of asymptotic structure at timelike infinity. Some remarks, however, should be made.

First, a definition of asymptotic flatness at timelike infinity has not yet been given. A detailed study of specific sources is expected to show essential differences in their behavior near $\mathscr{T}$ and indicate what new conditions should be added and which of the conditions for AAFS should be relaxed on some part of $\mathscr{T}$. At the same time this study is expected to indicate criteria for classifying asymptotically flat space–times depending on their behavior near $\mathscr{T}$. Some of the quantities to be examined in such a study have been already pointed out in this work. Many related questions, such as the dependence of tensor fields at points of $\mathscr{T}$ *on the direction* of approach to $\mathscr{T}$ should be examined.

Three major classes of asymptotically flat space–times are expected depending on whether the conditions of the definition of AAFS (Sec. 3) are (a) fulfilled throughout $\mathscr{T}^-$ and $\mathscr{T}^+$, (b) violated at some point (or a set of points with zero volume) of $\mathscr{T}$, and (c) violated at some set of points of $\mathscr{T}$ with nonzero volume. In case (a) (e.g., a field of pure radiation whose energy escapes eventually to $\mathscr{N}^+$) the energy-momentum tensor of the source will not be strong enough to destroy asymptotic flatness as $\mathscr{T}$. In case (b) (e.g., the space–time of a dead star or a black hole) the source will touch $\mathscr{T}$ at a point (i.e., the end of the source's world tube). The dependence of the energy-momentum tensor and other fields on the direction of approach to that point will provide criteria for subclassification.

Second, the relation of asymptotic flatness at all three infinities should be studied. The emergence of conformal mapping in Theorem 1 for that part of $\mathscr{B}$ called $\mathscr{N}$ indicates a close relation between Penrose's formulation for null infinity and the present formulation for timelike infinity. In a future paper we will relate this observation with the corresponding one made for spatial infinity[8] and present a unified formulation for the whole boundary $\mathscr{B}$.

Third, the essential question of uniqueness of the boundary should be examined. Since we are interested in the physical fields which register on $\mathscr{T}$ and in general on $\mathscr{B}$, we have to know to what extent the informations contained in the extension of the fields to $\mathscr{B}$ describe properties of the physical space–time independently of the completion.

After answering the above questions we will be able to attack new problems related to the stability of asymptotic flatness, the multipole moments, the Cauchy problem, etc. The main difference in a reformulation of the Cauchy problem is expected to arise from the fact that $\mathscr{T}^-$ and $\mathscr{N}^-$ (i.e., the hypersurfaces on which initial data will be specified) have now well-known and relatively simple intrinsic geometries, contrary to the case where an arbitrary spacelike hypersurface is used.

ACKNOWLEDGMENT

I would like to thank Dr. B. Xanthopoulos for useful discussions and comments on the manuscript.

[1]R. Penrose, Phys. Rev. Lett. **10**, 66 (1963); Proc. R. Soc. London Ser. A **284**, 159 (1965).
[2]R. Geroch, J. Math. Phys. **13**, 956 (1972).
[3]A. Ashtekar and R. O. Hansen, J. Math. Phys. **19**, 1542 (1978).
[4]P. Sommers, J. Math. Phys. **19**, 549 (1978).
[5]D. Eardley and R. K. Sachs, J. Math. Phys. **14**, 209 (1973).
[6]B. G. Schmidt, Commun. Math. Phys. **36**, 73 (1974).
[7]S. Persides, J. Math. Phys. **21**, 135 (1980).
[8]S. Persides, J. Math. Phys. **21**, 142 (1980).
[9]S. Persides, J. Math. Phys. **20**, 1731 (1979).

A unified formulation of timelike, null and spatial infinity

S. Persides
Astronomy Department, University of Thessaloniki, Thessaloniki, Greece

(Received 15 July 1980; accepted for publication 25 September 1980)

A formulation is presented for studying simultaneously the timelike, null and spatial infinities of space-times which resemble asymptotically Minkowski's space-time. For this a relation of the form $f(\mathbf{g}, \hat{\mathbf{g}}, \Omega) = 0$ is determined so that given a space-time $(\mathscr{M}, \mathbf{g})$ a space $(\hat{\mathscr{M}}, \mathbf{g})$ with boundary $\mathscr{B}$ can be found with $\mathscr{M}$ imbedded in $\hat{\mathscr{M}}$ and $\hat{\mathbf{g}}$ and Ω C^∞ fields on $\hat{\mathscr{M}}$. A space-time $(\mathscr{M}, \mathbf{g})$ for which this is possible is called (globally)asymptotically simple. Then the conditions for $\mathscr{B}$ to resemble the boundary of Minkowski's space-time are determined. Thus the concept of a (globally) almost asymptotically flat space-time is defined.

PACS numbers: 04.20.Cv

1. INTRODUCTION

In the last two decades substantial progress has been made towards understanding the asymptotic structure of space-times which resemble Minkowski's space-time at large distances from the source.[1-11] Although there are many different attitudes and approaches in the numerous papers on timelike,[9,12] null[1,2,13] and spatial[4,5,7,8,14,15] infinity, there are some common features. First, there is a tendency to try to bring the asymptotic region closer by imbedding the physical space-time in a larger manifold with boundary, this boundary representing "infinity." Second, each asymptotic region is studied separately and in most cases by different techniques.

In a recent formulation, however, of spatial infinity[15] an unexpected feature arose: a definiton of asymptotic simplicity at infinite spacelike distances gave as a case to be excluded the conformal mapping proposed by Penrose for null infinity. A similar feature appeared in a study of timelike infinity[12] by the same technique. Thus a question is raised: Is there a unified formulation of asymptotic structure possible?

This paper presents a positive answer to this question. Starting from observations made on the two independent but similar formulations of spatial and timelike infinity we give a global definition of asymptotic simplicity. Then depending on the relations satisfied by the physical and unphysical metrics on each part of the boundary we define the timelike, null and spatial infinities. Demanding that the intrinsic geometry of the whole boundary is identical to the intrinsic geometry of the boundary of Minkowski's space-time we have a global definition of asymptotic flatness. Thus a common base is established for studying the asymptotic structure of asymptotically flat space-times at timelike, null and spatial infinity.

There are perhaps three arguments supporting the usefulness of a unified formulation of asymptotic flatness. First, it makes the whole formulation more elegant and transparent and brings out the close analogy between the threatments of different asymptotic regimes. This is accomplished by giving a single definition of the global boundary which is then divided into pieces (timelike, null and spatial boundaries) depending on the particular properties of each piece. Second, such a unified formulation is expected to be very useful in relating the physical fields of one piece to the physical fields of another piece (e.g. the Bondi mass at null infinity to the ADM mass at spatial infinity). A direct link between two separate pieces of the boundary does not seem to be possible and a unified formulation is expected to provide a possibility of going from one piece of the boundary to the interior of the space-time and then to the other piece of the boundary. Third, a unified formulation emphasizies the existence of a boundary and its C^∞ structure, leaving flatness as a secondary feature. This raises the possibility of defining boundaries for space-times which do not resemble asymptotically the Minkowskian space-time (e.g. for cosmological models). Obviously, there are many problems to be solved at spatial and (particularly) timelike infinity (e.g. uniqueness of the boundary, definition of physical fields, etc.). The presentation of a unified formulation before the solution of these problems will help us to have a complete and coherent picture of all fronts and thus be able to coordinate more effectively the attacks on the several problems.

This work is based on many and lengthly results obtained in three previous papers refering to null,[13] spatial[15] and timelike[12] infinities. These results are considered known and are not repeated here. Furthermore the notation of these papers is closely followed.

2. GLOBAL ASYMPTOTIC SIMPLICITY

Asymptotic simplicity, as defined separately for null, spatial and timelike infinity, is a concept which guarantees the possibility of imbedding the physical space-time $(\mathscr{M}, \mathrm{g})$ into a space $\hat{\mathscr{M}}$ with a boundary and a C^∞ metric $\hat{\mathrm{g}}$ on $\hat{\mathscr{M}}$ (including the boundary). In each case, i.e. for timelike, null and spatial infinity, a different tensor relation relates g and $\hat{\mathrm{g}}$. For timelike[12] infinity this relation is

$$\hat{g}^{\mu\nu} - \hat{g}^{\mu\lambda}\hat{g}^{\nu\rho}\Omega_{|\lambda}\Omega_{|\rho} = \Omega^{-2}g^{\mu\nu} - \Omega^{-4}g^{\mu\lambda}g^{\nu\rho}\Omega_{;\lambda}\Omega_{;\rho} \tag{1}$$

for null[2,13] infinity

$$\hat{g}^{\mu\nu} = \Omega^{-2}g^{\mu\nu} \tag{2}$$

and for spatial[15] infinity

$$\hat{g}^{\mu\nu} + \hat{g}^{\mu\lambda}\hat{g}^{\nu\rho}\Omega_{|\lambda}\Omega_{|\rho} = \Omega^{-2}g^{\mu\nu} + \Omega^{-4}g^{\mu\lambda}g^{\nu\rho}\Omega_{;\lambda}\Omega_{;\rho}. \tag{3}$$

The first objective, in order to give a single formulation appropriate for all three infinities, is to unify the above formulas into a single formula. For this we observe that for Minkowski's space-time the component of the physical metric which diverges on the boundary too fast (in the appropriate coordinate system) is g_{11} and Eqs. (1)–(3) serve exactly this purpose, i.e. they replace g^{11} by $\hat{g}^{11}$ so that its inverse $\hat{g}_{11}$ does not diverge on the boundary. This is accomplished by the relations ($\tilde{g}^{\mu\nu} \equiv \Omega^{-2} g^{\mu\nu}$)

$$\hat{g}^{11} = 1 - \tilde{g}^{11} \quad \text{(timelike infinity)} \tag{4}$$

$$\hat{g}^{11} = \tilde{g}^{11} \quad \text{(null infinity)} \tag{5}$$

$$\hat{g}^{11} = -1 - \tilde{g}^{11} \quad \text{(spatial infinity)} \tag{6}$$

which are obtained from the tensor formulas (1)–(3) for $\mu = \nu = 1$ in appropriate coordinates. Equations (4)–(6) are algebraic equations of first degree with respect to $\hat{g}^{11}$ and $\tilde{g}^{11}$ [while (1)–(3) are quadratic]. The lowest degree algebraic equation which contains (4)–(6) as special cases is

$$(\hat{g}^{11} - \tilde{g}^{11})(\hat{g}^{11} + \tilde{g}^{11} + 1)(\hat{g}^{11} + \tilde{g}^{11} - 1) = 0 \tag{7}$$

or

$$\hat{g}^{11} - \tilde{g}^{11} - (\hat{g}^{11})^3 - \tilde{g}^{11}(\hat{g}^{11})^2 + \hat{g}^{11}(\tilde{g}^{11})^2 + (\tilde{g}^{11})^3 = 0 . \tag{8}$$

This equation should be obtained from a tensor relation in coordinates x^μ where $\Omega = x^1$ is the scalar field. From (8) we have the tensor relation

$$\hat{g}^{\mu\nu} - \tilde{g}^{\mu\nu} - (\hat{g}^{\sigma\tau}\Omega_{|\sigma}\Omega_{|\tau})\hat{g}^{\mu\lambda}\hat{g}^{\nu\rho}\Omega_{|\lambda}\Omega_{|\rho} - (\tilde{g}^{\sigma\tau}\Omega_{;\sigma}\Omega_{;\tau})\hat{g}^{\mu\lambda}\hat{g}^{\nu\rho}\Omega_{|\lambda}\Omega_{|\rho} + (\hat{g}^{\sigma\tau}\Omega_{|\sigma}\Omega_{|\tau})\tilde{g}^{\mu\lambda}\tilde{g}^{\nu\rho}\Omega_{;\lambda}\Omega_{;\rho} + (\tilde{g}^{\sigma\tau}\Omega_{;\sigma}\Omega_{;\tau})\tilde{g}^{\mu\lambda}\tilde{g}^{\nu\rho}\Omega_{;\lambda}\Omega_{;\rho} = 0 . \tag{9}$$

This is an equation of the form $f(\mathrm{g}, \hat{\mathrm{g}}, \Omega) = 0$. It is the basic equation which determines $\hat{\mathrm{g}}$ from g. It replaces all three equations (1)–(3) and it will be the essential condition in a global definition of asymptotic simplicity.

Definition: A space-time $(\mathscr{M}, \mathrm{g})$ is (globally) *asymptotically simple* iff there exist

(a) a space $\hat{\mathscr{M}}$ with a nonempty boundary $\mathscr{B}$ ($\mathscr{B} \subset \hat{\mathscr{M}}$),

(b) a diffeomorphism f: $U \to \hat{U} - \mathscr{B}$ from an open subset U of $\mathscr{M}$ to $\hat{U} - \mathscr{B}$, where $\hat{U}$ is an open neighborhood of $\mathscr{B}$,

(c) a set of disjoint submanifolds $\mathscr{B}_i$ of $\mathscr{B}$ with $\cup_i \mathscr{B}_i = \mathscr{B}$, such that on some open neighborhood $\hat{U}_i$ of each $\mathscr{B}_i$ a C^∞ metric $\hat{\mathrm{g}}$ and a C^∞ scalar field Ω can be defined with $\Omega = 0$ on $\mathscr{B}_i$, $\Omega > 0$ on $\hat{U}_i - \mathscr{B}_i$ and satisfying condition (9) on $\hat{U}_i$.

The concept of asymptotic simplicity can be generalized to include cosmological models. The only essential modification will be the replacement of Eq. (9) by some other equations, depending on the asymptotic behavior of space-times we want to consider. Since here we are interested in space-times which at large distances resemble Minkowski's space-time, we give the following definitions motivated from studies of asymptotic structure at the corresponding case:

Timelike infinity $\mathscr{T}$ of an asymptotically simple space-time is the union of all $\mathscr{B}_i$'s for which

$$\Omega^{-4} g^{\mu\nu}\Omega_{;\mu}\Omega_{;\nu} \hat{=} 1 \quad \text{and} \quad \hat{g}^{\mu\nu}\Omega_{|\mu}\Omega_{|\nu} \hat{=} 1 . \tag{10}$$

Null infinity $\mathscr{N}$ of an asymptotically simple space-time is the union of all $\mathscr{B}_i$'s for which

$$\Omega^{-4} g^{\mu\nu}\Omega_{;\mu}\Omega_{;\nu} \hat{=} -1 \quad \text{and} \quad \hat{g}^{\mu\nu}\Omega_{|\mu}\Omega_{|\nu} \hat{=} 0 . \tag{11}$$

Spatial infinity $\mathscr{S}$ of an asymptotically simple space-time is the union of all $\mathscr{B}_i$'s for which

$$\Omega^{-4} g^{\mu\nu}\Omega_{;\mu}\Omega_{;\nu} \hat{=} -1 \quad \text{and} \quad \hat{g}^{\mu\nu}\Omega_{|\mu}\Omega_{|\nu} \hat{=} -1 . \tag{12}$$

The question of existence of (globally) asymptotically simple space-times with one or more of $\mathscr{T}$, $\mathscr{N}$, $\mathscr{S}$ (connected or not) will not be raised here since we are interested in space-times which have a particular composition of $\mathscr{T}$, $\mathscr{N}$ and $\mathscr{S}$. It should be noted, however, that if $\mathscr{T}$, $\mathscr{N}$ and/or $\mathscr{S}$ exist, then they are respectively spacelike, null and timelike hypersurfaces of $\hat{\mathscr{M}}$ with respect to $\hat{\mathrm{g}}$.

The usefulness of the relations (10)–(12) of the previous definitions lies in the fact that together with Eq. (9) determine uniquely $\hat{\mathrm{g}}$ from g near $\mathscr{T}$, $\mathscr{N}$ and $\mathscr{S}$. This is expressed by the following theorem.

Theorem 1: Let $\mathscr{T}$, $\mathscr{N}$ and $\mathscr{S}$ be the timelike, null and spatial infinities of a (globally) asymptotically simple space-time. Then in a coordinate system x^μ with $\Omega = x^1$ we have $(i, j = 0, 2, 3)$

(a) on an open neighborhood $\hat{U}_T$ of $\mathscr{T}$

$$\hat{g}^{11} = 1 - \tilde{g}^{11} , \tag{13}$$

$$\hat{g}^{1i} = \tilde{g}^{1i}(-1 + 1/\tilde{g}^{11}) , \tag{14}$$

$$\hat{g}^{ij} = \tilde{g}^{ij} + \tilde{g}^{1i}\tilde{g}^{1j}[1 - 4\tilde{g}^{11} + 2(\tilde{g}^{11})^2] , \tag{15}$$

(b) on an open neighborhood $\hat{U}_N$ of $\mathscr{N}$

$$\hat{g}^{\mu\nu} = \tilde{g}^{\mu\nu} , \tag{16}$$

(c) on an open neighborhood $\hat{U}_S$ of $\mathscr{S}$

$$\hat{g}^{11} = -1 - \tilde{g}^{11} , \tag{17}$$

$$\hat{g}^{1i} = \hat{g}^{1i}(-1 - 1/\tilde{g}^{11}) , \tag{18}$$

$$\hat{g}^{ij} = \tilde{g}^{ij} + \tilde{g}^{1i}\tilde{g}^{1j}(-1 - 2\tilde{g}^{11}) . \tag{19}$$

Proof: In a coordinate system x^μ in which $\Omega = x^1$ we have $\Omega_{;\mu} = \Omega_{|\mu} = \delta^1_\mu$ and Eq. (9) becomes

$$\hat{g}^{\mu\nu} - \tilde{g}^{\mu\nu} - \hat{g}^{11}\hat{g}^{1\mu}\hat{g}^{1\nu} - \tilde{g}^{11}\hat{g}^{1\mu}\hat{g}^{1\nu} + \hat{g}^{11}\tilde{g}^{1\mu}\tilde{g}^{1\nu} + \tilde{g}^{11}\tilde{g}^{1\mu}\tilde{g}^{1\nu} = 0 . \tag{20}$$

For $\mu = \nu = 1$ this equation reduces to (7) and gives three possible solutions expressed by Eqs. (4)–(6). On $\mathscr{T}$ we have from Eqs. (10) $\tilde{g}^{11} \hat{=} 0$ and $\hat{g}^{11} \hat{=} 1$ which are satisfied only if we accept (4) and reject (5) and (6). Thus Eq. (13) has been established. Using this equation and (9) we obtain (14) and (15) for $\mu = 1$, $\nu = i$ and $\mu = i$, $\nu = j$ respectively. On $\mathscr{N}$ we have from (11) $\tilde{g}^{11} \hat{=} 0$ and $\hat{g}^{11} \hat{=} 0$ which are satisfied only if we accept (5). Then using (20) we have Eqs. (16). Finally, on $\mathscr{S}$ we have from (12) $\tilde{g}^{11} \hat{=} 0$, $\hat{g}^{11} \hat{=} -1$ which are satisfied only if we accept (6). Thus we have (17) and from (20) the remaining Eqs. (18) and (19).

Equations (13)–(19) give explicitly the contravariant components of the unphysical metric tensor in coordinates x^μ with $\Omega = x^1$ in an open neighborhood of the corresponding part of the boundary. Thus given $(\mathscr{M}, \mathrm{g})$ we can calculate $\tilde{g}_{\mu\nu}$ and $\hat{g}_{\mu\nu}$ near $\mathscr{T}$, $\mathscr{N}$ and/or $\mathscr{S}$ (if they exist) and examine whether they are C^∞ functions of x^μ (essentially of x^1) on the boundary. It should be emphasized, however, that the formulas (13)–(19) are not identical to those given in the previous studies of asymptotic structure. Specifically, instead of (15) and (19) we have found respectively[15,12]

$$\hat{g}_T^{ij} = \tilde{g}^{ij} + \tilde{g}^{i1}\tilde{g}^{j1}[(-1+1/\tilde{g}^{11})^2 - 1] \tag{21}$$

and

$$\hat{g}_S^{ij} = \tilde{g}^{ij} + \tilde{g}^{i1}\tilde{g}^{j1}[1 - (1+1/\tilde{g}^{11})^2], \tag{22}$$

the remaining of (13)–(19) being identical. Thus a question is raised: It is possible to find that a space-time is asymptotically simple (at $\mathscr{T}$, $\mathscr{N}$ and/or $\mathscr{S}$) using the present unified formulation and *not* asymptotically simple using the separate formulation for $\mathscr{T}$, $\mathscr{N}$ and/or $\mathscr{S}$? In other words, are the two formulations equivalent? Fortunately, the two formulations are equivalent, as it will be shown in the next section.

3. GLOBAL ASYMPTOTIC FLATNESS

The boundary $\mathscr{B}_M$ of Minkowski's space-time consists of (a) the past $\mathscr{T}_M^-$ and future $\mathscr{T}_M^+$ timelike infinities, each one isometric to the spacelike hyperboloid, (b) the past $\mathscr{N}_M^-$ and future $\mathscr{N}_M^+$ null infinities (or $\mathscr{I}^-$ and $\mathscr{I}^+$ in Penrose's notation) and (c) the spatial boundary $\mathscr{S}_M$, isometric to the timelike hyperboloid. Thus the timelike infinity is $\mathscr{T}_M = \mathscr{T}_M^- \cup \mathscr{T}_M^+$, the null infinity $\mathscr{N}_M = \mathscr{N}_M^- \cup \mathscr{N}_M^+$ and the spatial infinity $\mathscr{S}_M$. The whole boundary $\mathscr{B}_M = \mathscr{T}_M \cup \mathscr{N}_M \cup \mathscr{S}_M$ is a three-dimensional well-known disconnected manifold which is the direct sum of five disjoint submanifolds with a three-metric on each. We relate now such a boundary with the concept of asymptotic simplicity.

Theorem 2: For a (globally) asymptotically simple space-time the following conditions are equivalent:

(a) The boundary $\mathscr{B}$ is isometric to $\mathscr{B}_M$ and on $\mathscr{T}$, $\mathscr{N}$ and $\mathscr{S}$ $(\mathscr{B} = \mathscr{T} \cup \mathscr{N} \cup \mathscr{S})$ the conditions (10), (11) and (12) respectively hold.

(b) The boundary $\mathscr{B}$ is direct sum of five disjoint submanifolds $\mathscr{T}^-, \mathscr{T}^+, \mathscr{N}^-, \mathscr{N}^+$ and $\mathscr{S}$. On some open neighborhood $\hat{U}_T^-, \hat{U}_T^+, \hat{U}_N^-, \hat{U}_N^+, \hat{U}_S$ of each of the above submanifolds there exists a coordinate system x^μ with $\Omega = \omega \equiv x^1$ in which the following hold:

(b1) On $\hat{U}_T^-$ and $\hat{U}_T^+$ we have $\tilde{g}^{11} = \omega^2 + O_3$ and ($x^0 = \chi$, $x^2 = \theta, x^3 = \varphi$, $-\infty < \chi < \infty$, θ, φ as usual)

$$\hat{g}_{\mu\nu} = \begin{pmatrix} -1+O_1 & \alpha+O_1 & O_1 & O_1 \\ \alpha+O_1 & \beta+O_1 & \gamma+O_1 & \delta+O_1 \\ O_1 & \gamma+O_1 & -\sinh^2\chi+O_1 & O_1 \\ O_1 & \delta+O_1 & O_1 & -\sinh^2\chi\sin^2\theta+O_1 \end{pmatrix} \tag{23}$$

where α, γ, δ are arbitrary functions of χ, θ, φ and

$$\beta = 1 - \alpha^2 - \gamma^2\sinh^{-2}\chi - \delta^2\sinh^{-2}\chi\sin^{-2}\theta. \tag{24}$$

(b2) On $\hat{U}_N^-$ and $\hat{U}_N^+$ we have $\tilde{g}^{11} = -\omega^2 + O_3$ and ($x^0 = u$, $x^2 = \theta, x^3 = \varphi$, $-\infty < u < +\infty$, θ, φ as usual)

$$\hat{g}_{\mu\nu} = \begin{pmatrix} \lambda\omega^2+O_3 & -1+O_1 & \mu\omega+O_2 & \nu\omega+O_2 \\ -1+O_1 & O_0 & O_0 & O_0 \\ \mu\omega+O_1 & O_0 & -1+O_1 & O_1 \\ \nu\omega+O_1 & O_0 & O_1 & -\sin^2\theta+O_1 \end{pmatrix} \tag{25}$$

where μ, ν are arbitrary functions of u, θ, φ and

$$\lambda = 1 - \mu^2 - \nu^2\sin^{-2}\theta. \tag{26}$$

(b3) On $\hat{U}_S$ we have $\tilde{g}^{11} = -\omega^2 + O_3$ and ($x^0 = \chi, x^2 = \theta$, $x^3 = \varphi$, $-\infty < \chi < +\infty$, θ, φ as usual)

$$\hat{g}_{\mu\nu} = \begin{pmatrix} 1+O_1 & \alpha+O_1 & O_1 & O_1 \\ \alpha+O_1 & \beta+O_1 & \gamma+O_1 & \delta+O_1 \\ O_1 & \gamma+O_1 & -\cosh^2\chi+O_1 & O_1 \\ O_1 & \delta+O_1 & O_1 & -\cosh^2\chi\sin^2\theta+O_1 \end{pmatrix} \tag{27}$$

where α, γ, δ are arbitrary functions of χ, θ, φ and

$$\beta = -1 + \alpha^2 - \gamma^2\cosh^{-2}\chi - \delta\cosh^{-2}\chi\sin^{-2}\theta. \tag{28}$$

Proof: Let (a) be true. Then $\mathscr{T}^-$ (or $\mathscr{T}^+$) has a neighborhood $\hat{U}_T^-$ (or $\hat{U}_T^+$) in which $\Omega = \omega$ and $\hat{g}_{\mu\nu}$ is given by (23) in appropriate coordinates $\chi, \omega, \theta, \varphi$. From (10) we have $\tilde{g}^{11} = \omega^2 + O_3$ and $\hat{g}^{11} = 1 + O_1$. Comparison with $\hat{g}^{11}$ obtained from (23) gives Eq. (24). Thus we have proved the case (b1). Similarly $\mathscr{N}^-$ (or $\mathscr{N}^+$) has a neighborhood $\hat{U}_N^-$ (or $\hat{U}_N^+$) in which $\Omega = \omega$ and $\hat{g}_{\mu\nu}$ is given by (25) but with $\hat{g}_{00} = \kappa\omega + \lambda\omega^2 + O_3$. Calculating $\hat{g}^{11}$ and setting it equal to $-\omega^2 + O_3$, as (11) suggest, we find $\kappa = 0$ and λ given by (26). Thus we have proved the case (b2). Similarly we prove case (b3). The proof of (a) from (b) is simple.

It should be noted that there are three more equivalent expressions[12,15] of the conditions of the previous theorem similar to (b). These expressions involve respectively $\hat{g}^{\mu\nu}$, $\tilde{g}_{\mu\nu}$ and $\tilde{g}^{\mu\nu}$. Thus the tensor conditions of (a) can be expressed equivalently by the existence of coordinate systems where the unphysical or the conformal metric have specific forms.

In a definition of asymptotic flatness the conditions on the boundary may involve the physical fields which have not been examined yet for spatial and timelike infinity. At null infinity, however, the conditions $\Omega_{;\mu\nu} \hat{=} 0$ and $C_{\lambda\mu\rho\nu} \hat{=} 0$ have been included in the definition.[13] Thus it seems wise at this point to define a weaker concept and leave open the question of whether or not more conditions are needed on $\mathscr{T}$ and $\mathscr{S}$. Another reason justifying this attitude is the possibility that some of the space-times we want to call asymptotically flat may violate some of the conditions on some compact submanifold of $\mathscr{T}$. Thus we propose the following definition:

Definition: A space-time is (globally) *almost asymptotically flat* (or Minkowskian) iff it is globally asymptotically simple, satisfies (a) or (b) of Theorem 2 and on $\mathscr{N}$ we have

$$\Omega_{;\mu\nu} \hat{=} 0, \quad \tilde{C}_{\lambda\mu\rho\nu} \hat{=} 0. \tag{29}$$

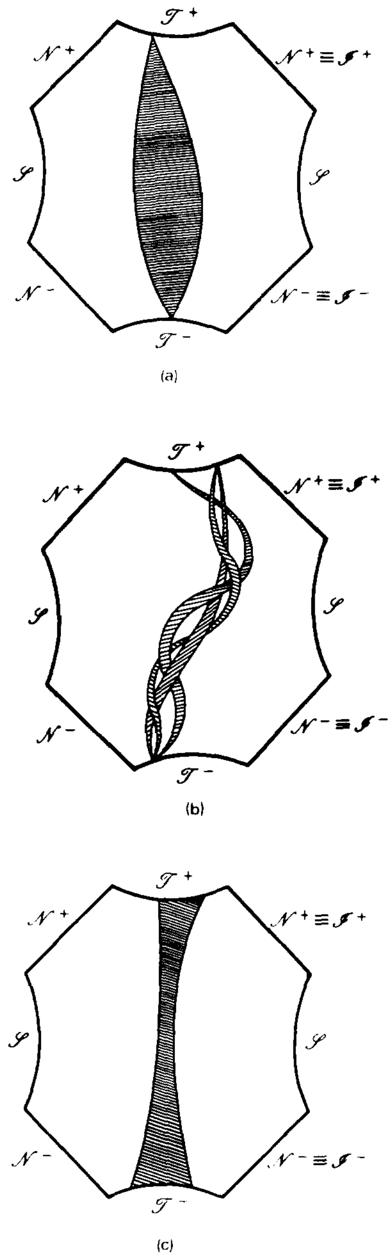

FIG. 1.

Using expressions (23), (25) and (27) we can obtain the general expressions of the physical metric near $\mathscr{T}$, $\mathscr{N}$ or $\mathscr{S}$. These expressions are identical with the general expressions of the physical metric obtained in the separate formulations of timelike,[12] null,[13] and spatial[15] infinities. This is important since the formulas (15) and (19) of the present formulation are different from the formulas (21) and (22) and proves the equivalence of the two formulations. A more direct proof can be given by taking the difference (15) and (21) near $\mathscr{T}$ and the difference of (19) and (22) near $\mathscr{S}$. Using expressions (23) and (27) we find after some calculations that these differences are of higher order in ω and consequently do not affect asymptotic simplicity.

4. GENERAL REMARKS

In this paper we have proposed a formulation for studying the asymptotic structure of a space-time simultaneously at timelike, null and spatial infinities. The space-times we considered resemble Minkowski's space-time and have been called almost asymptotically flat. Future work will show whether the word "almost" can be eliminated after adding some additional conditions or should be simply omitted without any additional conditions. In any case we can draw a diagram for an almost asymptotically flat space-time. This diagram is similar to Penrose's diagram in the case of conformal completion and shows the different parts of the boundary in a characteristic way. In the following figures we have drawn the space-times generated by (a) a star, (b) three bodies of which one escapes eventually to infinite distance from the other two and (c) a collapsing and then exploding dust cloud. Note that $\mathscr{S}$ is a single hypersurface. The same is true for $\mathscr{N}^+$ and $\mathscr{N}^-$ separately. We can imagine as boundary the surface generated by rotation of each figure about an axis passing through the middle of $\mathscr{T}^+$ and $\mathscr{T}^-$ (see Fig. 1).

ACKNOWLEDGMENTS

I would like to thank Dr. B. Xanthopoulos for useful discussions and comments on the manuscript.

[1]H. Bondi, M. G. J. van der Burg, and A. W. K. Metzner, Proc. R. Soc. London Ser. A **269**, 21 (1962).
[2]R. Penrose, Proc. R. Soc. A. London Ser. A **284**, 159 (1965).
[3]R. K. Sachs, Proc. R. Soc. London Ser. A **270**, 103 (1962).
[4]R. Arnowitt, S. Deser, and C. W. Misner, Phys. Rev. **117**, 1595 (1960); **118**, 1100 (1960); **121**, 1556 (1961); **122**, 997 (1961).
[5]R. Geroch, J. Math. Phys. **13**, 956 (1972).
[6]R. Geroch, in *Asymptotic Structure of Space-Time*, edited by F. P. Esposito and L. Witten (Plenum, New York, 1977).
[7]A. Ashtekar,and R. O. Hansen, J. Math. Phys. **19**, 1542 (1978).
[8]P. Sommers, J. Math. Phys. **19**, 549 (1978).
[9]D. Eardley and R. K. Sachs, J. Math. Phys. **14**, 209 1973; B. G. Schmidt, Commun. Math. Phys. **36**, 73 (1974).
[10]R. Geroch and B. C. Xanthopoulos, J. Math. Phys. **19**, 714 (1978).
[11]R. Geroch and G. T. Horowitz, Phys. Rev. Lett. **40**, 203 (1978).
[12]S. Persides, "Timelike Infinity," J. Math. Phys. **23**, 283(1982).
[13]S. Persides, J. Math. Phys. **20**, 1731 (1979).
[14]S. Persides, J. Math. Phys. **21**, 135 (1980).
[15]S. Persides, J. Math. Phys. **21**, 142 (1980).

Electromagnetic energy tensors and the Lorentz equation of motion for fields with electric and magnetic charge distributions

J. M. Nevin
Department of Mathematics, University of Natal, P.O. Box 375, Pietermaritzburg, 3200 Natal, South Africa

(Received 24 June 1981; accepted for publication 11 September 1981)

In special or general relativity the electromagnetic energy tensor is usually taken to be $\theta^{ab} = (1/4\pi)(F^a{}_c F^{bc} - \frac{1}{4}g^{ab}F_{cd}F^{cd})$. This expression may also be used in the generalized theory which allows magnetic as well as electric charge. Rund [J. Math. Phys. **18**, 84 and 1312 (1977)] has suggested a new approach to the generalized theory with an alternative form for the energy tensor. We show that in Rund's theory there are other possible definitions for the energy tensor. However, there is a strong indication that a particular energy tensor gives rise in a definite way to a corresponding Lorentz equation of motion. This equation is derived for each of the energy tensors and it is found that only θ^{ab} gives the Lorentz equation which is usually assumed in the generalized theory. Furthermore, the Lorentz equations arising from the other energy tensors will not give charge quantization.

PACS numbers: 04.20.Me, 04.40. + c, 03.50.De

1. ELECTROMAGNETIC FIELD TENSORS

We use the notation[1] of general relativity. g_{ab} is the gravitational metric tensor with normal form diag $(-1, +1, +1, +1)$. A semicolon denotes covariant differentiation with respect to this metric. For a skew symmetric tensor t_{ab} the dual[2] tensor $^*t_{ab}$ is defined by $^*t_{ab} = \frac{1}{2}e_{abcd}t^{cd}$. It follows that $^{**}t_{ab} = -t_{ab}$.

F_{ab} denotes the electromagnetic field tensor which satisfies the generalized Maxwell equations:

$$F^{ab}{}_{;b} = 4\pi j^a, \quad {}^*F^{ab}{}_{;b} = -4\pi s^a, \tag{1.1}$$

where j^a and s^a are the electric and magnetic current 4-vectors, respectively. In particular, for a fluid of dual-charged particles having 4-velocity field U^a and electric and magnetic charge densities ϵ, γ, respectively, $j^a = \epsilon U^a$ and $s^a = \gamma U^a$.

In Rund's theory there are two underlying fields f_{ab} and C_{ab} both skew-symmetric and with the following properties[3]:

$$f^{ab}{}_{;b} = 4\pi j^a, \quad {}^*f^{ab}{}_{;b} = 0, \tag{1.2}$$

$$C^{ab}{}_{;b} = 4\pi s^a, \quad {}^*C^{ab}{}_{;b} = 0, \tag{1.3}$$

$$F_{ab} = f_{ab} + {}^*C_{ab}. \tag{1.4}$$

Equations (1.2)–(1.4) together imply (1.1). Furthermore, (1.2) and (1.3) imply that electric and magnetic charge distributions give rise independently to the fields f_{ab} and C_{ab}, respectively. We regard f_{ab} and C_{ab} as fundamental tensors describing the electromagnetic field rather than F_{ab}.

Equations (1.2) (ii) and (1.3) (ii) imply the existence of vector fields ϕ_a, ψ_a such that

$$f_{ab} = \psi_{b;a} - \psi_{a;b}, \quad C_{ab} = \phi_{b;a} - \phi_{a;b}. \tag{1.5}$$

2. ELECTROMAGNETIC ENERGY TENSORS AND THE FIELD OF A DUAL-CHARGED MASS PARTICLE

Let

$$\theta^{ab}(F) = (1/4\pi)(F^a{}_e F^{be} - \tfrac{1}{4}g^{ab}F\cdot F),$$

where

$$F\cdot F = F_{ed}F^{ed}. \tag{2.1}$$

This has the same form as the classical electromagnetic energy tensor. Rund has suggested as an alternative[4]

$$T^{ab}(f,c) = (1/4\pi)(f^a{}_e F^{be} + C^a{}_e {}^*F^{be} - \tfrac{1}{4}g^{ab}F\cdot F). \tag{2.2}$$

T^{ab} has the following properties: (i) If $C_{ab} = 0$, then $T^{ab} = \theta^{ab}(F) = \theta^{ab}(f)$. (ii) Let $L = R - F\cdot F + 16\pi(\psi_h j^h - \theta_h s^h)$, where R is the curvative scalar. L may be expressed explicitly in terms of g_{ab}, first and second partial derivatives of g_{ab}, ψ_a, ϕ_a, $\psi_{a;b}$, $\phi_{a;b}$, J^a, and S^a, where $J^a = (-g)^{1/2}j^a$ and $S^a = (-g)^{1/2}s^a$. Let $I = \int_{\mathscr{D}} L\, dv$, where $\mathscr{D}$ is a 4-dimensional region of space-time.

Then: (a) Extremizing I with respect to variations in the g_{ab} subject to certain boundary conditions[5] leads to field equations[6] in the form $G^{ab} = 8\pi T^{ab}$. (b) Extremizing I with respect to variations in ψ_a or ϕ_a subject to certain boundary conditions leads to the generalized Maxwell equations (1.1).[9]

There are other tensors which have similar properties to T^{ab}. For example, let $\bar{L} = R - (f\cdot f + \beta c\cdot c) + 16\pi(\psi_h j^h + \beta\phi_h s^h)$, where β is some constant. Extremizing $\int \bar{L}\, dv$ with respect to variations in g_{ab} leads to a corresponding energy tensor

$$E^{ab}(f,c,\beta) = \theta^{ab}(f) + \beta\theta^{ab}(c), \tag{2.3}$$

and extremizing $\int \bar{L}\, dv$ with respect to variations in ψ_a and ϕ_a again leads to the generalized Maxwell equations (1.1).

For each of the energy tensors θ^{ab}, T^{ab}, and E^{ab} one may determine a static spherically symmetric asymptotically flat solution of the coupled Maxwell–Einstein equations:

$$F^{ab}{}_{;b} = 0, \quad {}^*F^{ab}{}_{;b} = 0, \tag{2.4}$$

$$G^{ab} = 8\pi U^{ab}, \tag{2.5}$$

where U^{ab} stands for any of the above electromagnetic energy tensors.

This has been done for θ^{ab} and T^{ab} in Refs. 7 and 8, where it is shown that in a coordinate system (t, r, θ, ϕ)

$$g_{ab} = \mathrm{diag}(e^\nu, -e^\lambda, -r^2, -r^2\sin^2\theta), \tag{2.6}$$

 0022-2488/82/020293-03$02.50

$$f_{10} = -f_{01} = \epsilon r^{-2} e^{\frac{\lambda+\nu}{2}}, \quad \text{other } f_{ab} = 0, \tag{2.7}$$

$$c_{10} = -c_{01} = -\gamma r^{-2} e^{\frac{\lambda+\nu}{2}}, \quad \text{other } c_{ab} = 0, \tag{2.8}$$

$$8\pi\theta_{ab} = \frac{(\epsilon^2+\gamma^2)}{r^4}\Delta, \quad 8\pi T_{ab} = \frac{\epsilon^2-\gamma^2}{r^4}\Delta,$$

where $\Delta = \text{diag}\,(e^\nu, -e^\lambda, r^2, r^2\sin^2\theta)$ and ϵ and γ are arbitrary constants which may be regarded as the electric and magnetic charges, respectively, of the mass particle. Equations (2.6)–(2.8) are obtained as a consequence of the assumed symmetry and Eqs. (2.4) alone. Substitution into (2.5) now leads to

$$\lambda = -\nu, \quad e^\nu = 1 - 2m/r - k/r^2, \tag{2.9}$$

$$\text{where } k = \begin{cases} \epsilon^2+\gamma^2 & \text{for } \theta_{ab}, \\ \epsilon^2-\gamma^2 & \text{for } T_{ab}. \end{cases}$$

Turning to E_{ab}, we observe that (2.6)–(2.8) still hold. Since $E_{ab}(f,c,\beta) = \theta_{ab}(f) + \beta\theta_{ab}(c)$ and $8\pi\theta_{ab}(f) = \epsilon^2\Delta/r^4$, we find that $8\pi E_{ab} = [(\epsilon^2+\beta\gamma^2)/r^4]\Delta$ and (2.9) holds with $k = \epsilon^2 + \beta\gamma^2$. We see that for the field of a dual charged mass particle

$$E_{ab} = T_{ab} \text{ if } \beta = -1, \quad E_{ab} = \theta_{ab} \text{ if } \beta = +1.$$

3. THE LORENTZ EQUATION OF MOTION

Chase[10] has shown, using a method of Infeld and Schild,[11] that in a certain sense, the Lorentz equation of motion

$$\rho \dot{u}^a = \epsilon F^a{}_b u^b \tag{3.1}$$

for a test particle of mass ρ, electric charge ϵ, in a field F^{ab} without magnetic sources is a consequence of the Maxwell–Einstein equations (2.4) and (2.5). In his derivation Chase uses θ_{ab} for the electromagnetic energy tensor and the Coulomb potential for a point charge. It is quite possible that the method of Chase can be generalized to dual charged particles using any of the above mentioned energy tensors and the Coulomb potentials which correspond to (2.7) and (2.8). Each choice of energy tensor should yield a corresponding equation of motion.

Since the method of Chase is quite involved and the approximation procedure is difficult to justify, we present instead an elementary argument resting on different assumptions. Consider the classical theory of an isentropic perfect charged fluid.[12] It is assumed in this theory that mass density is conserved meaning

$$(\rho u^a)_{;a} = 0 \tag{3.2}$$

and that the energy tensor has the form

$$S^{ab} = (\mu + p)u^a u^b + p g^{ab} + \theta^{ab}, \tag{3.3}$$

where ρ is mass density, μ is energy density, and p is pressure.

Suppose that pressure is negligible so that we have a charged "dust" fluid. Then $\mu = \rho$ and (3.3) becomes

$$S^{ab} = \rho u^a u^b + U^{ab}, \tag{3.4}$$

where $U^{ab} = \theta^{ab}$.

We will assume that, for a dual charged dust fluid, the energy tensor has the form (3.4), where U^{ab} is one of the electromagnetic energy tensors θ^{ab}, T^{ab}, and E^{ab}. We also assume that (3.2) holds. The field equations imply that

$$S^{ab}{}_{;b} = 0. \tag{3.5}$$

From (3.4) $S^{ab}{}_{;b} = (\rho u^b)_{;b} u^a + \rho u^b(u^a{}_{;b}) + u^{ab}{}_{;b}$.

Using (3.2) and (3.5) we find

$$\rho \dot{u}^a + u^{ab}{}_{;b} = 0. \tag{3.6}$$

Now the following expressions may be derived for $U^{ab}{}_{;b}$ (see Appendix):

$$\text{If } U^{ab} = \begin{cases} \theta^{ab} \\ T^{ab} \\ E^{ab} \end{cases} \text{ then } U^{ab}{}_{;b} = \begin{cases} \frac{1}{4\pi}\{F^a{}_e F^{be}{}_{;b} + {}^*F^a{}_e {}^*F^{be}{}_{;b}\} & (3.7) \\ \frac{1}{4\pi}\{f^a{}_e f^{be}{}_{;b} - C^a{}_e C^{be}{}_{;b}\} & (3.8) \\ \frac{1}{4\pi}\{f^a{}_e f^{be}{}_{;b} + \beta C^a{}_e C^{be}{}_{;b}\} & (3.9) \end{cases}.$$

Substituting (3.7), (3.8), and (3.9), respectively, into (3.6), using (1.1), (1.2) (i), (1.3) (i) and $j^a = \epsilon u^a$, $s^a = \gamma u^a$ we obtain the following equations of motion for the dual-charged dust fluid:

$$\rho \dot{u}^a = \begin{cases} (\epsilon F^a{}_e - \gamma {}^*F^a{}_e)u^e & (3.10) \\ (\epsilon f^a{}_e + \gamma C^a{}_e)u^e. & (3.11) \\ (\epsilon f^a{}_e - \gamma\beta C^a{}_e)u^e & (3.12) \end{cases}$$

Imagine that a small blob of dual charged dust is introduced into an existing background gravitational *cum* electromagnetic field in a region which was previously devoid of matter. It seems reasonable to suppose that as $\rho, \epsilon, \gamma \to 0$ the fields F_{ab} and g_{ab} at events inside the blob will approach the background fields and that we obtain in the limit laws of motion for a test particle which have the same form as (3.10)–(3.12) except that ρ, ϵ, and γ are interpreted as total mass, electric charge, and magnetic charge of the test particle, f_{ab}, c_{ab}, and F_{ab} belong to the background field, and covariant differentiation is with respect to the background field.

In the classical theory where $\gamma = 0$ and $F^a{}_e = f^a{}_e$, (3.10)–(3.12) all reduce to the usual Lorentz equation of motion which is thought to describe the motion of an electron in an external field provided the electron is not radiating significantly. Hence there is some reason to believe that one of (3.10)–(3.12) may apply to dual charged elementary particles.

4. CHARGE QUANTIZATION

In Sec. 2 we described the gravitational and electromagnetic fields of a dual charged particle. Let us assume that r is sufficiently large so that we can replace (2.9) by $e^{\nu} = 1$. Then (2.6) is the metric of flat Minkowski space in spherical polars. We make the usual spherical polar transformation from (t, r, θ, ϕ) to an inertial frame (t, x^1, x^2, x^3). Applying the tensor transformation laws we find from (2.7) and (2.8) that in this inertial frame,

$$(f_{10}, f_{20}, f_{30}) = \epsilon r/r^3, \quad (c_{10}, c_{20}, c_{30}) = -\gamma r/r^3. \tag{4.1}$$

From (1.4) and (4.1) we then find

$$F^{ab} = \begin{bmatrix} 0 & E_1 & E_2 & E_3 \\ -E_1 & 0 & B_3 & -B_2 \\ -E_2 & -B_3 & 0 & B_1 \\ -E_3 & B_2 & -B_1 & 0 \end{bmatrix}, \tag{4.2}$$

where $\mathbf{E} = \epsilon \mathbf{r}/r^3$, $\mathbf{B} = \gamma \mathbf{r}/r^3$.

Suppose we have a test particle of mass m, electric charge e, and magnetic charge g moving in the above field. If the equations of motion (3.10)–(3.12) with $a = 1,2,3$ are expressed in terms of $\mathbf{E}$ and $\mathbf{B}$ they become, respectively,

$$\frac{d}{dt}\left(\frac{m\mathbf{v}}{(1-v^2)^{1/2}}\right) = \begin{cases} e(\mathbf{E} + \mathbf{v}\times\mathbf{B}) + g(\mathbf{B} - \mathbf{v}\times\mathbf{E}) & (4.3) \\ e\mathbf{E} + g\mathbf{B} & (4.4) \\ e\mathbf{E} - \beta g\mathbf{B}; & (4.5) \end{cases}$$

(4.3) is the equation of motion which is usually assumed in the generalized theory.

Schwinger[13] has given an account of the Dirac quantization procedure which uses (4.2) and (4.3). His method rest on the fact that, according to (4.3), the Lorentz force on a dual charged test particle moving in the field of a fixed dual charged particle is not along the line of the join of these two particles. If (4.4) or (4.5) were used instead of (4.3) the Lorentz force would be along the line of the join and the argument would break down. Thus if the quantization procedure is regarded as a necessary part of the theory (4.4) and (4.5) must be rejected together with their corresponding energy tensors (2.2) and (2.3).

APPENDIX

Proof of (3.7):

$$4\pi\theta_a{}^b{}_{;b} = F_{ae}F^{be}{}_{;b} + F_{ae;b}F^{be} - \tfrac{1}{4}(F\cdot F)_{;a}, \tag{A1}$$

$$\begin{aligned} {}^*F_{ae}{}^*F^{be}{}_{;b} &= \tfrac{1}{2}e_{aepq}F^{pq}(\tfrac{1}{2}e^{bemn}F_{mn})_{;b} \\ &= -\tfrac{1}{2}F^{pq}F_{pq;a} + F^{pq}F_{aq;p} \\ &\quad \text{(using the skew symmetry of } F) \\ &= -\tfrac{1}{4}(F\cdot F)_{;a} + F^{pq}F_{aq;p}, \end{aligned} \tag{A2}$$

(A1) and (A2) give (3.7).

Proof of (3.8):

$$4\pi T_a{}^b{}_{;b} = f_{ae;b}F^{be} + f_{ae}f^{be}{}_{;b} + C_{ae;b}{}^*F^{be} - C_{ae}C^{be}{}_{;b} - \tfrac{1}{4}(F\cdot F)_{;a}, \tag{A3}$$

$$\begin{aligned} (F\cdot F)_{;a} &= 2F^{pq}F_{pq;a} \\ &= 2F^{pq}F_{pq;a} + 2F^{pq*}C_{pq;a} \\ &= 2F^{pq}f_{pq;a} + 2{}^*F^{pq}C_{pq;a} \\ &= 4(F^{pq}\psi_{q;pa} + {}^*F^{pq}\phi_{q;pa}). \end{aligned} \tag{A4}$$

Substituting (A4) in (A3) we find

$$4\pi T_a{}^b{}_{;b} = f_{ae;b}F^{be} - c_{ae}c^{be}{}_{;b} + F^{be}(f_{ae;b} - \psi_{e;ba}) + {}^*F^{be}(c_{ae;b} - \phi_{e;ba}), \tag{A5}$$

$$\begin{aligned} F^{be}(f_{ae;b} - \psi_{e;ba}) &= F^{be}(\psi_{e;ab} - \psi_{a;eb} - \psi_{e;ba}) \\ &= \tfrac{1}{2}F^{be}\{(\psi_{e;ab} - \psi_{e;ba}) + (\psi_{a;be} - \psi_{a;eb}) + (\psi_{b;ea} - \psi_{b;ae})\}; \end{aligned}$$

using the skew symmetry of F^{ab}

$$\begin{aligned} &= \tfrac{1}{2}F^{be}\psi^c(R_{ecab} + R_{acbe} + R_{bcea} \\ &= 0. \end{aligned}$$

Similarly ${}^*F^{be}(c_{ae;b} - \phi_{e;ba}) = 0$.

Equation (A5) now gives the required result. Equation (3.9) follows directly from (2.3), (3.7), (1.2) (ii), (1.3) (ii), and skew symmetry of f^{ab}, c^{ab}.

[1] S. W. Hawking and G. F. R. Ellis, *The large scale structure of space-time* (Cambridge University, Cambridge, England, 1973).

[2] R. Adler, M. Bazin, and M. Schiffer, *Introduction to general relativity* (McGraw-Hill, New York, 1975), pp. 92–93.

[3] Equations (1.2)–(1.4) are equivalent to Eqs. (2.33), (2.36), (2.38), (2.35), and (2.4) of Ref. 7. We have c_{ab} instead of $-i{}^*b_{ab}$, also $J^a = (-g)^{1/2}j^a$. Equations (1.5) are equivalent to (2.31) and (2.32).

[4] Reference 7, Eq. (3.10).

[5] Reference 1, p. 75.

[6] The general theory leads to T^{ab} in the form

$$\begin{aligned} 8\pi T^{ab} &= \frac{1}{(-g)^{1/2}}\frac{\partial}{\partial g_{ab}}\{(-g)^{1/2}(F\cdot F - 16\pi(\psi_h j^h - \phi_h s^h)\} \\ &= \frac{1}{(-g)^{1/2}}\frac{\partial}{\partial g_{ab}}((-g)^{1/2}F\cdot F). \end{aligned}$$

This can then be put in the form (2.2) as indicated in Ref. 7.

[7] H. Rund, J. Math. Phys. **18**, 84 (1977).

[8] H. Rund, J. Math. Phys. **18**, 1312 (1977).

[9] These equations are easily obtained by applying (3.4), p. 65 of Ref. 1, and using the fact that $({}^*F\cdot{}^*F) = -F\cdot F$.

[10] D. M. Chase, Phys. Rev. **95**, 243 (1954).

[11] L. Infeld and A. Schild, Phys. Rev. **21**, 408 (1949).

[12] Reference 1, p. 69–71.

[13] J. Schwinger, Science **165**, 757 (1969).

Coupled translational and rotational diffusion in liquids

Ulrich Steiger and Ronald F. Fox
School of Physics, Georgia Institute of Technology, Atlanta, Georgia 30332

(Received 29 January 1981; accepted for publication 11 August 1981)

The equations for coupled translational and rotational diffusion of asymmetric molecules immersed in a fluid are obtained. The method used begins with the Kramers–Liouville equation and leads to the generalized Smoluchowski equation for diffusion in the presence of potentials. Both external potentials and intermolecular potentials are considered. The contraction of the description from the Kramers–Liouville equation to the Smoluchowski equation is achieved by using a combination of operator calculus and cumulants. Explicit solutions to these equations are obtained for the two-dimensional case. Comparison of our results with earlier literature is also presented.

PACS numbers: 05.40. + j, 51.90. + r

I. INTRODUCTION

In this paper we study the translational and rotational motion of molecules immersed in a fluid. The molecules experience translational and rotational Brownian motion as a result of the bombardment by fluid molecules. The description of this essentially stochastic process in terms of the probability-distribution function $P(t,x)$ leads to a diffusion equation

$$\frac{\partial}{\partial t} P(t,x) = \sum_{ij} a_{ij}(x) \frac{\partial^2}{\partial x_i \partial x_j} P(t,x) + \sum_i b_i(x) \frac{\partial}{\partial x_i} P(t,x) \equiv AP(t,x) \quad (1)$$

for all times $t \geqslant 0$ and all points x, $x = (q_1, q_2, q_3, \phi, \theta, \psi)$. $q = (q_1, q_2, q_3)$ describe the position and the Euler angles $\alpha = (\phi, \theta, \psi)$ fix the orientation. The differential operator A is a diffusion operator. All eigenvalues of the symmetric matrix $[a_{ij}(x)]$ are non-negative. For translational diffusion A_T is simply a diffusion constant multiplied by the Laplace operator. Favro[1] derived the diffusion equation for rotational Brownian motion and was able to solve it for axial symmetric molecules using the fact that the diffusion operator A_R has the same form as the quantum mechanical Hamilton operator for a rigid body,[2] the properties of which are well known. In general the translational and rotational motions are coupled in a complicated way.

Already 50 years ago, Kolmogorov showed that under very general conditions a Markov process defined in terms of the transition probability $F(t,x,x')dx'$ of finding a particle initially at point x in the infinitesimal small set dx' after a lapse of time t, leads to a diffusion equation. The probability density

$$P(t,x) \equiv \int_{S_x} F(t,x',x) P(0,x')\, dx' \quad (2)$$

satisfies Eq. (1). S_x is the space containing all points x. $P(0,x)$ is the initial distribution at time $t = 0$.

The concept of a Markov process is an idealization of the underlying physical reality. For a complete dynamical description, it is necessary to consider the distribution function $f_c(t,x_c,y_c)$ defined on the phase space $S_{x_c} \times S_{y_c}$ consisting of all points (x_c, y_c) with $x_c = (q_1, q_2, q_3, \phi, \theta, \psi)$ and the canonically conjugate momenta $y_c = (p_1, p_2, p_3, p_\phi, p_\theta, p_\psi)$. The distribution function $f_c(t,x_c,y_c)$ satisfies the Kramers–Liouville equation[3,4]

$$\frac{\partial}{\partial t} f_c(t,x_c,y_c) = (L + K) f_c(t,x_c,y_c). \quad (3)$$

L is Liouville's operator and K denotes Kramers operator, which describes the effect of all random forces acting on the Brownian particle. If Eq. (3) can be solved for some initial distribution $f_c(0,x_c,y_c)$ then it is possible to find an operator $G(t,x_c)$ such that the averaged distribution $P(t,x_c)$ defined by

$$P(t,x_c) \equiv \int_{S_{y_c}} dy_c\, f_c(t,x_c,y_c) \quad (4)$$

fulfills the first order differential equation in time:

$$\frac{\partial}{\partial t} P(t,x_c) = G(t,x_c) P(t,x_c). \quad (5)$$

In general nothing is gained, since $G(t,x_c)$ might be a very complicated operator. We will use the cumulant expansion[5,6] to approximate the operator $G(t,x_c)$.

$$G(t,x_c) = \sum_{n=1}^{\infty} G^{(n)}(t,x_c). \quad (6)$$

It turns out, that the diffusion operator A is the first nonvanishing term in the expansion (6). Equation (1), where A is now replaced by the second cumulant $G^{(2)}(t,x_c)$ $[G^{(1)}(t,x_c) = 0]$, is a very good approximation of (5). K describes the time evolution of the distribution of the momenta due to random forces. The momenta $y_c(t)$ can be considered as random variables, which very quickly become independent. $y_c(t)$ is independent of $y_c(t + \Delta t)$ if the lapse of time Δt is large compared with the correlation time τ_k. It can be shown,[7] that the nth cumulant is proportional to

$$G^{(n)} \sim \hat{\tau}^{n-1}. \quad (7)$$

$\hat{\tau}$ is a dimensionless quantity. $\hat{\tau} \equiv \tau_k / \tau$. τ is some typical macroscopic time unit.

Intuitively, it is clear that we obtain a Markov process on S_x described by (1) if the correlation time τ_k of the momenta $y_c(t)$ becomes very small. It is the short correlation time which makes the higher order contributions small.

The idea of deriving the diffusion operator A as the lowest order of a cumulant expansion (6) is not new. The actual calculation of the operators $A, G^{(3)}, \ldots$, is complicated by the nonlinearity of the equation of motion for a rigid body. The time derivatives of the angular momentum L' and translational momentum p' expressed in an orthogonal coordinate frame attached to the moving particle are

$$\dot{L}' = L' \times I^{-1} L' + N', \qquad \dot{p}' = p' \times I^{-1} L' + F'. \tag{8}$$

N' and F' are the torques and the forces acting on the particle. The prime denotes vectors in the body fixed coordinate frame. I is the tensor of inertia. It is necessary to choose body fixed coordinates for both L' and p' since otherwise the friction tensor C depends on the orientation [see (70)].[8]

The purpose of this work is to analyze the rotational and translational diffusion in the most general case using a mathematically transparent method. We will show that

(i) The generalized Smoluchowski equation is the lowest order contribution of $G(t,x_c)$. Starting off with a Maxwell distribution at time $t=0$ the diffusion tensor is time dependent. For $t<\tau_k$ the diffusion tensor depends on the mass and the moments of inertia, and becomes stationary for $t \gg \tau_k$.

(ii) The diffusion equation couples the translational and rotational degrees of freedom even in the simplest case.[8] As an illustration, the two dimensional diffusion equation is solved. The solutions are obtained in terms of exponential and Mathieu functions. (Sec. V).

(iii) A suspension of N interacting Brownian particles leads to a diffusion equation for the N particle density $P(t,x_c^{(1)},x_c^{(2)},\dots,x_c^{(N)})$. (Sec. IV).

In Sec. II the operator calculus used later is introduced and applied to the translational motion. Section III treats coupled translational and rotational diffusion.

II. OPERATOR CALCULUS, TRANSLATIONAL DIFFUSION

The starting point of the theory is the Kramers–Liouville equation.[3,4]

$$\frac{\partial}{\partial t} f(t,q,p) = B f(t,q,p) = (L+K) f(t,q,p). \tag{9}$$

q are the coordinates describing the position, $q=(q_1,q_2,q_3)$ and p are the conjugate momenta. Liouville's operator is

$$Lf = -m^{-1} p \cdot \frac{\partial}{\partial q} f + \frac{\partial U}{\partial q} \cdot \frac{\partial}{\partial p} f. \tag{10}$$

U denotes the potential. Kramers operator is

$$Kf = \alpha \frac{\partial}{\partial p} \cdot \left(m^{-1} p + kT \frac{\partial}{\partial p} \right) f. \tag{11}$$

It is convenient[3] to work in the "interaction picture"

$$f \equiv e^{tK} \tilde{f}. \tag{12}$$

The exponential e^{tK} is defined by a formal power series in tK and acts on the new function $\tilde{f}$ which is assumed to be smooth enough, such that the series $e^{tK}\tilde{f} \equiv \Sigma_{n=0}^{\infty} [(tK)^n/n!]\tilde{f}$ converges. The smoother $\tilde{f}$ the smaller the contribution of $(tK)^n$ which is a differential operator of order $2n$ in the variable p. The time evolution for $\tilde{f}$ is governed by the Kramers–Liouville equation in the "interaction picture".

$$\frac{\partial}{\partial t} \tilde{f} = e^{-tK} L e^{tK} \tilde{f} \equiv \tilde{L}(t) \tilde{f}. \tag{13}$$

The operator $\tilde{L}(t)$ can be expressed in terms of the differential operators $\partial/\partial q$ and $\partial/\partial p$ using the identity

$$e^{-tK} L e^{tK} = e^{-[K,\cdot]t} L. \tag{14}$$

The proof of this equation is found in Ref. 5. The operator on the right hand side is by definition

$$e^{-[K,\cdot]t} L \equiv L + \sum_{n=1}^{\infty} [K,\cdot]^n L \, [(-t)^n/n!]. \tag{15}$$

The commutators $[K,\cdot]^n L$ can be defined by recursion,

$$\begin{aligned} [K,\cdot]^1 L &\equiv [K,L], \\ [K,\cdot]^2 L &\equiv [K,[K,L]], \\ [K,\cdot]^n L &\equiv [K,\cdot]([K,\cdot]^{n-1} L). \end{aligned} \tag{16}$$

We can calculate all terms in the infinite sum (15). Applying the commutator algebra discussed in Ref. 3 leads to

$$\tilde{L}(t) = -e^{-(\alpha/m)t} \frac{\partial}{\partial q} \cdot \left(\frac{p}{m} + kT \frac{\partial}{\partial p} \right) + e^{(\alpha/m)t} \frac{\partial}{\partial p} \cdot \left(\frac{\partial U}{\partial q} + \frac{\partial}{\partial q} \right). \tag{17}$$

In Sec. III the corresponding expression for translational and rotational motion is derived in great detail.

Formally, the solution of (13) can be written

$$\tilde{f}(t) = E(t)\tilde{f}_0 \equiv \overleftarrow{T} \exp \int_0^t ds \, \tilde{L}(s) \tilde{f}_0, \tag{18}$$

in which $\overleftarrow{T}$ exp is the time ordered exponential.[5] $\tilde{f}_0$ is the initial distribution. The time ordered exponential must be used because $\tilde{L}(t_1)$ does not commute with $\tilde{L}(t_2)$ if $t_1 \neq t_2$. We would like to derive the time evolution for the averaged distribution $P(t,q)$,

$$P(t,q) \equiv \int d^3p \, f(t,q,p) = \int d^3p \, e^{tK} \tilde{f}(t,q,p) = \int d^3p \, \tilde{f}(t,q,p) \equiv \langle \tilde{f}(t,q) \rangle. \tag{19}$$

The third equality can be proved by expanding the exponential e^{tK}. After integrating by parts, all but the lowest order term, which is $\tilde{f}$, vanish. We can assume that $\tilde{f}(t,q,p)|_{p_i=\infty} = 0$.

We write the initial condition $\tilde{f}(0,q,p) \equiv \tilde{f}_0(q,p) = f_0(q,p)$ in the form

$$f_0(q,p) = g(q,p) P_0(q), \qquad P_0(q) = \langle f_0(q) \rangle. \tag{20}$$

With Eqs. (18)–(20) one obtains

$$P(t,q) = \int d^3p \, \tilde{f}(t,q,p) = \int d^3p \, E(t) g(q,p) P_0(q) \equiv \langle E(t) \rangle_g P_0(q). \tag{21}$$

The operator $\langle E(t)\rangle_g$ is obtained by multiplying $g(q,p)$ from the left with $E(t)$ and integrating over the momenta p. Differentiating Eq. (21) with respect to t gives the time evolution equation

$$\frac{\partial}{\partial t}P(t,q)=\Big(\frac{\partial}{\partial t}\langle E(t)\rangle_g\Big)\langle E(t)\rangle_g^{-1}P(t,q). \tag{22}$$

We expect that the inverse $\langle E(t)\rangle_g^{-1}$ exists at least for small times. It may be obtained by the Neumann series[9] $A^{-1}=\Sigma_{n=0}^{\infty}(1-A)^n$. The operator

$$G(t,q)\equiv\Big(\frac{\partial}{\partial t}\langle E(t)\rangle_g\Big)\langle E(t)\rangle_g^{-1},$$
$$\frac{\partial}{\partial t}P(t,q)=G(t,q)P(t,q), \tag{23}$$

depends on q since $g(q,p)$ is a function on q and p. But in most physical applications the initial distribution of the momenta does not depend on the position q. In this case the operator G depends only on t.

In order to calculate $G(t)$ we use the cumulant expansion,[5–7] which is obtained by reordering the expression

$$G(t)=\sum_{n=0}^{\infty}\Big\langle\tilde L(t)T\exp\int_0^t\tilde L(s)ds\Big\rangle_g\Big\langle 1-T\exp\int_0^t\tilde L(s)ds\Big\rangle_g^n,$$
$$G(t)=\sum_{l=1}^{\infty}G^{(l)}. \tag{24}$$

Compare (18), (22), (23). $G^{(l)}$ contains all terms of the sum in (24) which are of order l in the operator $\tilde L(s)$. The two lowest order terms are

$$G^{(1)}(t)=\langle\tilde L(t)\rangle_g=\int d^3p\,\tilde L(t)g(p),$$
$$G^{(2)}(t)=\int_0^t ds\langle\tilde L(t)\tilde L(s)\rangle_g-\int_0^t ds\langle\tilde L(t)\rangle_g\langle\tilde L(s)\rangle_g$$
$$=\int_0^t ds\int d^3p\,\tilde L(t)\tilde L(s)g(p)-\int_0^t ds\int d^3p\,\tilde L(t)g(p)\int d^3p'\tilde L(s)g(p'). \tag{25}$$

The higher order terms are given in Sec. VI.

We assume that the distribution in the momenta is initially a Maxwell distribution

$$g(p)=(2\pi mkT)^{-3/2}\exp(-p^2/2mkT). \tag{26}$$

In this case, it is easy to verify that the first cumulant $G^{(1)}(t)$ vanishes for all times $t\geqslant 0$. The second cumulant is

$$G^{(2)}(t)=\frac{kT}{\alpha}\frac{\partial}{\partial q}\cdot\Big(\frac{1}{kT}\frac{\partial U}{\partial q}+\frac{\partial}{\partial q}\Big)(1-e^{-(\alpha/m)t}). \tag{27}$$

The time evolution equation (23) is, to second order in $\tilde L$, the Smoluchowski equation with time-dependent diffusion "constant",

$$\Lambda(t)=\frac{kT}{\alpha}(1-e^{-(\alpha/m)t}),$$
$$\frac{\partial}{\partial t}P(t,q)=\frac{\partial}{\partial q}\cdot\Lambda(t)\Big(\frac{1}{kT}\frac{\partial U}{\partial q}+\frac{\partial}{\partial q}\Big)P(t,q). \tag{28}$$

At $t=0$ the diffusion constant vanishes since by assumption the distribution in p was given by a symmetric function, the Maxwell distribution. After a short time of order m/α the particles start moving until finally the Boltzmann distribution is reached. In order to illustrate the meaning of the time-dependent diffusion constant $\Lambda(t)$ we calculate the first and second cumulant with the initial distribution $g(p)=\delta(p-p_0)$. All particles have the same momentum p_0 at $t=0$. In this case the first cumulant does not vanish:

$$G_\delta^{(1)}=-e^{(-\alpha/m)t}p_0m^{-1}\cdot\frac{\partial}{\partial q}, \tag{29}$$

$$G_\delta^{(2)}=\frac{1}{\alpha}(e^{-(\alpha/m)t}-e^{-2(\alpha/m)t})\times\Big\{\frac{1}{m}\Big(\frac{\partial}{\partial q}\cdot p_0\Big)^2-kT\Big(\frac{\partial}{\partial q}\Big)^2\Big\}+\frac{1}{\alpha}(1-e^{-(\alpha/m)t})\frac{\partial}{\partial q}\cdot\Big(\frac{\partial U}{\partial q}+kT\frac{\partial}{\partial q}\Big). \tag{30}$$

In the limit $t\to\infty$ both expressions (27), and (29) and (30) agree, as they should. The operator $G(t)$ is independent of the initial condition for large times. The larger α/m, the faster $G(t)$ approaches the constant expression. For very large values of α/m the dynamics governed by (23) approaches a Markov process. Formally the Markovian limit is obtained by first rescaling the time $\tau=\alpha^{-1}t$ and taking the limit $\alpha\to\infty$. In this limit all higher cumulants vanish since they are proportional to higher powers of $1/\alpha$.

III. COUPLED TRANSLATIONAL AND ROTATIONAL DIFFUSION

We consider particles of arbitrary shape in a fluid. The friction forces depend on the orientation. We will describe a proper choice for the variables. In Refs. 10 and 11 inconsistent definitions which lead to wrong results are used.

The position and orientation of each particle is determined by the six variables comprised in the sextuple x,

$$x=(q_1,q_2,q_3,\phi,\theta,\psi). \tag{31}$$

0 is an arbitrary origin and C the center of mass. q_1,q_2,q_3 are the coordinates of the vector $0C$ in the laboratory frame where $\hat e_1,\hat e_2,\hat e_3$ are three arbitrary orthogonal vectors of length one such that $\hat e_1\times\hat e_2=\hat e_3$, et cyclic. It is convenient, to choose the Euler angles $\alpha=(\phi,\theta,\psi)$ to describe the orientation.[12] We will also use the body fixed coordinate frame $\hat e_1',\hat e_2',\hat e_3'$ such that the tensor of inertia I becomes diagonal. The components of the vector $\hat e_i'$ expressed in the laboratory fixed frame $\hat e_1,\hat e_2,\hat e_3$ are

$$(\hat e_i')_l\equiv R_{li}(\phi,\theta,\psi),$$
$$R(\phi,\theta,\psi)\equiv[R_{li}(\phi,\theta,\psi)]. \tag{32}$$

The Euler angles are defined by

$$R(\phi,\theta,\psi)\equiv R_z(\phi)R_x(\theta)R_z(\psi). \tag{33}$$

$R_z(\phi)$ and $R_z(\psi)$ are counterclockwise rotations of a vector about the $\hat e_3$ axis. $R_x(\theta)$ is a rotation about the $\hat e_1$ axis.

$$R_z(\phi) = e^{\phi T_3},$$
$$R_x(\theta) = e^{\theta T_1}, \tag{34}$$
$$R_z(\psi) = e^{\psi T_3}.$$

The 3×3 matrices T_1, T_2, T_3 are defined

$$(T_i)_{lm} = \epsilon_{ilm}. \tag{35}$$

ϵ_{ilm} is the completely antisymmetric Levi–Civita tensor. Besides the position x (31) we need the momenta y,

$$y = (p'_1, p'_2, p'_3, L'_1, L'_2, L'_3). \tag{36}$$

Both the translational momenta p' and the angular momenta L' are expressed in the body fixed coordinate frame. The tensor of inertia and the friction tensor depend only on the mass distribution and shape of the particle. They are independent of the orientation if body fixed coordinates are used. According to (32) the vector p' and $p \equiv m\dot{q}$, where m is the mass and the dot denotes the time derivative, are related in the following way:

$$\begin{aligned} p' &= R^{\dagger}(\phi,\theta,\psi)p \\ &= R^{-1}(\phi,\theta,\psi)p. \end{aligned} \tag{37}$$

The angular momentum L' is the product of the angular velocity ω' and the tensor of inertia I,

$$L' = I\omega'. \tag{38}$$

With Eq. (37) the skewsymmetric angular velocity matrix Ω[13] expressed in the body fixed frame is

$$\Omega = R^{-1}\dot{R}. \tag{39}$$

The matrix Ω and the pseudovector ω' are related:

$$\Omega = \sum_{i=1}^{3} \omega'_i T_i. \tag{40}$$

In order to obtain Ω in terms of the Euler angles $\alpha = (\phi,\theta,\psi)$ and their time derivatives we substitute in (39) for the rotation R the expressions (33) and (34). Evaluating the time derivative in (39) and multiplying $\dot{R}$ from the left with R^{-1} leads to

$$\begin{aligned} \Omega = {} & \dot{\phi}\, e^{-\psi T_3} e^{-\theta T_1} T_3 e^{\theta T_1} e^{\psi T_3} \\ & + \dot{\theta}\, e^{-\psi T_3} T_1 e^{\psi T_3} + \dot{\psi} T_3. \end{aligned} \tag{41}$$

We compare this expression with (40). Equation (41) can be simplified using the commutator algebra $[T_i, T_j] = \epsilon_{ijk} T_k$.[2,12] One obtains for the angular velocity ω'

$$\begin{aligned} \omega'_1 &= \dot{\phi} \sin\theta \sin\psi + \dot{\theta}\cos\psi, \\ \omega'_2 &= \dot{\phi} \sin\theta \cos\psi - \dot{\theta}\sin\psi, \\ \omega'_3 &= \dot{\psi} + \dot{\phi}\cos\theta. \end{aligned} \tag{42}$$

Now we are able to describe the motion of the particle completely. The phase space $S_x \times S_y$ consists of all pairs $z = (x, y)$ defined by (31), (36), (37), (38), and (42).

A. Liouville's equation

The motion of the rigid body is a solution of the canonical equations[8]

$$\dot{x}_c = \frac{\partial H}{\partial y_c}, \quad \dot{y}_c = -\frac{\partial H}{\partial x_c},$$
$$H(x_c, y_c) = \frac{1}{2m} \| p \|^2 + \tfrac{1}{2} L' \cdot I^{-1} L' + U(x_c). \tag{43}$$

The canonical conjugate variables x_c and y_c are $x_c = x$ and $y_c = (p_1, p_2, p_3, p_\phi, p_\theta, p_\psi)$. The canonical conjugate momenta for the angle variables $\alpha = (\gamma,\theta,\psi)$ are given by $p_\alpha = \partial T/\partial\dot{\alpha}$ with $T \equiv \frac{1}{2} L' \cdot I^{-1} L'$.

$$\begin{aligned} p_\phi &= L'_1 \sin\theta \sin\psi + L'_2 \sin\theta\cos\psi + L'_3\cos\theta, \\ p_\theta &= L'_1 \cos\psi - L'_2 \sin\psi, \\ p^\psi &= L'_3. \end{aligned} \tag{44}$$

For every solution $z_c(t) \equiv (x_c(t), y_c(t))$ of Eq. (43) Liouville's theorem holds,

$$\frac{\partial}{\partial t} f_c(t,z_c) + \dot{z}_c \cdot \frac{\partial}{\partial z_c} f_c(t,z_c) = 0. \tag{45}$$

It would be more convenient to express the particle density distribution f_c as a function of the variables $z = (x, y)$ defined earlier, instead of as a function of $z_c = (x_c, y_c)$. We define a new density

$$f(t,z) \equiv f_c(t, z_c(z)). \tag{46}$$

With the following identities, one obtains the Liouville equation (48) for the new density $f(t,z)$.

$$\begin{aligned} \frac{\partial}{\partial z_c} &= \frac{\partial z}{\partial z_c}\frac{\partial}{\partial z}, \\ \dot{z} &= \frac{d}{dt} z(t) \equiv \frac{d}{dt} z(z_c(t)) = \frac{\partial z}{\partial z_c} \dot{z}_c, \\ \frac{\partial z}{\partial z_c}\frac{\partial z_c}{\partial z} &= \mathbb{1}_{12} \end{aligned} \tag{47}$$

$\mathbb{1}_{12}$ is the 12 dimensional identity matrix. We get

$$\frac{\partial}{\partial t} f(t,z) + \dot{z}\frac{\partial}{\partial z} f(t,z) = 0. \tag{48}$$

The transformation $z_c = z_c(z)$ is given by Eqs. (37) and (44). The Jacobian determinant is $-\sin\theta$. For any observable $O = O(z_c)$ the expectation value $EO \equiv \int dz_c O(z_c) f_c(t,z_c)$ can also be expressed in the new variables $z = (q,\alpha, p', L')$,

$$\begin{aligned} EO &= \int dz \left| \operatorname{Det}\left(\frac{\partial z_c}{\partial z}\right) \right| O(z_c(z)) f(t,z) \\ &= \int d^3q\, d\phi\, d\theta \sin\theta\, d\psi\, d^3p'\, d^3L' \\ &\quad \times O(q,\phi,\theta,\psi, p', L') f(t,q,\phi,\theta,\psi,p',L'). \end{aligned} \tag{49}$$

Equations (45) and (48) are formally the same but the meaning of the differential operators $\partial/\partial z_c$ and $\partial/\partial z$ are very different.

$$\frac{\partial}{\partial z_c} = \left(\frac{\partial}{\partial x_c}, \frac{\partial}{\partial y_c}\right), \quad \frac{\partial}{\partial z} = \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}\right). \tag{50}$$

The gradient $\partial/\partial x_c$ is evaluated with the canonical conjugate momenta $y_c = (p_1, p_2, p_3, p_\phi, p_\psi)$ fixed. When $\partial/\partial x$ operates, the momenta $y = (p'_1, p'_2, p'_3, L'_2, L'_3)$ are fixed.

$$\frac{\partial}{\partial x^c} = \left[\left(\frac{\partial}{\partial q_1}\right)_{\substack{q_2,q_3,\phi,\theta,\psi \\ p_1,p_2,p_3,p_\phi,p_\theta,p_\psi}}, \ldots, \left(\frac{\partial}{\partial \psi}\right)_{\substack{q_1,q_2,q_3,\phi,\theta \\ p_1,p_2,p_3,p_\phi,p_\theta,p_\psi}} \cdots \right],$$
$$\frac{\partial}{\partial x} = \left[\left(\frac{\partial}{\partial q_1}\right)_{\substack{q_2,q_3,\phi,\theta,\psi \\ p'_1,p'_2,p'_3,L'_1,L'_2,L'_3}}, \ldots, \left(\frac{\partial}{\partial \psi}\right)_{\substack{q_1,q_2,q_3,\phi,\theta \\ p'_1,p'_2,p'_3,L'_1,L'_2,L'_3}} \cdots \right]. \tag{51}$$

Rather than using (47) to calculate $\dot{z}$ we got back to Euler's equation.

$$\frac{\partial}{\partial x_c}\mathscr{L} - \frac{d}{dt}\frac{\partial}{\partial \dot{x}_c}\mathscr{L} = 0. \tag{52}$$

The Lagrange function $\mathscr{L}$ is $\mathscr{L} = \frac{1}{2}y^\dagger M^{-1}y - U(x_c)$. M is the generalized inertia matrix

$$M = \begin{pmatrix} m\mathbf{1}_3 & 0 \\ 0 & I \end{pmatrix}. \tag{53}$$

M is a symmetric 6×6 matrix. Keeping in mind that $y = y(x_c,\dot{x}_c)$ Eq. (52) can be written

$$\frac{d}{dt}y^\dagger M^{-1}\frac{\partial y}{\partial \dot{x}_c} - y^\dagger M^{-1}\frac{\partial y}{\partial x_c} + \frac{\partial U(x_c)}{\partial x_c} = 0. \tag{54}$$

The derivatives $\partial y/\partial x_c$ and $\partial y/\partial \dot{x}_c$ are 6×6 matrices. Evaluating the time derivative gives

$$\dot{y}^\dagger M^{-1}\frac{\partial y}{\partial \dot{x}_c} + y^\dagger\left(M^{-1}\frac{d}{dt}\frac{\partial y}{\partial \dot{x}_c} - M^{-1}\frac{\partial y}{\partial x_c}\right) + \frac{\partial U}{\partial x_c}(x_c) = 0. \tag{55}$$

The following definitions are useful:

$$A_{lk} \equiv \sum_{m=1}^{6} M^{-1}_{lm}\frac{\partial y_m}{\partial \dot{x}_{ck}},$$
$$B_{lk} \equiv \sum_{m=1}^{6} M^{-1}_{lm}\left(\frac{d}{dt}\frac{\partial y_m}{\partial \dot{x}_{ck}} - \frac{\partial y_m}{\partial x_{ck}}\right). \tag{56}$$

Equation (54) can be solved for $\dot{y}$. The result is

$$\dot{y}^\dagger = -\frac{\partial U(x)}{\partial x}A^{-1} - y^\dagger BA^{-1}. \tag{57}$$

$\partial U/\partial x = \partial U/\partial x_c$ in agreement with (51) since the potential U does not depend on the momenta. From the transformation $y = y(x_c,\dot{x}_c)$ given by (37) and (44) one obtains for the matrix A

$$A = \begin{pmatrix} R^{-1}(\phi,\theta,\psi) & & 0 & \\ & \sin\theta\sin\psi & \cos\psi & 0 \\ 0 & \sin\theta\cos\psi & -\sin\psi & 0 \\ & \cos\theta & 0 & 1 \end{pmatrix}. \tag{58}$$

The inverses of this matrix is

$$A^{-1} = \begin{pmatrix} R(\phi,\theta,\psi) & & 0 & \\ & \frac{1}{\sin\theta}\sin\psi & \frac{1}{\sin\theta}\cos\psi & 0 \\ 0 & \cos\psi & -\sin\psi & 0 \\ & -\cot\theta\sin\psi & -\cot\theta\cos\psi & 1 \end{pmatrix}. \tag{59}$$

We can write the matrix A and B in block form,

$$A = \begin{pmatrix} R^{-1} & 0 \\ 0 & A' \end{pmatrix}, \quad B = \begin{pmatrix} B_1 & B_2 \\ B_3 & B_4 \end{pmatrix},$$
$$BA^{-1} = \begin{pmatrix} B_1R & B_2A'^{-1} \\ B_3R & B_4A'^{-1} \end{pmatrix}. \tag{60}$$

Comparison of (37) and (44) with (58) gives

$$y = MA(x_c)\dot{x}_c. \tag{61}$$

With (61) the matrix B can be expressed in terms of A. $B = (d/dt)A - (\partial/\partial x_c)A\dot{x}_c$. The matrix B_1R is therefore equal to $((d/dt)R^{-1})R = -\Omega$. By direct calculation we find that also $B_4A'^{-1}$ is equal to $-\Omega$. The matrix B_3 vanishes. This leads to

$$BA^{-1} = -\begin{bmatrix} \Omega & \left(\frac{\partial}{\partial \alpha_j}\sum_{l,k} R^{-1}_{ik}\dot{q}_k A'^{-1}_{il}\right) \\ 0 & \Omega \end{bmatrix}. \tag{62}$$

We used the fact that the following contribution vanishes:

$$\sum_{l,k} R^{-1}_{il}\dot{q}_l\frac{\partial}{\partial \alpha_j}(R^{-1}_{ik}\dot{q}_k) = \frac{1}{2}\sum_{l,k}\frac{\partial}{\partial \alpha_j}(R^{-1}_{il}\dot{q}_l R^{-1}_{ik}\dot{q}_k) = \frac{1}{2}\frac{\partial}{\partial \alpha_j}\|R^{-1}\dot{q}\|^2 = \frac{1}{2}\frac{\partial}{\partial \alpha_j}\|\dot{q}\|^2 = 0.$$

We define the differential operator D_x,

$$D_x \equiv A^{-1\dagger}\frac{\partial}{\partial x}. \tag{63}$$

According to (57) $\dot{y}$ is

$$\dot{y} = -D_xU(x) + \begin{pmatrix} p'x\omega' \\ L'x\omega' \end{pmatrix}. \tag{64}$$

Equation (64) is Euler's equation of motion for a rigid body. The differential operator D_x is explicitly given by Eqs. (111) and (112). In the following it is more convenient to write the last term in Eq. (64) as a quadratic form in y,

$$(\dot{y})_n = -(D_xU(x))_n + \sum_{l,m} a_{lmn}y_ly_m,$$
$$a_{lmn} = \tfrac{1}{2}(C^{(n)}M^{-1} + M^{-1}C^{(n)\dagger})_{lm},$$
$$C^{(n)} = \begin{pmatrix} 0 & T_n \\ 0 & 0 \end{pmatrix},$$
$$C^{(n+3)} = \begin{pmatrix} 0 & 0 \\ 0 & T_n \end{pmatrix}, \quad n = 1,2,3. \tag{65}$$

The tensor a_{lmn} is defined such that $a_{lmn} = a_{mln}$. With these definitions we obtain Liouville's equation (48) in the form we will use it in the following.

$$\frac{\partial}{\partial t} f(t,x,y) = \left\{ -yM^{-1}\cdot D_x + (D_x U(x))\cdot\nabla - \sum_{l,m,n} a_{lmn} y_l y_m \nabla_n \right\} f(t,x,y). \tag{66}$$

The operator $\dot{x}\cdot\partial/\partial x$ in (48) and (51) is equal to $yM^{-1}\cdot D_x$ since $y = MA(x)\dot{x}$ [(61),(63)]. ∇ denotes the gradient with respect to y with components $\nabla_n \equiv \partial/\partial y_n$.

B. Kramers–Liouville equation

The motion of the particle is influenced by an external potential U and a "Brownian fluid," which is composed of molecules which exert fluctuating forces and torques,

$$\tilde{h}(t) = (\tilde{F}(t), \tilde{N}(t)). \tag{67}$$

In the absence of an external potential the equation of motion is

$$\dot{y} = -\int_{-\infty}^{t} ds\, \Gamma(t-s)y(s) + \tilde{h}(t). \tag{68}$$

For a derivation of the generalized Langevin equation (68) see Ref. 14. The friction tensor $\Gamma(t)$ is proportional to the correlation of the fluctuating forces $\tilde{h}(t)$,

$$\Gamma(t) = \frac{1}{kT} \langle \tilde{h}(0), \tilde{h}(t) \rangle. \tag{69}$$

The symmetric tensor $\Gamma(t)$ is independent of the momenta y for heavy solute molecules. In the following we will use the "Markovian limit".

$$\dot{y} = -Cy + \tilde{h}(t), \quad C \equiv \int_0^\infty \Gamma(s) ds. \tag{70}$$

The following discussion can be generalized simply by replacing the 6×6 matrix C with the corresponding expression in (68) in all equations.

In Ref. 14, Eq. (68) was derived from a linearized set of the equation of motion. Therefore one does not have to distinguish between the laboratory and the body fixed coordinate frames. The difference consists of quadratic terms $L'\times\omega'$ and $p'\times\omega'$. The idea is that over a short time of the order of the relaxation time both frames do not differ very much. After combining the stochastic equation (70) with Newton's equation, we can follow the orbit over an arbitrary long time and must therefore distinguish between both coordinate frames. The equation of motion containing the forces due to the fluid and the external forces is

$$\dot{y} = -Cy + D_x U + \begin{pmatrix} p'\times\omega' \\ L'\times\omega' \end{pmatrix} + \tilde{h}(t) \tag{71}$$

In Refs. 10 and 11 the term $p'\times\omega'$ is omitted. The generalization of Liouville's equation including stochastic forces can be obtained from (71).[5] The result is the Kramers–Liouville equation

$$\frac{\partial}{\partial t} f = (L+K) f,$$

$$Lf = -yM^{-1}\cdot D_x f + (D_x U)\cdot\nabla f - \sum_{l,m,n} a_{lmn} y_l y_m \nabla_n f, \tag{72}$$

$$Kf = \nabla\cdot C(M^{-1}y + kT\nabla) f.$$

The operator K is known as Kramers operator.

C. The operator $\tilde{L}$

In the translational case it proved very useful to go to the "interaction picture".

$$f = e^{tK}\tilde{f},$$

$$\tilde{L}(t) \equiv e^{-tK} L e^{tK} = L + \sum_{n=1}^{\infty} [K,\cdot]^n L \left(\frac{(-t)^n}{n!}\right). \tag{73}$$

The operator L consists of three terms.

$$L = L_0 + L_f + L_q,$$
$$L_0 = -y\cdot M^{-1} D_x,$$
$$L_f = (D_x U)\cdot\nabla, \tag{74}$$
$$L_q = -\sum_{lmn} a_{lmn} y_l y_m \nabla_n.$$

The calculation of the operators $\tilde{L}_0$ and $\tilde{L}_f$ does not pose any difficulties. However, for $\tilde{L}_q$ the situation is different since $\tilde{L}_q$ contains quadratic terms in q. The commutators with K become more complicated.

All operators needed in (74) are contained in the algebra generated by $x_i, y_m, \nabla_n, \partial/\partial x_i$. The position and momenta are independent. From the definition (51) we obtain $[\nabla_n, x_l] = 0$ and $[\partial/\partial x_i, y_m] = 0$. The partial derivative $\partial/\partial x_i$ is evaluated with the momenta $y = (p', L')$ held constant. The differential operator $(D_x)_i$ (63) also commutes with y_m and ∇_n for all components i,m,n. The only nonvanishing commutator needed for the calculation of $\tilde{L}$ is

$$[\nabla_n, y_m] = \delta_{nm}, \quad n,m = 1,\ldots,6. \tag{75}$$

The operator $\tilde{L}_0(f)$ is given by the infinite sum $\tilde{L}_0(t) = L_0 + \Sigma_{n=1}^{\infty} [K,\cdot]^n L_0((-t)^n/n!)$. In order to simplify the notation we introduce the matrices $\bar{C}$ and $\tilde{C}$ and the operator $\bar{D}_x$,

$$\bar{C} \equiv CM^{-1}, \quad \tilde{C} \equiv CkT, \quad \bar{D}_x \equiv -M^{-1}D_x. \tag{76}$$

Kramers operator becomes

$$K = \nabla\cdot\bar{C}y + \nabla\cdot\tilde{C}\nabla. \tag{77}$$

The operator L_0 is $L_0 = y\cdot\bar{D}_x$. The first time-dependent term in the expression for $\tilde{L}_0(t)$ is equal to $-t[K,L_0]$. This commutator is

$$[K,L_0] = [\nabla\cdot\bar{C}y, y\cdot\bar{D}_x] + [\nabla\cdot\tilde{C}\nabla, y\cdot\bar{D}_x]$$
$$= \sum_{n,l,m} \bar{C}_{nl} (\bar{D}_x)_m [\nabla_n y_l, y_m] \tag{78}$$
$$+ \sum_{n,l,m} \tilde{C}_{nl} (\bar{D}_x)_m [\nabla_n \nabla_l, y_m].$$

The following identities hold for arbitrary operators A,B,C:

$$[A,BC] = [A,B]C + B[A,C],$$
$$[AB,C] = A[B,C] + [A,C]B. \tag{79}$$

With (79), (78) becomes

$$[K,L_0] = \sum_{n,l,m} \bar{C}_{nl}(\bar{D}_x)_m \{\nabla_n [y_l, y_m] + [\nabla_n, y_m] y_l\}$$

$$+ \sum_{n,l,m} \tilde{C}_{nl}(\bar{D}_x)_m \{\nabla_n [\nabla_l, y_m] + [\nabla_n, y_m] \nabla_l\}.$$

With (75) and using the fact that the matrix $\tilde{C} = CkT$ is symmetric,[14] leads to

$$[K,L_0] = \bar{D}_x \cdot \bar{C} y + 2\bar{D}_x \cdot \tilde{C} \nabla. \tag{80}$$

For the higher order commutators one obtains

$$[K,\cdot]^n L_0 = \bar{D}_x \cdot \bar{C}^n y + 2 \sum_{m+l=n-1} \bar{D}_x \cdot \bar{C}^m \tilde{C}(-\bar{C}^\dagger)^l \nabla. \tag{81}$$

This equation can be proved by induction on n. The calculation is similar to the calculation of $[K,L_0]$. We observe that the matrix $\bar{C}^m \tilde{C}$ is symmetric for all $m \geqslant 0$ since

$$\bar{C}^m \tilde{C} = CM^{-1}CM^{-1} \cdots CM^{-1}CkT$$

$$= (\bar{C}^m \tilde{C})^\dagger = \tilde{C}\bar{C}^{\dagger m}.$$

C and M are symmetric. Using this property the last term in (81) becomes $2\Sigma_{m+l=n-1} \bar{D}_x \cdot \bar{C}^m (\bar{C})^l \tilde{C} \nabla$. The sum vanished for even n. For odd n it is equal to $2\bar{D}_x \cdot \bar{C}^{n-1} \tilde{C} \nabla$.

$$[K,\cdot]^n L_0 = \begin{cases} \bar{D}_x \cdot \bar{C}^n y, & n \text{ even} \\ \bar{D}_x \cdot \bar{C}^n y + 2\bar{D}_x \cdot \bar{C}^{n-1} \tilde{C} \nabla, & n \text{ odd.} \end{cases} \tag{82}$$

The final result for the operator $\tilde{L}_0(t)$ is

$$\tilde{L}_0(t) = \sum_{n=0}^{\infty} [K,\cdot]^n L_0 [(-t)^n/n!]$$

$$= \bar{D}_x \cdot e^{-t\bar{C}} y + \bar{D}_x \cdot (e^{t\bar{C}} - e^{-t\bar{C}}) \bar{C}^{-1} \tilde{C} \nabla, \tag{83}$$

and with the definitions of $\bar{C}$, $\tilde{C}$, and $\bar{D}_x$ [(76)] one obtains

$$\tilde{L}_0(t) = -y \cdot M^{-1} E(-t) D_x$$

$$+ kT\nabla \cdot [E(t) - E(-t)] D_x. \tag{84}$$

The matrix $E(t)$ is the exponential

$$E(t) \equiv e^{tCM^{-1}}. \tag{85}$$

The corresponding expression used earlier for the translational motion

$$\frac{-p}{m} \cdot \frac{\partial}{\partial q} e^{-(\alpha/m)t} + 2kT \sinh\left(\frac{\alpha}{m} t\right) \frac{\partial}{\partial p} \cdot \frac{\partial}{\partial q}$$

is a special case of (84). It is remarkable that no higher than second order derivatives appear in $\tilde{L}_0(t)$!

The calculation of the operator $\tilde{L}_f$ is similar. One obtains

$$\tilde{L}_f(t) = \nabla \cdot E(t) [D_x U(x)]. \tag{86}$$

In the final step we calculate the operator $\tilde{L}_q$ which is quadratic in the momenta y. This leads to major complications, but it turns out that the operator $\tilde{L}_q(t)$ contains no higher order derivatives than a third order derivative in the momenta q.

We will write L_q as the scalar product of two vectors with $6^3 = 216$ components [(74)]:

$$L_q \equiv -a \cdot (y \otimes y \otimes \nabla). \tag{87}$$

In order to find the commutators $[K,\cdot]^n L_q$ we make the ansatz that there exist some vectors $W^{(n)}$, $X^{(n)}$, $Y^{(n)}$, $Z^{(n)}$ such that

$$[K,\cdot]^n L_q = W^{(n)} \cdot (y \otimes y \otimes \nabla) + X^{(n)*} \cdot \nabla$$

$$+ Y^{(n)}(y \otimes \nabla \otimes \nabla) + Z^{(n)} \cdot (\nabla \otimes \nabla \otimes \nabla). \tag{88}$$

The vector $X^{(n)*} \in \mathbb{R}^6$ is defined $X_k^{(n)*} \equiv \Sigma_l X_{llk}^{(n)}$. The definition of the nth commutator [(16)] $[K,\cdot]^n L_q = [K,\cdot]([K,\cdot]^{n-1} L_q)$ allows us to derive recursion relations for the vectors $W^{(n)}$, $X^{(n)}$, $Y^{(n)}$, $Z^{(n)}$.

Lemma: $W^{(0)} = -a$, $X^{(0)} = 0$, $Y^{(0)} = 0$, $Z^{(0)} = 0$,

$$\begin{aligned} W^{(n+1)} &= W^{(n)} \Sigma \\ X^{(n+1)} &= X^{(n)} \Omega + W^{(n)} \Xi \\ Y^{(n+1)} &= Y^{(n)} \Upsilon + W^{(n)} \Psi \\ Z^{(n+1)} &= Z^{(n)} \Phi + W^{(n)} \Xi \end{aligned} \tag{89}$$

The 216×216 matrices Υ, Φ, Ξ, Ψ, Σ are defined

$$\begin{aligned} \Sigma &= \bar{C} \otimes \mathbb{1} \otimes \mathbb{1} + \mathbb{1} \otimes \bar{C} \otimes \mathbb{1} - \mathbb{1} \otimes \mathbb{1} \otimes \bar{C}^\dagger, \\ \Upsilon &= \bar{C} \otimes \mathbb{1} \otimes \mathbb{1} - \mathbb{1} \otimes \bar{C}^\dagger \otimes \mathbb{1} - \mathbb{1} \otimes \mathbb{1} \otimes \bar{C}^\dagger, \\ \Phi &= -\bar{C}^\dagger \otimes \mathbb{1} \otimes \mathbb{1} - \mathbb{1} \otimes \bar{C}^\dagger \otimes \mathbb{1} - \mathbb{1} \otimes \mathbb{1} \otimes \bar{C}^\dagger, \\ \Xi &= 2\tilde{C} \otimes \mathbb{1} \otimes \mathbb{1}, \\ \Psi &= 4\mathbb{1} \otimes \tilde{C} \otimes \mathbb{1}, \\ \Omega &= -\mathbb{1} \otimes \mathbb{1} \otimes \bar{C}^\dagger. \end{aligned} \tag{90}$$

$\mathbb{1}$ is the 6×6 identity matrix. $W_{klm}^{(n)}$ is symmetric in the first two indices $W_{klm}^{(n)} = W_{lkm}^{(n)}$ for all $n = 0,1,2,\ldots$.

Proof: All these relations follow directly from the definition of $X^{(n)}$, $Y^{(n)}$, $W^{(n)}$, $Z^{(n)}$ [(88)] and the definition of the commutator $[K,\cdot]^n$ [(16)].

The following equations are true for arbitrary vectors $X^{(n)}$, $Y^{(n)}$, $W^{(n)}$, $Z^{(n)}$ with the only restriction that $W^{(n)}$ is symmetric in the first two indices.

$$W_{klm}^{(n)} = W_{lkm}^{(n)}. \tag{91}$$

(1) $[\nabla \cdot \bar{C}y, X^{(n)*} \cdot \nabla] = (X^{(n)} \Omega)^* \cdot \nabla$,

(2) $[\nabla \cdot \bar{C}y, W^{(n)} \cdot (y \otimes y \otimes \nabla)] = W^{(n)} \Sigma \cdot (y \otimes y \otimes \nabla)$,

(3) $[\nabla \cdot \bar{C}y, Y^{(n)} \cdot (y \otimes \nabla \otimes \nabla)] = Y^{(n)} \Upsilon \cdot (y \otimes \nabla \otimes \nabla)$,

(4) $[\nabla \cdot \bar{C}y, Z^{(n)} \cdot (\nabla \oplus \nabla \oplus \nabla)] = Z^{(n)} \Phi \cdot (\nabla \otimes \nabla \oplus \nabla)$, (92)

(5) $[\nabla \cdot \tilde{C}\nabla, W^{(n)} \cdot (y \otimes y \otimes \nabla)] = (W^{(n)} \Xi)^* \cdot \nabla + W^{(n)} \Psi \cdot (y \otimes \nabla \otimes \nabla)$,

(6) $[\nabla \cdot \tilde{C}\nabla, Y^{(n)} \cdot (y \otimes \nabla \otimes \nabla)] = Y^{(n)} \Xi \cdot (\nabla \oplus \nabla \otimes \nabla)$,

(7) $[\nabla \cdot \tilde{C}\nabla, Z^{(n)} \cdot (\nabla \otimes \nabla \otimes \nabla)] = 0$.

The proof of these equations is mostly straightforward. For instance, the first equation is

$$[\nabla \cdot \bar{C}y, X^{(n)*} \cdot \nabla] = \sum_{\alpha,\beta,\gamma} \bar{C}_{\alpha\beta} X_\gamma^{(n)*} [\nabla_\alpha y_\beta, \nabla_\gamma]$$

$$= \sum_{\alpha,\beta,\gamma} \bar{C}_{\alpha\beta} X_\gamma^{(n)*} \nabla_\alpha (-\delta_{\alpha\beta}) = X^{(n)*}(-\bar{C}^\dagger) \cdot \nabla$$

$$= (X^{(n)})^* \cdot \nabla.$$

The fifth equation is different since there are two different types of terms:

$$[\nabla\cdot\tilde{C}\nabla, W^{(n)}\cdot(y\otimes y\otimes\nabla)]$$

$$= \sum_{\alpha,\beta,\gamma,\delta,\epsilon} \tilde{C}_{\alpha\beta} W^{(n)}_{\gamma\delta\epsilon}\{(\nabla_\alpha[\nabla_\beta, y_\gamma] + y_\delta\nabla_\epsilon) + [\nabla_\alpha, y_\gamma]\nabla_\beta y_\delta\nabla_\epsilon + y_\gamma\nabla_\alpha[\nabla_\beta, y_\delta]\nabla_\epsilon + y_\gamma[\nabla_\alpha, y_\delta]\nabla_\beta\nabla_\epsilon\}.$$

By assumption $W^{(n)}_{\gamma\delta\epsilon} = W^{(n)}_{\delta\gamma\epsilon}$ and $[y_\alpha,\nabla_\beta] = -\delta_{\alpha\beta}$. This gives the result

$$[\nabla\cdot\tilde{C}\nabla, W^{(n)}\cdot(y\otimes y\otimes\nabla)] = (W^{(n)}\Xi)^*\cdot\nabla + W^{(n)}\Psi\cdot(y\otimes\nabla\otimes\nabla).$$

The proof of the other equations is similar.

We define the vector valued function $W(t):\mathbb{R}\to\mathbb{R}^{216}$,

$$W(t) \equiv \sum_{n=0}^{\infty} ((-t)^n/n!) W^{(n)} \tag{93}$$

and similarly $X^*(t)$, $Y(t)$, and $Z(t)$. The recursion relations (89) for $W^{(n)}$, $X^{*(n)}$, $Y^{(n)}$, and $Z^{(n)}$ lead to the differential equations

$$\begin{aligned}
&W(0) = -a, \quad X(0) = 0, \quad Y(0) = 0, \quad Z(0) = 0,\\
&\dot{W}(t) = -W(t)\Sigma,\\
&\dot{X}(t) = -X(t)\Omega - W(t)\Xi,\\
&\dot{Y}(t) = -Y(t)\Upsilon - W(t)\Psi,\\
&\dot{Z}(t) = -Z(t)\phi - Y(t)\Xi.
\end{aligned} \tag{94}$$

These differential equations can be integrated and the results are

$$\begin{aligned}
&W(t) = -a\exp(-t\Sigma),\\
&X(t) = a\int_0^t ds\exp(-s\Sigma)\Xi\exp([s-t]\Omega),\\
&Y(t) = a\int_0^t ds\exp(-s\Sigma)\Psi\exp([s-t]\Upsilon),\\
&Z(t) = -\int_0^t ds Y(s)\Xi\exp([s-t]\Phi).
\end{aligned} \tag{95}$$

With these expressions the final result for the operator $\tilde{L}(t)$ is with (84), (86), (88), (95):

$$\begin{aligned}
\tilde{L}(t) = &-y\cdot M^{-1}E(t)D_x\\
&+ kT\nabla\cdot[E(t) - E(-t)]D_x\\
&+ \nabla\cdot E(t)[D_x U(x)]\\
&+ W(t)\cdot(y\otimes y\otimes\nabla) + X^*(t)\cdot\nabla\\
&+ Y(t)\cdot(y\otimes\nabla\otimes\nabla) + Z(t)\cdot(\nabla\otimes\nabla\otimes\nabla).
\end{aligned} \tag{96}$$

This is the Liouville operator in the interaction picture. The quadratic term L_q caused all the additional terms. Even if they are not explicitly known, we will be able to show that they do not contribute to the first and second cumulants.

D. First cumulant

We calculate the cumulants under the assumption that initially the distribution in the momenta y is a Maxwell distribution,

$$g(y) = \frac{1}{(2\pi kT)^3(\det M)^{1/2}} e^{-y\cdot M^{-1}y/2kT}. \tag{97}$$

The first cumulant is according to (25)

$$G^{(1)}(t) = \int d^6y\,\tilde{L}(t)\,g(y). \tag{98}$$

We use expression (96) of $\tilde{L}(t)$ and integrate by parts. The contribution at the boundaries vanish. The remaining terms are integrals over odd functions in y_m, which vanish. The first cumulant is identically zero for all times $t\geqslant 0$,

$$G^{(1)}(t)P(t,x) = 0. \tag{99}$$

E. Second cumulant

The second cumulant gives the first nonvanishing contribution,

$$G^{(2)}(t) = \int_0^t ds\int d^6y\,\tilde{L}(t)\,\tilde{L}(s)\,g(y) \tag{100}$$

with [(96)]

$$\begin{aligned}
G^{(2)}(t) = -&\int_0^t ds\int d^6y\,y\cdot M^{-1}E(-t)D_x\\
&\times[-y\cdot M^{-1}E(-s)D_x\\
&+ kT\nabla\cdot\{E(s) - E(-s)\}D_x\\
&+ \nabla\cdot E(s)[D_x U(x)] + W(s)\cdot(y\otimes y\otimes\nabla)\\
&+ X^*(s)\cdot\nabla + Y(s)\cdot(y\otimes\nabla\otimes\nabla)]g(y).
\end{aligned} \tag{101}$$

The remaining terms of the product $\tilde{L}(t)\tilde{L}(s)$ vanish after integrating by parts. The only term left from the operator $\tilde{L}(t)$ is $-y\cdot M^{-1}E(-t)D_x$. Also the term $Z(s)\cdot(\nabla\otimes\nabla\otimes\nabla)$ vanished after integrating by parts three times.

At first we can show that the contribution due to the terms $W(s)\cdot(y\otimes y\otimes\nabla)$, $X^*(s)\cdot\nabla$, and $Y(s)\cdot(y\otimes\nabla\otimes\nabla)$ cancel each other. We will show that the following integral vanishes for $k = 1,2,...,6$ and all times $s\geqslant 0$:

$$J_k^{(s)} = \int d^6y\,y_k[W(s)\cdot(y\otimes y\otimes\nabla) + X^*(s)\cdot\nabla + Y(s)\cdot(y\otimes\nabla\otimes\nabla)]g(y). \tag{102}$$

We recall that $\int d^6y\,y_i\,y_j\,g(y) = M_{ij}kT$. Again integrating by parts (102) becomes

$$J_k(s) = -\sum_{n,m}(kTW_{nmk}(s)M_{nm} + X_{nnk}(s) - Y_{nnk}(s)). \tag{103}$$

The function $J_k(s)$ may be written as $J_k(s) = \Sigma_{n=0}^{\infty} J_k^{(n)}(-s)^n/n!$. For the constants $J_k^{(n)}$ one obtains, according to (93),

$$J_k^{(l)} = -\sum_{n,m}(kTW^{(l)}_{nmk}M_{nm} + X^{(l)}_{nnk} - Y^{(l)}_{nnk}). \tag{104}$$

The recursion relations (89) allow us to define $J_k^{(l)}$ in terms of $J_k^{(l-1)}$,

$$J_k^{(l)} = \sum_{l',k',m'}\Big(kTW^{(l-1)}_{k'l'm'}M_{k'l'}\bar{C}^\dagger_{m'k} + X^{(l-1)}_{k'k'm'}\bar{C}_{m'k} + Y^{(l-1)}_{k'k'm'}\bar{C}^\dagger_{m'k}\Big).$$

Comparing this expression with (103) shows

$$J_k^{(l)} = -\sum_{k'} J_{k'}^{(l-1)} \bar{C}^{\dagger}_{k'k} = (J^{(l-1)}\bar{C}^{\dagger})_k. \tag{106}$$

The vector $J_k^{(0)}$ vanishes because $X^{(0)} = y^{(0)} = 0$ and $\Sigma_{k'l'} W^{(0)}_{k'l'm} M_{k'l'} = -\Sigma_{k'l'} a_{k'l'm} M_{k'l'}$ $= -\frac{1}{2}\,\mathrm{Tr}(C^{(m)} + M^{-1}C^{(m)\dagger}M) = 0$ [(65), (89)]. This shows that $J_k^{(l)} = 0$ for all l and k. Therefore

$$J_k(s) = 0, \quad s \geqslant 0. \tag{107}$$

The integration of the remaining four terms in (101) is straightforward. One has to keep in mind that the matrix $M^{-1}E(t)$ is symmetric.

The final result is

$$\frac{d}{dt}P(t,x) \cong G^{(2)}(t)P(t,x)$$
$$= D_x \cdot \Lambda(t)\left(D_x + \frac{1}{kT}(D_x U(x))\right)P(t,x). \tag{108}$$

The time-dependent diffusion matrix is

$$\Lambda(t) = kTC^{-1}(1 - e^{-tCM^{-1}}), \quad t \geqslant 0. \tag{109}$$

Equation (108) is the generalized Smoluchowski equation for coupled translational and rotational diffusion. Since we started with a Maxwell distribution at $t = 0$, the diffusion matrix $\Lambda(t)$ is time dependent. Equation (108) includes as a special case the translational diffusion and the rotational diffusion discussed in Ref. 1. The operator D_x depends on the orientation $\alpha = (\phi,\theta,\psi)$.

$$D_x \equiv \begin{pmatrix} D_{q'} \\ D_\alpha \end{pmatrix}, \tag{110}$$

$$D_{q'} = R^{\dagger}(\phi,\theta,\psi)\begin{pmatrix} \partial/\partial q_1 \\ \partial/\partial q_2 \\ \partial/\partial q_3 \end{pmatrix}, \tag{111}$$

$$D_\alpha = \begin{pmatrix} \cos\psi\dfrac{\partial}{\partial\theta} + \sin\psi\dfrac{1}{\sin\theta}\dfrac{\partial}{\partial\phi} - \cot\theta\sin\psi\dfrac{\partial}{\partial\psi} \\ -\sin\psi\dfrac{\partial}{\partial\theta} + \cos\psi\dfrac{1}{\sin\theta}\dfrac{\partial}{\partial\phi} - \cot\theta\cos\psi\dfrac{\partial}{\partial\psi} \\ \dfrac{\partial}{\partial\psi} \end{pmatrix}. \tag{112}$$

The rotation $R(\phi,\theta,\psi)$ is defined in (33) and (34). The expression for D_α follows from (59) and (63). Usually the friction tensor C is split into four 3×3 matrices.

$$C = \begin{pmatrix} C_{TT} & C_{TR} \\ C_{RT} & C_{RR} \end{pmatrix}. \tag{113}$$

For axialsymmetric molecules it is easy to show that $C_{TR} = C_{RT} = 0$.[8] In this case the diffusion equation is

$$\frac{\partial}{\partial t}P(t,q,\alpha) = \Big\{D_{q'}\cdot\Lambda_T\Big(D_{q'} + \frac{1}{kT}[D_{q'}U(q,\alpha)]\Big) \tag{114}$$
$$+ D_\alpha\cdot\Lambda_r\Big(D_\alpha + \frac{1}{kT}[D_\alpha U(q,\alpha)]\Big)\Big\}P(t,q,\alpha),$$

with

$$\Lambda_t = kTC_{TT}^{-1}(1 - e^{-tC_{TT}m^{-1}}),$$
$$\Lambda_R = kTC_{RR}^{-1}(1 - e^{-tC_{RR}I^{-1}}), \quad t \geqslant 0.$$

The diffusion of translational and rotational degrees of freedom is still coupled even if the potential U vanishes, since $D_{q'}$, depends on α. In Sec. V we will solve (114) in two dimensions for $U(q,\alpha) = 0$.

In Refs. 10 and 11, different expressions for the operators corresponding to $D_{q'}$ and D_α, which are wrong in our opinion, are used. Instead of D_α the operator $J \equiv -iqx(\partial/\partial q)$ was used. J is, up to a constant factor, the quantum mechanical angular momentum operator for a rotating *point particle*. Both operators D_α and J have the same commutator algebra since they are both infinitesimal generators of a representation of SO(3). D_α and J correspond to two different representations; see (136). A connection between $J = -iqx(\partial/\partial q)$ and the three Euler angles (ϕ,θ,ψ) also used in Refs. 10 and 11 is not obvious.

For axially symmetric particles one can factorize the angular dependence of $P(t,q,\phi,\theta,\psi)$ in ψ. The operator D_α^2 is in general not equal to $\Delta|_{r=1}$, the Laplace operator in spherical coordinates on the unit sphere. This is only true if we set $\partial/\partial\psi = 0$. If we consider only axial symmetric molecules and do not distinguish between two orientations which differ only by a rotation about the axis of symmetry, then we may use $D_\alpha^2|_\psi = \Delta|_{r=1}$; see (136). Reference 10 obtained wrong results by setting $J^2 = \Delta$.

It is important to keep in mind that the operator $D_{q'}$ depends on the orientation. $D_{q'}$ is the gradient along the body fixed coordinate axis. If $D_{q'}$ is replaced by $D_q = \partial/\partial q$ one obtains wrong results.[10,11] The coupling of translational and rotational diffusion of the two dimensional model discussed in Sec. V is a consequence of the α dependence of $D_{q'}$, only.

These claims will be justified in detail in Sec. V.

IV. *N* PARTICLE DIFFUSION

We consider N particles moving in a fluid interacting via arbitrary forces. In general the N particle density $P(t,x^{(1)},x^{(2)},...,x^{(N)})$ is not the product of the distributions $P(t,x^{(i)})$, where $x^{(i)}$ denotes the six coordinates of the ith particle $x^{(i)} = (q^{(1)},\alpha^{(i)})$. The N particles are correlated. The interaction energy is

$$U(x^{(1)},x^{(2)},...,x^{(N)}).$$

For an arbitrary observable $O(x^{(1)},x^{(2)},...,x^{(N)})$ depending on the position and orientation of the particles $1,...,N$ the expectation value is defined

$$EO(t) \equiv \int d\mu_{\mathbf{x}}\,P(t,\mathbf{x})O(\mathbf{x}) \tag{115}$$

with $\mathbf{x} \equiv (x^{(1)},x^{(2)},...,x^{(N)})$. The volume element $d\mu_{\mathbf{x}}$ is the product measure

$$d\mu_{\mathbf{x}} = \prod_{i=1}^{N} dq_1^{(i)}dq_2^{(i)}dq_3^{(i)}d\phi^{(i)}\sin\theta^{(i)}d\theta^{(i)}d\psi^{(i)}. \tag{116}$$

The objectives of this section is to derive the evolution equation for the N particle density $P(t,\mathbf{x})$ based on the Kramers–Liouville equation for the N particle motion. For the complete description of the N particle dynamics all positions $x^{(i)}$ and all momenta $y^{(i)}$ are required.

$$z^{(i)} \equiv (x^{(i)}, y^{(i)}),$$
$$\mathbf{z}(t) \equiv (z^{(1)}(t), z^{(2)}(t), \ldots, z^{(N)}(t)). \tag{117}$$

These variables are connected with the canonical variables $\mathbf{z}_c(t)$ through the transformation (37) and (44) applied on every single coordinate $z^{(i)}$, $i = 1,\ldots,N$,

$$\mathbf{z}(t) = \mathbf{z}(\mathbf{z}_c(t)) \equiv [z^{(1)}(z_c^{(1)}(t)),\ldots,z^{(N)}(z_c^{(N)}(t))]. \tag{118}$$

Liouville's equation holds for the density $f_c(t,\mathbf{z}_c)$ since the determinant of the Jacobian matrix of the flux $\mathbf{z}_c(t)$ is equal to 1 as a consequence of Hamilton's equation.

$$(\dot{x}_c^{(i)})_k = \frac{\partial H}{\partial (y_c^{(i)})_k}, \quad (\dot{y}_c^{(i)})_k = -\frac{\partial H}{\partial (x_c^{(i)})_k} \tag{119}$$

for $k = 1,2,\ldots,6$ and $i = 1,2,\ldots,N$. The Hamiltonian function is

$$H(\mathbf{x}, \mathbf{y}) = \frac{1}{2} \sum_{i=1}^{N} y^{(i)} \cdot M^{(i)-1} y^{(i)} + U(x^{(1)},\ldots,x^{(N)}).$$

The matrix $M^{(i)}$ is the generalized inertia matrix (53) of the ith particle. Liouville's equation is

$$\frac{\partial}{\partial t} f_c(t,\mathbf{z}_c) + \dot{\mathbf{z}}_c \cdot \frac{\partial}{\partial \mathbf{z}_c} f_c(t,\mathbf{z}_c) = 0. \tag{120}$$

$\dot{\mathbf{z}}_c$ is determined by (119). The expectation value of an observable $O(\mathbf{z}_c)$ is obtained by

$$EO(t) = \int d\mu_c\, f_c(t,\mathbf{z}_c) O(\mathbf{z}_c) \tag{121}$$

$d\mu_c$ is the volume element in the phase space $(S_{x_c} \times S_{y_c})^{xN}$.

$$d\mu_c = \prod_{i=1}^{N} \prod_{k=1}^{12} d(z_c^{(i)})_k. \tag{122}$$

Instead of the canonical variables $\mathbf{z}_c$ we use again $\mathbf{z}$. The transformation of the density f_c, the observable 0, and the measure $d\mu_c$ are

$$f(t,\mathbf{z}) \equiv f_c(t,\mathbf{z}_c(\mathbf{z})),$$
$$O(\mathbf{z}) \equiv O(\mathbf{z}_c(\mathbf{z})),$$
$$d\mu \equiv \left| \operatorname{Det} \frac{\partial \mathbf{z}_c}{\partial \mathbf{z}} \right| dz \tag{123}$$
$$= \prod_{i=1}^{N} \sin\theta^{(i)} \prod_{k=1}^{12} dz_k^{(i)}.$$

The expectation value of the function $O(\mathbf{z})$,

$$EO(t) = \int d\mu\, f(t,\mathbf{z}) O(\mathbf{z}), \tag{124}$$

agrees with the definition (121).

The Kramers–Liouville equation for the N particle problem has the form

$$\frac{\partial}{\partial t} f(t,\mathbf{z}) = -\dot{\mathbf{z}} \cdot \frac{\partial}{\partial \mathbf{z}} f(t,\mathbf{z}) + \sum_{i=1}^{N} K^{(i)} f(t,\mathbf{z}). \tag{125}$$

$K^{(i)}$ is the Kramers operator acting on the ith particle,

$$K^{(i)} \equiv \nabla^{(i)} \cdot C^{(i)} [M^{(i)-1} y^{(i)} + kT\nabla^{(i)}],$$
$$\nabla^{(i)} \equiv \frac{\partial}{\partial y^{(i)}}. \tag{126}$$

The Kramers operator is the direct sum of the individual operators $K^{(i)}$ acting on the ith particle. The forces due to the fluid are completely random and not correlated at different positions.[14] The correlation matrix of all components of all random forces and random torques, which is a $6^N \times 6^N$ matrix, is the direct sum of the correlation matrices $C^{(i)}$. Therefore Eq. (126) is justified. With $L^{(i)}$, the Liouville operator acting on the ith particle, the Kramers–Liouville equation (125) is the sum of N formally identical operators,

$$\frac{\partial}{\partial t} f(t,\mathbf{z}) = \sum_{i=1}^{N} (L^{(i)} + K^{(i)}) f(t,\mathbf{z}),$$
$$L^{(i)} = -y^{(i)} \cdot M^{(i)-1} D_{x(i)} + D_{x(i)} U(\mathbf{x}) \cdot \nabla^{(i)}$$
$$- a^{(i)} \cdot (y^{(i)} \oplus y^{(i)} \otimes \nabla^{(i)}). \tag{127}$$

All operators $L^{(i)}$ are connected through the potential $U(\mathbf{x})$. Equation (127) contains the complete N body dynamics.

Since $[K^{(i)}, L^{(j)}] = 0$ for $i \neq j$ we have

$$\exp\left(-t \sum_{i=1}^{N} K^{(i)}\right) \sum_{j=1}^{N} L^{(j)} \exp\left(t \sum_{i=1}^{N} K^{(i)}\right)$$
$$= \sum_{i=1}^{N} e^{-tK^{(i)}} L^{(i)} e^{tK^{(i)}}$$
$$= \sum_{i=1}^{N} \tilde{L}^{(i)}(t). \tag{128}$$

The operator $\tilde{L}^{(i)}(t)$ are given in Eq. (96) after replacing z by $z^{(i)}$, and M by $M^{(i)}$. The evolution equation for the density $\tilde{f}$ defined by $f \equiv e^{tK} \tilde{f}$ is therefore

$$\frac{\partial}{\partial t} \tilde{f}(t,\mathbf{z}) = \sum_{i=1}^{N} \tilde{L}^{(i)}(t) \tilde{f}(t,\mathbf{z}). \tag{129}$$

Suppose the momentum distribution is Gaussian initially,

$$g(\mathbf{y}) = \prod_{i=1}^{N} g(y^{(i)}),$$
$$g(y^{(i)}) = \frac{1}{(2\pi kT)^3 (\det M)^{1/2}} e^{-y^{(i)} \cdot M^{(i)-1} y^{(i)}/2kT}. \tag{130}$$

As in the one particle case the first cumulant vanishes.

$$G^{(1)}(t) P(t,\mathbf{z}) = \int \prod_{i=1}^{N} d^6 y^{(i)} \sum_{j=1}^{N} \tilde{L}^{(j)}(t) \prod_{k=1}^{N} g(y^{(k)}) P(t,\mathbf{x})$$
$$= \sum_{j=1}^{N} \int d^6 y^{(j)} \tilde{L}^{(j)}(t) g(y^{(j)}) P(t,\mathbf{x}) = 0. \tag{131}$$

The second cumulant is

$$G^{(2)}(t)P(t,\mathbf{x}) = \int_0^t ds \int \prod_{i=1}^N d^6y^{(i)} \sum_{l=1}^N \sum_{m=1}^N \tilde{L}^{(l)}(t)\tilde{L}^{(m)}(s) \prod_{j=1}^N g(y^{(j)})\, P(t,\mathbf{x})$$
$$= \int_0^t ds \sum_{l=1}^N \int \prod_{i=1}^N d^6y^{(i)}\tilde{L}^{(l)}(t)\tilde{L}^{(l)}(s) \prod_{j=1}^N g(y^{(j)})P(t,\mathbf{x})$$
$$+ \int_0^t ds \sum_{l\neq m}^N \int \prod_{i=1}^N d^6y^{(i)}\tilde{L}^{(l)}(t)\tilde{L}^{(m)}(s) \prod_{j=1}^N g(y^{(j)})P(t,\mathbf{x})$$
$$= \left\{ \sum_{l=1}^N G^{(2)(l)}(t) + \sum_{l\neq m} \int_0^t ds\, G^{(1)(l)}(t)\, G^{(1)(m)}(s)\right\} P(t,\mathbf{x}). \tag{132}$$

The second term vanishes because all first cumulants $G^{(1)(l)}$ $l = 1,...,N$ are zero. The remaining term is the sum of the cumulants calculated for the one particle dynamics. The N particle diffusion equation is

$$\frac{\partial}{\partial t} P(t,x^{(1)},...,x^{(N)}) = \sum_{i=1}^N D_{x(i)}\cdot \Lambda^{(i)}(t) \times\left(D_{x(i)} + \frac{1}{kT} D_{x(i)}\, U(x^{(1)},...,x^{(N)})\right) \times P(t,x^{(1)},...,x^{(N)}),$$
$$\Lambda^{(i)}(t) = kTC^{(i)-1}(\mathbb{1} - e^{-tC_{(i)}M^{(i)-1}}). \tag{133}$$

This is the generalization of the Smoluchowski equation for N interacting translating and rotating particles.

V. CORRELATIONS BETWEEN THE VARIABLES q AND α

We consider the one particle diffusion equation (114). In general the positions and orientations are correlated. The correlations are not only caused by the potential $U = U(x)$, $x = (q,\alpha)$ or by nonvanishing elements of the matrix $C_{TR} = C^{\dagger}_{RT}$. We will show that, if the positions q and the orientations α are uncorrelated at $t = t_0$ there are in general correlations for $t > t_0$ even if the potential vanishes and also $C_{TR} = 0$.

A. Axially symmetric particles

As an illustration we consider axially symmetric particles. In this case one can show that $C_{TR} = 0$.[15] If we identify the axis of symmetry with the e'_3 axis the matrices C_{TT}^{-1} and C_{RR}^{-1} are diagonal.

$$C_{TT}^{-1} = \begin{pmatrix} a & 0 & 0 \\ 0 & a & 0 \\ 0 & 0 & a_3 \end{pmatrix}, \quad C_{RR}^{-1} = \begin{pmatrix} b & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & b_3 \end{pmatrix}. \tag{134}$$

We assume that we know the distribution at time $t = t_0$, where t_0 is large compared with the translational and rotational relaxation time of the momenta.

$$t_0 \gg m\,\|C_{TT}^{-1}\| \text{ and } t_0 \gg \|C_{RR}^{-1}I\|,$$
$$\frac{\partial}{\partial t} P(t,q,\alpha) = kT\,[aD_{q'}^2 + (a_3 - a)(D_{q'})_3^2 + bD_\alpha^2 + (b_3 - b)(D_\alpha)_3^2\,]P(t,q,\alpha) \quad \text{for } t > t_0. \tag{135}$$

This equation is a special case of (114) where we used expression (134) for the friction tensor. We also used $\Lambda(t) \cong kTC^{-1}$ for $t \gg t_0$.

The differential operators $(D_{q'})^2$, $(D_{q'})_3^2$, D_α^2, and $(D_\alpha)_3^2$ are given by Eqs. (111) and (112).

$$(D_{q'})^2 = \Delta_q,$$
$$(D_{q'})_3^2 = \frac{\partial}{\partial q}\cdot B(\alpha)\frac{\partial}{\partial q}; \quad B_{ij}(\alpha) \equiv R_{3i}(\alpha)R_{3j}(\alpha),$$
$$(D_\alpha)^2 = \frac{\partial^2}{\partial^2\theta} + \frac{1}{\sin^2\theta}\left(\frac{\partial^2}{\partial^2\phi} + \frac{\partial^2}{\partial^2\psi}\right) - 2\frac{\cos\theta}{\sin^2\theta}\frac{\partial}{\partial\phi}\frac{\partial}{\partial\psi} + \cot\theta\frac{\partial}{\partial\theta}, \tag{136}$$
$$(D_\alpha)_3^2 = \frac{\partial^2}{\partial^2\psi}.$$

Δ_q is the Laplace operator in Cartesian coordinates.

We define the new density $P(t,q,\phi,\theta)$,

$$P(t,q,\phi,\theta) = \int d\psi\, P(t,q,\phi,\theta,\psi). \tag{137}$$

Integrating Eq. (135) on both sides with respect to ψ leads to

$$\frac{\partial}{\partial t} P(t,q,\phi,\theta) = kT\left[a\Delta_q + (a_3 - a)\frac{\partial}{\partial q}\cdot B(\phi,\theta)\frac{\partial}{\partial q} + b\Delta\,|_{r=1}\right]P(t,q,\phi,\theta). \tag{138}$$

The matrix $B(\alpha)$ defined in Eq. (136) does not depend on ψ.

$$B(\phi,\theta) = \begin{pmatrix} \sin\theta\cos\phi \\ \sin\theta\cos\phi \\ \cos\theta \end{pmatrix} \otimes \begin{pmatrix} \sin\theta\cos\phi \\ \sin\theta\cos\phi \\ \cos\theta \end{pmatrix}. \tag{139}$$

The contributions of Eq. (135) which contain a derivative with respect to ψ vanish after integrating by parts. Therefore the operator D_α^2 reduces to $\Delta\,|_{r=1}$, the Laplace operator in spherical coordinates on the unit sphere.

$$D_\alpha^2\,|_\psi = \Delta\,|_{r=1}. \tag{140}$$

We assume that the initial condition factorizes. For $t > t_0$ the solution of (138) has the form

$$P(t_0,q,\phi,\theta) = P_{0t}(q)P_{0R}(\phi,\theta),$$
$$P(t,q,\phi,\theta) = P_T(t,[P_R])P_R(t,\phi,\theta), \quad t > t_0. \tag{141}$$

The function $P_T(t)$ is also a functional of the distribution $P_R(t)$. $P_T(t)$ and $P_R(t)$ are probability densities, $\int d^3q\, P(t,q,[P_R]) = 1$ and $\int d\phi d\theta \sin\theta P_R(t,\phi,\theta) = 1$ for all times $t > t_0$. The boundary conditions are: $P_T(t,q,[P_R]) = 0$ if $q_i = \infty$ for some $i = 1,2,3$. Substituting (141) into Eq. (138) and integrating with respect to ϕ and θ (using the weight $\sin\theta$) leads to Eq. (142). Similarly one obtains (143) by integrating with respect to q.

$$\frac{\partial}{\partial t} P_T(t,q,[P_R]) = kT\Big\{a\Delta_q + (a_3 - a) \times \sum_{ij} \int d\phi \sin\theta d\theta B_{ij}(\phi,\theta) P_R(t,\phi,\theta) \frac{\partial}{\partial q_i}\frac{\partial}{\partial q_j}\Big\} \times P_T(t,q,[P_R]), \tag{142}$$

$$\frac{\partial}{\partial t} P_R(t,\phi,\theta) = kTb\Delta\Big|_{r=1} P_R(t,\phi,\theta) \quad \text{for } t \geqslant t_0. \tag{143}$$

The second equation describes the "Brownian motion on the unit sphere." The eigenfunction of $\Delta\ |_{r=1}$ are the spherical harmonics $Y_{1m}(\theta,\phi)$. Substituting a solution $P_R(t,\phi,\theta)$ of (143) into Eq. (142) one obtains an expression which is formally a diffusion equation with time-dependent diffusion coefficients. The off diagonal elements of the diffusion matrix vanish if the distribution $P_R(t,\phi,\theta)$ is uniform.

Similarly, one can show that for arbitrary molecules with $C_{TR} = 0$ a solution of the form (141) (including ψ) exists, if the positions and orientations are uncorrelated at time $t = t_0$ and if $U = 0$.

B. Diffusion in two dimensions

In two dimensions the diffusion equation without external potential can be solved for arbitrary initial conditions. Equation (108) reduces to

$$\frac{\partial}{\partial t} P(t,q_1,q_2,\phi) = AP(t,q_1,q_2,\phi), \tag{144}$$

$$A = \begin{pmatrix} \frac{\partial}{\partial q_1} \\ \frac{\partial}{\partial q_2} \end{pmatrix} \cdot \Lambda(\phi) \begin{pmatrix} \frac{\partial}{\partial q_1} \\ \frac{\partial}{\partial q_2} \end{pmatrix} + kT\gamma \frac{\partial^2}{\partial^2\phi}, \quad t \geqslant t_0$$

$$\Lambda(\phi) = kT\begin{pmatrix} \alpha\cos^2\phi + \beta\sin^2\phi & (\beta-\alpha)\sin\phi\cos\phi \\ (\beta-\alpha)\sin\phi\cos\phi & \alpha\sin^2\phi + \beta\cos^2\phi \end{pmatrix}. \tag{145}$$

$kT\alpha$, $kT\beta$, and $kT\gamma$ are the diffusion constants corresponding to the degrees of freedom q_1, q_2, and ϕ. We assume that $\alpha > \beta$. We use the following identities to simplify the matrix $\Lambda(\phi)$:

$$\alpha\cos^2\phi + \beta\sin^2\phi = \frac{\alpha+\beta}{2} + \frac{\alpha-\beta}{2}\cos 2\phi,$$
$$\alpha\sin^2\phi + \beta\cos^2\phi = \frac{\alpha+\beta}{2} - \frac{\alpha-\beta}{2}\cos 2\phi, \tag{146}$$
$$2\sin\phi\cos\phi = \sin 2\phi,$$

$$\Lambda(\phi) = kT\delta\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + kT\epsilon\begin{pmatrix} \cos 2\phi & -\sin 2\phi \\ -\sin 2\phi & -\cos 2\phi \end{pmatrix}. \tag{147}$$

δ is the average translational diffusion constant and ϵ is a measure for the asymmetry of the particle.

$$\delta \equiv \frac{\alpha+\beta}{2}, \quad \epsilon \equiv \frac{\alpha-\beta}{2}. \tag{148}$$

Without solving (144) explicitly it is already possible to make some statements about the lowest moments of q_1, q_2, and ϕ. One obtains the following differential equations for the expectation values $\langle\cdots\rangle_t = \int dq_1\, dq_2\, d\phi \cdots P(t,q_1,q_2,\phi)$:

$$\frac{d}{dt}\langle q_1\rangle_t = 0,$$
$$\frac{d}{dt}\langle q_1^2\rangle_t = 2kT\delta + 2kT\epsilon\langle\cos 2\phi\rangle_t,$$
$$\frac{d}{dt}\langle q_1q_2\rangle_t = -2kT\epsilon\langle\sin 2\phi\rangle_t, \tag{149}$$
$$\frac{d}{dt}\langle\cos 2\phi\rangle_t = -4\gamma kT\langle\cos 2\phi\rangle_t,$$
$$\frac{d}{dt}\langle\sin 2\phi\rangle_t = -4\gamma kT\langle\sin 2\phi\rangle_t.$$

This leads to

$$\langle\cos 2\phi\rangle_t = e^{-4\gamma kTt}\langle\cos 2\phi\rangle_{t_0},$$
$$\langle q_1^2\rangle_t = 2kT\delta t + \frac{\epsilon}{2\gamma}(1 - e^{-4\gamma kTt})\langle\cos 2\phi\rangle_{t_0} + \langle q_1^2\rangle_{t_0}, \tag{150}$$
$$\langle q_1q_2\rangle_t = \frac{-\epsilon}{2\gamma}(1 - e^{-4\gamma kTt})\langle\sin 2\phi\rangle_{t_0} + \langle q_1q_2\rangle_{t_0}.$$

The calculation of arbitrary expectation values $\langle 0\rangle_t$, $0 = 0(q_1,q_2,\phi)$ can be reduced to the problem of finding the eigenvectors and eigenvalues of the diffusion operator A in Eq. (144).

$$(A - \lambda_{k_1k_2l})\psi_{k_1k_2l} = 0. \tag{151}$$

For the symmetric case $\alpha = \beta$ the solutions of (151) are

$$\psi'_{k_1k_2l}(q_1,q_2,\phi) = \frac{1}{\pi^{1/2}}\frac{1}{L} e^{ik_1q_1} e^{ik_2q_2}\sin(l\phi),$$
$$\psi_{k_1k_2l}(q_1,q_2,\phi) = \frac{1}{\pi^{1/2}}\frac{1}{L} e^{ik_1q_1} e^{ik_2q_2}\cos(l\phi). \tag{152}$$

We choose a box of length L and assume periodic boundary conditions,

$$\psi(0,q_2,\phi) = \psi(L,q_2,\phi), \quad \psi(q_1,0,\phi) = \psi(q_1,L,\phi),$$
$$\psi(q_1,q_2,\phi) = \psi(q_1,q_2,\phi + 2\pi). \tag{153}$$

The possible values for k_1, k_2, and l are

$$k_1 = \pm\frac{2n\pi}{\lambda}, \quad k_2 = \pm\frac{2m\pi}{L}, \quad n,m \in N$$
$$l = 0,1,2,\ldots. \tag{154}$$

In the general case $\alpha > \beta$ we make the ansatz that the eigenfunctions can be written

$$\psi_{k_1k_2l}(q_1,q_2,\phi) = \frac{1}{\pi^{1/2}}\frac{1}{L} e^{ik_1q_1} e^{ik_2q_2} g_{k_1k_2l}(\phi). \tag{155}$$

One obtains the following differential equation for the unknown function $g_{k_1k_2l}(\phi)$ [(144), (147), (151)]:

$$\left[-\delta(k_1^2+k_2^2)-\epsilon(k_1^2-k_2^2)\cos 2\phi + 2\epsilon k_1k_2\sin 2\phi+\gamma\frac{\partial^2}{\partial\phi^2}-\frac{\lambda_{k_1k_2l}}{kT}\right]g_{k_1k_2l}(\phi)=0. \tag{156}$$

We define the complex wavenumber k',

$$k'\equiv k_1+ik_2, \qquad \psi\equiv\arctan(k_2/k_1). \tag{157}$$

k' can be written $k'=|k'|l^{i\psi}$. Equation (156) becomes

$$\left\{-\left(\delta|k'|^2+\frac{\epsilon}{2}|k'|^2e^{2i\phi+2i\psi}+\frac{\epsilon}{2}|k'|^2e^{-2i\phi-2i\psi}\right)+\gamma^2\frac{\partial^2}{\partial\phi^2}-\frac{\lambda_{k_1k_2l}}{kT}\right\}g_{k_1k_2l}(\phi)=0. \tag{158}$$

The exponentials can be combined to $\cos(2[\phi+\psi])$. Equation (158) is equivalent to Mathieu's equation.[16]

$$\frac{d}{dz}y_l(z)+(a_l(r)-2r\cos 2z)y_l(z)=0, \tag{159}$$

$$r=\frac{(k_1^2+k_2^2)(\alpha-\beta)}{4\gamma},$$

$$z=\phi+\arctan(k_2/k_1), \tag{160}$$

$$\lambda_{k_1k_2l}=-kT\left\{\gamma a_l(r)+\frac{\alpha+\beta}{2}(k_1^2+k_2^2)\right\},$$

$$g_{k_1k_2l}(\phi)=y_l[\phi+\arctan(k_2/k_1)].$$

The eigenvalues $a_l(r)$ of Mathieu's equation are negative for certain values of r and l,[16] but the eigenvalues $\lambda_{k_1k_2l}$ are always less or equal to zero for all k_1k_2, and l.

Equation (159) has a complete set of orthogonal solutions $ce_l(r,z)$ and $se_l(r,z)$ with the corresponding eigenvalues denoted by $a_l(r)$ and $b_l(r)$.[16] The eigenfunctions of (151) are

$$\psi'_{k_1k_2l}(q_1,q_2,\phi)=\frac{1}{\pi^{1/2}}\frac{1}{L}e^{ik_1q_1}e^{ik_2q_2}se_l(r,\phi+\arctan(k_2/k_1))$$

$$\psi_{k_1k_2l}(q_1,q_2,\phi)=\frac{1}{\pi^{1/2}}\frac{1}{L}e^{ik_1q_1}e^{ik_2q_2}ce_l(r,\phi+\arctan(k_2/k_1)). \tag{161}$$

since $\{\psi'_{k_1k_2l},\psi_{k_1k_2l}\}$ is a complete set of orthogonal eigenfunctions of the diffusion operator (151), the expectation value $\langle 0\rangle_t$ can be found by

$$\langle 0\rangle_t=\int dq_1\,dq_2 d\phi\,P(t,q_1,q_2,\phi)\,O(q_1,q_2,\phi)$$

$$=\sum_{k_1k_2l}e^{\lambda_{k_1k_2l}t}P_{k_1k_2l}\,O_{k_1k_2l}+\sum_{k_1k_2l}e^{\lambda'_{k_1k_2l}t}P'_{k_1k_2l}\,O'_{k_1k_2l}. \tag{162}$$

The coefficients $O_{k_1k_2l}, O'_{k_1k_2l}, P_{k_1k_2l}, P'_{k_1k_2l}$ are obtained from $O(q_1,q_2,\phi)$ and the initial distribution $P(t_0,q_1,q_2,\phi)$.

$$O_{k_1k_2l}=\int dq_1\,dq_2\,d\phi\,\psi^*_{k_1k_2l}(q_1,q_2,\phi)\,O(q_1,q_2,\phi)\equiv(\psi_{k_1k_2l},0),$$

$$O'_{k_1k_2l}=(\psi'_{k_1k_2l},0), \tag{163}$$

$$P_{k_1k_2l}=(P(t_0),\psi_{k_1k_2l}),$$

$$P'_{k_1k_2l}=(P(t_0),\psi'_{k_1k_2l}).$$

As an illustration we consider the following two observables:

$$O^s(q_1,\phi)\equiv\sin(k_1q_1)se_1(r,\phi),$$

$$O^c(q_1,\phi)\equiv\sin(k_1q_1)ce_1(r,\phi), \tag{164}$$

with $k_1=2\pi/L$ and $r=\pi^2(\alpha-\beta)/\gamma L^2$. We assume that the asymmetry is small. In this case $r\ll 1$ and the Mathieu functions se_1 and ce_1 are approximately

$$ce_1(r,\phi)\cong\cos(\phi)-\frac{r}{8}\cos(3\phi),$$

$$se_1(r,\phi)\cong\sin(\phi)-\frac{r}{8}\sin(3\phi). \tag{165}$$

The corresponding eigenvalues are

$$a_1(r)\cong 1+r,$$

$$b_1(r)\cong 1-r. \tag{166}$$

The eigenvalues $\lambda_{\pm k_101}$ and $\lambda'_{\pm k_101}$ are

$$\lambda_{\pm k_101}\cong -kT\left(\gamma+\frac{\pi^2}{L^2}(3\alpha+\beta)\right),$$

$$\lambda'_{\pm k_101}\cong -kT\left(\gamma+\frac{\pi^2}{L^2}(\alpha+3\beta)\right), \tag{167}$$

and for the expectation values of O^s and O^c one obtains

$$\langle O^c\rangle_t\cong ce^{-kT[\gamma+(\pi^2/L^2)(3\alpha+\beta)]t},$$

$$\langle O^s\rangle_t\cong c'e^{-kT[\gamma+(\pi^2/L^2)(\alpha+3\beta)]t}. \tag{168}$$

The constants c and c' can be written $c=(O^c,P(t_0))$ and $c'=(O^s,P(t_0))$.

The state O^c decays faster since we assumed $\alpha>\beta$. α corresponds to the diffusion along the e'_1 axis of the molecule. In the state O^c the molecule axis e'_1 is mainly parallel to the e_1 direction of the laboratory frame; in the state $O^s e'_1$ is mainly parallel to the e_2 axis. The average speed of the molecules in state O^c is bigger in the direction e_1; e_1 is also the direction of the spatial inhomogenity. Therefore O^C decays faster than O^s. This example is typical for the type of coupling of q_1,q_2, and ϕ, which occurs in the translational and rotational diffusion if the otential U vanishes and also $C_{TR}=0$.

VI. CONCLUDING REMARKS

We have shown that a "contraction of the description" is achieved when a Kramers–Liouville process is averaged with respect to its momenta variables. The second cumulant of an ordered time evolution cumulant expansion yields the generalized Smoluchowski equation as the contracted description. We have examined the details of the dynamical operator algera generated by the contraction procedure for translational and rotational degrees of freedom, and for as many as N distinct particles.

A more thorough description of the higher order cumulants, shown to be small here, will appear in a forthcoming paper.

ACKNOWLEDGMENT

NSF Grant 7921541 for theoretical physics has supported this research. This work is from the dissertation of U. Steiger, Georgia Tech, 1981.

[1]L. D. Favro, Phys. Rev. **119**, 53 (1966).
[2]R. Jost, *Quantenmechanik I* (Verein der Mathematiker und Physiker, ETH, Zürich, 1969).
[3]J. L. Davis and R. F. Fox, J. Stat. Phys. **22**, 627 (1980).
[4]S. Chandrasekhar, Rev. Mod. Phys. **15**, 1 (1943).
[5]R. F. Fox, Phys. Rep. **48**, 179 (1978).
[6]R. F. Fox, J. Math. Phys. **20**, 2467 (1979).
[7]R. F. Fox, J. Math. Phys. **17**, 1148 (1976).
[8]D. W. Condiff and J. S. Dahler, J. Chem. Phys. **44**, 3988 (1966).
[9]K. Yosida, *Functional Analysis* (Springer, Berlin, 1974).
[10]G. T. Evans, J. Chem. Phys. **65**, 3030 (1978).
[11]L.-P. Hwang and J. H. Freed, J. Chem. Phys. **63**, 118 (1975).
[12]H. Goldstein, *Classical Mechanics* (Addison-Wesley, Reading, Mass., 1959), Chaps. 4 and 5.
[13]V. I. Arnold, *Mathematical Methods of Classical Mechanics* (Springer, Berlin, 1978).
[14]E. H. Hauge and A. Martin-Löf, J. Stat. Phys. **7**, 259 (1973).
[15]J. Happel and H. Brenner, *Low Reynolds Number Hydrodynamics* (Prentice Hall, Englewood Cliffs, N. J., 1965).
[16]I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals and Series and Products* (Academic, New York, 1965), Sec. 8.6.

Two variable relativistic tensor harmonics

M. Gmitro
Institute of Nuclear Physics ČSAV, CS 250 68 Řež, Czechoslovakia and Laboratory of Theoretical Physics, JINR Dubna, P.O. Box 79, SU 101 000 Moscow, USSR

A. A. Ovchinnikova
Institute of Nuclear Physics, Moscow State University, Moscow District, Dubna, USSR

(Received 4 April 1980; accepted for publication 7 November 1980)

Three bases in the Hilbert space of tensor fields on the unit spheres associated with two independent vectors are discussed: the tensor spherical harmonics and the symmetric and unsymmetric tensor helicity harmonics. Under the conditions which we specify they form complete sets of independent Lorentz covariants which may serve the purpose of the analysis of reactions with several particles in the final state.

PACS numbers: 11.30.Cp, 03.65.Fd

1. INTRODUCTION

The Lorentz covariants written in the Cartesian basis have been the traditional theoretical tool to analyze the lepton- and meson-induced processes on nucleons and nuclei. Though manifestly covariant, the technique frequently does not fit the purposes of the physical investigations, e.g., already the basic problem of selecting the independent covariants may prove to be practically insoluble in many cases.

It is our experience that the tensor harmonics in the spherical and helicity bases constitute a convenient, highly flexible framework for the construction of the Lorentz covariants: The most important mathematical properties of the tensor harmonics follow directly from the well-known formulas of the angular-momentum algebra. The construction of the independent sets of covariants is straightforward. Orthogonality properties of the tensor bases make the calculations of rates and other physical quantities much easier that with the cumbersome Cartesian techniques. Besides these technical advantages two gratifying properties of the new formalism constitute its main merit and should be mentioned. First, the relativistic tensor harmonics allow a natural unification of the theoretical treatment of a big class of different physical processes. Second, the formalism, though fully equivalent to the covariant Cartesian expressions is actually very much similar to the familiar nonrelativistic multipole-expansion formulas. Therefore, the physical results may always be easily interpreted by a direct extrapolation to the domain of classical nuclear physics. Using the tensor harmonics, we need not perform any "nonrelativistic reduction" which is normally done, e.g., via the Foldy–Wouthuysen transformation. Namely this is the step which frequently makes the treatment of physical processes unwieldy and brings in the approximations which are usually difficult to control. The trick here is indeed in choosing the appropriate reference frame. It is the Breit frame which being fully appropriate physically, gives simultaneously an enormous simplification of the formulas.

The formalism of the relativistic sth order tensor spherical harmonics has been presented recently by Daumens and Minnaert.[1] For the corresponding analysis performed in the helicity basis we refer the reader to the paper by Akyeampong.[2] As a matter of fact the method was first introduced by Stech and Schülke[3] who have considered, however, only the specific case of nuclear beta-decay. Recently, Delorme[4] presented the application of the relativistic spherical tensor harmonics in the context of the so-called elementary-particle theory of nuclear currents. The treatment in Refs. 1–4 is always limited to the one-variable harmonics which correspond to the case of binary reactions.

Here we shall present our results concerning the two-variable Lorentz-covariant tensor harmonics in the spherical and helicity bases. It will be shown that they provide actually the most general description of the multivariable tensor fields, which may be needed in the analysis of any reaction of the type $a + b \rightarrow 1 + 2 + \cdots + n$.

In Sec. 2 we define the spherical tetrads and build up the second-order tensor spherical basis. Section 3 is devoted to the (scalar) spherical harmonics in two variables. There we display the reduction formula which permits an easy elimination of those harmonics which can be expressed as linear combinations (with scalar coefficients) of the harmonics which form the basic set. In Sec. 4 the second order tensor spherical harmonics in two variables are introduced and their most important properties are listed. In Sec. 5 two different forms of tensor harmonics in the helicity bases are deduced from the two-variable tensor spherical harmonics constructed in the preceding section. Finally, in Sec. 6, we indicate, using a particular example of the reaction with three particles in the final state, how the formalism of the present paper may be applied and indicate some of its merits in comparison with the Cartesian expressions.

2. TENSOR SPHERICAL BASIS

First we have to introduce a set of orthogonal 4-vectors on which to define the projections of the tensor fields. Following Daumens and Minnaert[1] we choose three spacelike vectors e_μ^{1n} $(n = \pm 1, 0)$ on the unit sphere $S^2(e)$ embedded in the subspace $E^3(e)$ orthogonal to the timelike vector $e_\mu^{00} \equiv e_\mu$. The complex vectors e_μ^{rn} satisfy the following conditions:

$$(e_\mu^{rn})^* = (-1)^{r+n+1} e_\mu^{r-n}, \tag{1}$$

$$e_\mu^{rn} (e_\mu^{r'n'})^* = \delta_{rr'} \delta_{nn'}. \tag{2}$$

We use the Pauli metrics (i.e., $a_\mu b_\mu = a \cdot b = \mathbf{a} \cdot \mathbf{b} - a_0 b_0$) and the usual summation convention for repeated Greek indices ($\mu = 0, 1, 2, 3$). Note that three vectors $e_\mu^{1n} (n = \pm 1, 0)$ form the ususal three dimensional spherical basis.

The spherical components of an arbitrary vector a_μ in the basis just introduced are given from the decomposition

$$a_\mu = \sum_m a^{1m} e_\mu^{1m*} + a^{00} e_\mu^{00*}. \tag{3}$$

This means

$$a^{00} = a_\mu e_\mu^{00}, \quad a^{1m} = a_\mu e_\mu^{1m}. \tag{4}$$

The construction of the tensor spherical bases of an arbitrary order has been discussed in detail by Daumens and Minnaert.[1] We shall deal with the second-order tensor basis

$$t_{\mu\lambda}^{(r_1 r_2) rn} = \sum_{n_1 n_2} \begin{bmatrix} r_1 & r_2 & r \\ n_1 & n_2 & n \end{bmatrix} e_\mu^{r_1 n_1} e_\lambda^{r_2 n_2} \tag{5}$$

only, where the symbol $\begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix}$ denotes the Clebsch–Gordan coefficient. The parity and orthogonality properties of basis (5) read as follows

$$P{:}t_{\mu\lambda}^{(r_1 r_2) rn} = (-1)^{r_1 r_2} t_{\mu\lambda}^{(r_1 r_2) rn}, \tag{6}$$

$$t_{\mu\lambda}^{(r_1 r_2) rn} \cdot t_{\mu\lambda}^{(r_1' r_2' r' n')} = \delta_{r_1 r_1'} \delta_{r_2 r_2'} \delta_{rr'} \delta_{-nn'} (-1)^{r+n}. \tag{7}$$

3. SPHERICAL HARMONICS IN TWO VARIABLES

The relativistic spherical harmonics may be introduced by taking the projections of an arbitrary unit 4-vector u_μ on the basis e_μ^{rn}. The construction as performed, e.g., in Ref. 1,

$$\begin{aligned} Y_m^l(u) &= (-1)^l \left[\frac{(2l+1)!!}{4\pi l!} \right]^{1/2} (u_{\mu_1} \otimes u_{\mu_2} \otimes \cdots \otimes u_{\mu_l}) \\ &\times \sum_{m_i n_i} \begin{bmatrix} 1 & 1 & 2 \\ n_1 & n_2 & m_2 \end{bmatrix} \begin{bmatrix} 2 & 1 & 3 \\ m_2 & n_3 & m_3 \end{bmatrix} \cdots \\ &\times \begin{bmatrix} l-1 & 1 & l \\ m_{l-1} & n_l & m \end{bmatrix} e_{\mu_1}^{1n_1} \cdots e_{\mu_l}^{1n_l} \end{aligned} \tag{8}$$

holds for the purely spacelike vectors (i.e., $u_0 e = 0$) only. This condition, however, is not restrictive, since we shall work in a reference system defined by a timelike vector $Q_\mu = (\mathbf{0}, Q_0)$. Then, choosing the spherical basis in such a way that $e_\mu^{00} = Q_\mu / (-Q^2)^{1/2}$, we may always instead of an arbitrary vector a_μ consider the vector

$$\tilde{a}_\mu = a_r - (Qa/Q^2) Q_\mu, \tag{9}$$

which is orthogonal to e_μ:

$$\tilde{a}_\mu \cdot e_\mu^{00} = 0, \tag{10}$$

and then define the unit 4-vector $u_\mu = \tilde{a}_\mu / (\tilde{a}^2)^{1/2}$. In this way, since the time components of u_μ vanish, the spherical harmonics $Y_m^l(u)$ defined in (8) on $S^2(e)$ become identical with the usual spherical harmonics as defined, e.g., in Ref. 5. In the simplest case $l = 1$ we have

$$u_\mu = \sqrt{\frac{4\pi}{3}} \sum_m Y_m^1(u) e_\mu^{1m*}. \tag{11}$$

Proceeding to the case of two variables u and v, we construct as usual the objects, which transform according to the irreducible representations of the rotation group:

$$\begin{aligned} \{(l_1 l_2) lm\} &= \{ Y_{l_1}(u) Y_{l_2}(v) \}_{lm} \\ &= \sum_{m_1 m_2} \begin{bmatrix} l_1 & l_2 & l \\ m_1 & m_2 & m \end{bmatrix} Y_{m_1}^{l_1}(u) Y_{m_2}^{l_2}(v). \end{aligned} \tag{12}$$

The bipolar harmonics[5,6] (12) form a complete orthonormal basis on the unit sphere embedded in E^3:

$$\begin{aligned} &\int d\Omega_u d\Omega_v \{ Y_{l_1}(u) Y_{l_2}(v) \}_{lm}^* \{ Y_{l_1'}(u) Y_{l_2'}(v) \}_{l'm'} \\ &= \delta_{l_1 l_1'} \delta_{l_2 l_2'} \delta_{ll'} \delta_{mm'}. \end{aligned} \tag{12a}$$

The product in the Hilbert space is

$$\begin{aligned} &\{(l_1 l_2) lm\} \{(l_1' l_2') l' m'\} \\ &= \frac{1}{4\pi} \sum_{LL' \mathscr{L} M} \hat{l}_1 \hat{l}_1' \hat{l}_2 \hat{l}_2' \hat{l} \hat{l}' \begin{bmatrix} l_1 & l_1' & L \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} l_2 & l_2' & L' \\ 0 & 0 & 0 \end{bmatrix} \\ &\times \begin{bmatrix} l & l' & \mathscr{L} \\ m & m' & M \end{bmatrix} \begin{Bmatrix} l_1 & l_2 & l \\ l_1' & l_2' & l' \\ L & L' & \mathscr{L} \end{Bmatrix} \{(LL') \mathscr{L} M\}, \end{aligned} \tag{12b}$$

with the usual[5] notation $\begin{Bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{Bmatrix}$ for the $9j$ symbol. The parity operator acts on the bipolar harmonics as follows

$$\hat{P}_r \{(l_1 l_2) lm\} = (-1)^{l_1 + l_2} \{(l_1 l_2) lm\}. \tag{12c}$$

For each given value of l we would like to keep only independent terms of the infinite-dimensional basis (12). By independent we mean such terms which cannot be expressed through the remaining ones as their linear combination with scalar (i.e., depending only on u^2, v^2, and $u \cdot v$, and their powers) coefficients. The basic identity which will be needed for the separation of the independent bipolar harmonics is easily obtained as the relation between the two coupling schemes $((st) l_1 (st') l_2; lm)$ and $((ss) 0 (tt') l; lm)$ of the four momenta s, t, s, and t':

$$\begin{aligned} &\{ Y_s(u) Y_s(v) \}_{00} \{ Y_t(u) Y_{t'}(v) \}_{lm} \\ &= \sum_{l_1 l_2} \hat{l}_1 \hat{l}_2 \hat{l} \begin{Bmatrix} s & s & 0 \\ t & t' & l \\ l_1 & l_2 & l \end{Bmatrix} \alpha(stl_1) \alpha(st' l_2) \{ Y_{l_1}(u_1) Y_{l_2}(v) \}_{lm}. \end{aligned} \tag{13}$$

In deriving (13) we have used the expansion

$$Y_{m_s}^s(u) Y_{m_t}^t(u) = \sum_{l_1 m_1} \alpha(stl_1) \begin{bmatrix} s & t & l_1 \\ m_s & m_t & m_1 \end{bmatrix} Y_{m_1}^{l_1}(u), \tag{14}$$

$$\alpha(stl_1) = \frac{\hat{s}\hat{t}}{\hat{l}_1 \sqrt{4\pi}} \begin{bmatrix} s & t & l_1 \\ 0 & 0 & 0 \end{bmatrix},$$

where $\hat{a} = (2a+1)^{1/2}$.

In order to construct a set of independent bipolar harmonics one should express (where possible) the given harmonics $\{(\bar{l}_1 \bar{l}_2) lm\}$ with "large" value of $\bar{l}_1 + \bar{l}_2$ through ones with $l_1 + l_2 < \bar{l}_1 + \bar{l}_2$. Using (13), it can be done in the following manner. Take $s > 0$ and define $t = \bar{l}_1 - s \geqslant 0$, $t' = \bar{l}_2 - s \geqslant 0$. Now Eq. (13) can be rewritten in the form

$$\begin{aligned} &\hat{\bar{l}}_1 \hat{\bar{l}}_2 \hat{l} \begin{Bmatrix} s & s & 0 \\ \bar{l}_1 - s & \bar{l}_2 - s & l \\ \bar{l}_1 & \bar{l}_2 & l \end{Bmatrix} \alpha(s\bar{l}_1 - s\bar{l}_1) \alpha(s\bar{l}_2 - s\bar{l}_2) \\ &\times \{ Y_{\bar{l}_1}(u) Y_{\bar{l}_2}(v) \}_{lm} \\ &= \{ Y_s(u) Y_s(v) \}_{00} \{ Y_{\bar{l}_1 - s}(u) Y_{\bar{l}_2 - s}(v) \}_{lm} \end{aligned} \tag{13'}$$

$$- \sum_{l_1 l_2}{}' \hat{l}_1 \hat{l}_2 \hat{l} \begin{Bmatrix} s & s & 0 \\ \bar{l}_1 - s & \bar{l}_2 - s & l \\ l_1 & l_2 & l \end{Bmatrix} \alpha(s\bar{l}_1 - sl_1)\alpha(s\bar{l}_2 - sl_2) \times \{ Y_{l_1}(u) Y_{l_2}(v) \}_{lm},$$

where the prime on the summation symbol reminds us that the term with $l_1 = \bar{l}_1$, $l_2 = \bar{l}_2$ (i.e., maximal value of $l_1 + l_2$) has been extracted and written explicitly on the lhs. Now it is easy to see that if there exists such value of $s > 0$ (certainly it is more suitable to choose s as large as possible) that $\bar{l}_1 + \bar{l}_2 - 2s \geqslant 1$, then the harmonics $\{\bar{l}_1 \bar{l}_2)lm\}$ can be expressed as requested above, with $\{(ss)00\}$ playing the role of scalar coefficient. Simulataneously one sees that Eq. (13′) breaks down in the other cases and there exist

$$l \text{ harmonics with } \bar{l}_1 + \bar{l}_2 = l + 1, \text{ and}$$
$$l + 1 \text{ harmonics with } \bar{l}_1 + \bar{l}_2 = l, \tag{15}$$

which can not be further decomposed. We have shown that the infinite-dimensional set (12) always contains just $2l + 1$ terms which are independent in the above sense.

Concerning the case of the spherical harmonics in n variables it should be noted, that further generalization of (12) is actually not needed. In physical applications which we have in mind we always deal with the four-dimensional Minkowski space. In this space there are only four independent vectors, that is the timelike vector $Q_\mu = (\mathbf{0}, iQ_0)$ and three spacelike vectors, e.g., u_μ, v_μ, and $\epsilon_{\mu\alpha\beta\gamma} Q_\alpha u_\beta v_\gamma$. Any other vector can be obtained as their linear combination with scalar coefficients. Therefore, the prescriptions (12) and (15) are sufficient to obtain the multivariable spherical harmonics as well.

4. TENSOR SPHERICAL HARMONICS

The second-order tensor spherical harmonics in two variables are defined as

$$T^{(l_1 l_2)lJM}_{(r_1 r_2)r\mu\lambda}(u,v,Q) = \sum_{mn} \begin{bmatrix} l & r & J \\ m & n & M \end{bmatrix} \{ Y_{l_1}(u) Y_{l_2}(v) \}_{lm} \times t^{(r_1 r_2)rn}_{\mu\lambda}(Q). \tag{16}$$

The generalization to the tensor harmonics of an arbitrary higher order is straightforward.

Using the properties (6) and (7) of the basis tensor $t^{(r_1 r_2)rn}_{\mu\lambda}$ and those of the two-variable spherical harmonics (12) we may easily see that

(i) the tensors $T^{(l_1 l_2)lJM}_{(r_1 r_2)r\mu\lambda}$ for $l_1 + l_2 = l, l + 1$ form a set of independent Lorentz covariants,

(ii) they satisfy the identity

$$T^{(l_1 l_2)lJM}_{(r_1 r_2)r\mu\lambda}(-u, -v, Q) = (-1)^{l_1 + l_2} T^{(l_1 l_2)lJM}_{(r_1 r_2)r\mu\lambda}(u,v,Q). \tag{17}$$

Now, recalling the definitions of tensor

$$V_{\mu\lambda}(u,v) \xrightarrow{P} (-1)^{\delta_{\mu 0} + \delta_{\lambda 0}} V_{\mu\lambda}(-u, -v)$$

and pseudotensor

$$A_{\mu\lambda}(u,v) \xrightarrow{P} (-1)^{\delta_{\mu 0} + \delta_{\lambda 0} + 1} A_{\mu\lambda}(-u, -v)$$

operators, we can easily see that the tensor harmonics (16) transform under the parity operation P-like tensors (pseudotensors) if the sum $l_1 + l_2 + r_1 + r_2$ is even (odd).

(iii) $T^{(l_1 l_2)lJM}_{(r_1 r_2)r\mu\lambda}$ are othonormal on the unit sphere $S^2(e)$ embedded into the space $E^3(e)$ orthogonal to the vector e^{00}_μ:

$$\int d\Omega_u d\Omega_v (T^{(l_1 l_2)lJM}_{(r_1 r_2)r\mu\lambda})^* T^{(l'_1 l'_2)l'J'M'}_{(r'_1 r'_2)r'\mu\lambda} = \delta_{JJ'} \delta_{MM'} \delta_{l_1 l'_1} \delta_{l_2 l'_2} \delta_{ll'} \delta_{r_1 r'_1} \delta_{r_2 r'_2} \delta_{rr'}. \tag{18}$$

(iv) the scalar product of $T^{\cdots JM}_{\cdots\mu\lambda}$ in the four-dimensional space is

$$T^{(l_1 l_2)lJM}_{(r_1 r_2)r\mu\lambda} \cdot T^{(l'_1 l'_2)l'J'M'}_{(r'_1 r'_2)r'\mu\lambda} = \frac{\hat{l}\hat{l}'\hat{J}\hat{J}'}{4\pi} \hat{l}_1 \hat{l}_2 \hat{l}'_1 \hat{l}'_2 \delta_{r_1 r'_1} \delta_{r_2 r'_2} \delta_{rr'} \sum_{\substack{stx \\ m_s m_t m_x}} \begin{bmatrix} l_1 & l'_1 & s \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} l_2 & l'_2 & t \\ 0 & 0 & 0 \end{bmatrix} \times \begin{bmatrix} s & t & x \\ m_s & m_t & m_x \end{bmatrix} \begin{bmatrix} J & J' & x \\ M & M' & m_x \end{bmatrix} \times W(lrxJ';Jl') \begin{Bmatrix} l_1 & l_2 & l \\ l'_1 & l'_2 & l' \\ s & t & x \end{Bmatrix} Y^s_{m_s}(u) Y^t_{m_t}(v), \tag{19}$$

where $W(\cdots;\cdot\cdot)$ denotes the Racah coefficient, and

(v) the number $N^{(s)}_J$ of independent tensor harmonics of an arbitrary order s is

$$N^{(s)}_J = \sum_r n^{(s)}_r \sum_{l = |J - r|}^{J + r} (2l + 1) = \sum_r n^{(s)}_r (2r + 1)(2J + 1), \tag{20}$$

where $n^{(s)}_r$ is the statistical weight of the corresponding basis tensor. E.g., for the second-order ($s = 2$) spherical harmonics (19) with $t^{(r_1 r_2)rn}_{\mu\lambda}$ we have $n_0{}^{(2)} = 2$, $n_1{}^{(2)} = 3$, and $n_2{}^{(2)} = 1$; therefore

$$N^{(2)}_J = 16(2J + 1). \tag{21}$$

Note that while it seems to be actually impossible to count the number of independent Lorentz covariants when working with the Cartesian forms, in the spherical basis the result (20) has been obtained in a very natural and elegant way.

5. TENSOR HARMONICS IN THE HELICITY BASIS

The helicity basis state analogous to those of Jacob and Wick[7] are introduced here in a slightly formal way, and therefore, the interpretation in terms of conserved quantities may be lost. Nevertheless, the technique has proved to be very helpful when a particular direction may be chosen according to the nature of the physical problem. We keep calling helicities the projections of the (internal) angular momentum (vectors e^{1n}_μ of the basic tetrads) on the directions $\vartheta_1 \varphi_1$ and $\vartheta_2 \varphi_2$ connected with the scalar harmonics (8)

$$Y^l_m(u) = \frac{\hat{l}}{\sqrt{4\pi}} D^{l*}_{m0}(\varphi_1 \vartheta_1 0), \tag{22}$$

and

$$Y^l_m(v) = \frac{\hat{l}}{\sqrt{4\pi}} D^{l*}_{m0}(\varphi_2 \vartheta_2 0) \tag{23}$$

considered above. The one-variable tensor harmonics in the helicity basis have been discussed in substantial detail by Akyeampong.[2] We shall follow his definitions and notation where possible.

In our case the second order helicity tensor basis may depend on three variables Q_λ, u_λ, and v_λ. [The dependence on Q_λ comes through ϵ_λ^{00} just as in the case of the spherical basis (5).] Indeed, we can write[5]

$$e_\lambda^{1n} = \sum_k D_{mk}^{1*}(\varphi_1\vartheta_1 0)\epsilon_\lambda^{1k}(u), \tag{24}$$

$$e_\lambda^{1n} = \sum_{k'} D_{nk'}^{1*}(\varphi_2\vartheta_2 0)\epsilon_\lambda^{1k'}(v), \tag{25}$$

and consider any combination of the tetrads $\epsilon_\lambda^{rk}(u)$ and $\epsilon_\lambda^{rk}(v)$. The particular choice of the tensor helicity basis certainly depends on the character of the physical problem to be solved. Taking the helicity basis in the form $\epsilon_\mu^{r_1k_1}(u)\epsilon_\lambda^{r_2k_2}(v)$, we can define the helicity-basis tensor harmonics

$$\begin{aligned} S_{j_2r_2k_2\mu\lambda}^{j_1r_1k_1JM}&(u,v,Q) \\ &= \frac{\hat{j}_1\hat{j}_2}{4\pi}\sum_{m_1'm_2'}\begin{bmatrix} j_1 & j_2 & J \\ m_1' & m_2' & M \end{bmatrix} \\ &\times D_{m_1'k_1}^{j_1*}(\varphi_1\vartheta_1 0)D_{m_2'k_2}^{j_2*}(\varphi_2\vartheta_2 0)\epsilon_\mu^{r_1k_1}(u)\epsilon_\lambda^{r_2k_2}(v), \end{aligned} \tag{26}$$

which are in the following way connected with the tensor spherical harmonics:

$$\begin{aligned} T_{(r_1r_2)r\mu\lambda}^{(l_1l_2)lJM} = \hat{l}_1\hat{l}_2\hat{l}\hat{r}\sum_{j_1k_1}&\begin{Bmatrix} l_1 & l_2 & l \\ r_1 & r_2 & r \\ j_1 & j_2 & J \end{Bmatrix}\begin{bmatrix} l_2 & r_2 & j_2 \\ 0 & k_1 & k_2 \end{bmatrix} \\ &\times\begin{bmatrix} l_2 & r_2 & j_2 \\ 0 & k_2 & k_2 \end{bmatrix} S_{j_2r_2k_2\mu\lambda}^{j_1r_1k_1JM}(u,v,Q). \end{aligned} \tag{27}$$

The tensor helicity basis can also be chosen to depend on two variables, say Q_λ and u_λ only: $\epsilon_\mu^{r_1k_1}(u)\epsilon_\lambda^{r_2k_2}(u)$. The corresponding tensor harmonics will have the following form

$$\begin{aligned} U_{(r_1r_2)rk\mu\lambda}^{jl_2\,JM}(u,v,Q) &= \frac{\hat{j}\hat{l}_2}{4\pi}\sum_{m'm_2}\begin{bmatrix} j & l_2 & J \\ m' & m_2 & M \end{bmatrix} \\ &\times D_{m'k}^{j*}(\varphi_1\vartheta_1 0)D_{m_20}^{l_2*}(\varphi_2\vartheta_2 0)\sum_{k_1k_2}\begin{bmatrix} r_1 & r_2 & r \\ k_1 & k_2 & k \end{bmatrix} \\ &\times\epsilon_\mu^{r_1k_1}(u)\epsilon_\lambda^{r_2k_2}(u). \end{aligned} \tag{28}$$

Now we list some useful properties of S and U:

(i) The harmonics (26) and (28) form the sets of independent (see Sec. 3) Lorentz covariants if just $2J+1$ different pairs j_1j_2 (jl_2) are taken for each combination of $r_1k_1r_2k_2(r_1r_2rk)$. The reduction formula which allows us to fix these restrictions and thus to deal with the independent covariants only can be easily obtained in the same manner as Eq. (13). It reads

$$\begin{aligned} &\frac{4\pi}{\hat{s}^2}\{(ss)00\}\sum_{\tau_1\tau_2}\begin{bmatrix} t_1 & t_2 & J \\ \tau_1 & \tau_2 & M \end{bmatrix}D_{\tau_1\beta_1}^{t_1*}(\varphi_1\vartheta_1 0)D_{\tau_2\beta_2}^{t_2*}(\varphi_2\vartheta_2 0) \\ &\quad = \sum_{j_1j_2}\hat{j}_1\hat{j}_2\hat{J}\begin{Bmatrix} s & s & 0 \\ t_1 & t_2 & J \\ j_1 & j_2 & J \end{Bmatrix}\begin{bmatrix} s & t_1 & j_1 \\ 0 & \beta_0 & \beta_1 \end{bmatrix}\begin{bmatrix} s & t_2 & j_2 \\ 0 & \beta_2 & \beta_2 \end{bmatrix}\sum_{m_1m_2}\begin{bmatrix} j_1 & j_2 & J \\ m_1 & m_2 & M \end{bmatrix}D_{m_1\beta_1}^{j_1*}(\varphi_1\vartheta_1 0)D_{m_2\beta_2}^{j_2*}(\varphi_2\vartheta_2 0). \end{aligned} \tag{29}$$

(ii) The parity operator acts on them according to

$$S_{j_2r_2k_2\mu\lambda}^{j_1r_1k_1JM}(-u,-v,Q) = (-1)^{j_1+j_2+r_1+r_2}S_{j_2r_2-k_2\mu\lambda}^{j_1r_1-k_1JM}(u,v,Q), \tag{30}$$

$$U_{(r_1r_2)rk\mu\lambda}^{jl_2\,JM}(-u,-v,Q) = (-1)^{j+l_2+r}U_{(r_1r_2)r-k\mu\lambda}^{jl_2\,JM}(u,v,Q). \tag{31}$$

(iii) The S and U harmonics are orthonormal if integrated with $d\Omega_i = \sin\vartheta_i d\vartheta_i d\varphi_i$; e.g.,

$$\int d\Omega_{\mathbf{u}}d\Omega_{\mathbf{v}}(S_{j_2r_2k_2\mu\lambda}^{j_1r_1k_1JM})^* S_{j_2'r_2'k_2'\mu\lambda}^{j_1'r_1'k_1'J'M'} = \delta_{JJ'}\delta_{MM'}\delta_{j_1j_1'}\delta_{j_2j_2'}\delta_{r_1r_1'}\delta_{r_2r_2'}\delta_{k_1k_1'}\delta_{k_2k_2'}, \tag{32}$$

(iv) The scalar product is

$$\begin{aligned} &S_{j_2r_2k_2\mu\lambda}^{j_1r_1k_1JM}\cdot S_{j_2'r_2'k_2'\mu\lambda}^{j_1'r_1'k_1'J'M'} \\ &\quad = \frac{\hat{J}\hat{J}'}{4\pi}\hat{j}_1\hat{j}_2\hat{j}_1'\hat{j}_2'(-1)^{r_1+r_2+k_1+k_2}\delta_{r_1r_1'}\delta_{r_2r_2'}\delta_{-k_1k_1'}\delta_{-k_2k_2'} \\ &\quad\quad\times\sum_{\substack{J_iM_i \\ xm_x}}\begin{bmatrix} J_1 & J_2 & x \\ M_1 & M_2 & m_x \end{bmatrix}\begin{bmatrix} J & J' & x \\ M & M' & m_x \end{bmatrix}\begin{bmatrix} j_1 & j_1' & J_1 \\ k_1 & -k_1 & 0 \end{bmatrix}\begin{bmatrix} j_2 & j_2' & J_2 \\ k_2 & -k_2 & 0 \end{bmatrix}\begin{Bmatrix} j_1 & j_2 & J \\ j_1' & j_2' & J' \\ J_1 & J_2 & x \end{Bmatrix}Y_{M_1}^{J_1}(u)Y_{M_2}^{J_2}(v). \end{aligned} \tag{33}$$

The scalar product of the U harmonics has a similar expression; we do not display it.

6. CONCLUSION

In the present paper an extension of the relativistic tensor harmonics to the case of two variables has been suggested. They appear to be rather compact and very flexible in comparison with the Cartesian forms which were usually employed in the applications. To give just one example, consider the reaction of the radiative muon capture on atomic nuclei[8,9]:

$$\mu^- + A(J_i,p_i) \to \nu + \gamma(k) + B(J_f,p_f).$$

Introducing the momenta $Q_\lambda = (p_f + p_i)_\lambda$ and $q_\lambda = (p_f - p_i)_\lambda$ we can fix the general form of the corresponding weak hadronic matrix elements in the elegant form which is also easy to manipulate

$$\begin{aligned} T_{\mu\lambda} = \sum_{l_irlrJM}\hat{J}_i^{-1}&\begin{bmatrix} J_i & J & J_f \\ M_i & M & M_f \end{bmatrix} \\ &\times F_{(r_1r_2)r}^{(l_1l_2)lJ}(k,q,Q)T_{(r_1r_2)r\mu\lambda}^{(l_1l_2)lJM}(k,q,Q), \end{aligned} \tag{34}$$

where $F^{(l_1 l_2)lJ}_{(r_1 r_2)r}(k,q,Q)$ are form factors. The structure of this tensor in the Cartesian covariants is considerably more complicated and for different values of J_i and J_f must be constructed individually. In addition it is not always easy to eliminate the dependent covariants in the Cartesian basis. For example, for the reaction $\mu^- + p \rightarrow \nu + \gamma + n$ Hwang and Primakoff[9] write

$$T_{\mu\lambda} = -\bar{u}(p_f)\{F_1\gamma_\mu\gamma_\lambda + \gamma_\mu\gamma_\nu k_\nu (k_\lambda F_2 + q_\lambda F_3 + Q_\lambda F_4) + \dots + \gamma_5[F_{35}\gamma_\mu\gamma_\lambda + \gamma_\mu\gamma_\nu k_\nu (k_\lambda F_{36} + q_\lambda F_{37} + Q_\lambda F_{38}) + \dots]\}\, u(p_i), \quad (35)$$

where 68 covariants appear which contain the Dirac matrices γ_λ, and the momenta $k_\lambda q_\lambda$, and Q_λ. $F_i(k,q,Q)$ are form factors. Comparing with Eq. (21) one realizes that only 64 independent covariants exist in this case. To eliminate superfluous Cartesian covariants in (35) one has to derive the corresponding four coupling equations. The present authors have obtained them[10] using the symbol manipulating computer program SCHOONSCHIP. The relations are very complicated; each of four equations couples 26 individual covariants. The further possible applications of the tensor harmonics for the description of the radiative muon capture reaction are presently being studied.

ACKNOWLEDGMENTS

We are indebted to J. A. Smordinskij and R. A. Eramzhyan for discussions and for reading the manuscript. We wish to thank the referee for valuable comments and for bringing papers of Ref. 6 to our attention.

[1]M. Daumens and P. Minnaert, J. Math. Phys. **17**, 2085 (1976).
[2]D. A. Akyeampong, J. Math. Phys. **20**, 505 (1979).
[3]B. Stech and L. Schülke, Z. Phys. **179**, 314 (1964).
[4]J. Delorme, in *Mesons in Nuclei*, Vol. I, edited by M. Rho and D. Wilkinson (North-Holland, Amsterdam, 1979), pp. 107–147.
[5]D. M. Brink and G. R. Satchler, *Angular Momentum* (Clarendon, Oxford, 1968); E. Wigner, *Group Theory* (Academic, New York, 1959); D. A. Varšalovič, A. N. Moskalov, and V. K. Khersonskij, *Quantum Theory of Angular Momenta* (Nauka, Leningrad, 1975), in Russian.
[6]L. C. Biedenharn, J. Math. and Phys. (Cambridge, Mass.) **31**, 287 (1953); R. J. Buehler and J. O. Hirschfelder, Phys. Rev. **83**, 628 (1951).
[7]M. Jacob and G. C. Wick, Ann. Phys. (NY) 7, 404 (1959).
[8]M. Gmitro and A. A. Ovchinnikova, JINR Reports E2-12639 and E2-12640, Dubna, 1979.
[9]W. Y. P. Hwang and H. Primakoff, Phys. Rev. C **18**, 414 (1978).
[10]M. Gmitro and A. A. Ovchinnikova, JINR Preprint E2-80-123, Dubna, 1980.

A 4-vector generalization of the sine–Gordon equation

D. G. Swanson
Auburn University, Auburn, Alabama 36849

(Received 14 April 1981; accepted for publication 10 August 1981)

Using the differential operators from the Dirac equations, an algebra is developed which leads to 4-vector functions and the generalization of many scalar functions. Assuming particles to be described by the potential solutions of the generalized sine–Gordon equation (a set of four coupled nonlinear equations), a single soliton is shown to be localized within a light sphere and have intrinsic properties of group velocity, phase velocity, angular momentum, and wave-particle duality.

PACS numbers: 11.30.Na, 12.20.Hx

In the search for a possible soliton description of the particles of physics, the form of the Dirac equations offers a tantalizing suggestion that they are part of an inverse scattering scheme, but they offer no firm direction about the underlying soliton equation. Examining the form of the basic operators of the Dirac equations may guide us, however, in the search for the soliton equation which may describe our universe. In what follows, the Dirac equations are written in a form which bears a superficial resemblance to the AKNS equations,[1] and the operators lead to a set of 4-vector variables and 4-vector functions which suggest a 4-vector generalization of the sine–Gordon equation which appears to have enough interesting properties to merit close scrutiny.

In a fashion similar to what is used with the one-dimensional sine–Gordon equation, the basic variables are sums and differences of the $\mathbf{r},t$ variables such that (with $c = 1$) $\xi = \frac{1}{2}(z + t)$, $\tau = \frac{1}{2}(z - t)$, $\mu = \frac{1}{2}(x + iy)$, $\nu = \frac{1}{2}(x - iy)$. With these variables, the Dirac equations may be written as

$$(D^+ - iA^+)\psi = -im\phi, \qquad (D^- - iA^-)\phi = im\psi, \tag{1}$$

where $D^\pm$ are derivative operators given by

$$D^+ = \begin{pmatrix} \partial_\xi & \partial_\mu \\ \partial_\nu & \partial_\tau \end{pmatrix}, \quad D^- = \begin{pmatrix} \partial_\tau & \partial_\mu \\ \partial_\nu & \partial_\xi \end{pmatrix}, \tag{2}$$

and $A^\pm$ are representations of the normalized 4-vector potential

$$A^+ = \begin{pmatrix} A_\xi & A_\mu \\ A_\nu & -A_\tau \end{pmatrix},$$
$$A^- = \begin{pmatrix} A_\tau & A_\mu \\ A_\nu & -A_\xi \end{pmatrix} = \begin{pmatrix} A_z - A_4 & A_x - iA_y \\ A_x + iA_y & -A_z - A_4 \end{pmatrix}, \tag{3}$$

and A_4 is the scalar potential. The rest mass is m, and ψ and ϕ are column vectors. The form of Eqs. (1) are reminiscent of, but not equivalent to, the AKNS equations in that they involve a pair of first-order linear differential equations with an eigenvalue term (the mass) and a potential term, but the potential and mass terms are interchanged from the familiar AKNS equations where the potentials couple the equations. It appears a gauge transformation can eliminate the $A^\pm$ terms, and additional coupling potentials can be added on the right if they are taken to vanish rapidly enough away from a particle, but the eigenvalue coupling is still different.

The form of the differential operator suggests a new algebra, however, which may lead to a modified set of equations. The basic 4-vector variables associated with the operators $D^\pm$ are

$$R^+ = \begin{pmatrix} \xi & \nu \\ \mu & -\tau \end{pmatrix}, \quad R^- = \begin{pmatrix} \tau & \nu \\ \mu & -\xi \end{pmatrix} \tag{4}$$

and the "scalar product" is

$$R^+R^- = R^-R^+ \equiv R^2 I,$$

where $R^2 = \xi\tau + \mu\nu = x^2 + y^2 + z^2 - t^2$ and I is the unit 2×2 matrix. The differentiation rules are

$$\begin{aligned} &D^+R^+ = D^-R^- = 2I, \\ &D^+\tilde{R}^- = D^-\tilde{R}^+ = \tilde{D}^+R^- = \tilde{D}^-R^+ = 0, \\ &D^+\tilde{R}^+ = D^-\tilde{R}^- = D^+R^- = D^-R^+ = I, \\ &D^\pm R^2 I = R^\mp. \end{aligned} \tag{5}$$

The transformation laws for 4-vectors of the form (3) and (4) are

$$A^{\pm\prime} = T^\pm A^\pm T^\pm, \tag{6}$$

where for a frame moving in the z direction at speed $v = \beta c$,

$$T^\pm = \begin{pmatrix} f_\pm & 0 \\ 0 & f_\mp \end{pmatrix}, \tag{7}$$

and $f_\pm = [(1 \pm \beta)/(1 \mp \beta)]^{1/4}$ so $f_\pm^2 = \gamma(1 \pm \beta)$ and $f_+ f_- = 1$. Field vectors or axial 4-vectors such as

$$F_\pm = D^\pm A^\pm \tag{8}$$

transform as

$$F'_\pm = T^\pm F_\pm T^\mp. \tag{9}$$

In order to construct functions of $R^\pm$ which will transform properly, it is necessary to require that repeated products of $R^\pm$ must alternate, since $T^+T^- = I$, so that a function is of the form

$$f(R^+) = a_0 I + a_1 R^+ + a_2 R^+R^- + a_3 R^+R^-R^+ + a_4 R^+R^-R^+R^- + \cdots.$$

From the differentiation rules (5) it may be established that

$$D^\pm(R^{2n})I = nR^{2n-2}R^\mp, \qquad D^\pm(R^{2n}R^\pm) = (n+2)R^{2n}I. \tag{10}$$

These relations allow us to define a set of special functions which are analogous to one-dimensional functions, but not equivalent.

(1) $\mathrm{EXP}(R^+)$ is a solution of

$$D^+\mathrm{EXP}(R^+) = \mathrm{EXP}(R^-), \tag{11}$$

$$\mathrm{EXP}(R^+) = \frac{I}{0!1!} + \frac{R^+}{0!2!} + \frac{R^2I}{1!2!} + \frac{R^2R^+}{1!3!} + \cdots + \frac{R^{2n}I}{n!(n+1)!} + \frac{R^{2n}R^+}{n!(n+2)!} + \cdots. \tag{12}$$

(2) $\mathrm{SIN}(R^+)$ and $\mathrm{COS}(R^+)$ are solutions of

$$D^+D^-F(R^\pm) = -F(R^\pm), \tag{13}$$

$$\mathrm{SIN}(R^\pm) = R^\pm\left(\frac{1}{0!2!} - \frac{R^2}{1!3!} + \frac{R^4}{2!4!} - \cdots + \frac{(-1)^nR^{2n}}{n!(n+2)!} + \cdots\right), \tag{14}$$

$$\mathrm{COS}(R^\pm) = I\left(\frac{1}{0!1!} - \frac{R^2}{1!2!} + \frac{R^4}{2!3!} - \cdots + \frac{(-1)^nR^{2n}}{n!(n+1)!}\right), \tag{15}$$

so that

$$\mathrm{EXP}(iR^+) = \mathrm{COS}(R^+) + i\,\mathrm{SIN}(R^+), \tag{16}$$

$$\mathrm{EXP}(ER^+) = \mathrm{COS}(R^+) + E\,\mathrm{SIN}(R^+), \tag{17}$$

where E is the 2×2 "imaginary matrix,"

$$E = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \qquad EE = -I,$$

also

$$D^+\,\mathrm{SIN}(R^+) = \mathrm{COS}(R^+), \tag{18}$$

$$D^+\,\mathrm{COS}(R^+) = -\mathrm{SIN}(R^-).$$

These special functions may be written in terms of Bessel functions, such as

$$\mathrm{EXP}(iR^+) = IJ_1(2R)/R + iR^+J_2(2R)/R^2 \tag{19}$$

so they represent a four-dimensional extension of the sequence

$$\cos(x) \rightarrow J_0(\rho) \rightarrow (\sin r)/r \rightarrow J_1(2R)/R$$

as the dimensionality progresses from 1 through 4. The relationship (18) is apparently unique to one and four dimensions, however. This fact is the basis for the suggested form of a valid 4-vector soliton equation which follows.

(3) The $\mathrm{SIN}(R^+)$ and $\mathrm{COS}(R^+)$ functions are merely the first two of an infinite set of functions which are solutions of

$$\begin{aligned} &D^+D^-F_{2n} + [1 - n(n+1)/R^2]F_{2n} = 0, \\ &D^+D^-F_{2n+1} + [1 - n(n+2)/R^2]F_{2n+1} = 0, \end{aligned} \tag{20}$$

which satisfy the recursion formulas

$$\begin{aligned} D^+F_{2n}(R^+) &= (2n+1)^{-1} \times [nF_{2n-1}(R^-) - (n+1)F_{2n+1}(R^-)], \\ D^+F_{2n+1}(R^+) &= [2(n+1)]^{-1} \times [(n+2)F_{2n}(R^+) - nF_{2n+2}(R^+)], \end{aligned} \tag{21}$$

and the sum rule

$$I = \sum_{n=0}^{\infty}(n+1)^2F_n(R^+)F_n(R^-). \tag{22}$$

The solutions may be written

$$F_{2n}(R^+) = IR^{2n}\sum_{k=0}^{\infty}\frac{(-1)^kR^{2k}}{k!(2n+k+1)!} = IJ_{2n+1}(2R)/R,$$

$$F_{2n+1}(R^+) = R^+R^{2n}\sum_{k=0}^{\infty}\frac{(-1)^kR^{2k}}{k!(2n+k+2)!} \tag{23}$$

$$= R^+J_{2n+2}(2R)/R^2.$$

For more general arguments than $R^\pm$, we construct $X^\pm$, where X^+ is linear in ξ, τ, μ, ν. Then if U^+ is a general 4-vector, $f(X^2)$ is a scalar function of X^2, where $X^+X^- = IX^2$, then

$$\begin{aligned} &D^+fI = f'D^+X^2I, \\ &D^+[fU^+] = fD^+U^+ + f'(D^+X^2I)U^+, \end{aligned} \tag{24}$$

$$\begin{aligned} D^+D^-f &= f'D^+D^-X^2I + f''(D^+X^2I)(D^-X^2I), \\ D^+D^-(fX^+) &= f'\{D^+[(D^-X^2I)X^+] + (D^+X^2I)D^-X^+\} \\ &\quad + f''(D^+X^2)(D^-X^2)X^+. \end{aligned} \tag{25}$$

A sufficient condition that $D^+D^-U(X^+)$ "point" in the "direction" of X^+, i.e.,

$$D^+D^-[f(X^2)X^+] = g(X^2)X^+$$

is that $X^2 = m^2R^2$ with m a constant. Some nontrivial examples of $X^\pm$ are in Table I, where the constant 4-vector $P^\pm$ satisfies

$$P^+P^- + m^2I = 0, \tag{26}$$

and the $X_1^\pm$ in the last relations represent any of the first four expressions. Note that $X^- = E\tilde{X}^+E$.

We can then generalize Eqs. (20) such that

$$D^+D^-F + [m^2 - n(n+2)/R^2]F = 0 \tag{27}$$

TABLE I. Examples of nontrivial $X^\pm$ with the property $X^+X^- \propto IR^2$. $X_1^\pm$ represents any of the $X^\pm$ pairs in the first four rows with α, β scalars and $P^+P^- = -m^2I$.

X^+	X^-	X^2
P^+R^+	$-R^\pm P^\mp$	m^2R^2
$P^\pm\tilde{R}^\mp$	$-\tilde{R}^\pm P^\mp$	m^2R^2
$\tilde{P}^+R^\mp$	$-R^\pm\tilde{P}^\mp$	m^2R^2
$\tilde{P}^\pm\tilde{R}^\mp$	$-\tilde{R}^\pm\tilde{P}^\mp$	m^2R^2
$\alpha X_1^+ + \beta EX_1^+$	$\alpha X_1^- - \beta X_1^-E$	$X_1^2(\alpha^2+\beta^2)$
$\alpha X_1^+ + \beta X_1^+E$	$\alpha X_1^- - \beta EX_1^-$	$X_1^2(\alpha^2+\beta^2)$
$X_1^+ + \alpha EX_1^+ + \beta X_1^+E + \alpha\beta\tilde{X}_1^-$	$X_1^- - \alpha X_1^-E - \beta EX_1^- + \alpha\beta\tilde{X}_1^+$	$X_1^2(1+\alpha^2+\beta^2+\alpha^2\beta^2)$

has the solution

$$F = F_{2n+1}(X^+).$$

Since $D^+D^- = (\nabla^2 - \partial_t^2)I$, this is a 4-vector Klein–Gordon equation.

Maxwell's equations appear in an appealing form in this algebra. The field vectors are derived simply from the potentials by

$$F_\pm = D^\pm A^\pm. \tag{28}$$

and Maxwell's equations are

$$D^\pm F_\mp + J^\mp = 0 \tag{29}$$

with $J^\pm$ being the 4-vector current (normalized)

$$J^+ = \begin{pmatrix} J_\xi & J_\mu \\ J_\nu & -J_\tau \end{pmatrix} = \begin{pmatrix} J_z + \rho & J_x - iJ_y \\ J_x + iJ_y & -J_z + \rho \end{pmatrix}.$$

The $D^\pm$ operators have the (vector, scalar) properties

$$D^+A^\pm = \begin{cases} i\nabla\times\mathbf{A} \pm \nabla A_4 + \partial_t \mathbf{A} & \text{(vector portion)} \\ \nabla\cdot\mathbf{A} \pm \partial_t A_4 & \text{(scalar portion)}. \end{cases}$$

In view of the remarkable properties of the algebra, and the result (18), a generalized 4-vector form of the sine–Gordon equation

$$D^+D^-U^+ = -\mathrm{SIN}(U^+) \tag{30}$$

is proposed as a suitable soliton equation candidate. Among the reasons for the choice are (a) the sine–Gordon solutions are relativistically invariant; (b) the sine–Gordon equation is gauge invariant; (c) the result (18) appears to be a sufficient condition that an infinite number of conservation laws may be generated by a generalization of the method of Lamb,[2] using the Lagrangian

$$L = \tfrac{1}{2}U^+U^- + I - \mathrm{COS}\,U^\pm. \tag{31}$$

By gauge invariance, we mean that if the AKNS equations are written with diagonal potential terms,

$$\psi_z - iv^+\psi + i\zeta\psi = q\phi,$$

$$\phi_z - iv^-\phi - i\zeta\phi = r\psi,$$

where $v^\pm = v_1 \pm v_2$, then if $q = \frac{1}{2}u_z e^{i\theta}$ and $r = -q^*$, then u satisfies

$$u_{zt} = \sin u \tag{32}$$

if $\partial_t v^+ = 0$ and $\theta = 2\int^z v_2 dz'$. In other words, one may solve for u without regard to $v^\pm$ as long as $v^\pm$ satisfy certain gauge conditions.

Although no analytic solutions of Eq. (30) are known, a series solution of the form

$$\begin{aligned} U^+ = R^+P^-\Big[& \frac{1}{0!2!} - \frac{2}{1!3!}\left(\frac{mR}{2}\right)^2 + \frac{6}{2!4!}\left(\frac{mR}{2}\right)^4 \\ & - \frac{46}{3!5!}\left(\frac{mR}{2}\right)^6 + \frac{2\cdot 347}{4!6!}\left(\frac{mR}{2}\right)^8 \\ & - \frac{2\cdot 41\cdot 563}{3\cdot 5!\cdot 7!}\left(\frac{mR}{2}\right)^{10} + \frac{6\cdot 11\cdot 23\cdot 317}{6!8!}\left(\frac{mR}{2}\right)^{12} \\ & - \frac{2\cdot 30488957}{3\cdot 7!9!}\left(\frac{mR}{2}\right)^{14} + \cdots \Big] \end{aligned} \tag{33}$$

leads to localized solutions in a surprising sense. The function $Q_+ = D^+U^+$ is exponentially decaying outside the light sphere, i.e., $r^2 > t^2$, whereas Q_+ is oscillatory inside the light sphere with the wavelength given by the mass if P^- is the 4-vector momentum. Furthermore, the vector portion of U^+, namely, R^+P^-, has vector–scalar components given by

$$R^+P^- = \begin{cases} i\mathbf{r}\times\mathbf{p} + (\mathbf{p}t - p_4\mathbf{r}) & \text{vector} \\ \mathbf{p}\cdot\mathbf{r} - p_4 t & \text{scalar,} \end{cases} \tag{34}$$

so the real part of the vector portion describes the group velocity, the scalar portion describes the phase velocity, and the imaginary part of the vector portion describes the angular momentum (the spin angular momentum is of course imbedded in the fabric of the algebra). In this formalism, then, a soliton is localized within a light sphere and its effective character is described by a center of motion, a phase motion, and an angular momentum. This concept of a single-particle soliton guarantees causality, but allows a particle to absorb radiation from anywhere inside its light sphere or interact with any other particle inside its light sphere. In evaluating dynamics in this picture, an integral over space will describe the particle "center" as moving with the group velocity and its size will be defined by its mass since that defines its wavelength. More complicated structures may be described by other forms of X^+ as in Table I, which is not exhaustive, or by other solutions of Eq. (30).

In conclusion, it has not been proven that Eq. (30) is a true soliton equation, but it is an appropriate 4-vector generalization of the sine–Gordon equation, and the remarkable property (18) seems to imply that an infinite number of conservation laws are obtainable. These properties, along with the properties of the solution (33) which include causality, phase and group velocity, angular momentum and effective localization due to the oscillation wavelength, seem to demand further investigation. It is also apparent from the structure that this generalization of the sine–Gordon equation is unrelated to other higher dimensional generalizations[3–5] since they deal only with scalar functions which lead to essentially trivial results.[6] This representation is a 4-vector formulation and not simply three dimensions plus time with scalar functions.

ACKNOWLEDGMENT

This research was supported in part by the National Science Foundation under Grant No. ENG76-08290.

[1]M. J. Ablowitz *et al.*, Phys. Rev. Lett. **31**, 125 (1973).
[2]G. L. Lamb, Phys. Lett. A **32**, 251 (1970).
[3]R. L. Anderson *et al.*, Lett. Math. Phys. **3**, 351 (1979).
[4]G. Leibbrandt, Phys. Rev. Lett. **41**, 435 (1978).
[5]Z. Popowicz, Lett. Math. Phys. **3**, 431 (1979).
[6]R. Benguria and K. M. Case, Proc. Natl. Acad. Sci. USA **77**, 4393 (1980).

Linear response theory revisited III: One-body response formulas and generalized Boltzmann equations

M. Charbonneau,[a] K. M. van Vliet, and P. Vasilopoulos
Centre de recherche de mathématiques appliquées, and Départment de Physique, Université de Montréal, Montreal, Québec H3C 3J7, Canada

(Received 27 March 1981; accepted for publication 30 July 1981)

The many-body linear response expressions obtained in previous papers [J. Math. Phys. **19**, 1345 (1978); **20**, 2573 (1979)] are applied to systems of weakly interacting particles. General expressions for the susceptibility and conductivity in such systems are obtained. The diagonal parts depend on the scattering processes, for which we consider interactions with bosons with mass and electron-phonon interaction. For elastic collisions simple closed forms result. For general two-body collisions, the closed expressions are cumbersome, except when the current is due to collisional current through localized states, such as Landau orbits; in that event a generalized Adams-Holstein result is obtained. The nondiagonal electrical conductivity is shown to be of paramount importance for the quantum mechanical Hall effect. We also derive quantum mechanical Boltzmann equations, both for the diagonal occupancy operator $\langle \mathbf{n}_\zeta \rangle_t$ and for the nondiagonal operator $\langle c_{\zeta'}^{+} c_{\zeta''} \rangle_t$. The total Boltzmann equation is shown to be fully equivalent with the linear response results. Finally, in the last part we derive the Boltzmann equation for the Wigner function of inhomogeneous systems. In the classical limit this yields the usual Boltzmann transport equation. This equation has therefore been obtained by first principles from the von Neumann equation.

PACS numbers: 51.10. + y

1. Introduction

In a previous paper,[1] referred to as LRT I, we discussed the Kubo–Green formulas which relate transport coefficients to certain forms of the correlation function of fluctuations about an equilibrium state. It was argued that in Kubo's theory proper no dissipation occurs; this is reflected by the Heisenberg form for the time-dependent operator $B(t)$ of the system, and by zero entropy production. Dissipative behavior was introduced by writing the system Hamiltonian as $H = H^0 + \lambda V$, where H^0 represents the motion proper and λV is the cause of randomizing transitions, such as electron-phonon interactions in an electron-phonon gas. We considered the van Hove limit $\lambda \to 0$, $t \to \infty$, $\lambda^2 t$ finite, which led to an entirely different form of the time behavior for the reduced operators $B^R(t)$. In the subdynamics of H^0 there is now clearcut relaxation, as expressed by the reduced operators

$$K_d^R(t) = e^{-\Lambda_d t} K_d^R(0), \tag{1.1}$$

where $K_d^R(0) \equiv K_d^R = K_d^S \equiv K_d$ is the Schrödinger operator[2] and the subscript "d" denotes the diagonal part in the representation of H^0; Λ_d is the master superoperator in Liouville space, defined by

$$\Lambda_d K = -\sum_{\gamma\gamma''} |\gamma\rangle \langle\gamma| [W_{\gamma''\gamma} \langle\gamma''|K|\gamma''\rangle - W_{\gamma\gamma''} \langle\gamma|K|\gamma\rangle], \tag{1.2}$$

where $|\gamma\rangle$ are the eigenstates of H^0, with energy $\mathscr{E}_\gamma$, and where $W_{\gamma\gamma''}$ is given by the golden rule

$$W_{\gamma\gamma''} = (2\pi\lambda^2/\hbar) |\langle\gamma|V|\gamma''\rangle|^2 \delta(\mathscr{E}_\gamma - \mathscr{E}_{\gamma''}) = W_{\gamma''\gamma}. \tag{1.3}$$

One eigenvalue of Λ_d is zero, determining the equilibrium behavior (see LRT II Sec. 8); the other eigenvalues are positive definite (see Vigfussen[3]), thus governing the approach to equilibrium. The superoperator Λ_d is also written as

$$\Lambda_d K = \sum_\gamma |\gamma\rangle \langle\gamma| M \langle\gamma|K|\gamma\rangle, \tag{1.4}$$

where M is the master operator in the space of functions $F(\gamma)$:

$$MF(\gamma) = -\sum_{\gamma''} [W_{\gamma''\gamma} F(\gamma'') - W_{\gamma\gamma''} F(\gamma)]$$
$$= \sum_{\gamma''} W_{\gamma\gamma''} [F(\gamma) - F(\gamma'')]. \tag{1.5}$$

The response formulas in the subdynamics of H^0 can also be obtained without previous knowledge of the Kubo–Green formulas. To that purpose we applied projection operator techniques to the von Neumann equation for the full density operator; these results were laid down[4] in LRT II. Applying the van Hove limit, we arrived at an inhomogeneous master equation, which is a many-body equation, which does not only contain the relaxation terms of the Pauli master equation but also the coupling to an external field with field Hamiltonian $-AF(t)$, $F(t)$ being an applied generalized force and A the conjugate extensive operator. The solution of the inhomogeneous master equation gave the new response formulas. We also included the nondiagonal part of the many-body operators K in this treatment; the reduced operators, i.e., after the van Hove limit, were found to have the form

$$K^R(t) = e^{-(\Lambda_d - i\mathscr{L}^0)t} K, \tag{1.6}$$

where $\mathscr{L}^0$ is the interaction Liouville operator, $\mathscr{L}^0 K = \hbar^{-1}[H^0, K]$.

The inhomogeneous master equation referred to above

[a] Now at CAE Electronics Ltd., 8585 Côte de Liesse, Montréal, Québec, Canada.

 0022-2488/82/020318-19$02.50

is the many-body counterpart of the Boltzmann equation for one-particle distribution functions; like the Boltzmann equation it contains streaming terms, which represent the effects of an external field, and dissipative terms, which account for the influence of collisions. The main tenet of the new treatment is that the necessary "randomness conditions" are carried out on the many-body level. Thus, no new assumptions are to be introduced when we go to the one-body or two-body level, except closure relations [see LRT II, Eq. (8.1)]. In this respect our treatment differs in essence from the various one-body treatments in the literature which start from a one-particle von Neumann equation, cf. Kohn and Luttinger,[5] Adams and Holstein,[6] Kahn and Frederikse,[7] Argyres,[8] and Argyres and Roth.[9] The one-particle von Neumann equation is not very suitable for a perturbation approach since it is linear, so that it cannot properly arrive at the quadratic (or quartic) Boltzmann collision terms. The treatment of LRT II, on the contrary, led to a quantum mechanical Boltzmann equation with the full collision terms. We still note in this respect that the van Hove limit is equivalent with the first-order Born approximation used by others.[10]

In the present article we shall more fully be concerned with one-body results, derived from the many-body results of the previous articles. To that purpose we consider H^0 to represent the Hamiltonian of a fermion gas and boson gas; λV is the interaction between them, being of a binary nature. Thus,

$$H^0 = \sum_{i=1}^{n} h_f^0(r_i) + \sum_{j=1}^{N} H_b^0(R_j), \tag{1.7}$$

$$\lambda V = \sum_{ij} \lambda v(r_i - R_j). \tag{1.8}$$

We use the formalism of second quantization. So, let $\{|\zeta\rangle\}$ denote the set of quantum states of h_f^0 with eigenvalues $\{\epsilon_\zeta\}$, and let $\{|\eta\rangle\}$ denote the set of quantum states of H_b^0 with eignevalues $\{E_\eta\}$, we then have

$$H^0 = \sum_\zeta \mathbf{n}_\zeta \epsilon_\zeta + \sum_\eta \mathbf{N}_\eta E_\eta, \tag{1.9}$$

$$\lambda V = \sum_{\zeta''\zeta'\eta''\eta'} c^\dagger_{\zeta''} a^\dagger_{\eta''} (\zeta''\eta''|\lambda v|\zeta'\eta') a_{\eta'} c_{\zeta'}, \tag{1.10}$$

$$|\gamma\rangle = |\{n_\zeta\},\{N_\eta\}\rangle = |\{n_\zeta\}\rangle \otimes |\{N_\eta\}\rangle; \tag{1.11}$$

here $\mathbf{n}_\zeta = c^\dagger_\zeta c_\zeta$ are occupation operators and n_ζ is the occupation number; similarly for $\mathbf{N}_\eta = a^\dagger_\eta a_\eta$ and N_η; the c's and a's are the creation and annihilation operators for fermions and bosons, respectively, At some points we will indicate the changes if both gases are bosons or fermions or if the bosons are quasiparticles like phonons.

The present article has a threefold purpose. First we derive the one-body linear response results (Part A, Secs. 2–4). Next we derive a fully quantum mechanical Boltzmann equation both in diagonal and nondiagonal form; this is an extension of LRT II Sec. 8 (Part B, Secs. 5 and 6). These equations are shown to be fully equivalent to the one-body linear response results (Part B, Sec. 7). Finally, we consider inhomogeneous systems and derive a Boltzmann equation for the Wigner function corresponding to the one-particle distribution function (Part C, Sec. 8). From this equation the classical Boltzmann equation is easily recovered. In Fig. 1 we give a flow diagram of the various connections.

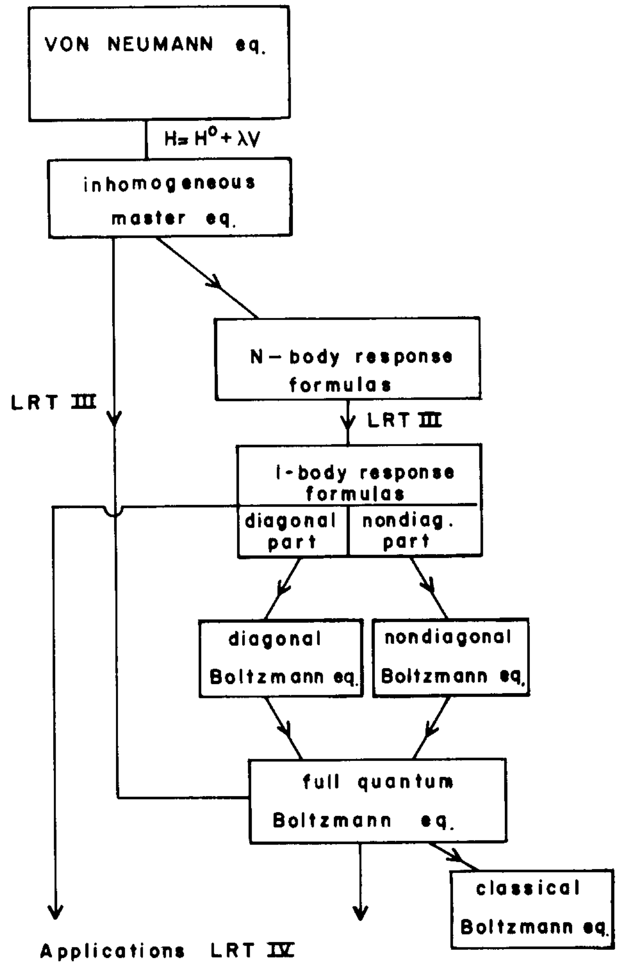

FIG. 1. Flow diagram of the various connections.

We still note that a fourth article, containing applications of the present developments for magnetic and other transport phenomena is in preparation.

A. ONE-BODY LINEAR RESPONSE RESULTS

2. The diagonal susceptibility and conductivity

The many-body forms for the diagonal part of the susceptibility of a variable B and for the diagonal part of the conductivity of a variable $\dot B$ were in LRT II given as

$$\chi^d_{BA}(i_\omega) = \beta \int_0^\infty dt\, e^{-i\omega t} \mathrm{Tr}[\rho_{eq} (\dot A^R)_d B^R_d(t)] \tag{2.1}$$

and

$$L^d_{BA}(i\omega) = \beta \int_0^\infty dt\, e^{-i\omega t} \mathrm{Tr}[\rho_{eq} (\dot A^R)_d (\dot B^R(t))_d], \tag{2.2}$$

where the superscript R stands for the reduced operator; $\beta = 1/kT$. The time dependence for $B^R_d(t)$ was given already in (1.1); the time dependence for $(\dot B^R(t))_d$ is likewise

$$(\dot B^R(t))_d = e^{-\Lambda_d t} \dot B^R_d; \tag{2.3}$$

however, $\dot B^R_d$ is more than the Schrödinger operator $\dot B_d$, see

LRT II, Eqs.(4.28) and (4.29),

$$J^R_{B,d} \equiv (\dot{B}^R)_d = -\Lambda_d B_d + (\dot{B})_d. \tag{2.4}$$

Similarly for $J_{A,d} \equiv (\dot{A}^R)_d$.

In case we deal with the electrical conductivity, the external field Hamiltonian is $-\mathbf{A}\cdot\mathbf{F}(t)$ with $\mathbf{F}(t) = q\mathbf{E}$, $\mathbf{A} = \Sigma_i(\mathbf{r}_i - \langle \mathbf{r}_i \rangle_{eq})$, where q is the charge of the carrier (including sign), $\mathbf{r}_i$ are the positions of the carriers, and $\langle \mathbf{r}_i \rangle_{eq}$ are the equilibrium positions prior to the switching on of the field. The electrical current Schrödinger operator is $\mathbf{J} = q\Sigma_i \mathbf{v}_i/\Omega = q\dot{A}/\Omega$, where Ω is the volume of the sample, see LRT I, Eq. (2.31). Thus we have, denoting by greek subscripts the vector and tensor components,

$$\sigma^d_{\mu\nu}(i\omega) = \beta\Omega \int_0^\infty dt\, e^{-i\omega t} \mathrm{Tr}\left[\rho_{eq} J^R_{d\nu} J^R_{d\mu}(t)\right]; \tag{2.5}$$

the reduced current is given by

$$\mathbf{J}^R_d = \frac{q}{\Omega}\left[-\Lambda_d \sum_i (\mathbf{r}_i - \mathbf{r}_i^{eq})_d + \sum_i \mathbf{v}_{id}\right], \tag{2.6}$$

the two parts representing collisional current and ponderomotive current, respectively; the former accounts for the many-body effects in the subdynamics of H^0. (In the full dynamics of H, this term is absent.)

We will develop the one-body form for (2.1). Since both $\dot{A}^R_d$ and B^R_d are extensive operators of the fermion system we have

$$B^R_d(t) = e^{\Lambda_d t} \sum_\zeta \mathbf{n}_\zeta (\zeta |b| \zeta), \tag{2.7}$$

$$(\dot{A}^R)_d = \sum_\zeta [-\Lambda_d \mathbf{n}_\zeta (\zeta|a|\zeta) + \mathbf{n}_\zeta(\zeta|\dot{a}|\zeta)], \tag{2.8}$$

with lower case symbols denoting one-body operators. Thus (2.1) becomes

$$\chi^d_{BA}(i\omega) = \beta\int_0^\infty dt e^{-i\omega t}\mathrm{Tr}\{\rho_{eq}\sum_{\zeta'}[-\Lambda_d \mathbf{n}_{\zeta'}(\zeta'|a|\zeta') + \mathbf{n}_{\zeta'}(\zeta'|\dot{a}|\zeta')]\sum_{\zeta''} e^{\Lambda_d t}\mathbf{n}_{\zeta''}(\zeta''|b|\zeta'')\}. \tag{2.9}$$

We take the operation in the representation $\{|\gamma\rangle\} = |\{n_\zeta\}\rangle \otimes |\{N_\eta\}\rangle$ and we develop the exponential

$$\chi^d_{BA}(i\omega) = \beta\int_0^\infty dt\, e^{-i\omega t} \sum_{\{n_\zeta\}}\sum_{\{N_\eta\}} \{p_{eq}(\{n_\zeta\},\{N_\eta\})\langle\{n_\zeta\}| \times \sum_{\zeta'}[-\Lambda_d \mathbf{n}_{\zeta'}(\zeta'|a|\zeta') + \mathbf{n}_{\zeta'}(\zeta'|\dot{a}|\zeta')] \times \sum_{\zeta''}\sum_{k=0}^\infty \frac{(-t)^k}{k!}(\Lambda_d)^k \mathbf{n}_{\zeta''}(\zeta''|b|\zeta'')|\{n_\zeta\}\rangle\Big\}, \tag{2.10}$$

$p_{eq} = \langle\{n_\zeta\},\{N_\eta\}|\rho_{eq}|\{n_\zeta\},\{N_\eta\}\rangle$.[12] The standard adiabatic assumption is made that the boson average can be made separately, i.e.,

$$\sum_{\{n_\zeta\}}\sum_{\{N_\eta\}} p_{eq}(\{n_\zeta\},\{N_\eta\})\cdots = \sum_{\{n_\zeta\}} p_{eq}(\{n_\zeta\})\langle\cdots\rangle_b, \tag{2.11}$$

where the latter average is an equilibrium average over the boson distribution.

We need the following two theorems (also stated in LRT II but not proven there) which are the main link of the many-body and one-body descriptions.

Theorem 1:

$$\langle M n_{\zeta'}\rangle_b = \mathscr{B}_{\zeta'} n_{\zeta'}. \tag{2.12}$$

Here M is the linear master operator in function space, representing the many-body aspects, while $\mathscr{B}_\zeta$ is the nonlinear Boltzmann operator of the one-body description,[13] given by

$$\mathscr{B}_\zeta f(\zeta) = \sum_{\bar{\zeta}}\{w_{\zeta\bar{\zeta}} f(\zeta)[1-f(\bar{\zeta})] - w_{\bar{\zeta}\zeta} f(\bar{\zeta})[1-f(\zeta)]\}. \tag{2.13}$$

The fermion transition rates are given by

$$w_{\zeta''\zeta'} = \sum_{\eta'\eta''} Q(\zeta''\eta'';\zeta',\eta')\langle N_{\eta''}(1+N_{\eta'})\rangle_{eq}$$

$$\approx \sum_{\eta'\eta''} Q(\zeta''\eta'';\zeta',\eta')\langle N_{\eta''}\rangle_{eq}(1+\langle N_{\eta'}\rangle_{eq}); \tag{2.14}$$

the latter equality is based on the truncation rule or closure property of LRT II, Eq. (8.1); it is exact in the grand canonical ensemble. The Q's are the binary transition rates [see LRT II, Eq (8.18)]. Whereas the two-body transition rates Q are reciprocal, the one-body rates w are not; from the quilibrium Bose–Einstein distribution one finds

$$w_{\zeta''\zeta'} = \sum_{\eta'\eta''} Q(\zeta''\eta'';\zeta'\eta')e^{-\beta E_{\eta''}}(1+\langle N_{\eta''}\rangle_{eq})e^{\beta E_{\eta'}}\langle N_{\eta'}\rangle_{eq}$$

$$= e^{\beta(\epsilon_{\zeta''} - \epsilon_{\zeta'})}\sum_{\eta'\eta''} Q(\zeta'\eta';\zeta''\eta'')(1+\langle N_{\eta''}\rangle_{eq})\langle N_{\eta'}\rangle_{eq},$$

where we used the delta property $\delta(E_{\eta''} + \epsilon_{\zeta''} - E_{\eta'} - \epsilon_{\zeta'})$ in the definition of the Q's. Thus we have

$$w_{\zeta''\zeta'} = w_{\zeta'\zeta''} e^{\beta(\epsilon_{\zeta''} - \epsilon_{\zeta'})}. \tag{2.15}$$

Theorem 2:

$$\langle \Lambda_d \mathbf{n}_{\zeta'}\rangle_b = \sum_{\{n_\zeta\}} |\{n_\zeta\}\rangle\langle\{n_\zeta\}|\mathscr{B}_{\zeta'} n_{\zeta'}; \tag{2.16}$$

this gives the connection between the master equation in the Liouville space and the Boltzmann operator. The theorem follows from (2.12) by multiplying $Mn_{\zeta'}$ by the projector $|\{n_\zeta\},\{N_\eta\}\rangle\langle\{n_\zeta\},\{N_\eta\}|$, summing over all many-body states, applying Eq. (1.4), and performing a boson average.

The proof of (2.12) is straightforward. For $w_{\gamma\bar{\gamma}}$ we have from (1.3), (1.10), and (1.11),

$$W_{\gamma\bar{\gamma}} = \frac{2\pi\lambda^2}{\hbar}\sum_{\zeta'\zeta''\eta'\eta''} |\langle\{n_\zeta\}\{N_\eta\}|c^\dagger_{\zeta''}a^\dagger_{\eta''}a_{\eta'}c_{\zeta'}|\{\bar{n}_\zeta\}\{\bar{N}_\eta\}\rangle|^2 \times |(\zeta''\eta''|v|\zeta'\eta')|^2\delta(\epsilon_{\zeta''} - \epsilon_{\zeta'} + E_{\eta''} - E_{\eta'}). \tag{2.17}$$

One easily finds that the only connected state for given $|\gamma\rangle$ and given $\zeta'\zeta''\eta'\eta''$ is $|\bar{\gamma}\rangle \equiv |\bar{\gamma}_{\zeta''\zeta'\eta''\eta'}\rangle$, with [LRT II, Eqs. (8.19), (8.13), and (8.14)]

$$W_{\gamma\bar{\gamma}_{\zeta''\zeta'\eta''\eta'}} = Q(\zeta'\eta';\zeta''\eta'')(1-n_{\zeta''})n_{\zeta'}(1+N_{\eta''})N_{\eta'}, \tag{2.18}$$

where

$$\bar{n}_\zeta = n_\zeta(1-\delta_{\zeta\zeta'} - \delta_{\zeta\zeta''}) + (1-n_\zeta)(\delta_{\zeta\zeta'} + \delta_{\zeta\zeta''}), \tag{2.19}$$

$$\bar{N}_\eta = N_\eta - \delta_{\eta\eta'} + \delta_{\eta\eta''}. \tag{2.20}$$

We now make the standard adiabatic assumption, (2.11); then in calculating $Mn_{\zeta''}$, employing (1.3) and (2.18), we perform a boson average; the result is

$$\begin{aligned}\langle Mn_{\zeta''}\rangle_b &= \sum_{\zeta'\zeta''} w_{\zeta'\zeta''}(1-n_{\zeta''})n_{\zeta'}(n_{\zeta''}-\overline{n}_{\zeta''})\\ &\overset{(2.19)}{=} \sum_{\zeta'\zeta''} w_{\zeta'\zeta''}(1-n_{\zeta''})n_{\zeta'}(2n_{\zeta''}-1)(\delta_{\zeta''\zeta'}+\delta_{\zeta''\zeta''})\\ &= \sum_{\zeta'}[w_{\zeta''\zeta'}n_{\zeta''}(1-n_{\zeta'})(2n_{\zeta''}-1)\\ &\quad + w_{\zeta'\zeta''}(1-n_{\zeta''})n_{\zeta'}(2n_{\zeta''}-1)]\\ &= \mathscr{B}_{\zeta''}n_{\zeta''},\end{aligned} \tag{2.21}$$

where we still used, noticing $n^2_{\zeta''}=n_{\zeta''}$,

$$\begin{aligned} n_{\zeta''}(2n_{\zeta''}-1) &= n_{\zeta''},\\ (1-n_{\zeta''})(2n_{\zeta''}-1) &= -(1-n_{\zeta''}).\end{aligned} \tag{2.22}$$

2.1. *Linear collision operator*

A simple closed expression for χ, L, and σ can only be found when the collision operator is linear. This occurs in two cases. First we may have elastic or near-elastic collisions. Then $w_{\zeta'\zeta''}\approx w_{\zeta''\zeta'}$. The linear Boltzmann operator then is

$$\mathscr{B}^l_\zeta f(\zeta)=\sum_{\zeta'} w_{\zeta\zeta'}[f(\zeta)-f(\zeta')]. \tag{2.23}$$

Electron collisions with acoustical phonons is an example of near elastic collisions (see Sec. 4). Strictly elastic collisions occur when the scattering involves heavy obstacles (one-body collisions), such as in impurity scattering. Then by (2.14) and (2.15) since $N_\eta \ll 1$,

$$\begin{aligned} w_{\zeta'\zeta''} &\approx \sum_{\eta'\eta''} Q(\zeta',\eta';\zeta'',\eta'')\langle N_{\eta'}\rangle_b\\ &\approx \sum_{\eta''} Q(\zeta',\eta'';\zeta'',\eta'')\sum_{\eta'}\langle N_{\eta'}\rangle_b\\ &= NQ^0(\zeta',\zeta''),\end{aligned} \tag{2.24}$$

$$Q^0(\zeta',\zeta'')=2\pi(\lambda^2/\hbar)|(\zeta'|v|\zeta'')|^2\delta(\epsilon_{\zeta'}-\epsilon_{\zeta''}), \tag{2.25}$$

indicating one-body collisions.

Secondly, the Boltzmann operator is linear when we deal with nondegenerate systems such that $f(\zeta)\ll 1$. In that case we have from (2.13), for the collision operator,

$$\mathscr{B}^l_\zeta f(\zeta)=\sum_{\zeta'}[w_{\zeta\zeta'}f(\zeta)-w_{\zeta'\zeta}f(\zeta')]. \tag{2.26}$$

In contrast to the case of Eq. (2.23), now generally $w_{\zeta'\zeta''}\neq w_{\zeta''\zeta'}$. We shall therefore use the form (2.26) since it encompasses both cases.

It is now possible to compound the M operator; first, we will show that

$$\langle M^2\cdots\rangle_b=\langle M\langle M\cdots\rangle_b\rangle_b. \tag{2.27}$$

For the boson average of the left-hand side we have terms like

$$\sum_{\eta'\eta''}\sum_{\overline{\eta}'\overline{\eta}''}\langle Q(\zeta''\eta'';\zeta'\eta')Q(\overline{\zeta}''\overline{\eta}'';\overline{\zeta}'\overline{\eta}')(1+N_{\eta''})N_{\eta'} \times(1+N_{\overline{\eta}''})N_{\overline{\eta}'}\cdots\rangle_{\text{eq}}. \tag{2.28}$$

In this series we first pick the terms with $\overline{\eta}'\neq\eta'$ and $\overline{\eta}''\neq\eta''$. We can then use the truncation rule for the boson average,

$$\begin{aligned}&\langle(1+N_{\eta''})N_{\eta'}(1+N_{\overline{\eta}''})N_{\overline{\eta}'}\rangle_{\text{eq}}\\ &\approx\langle(1+N_{\eta''})N_{\eta'}\rangle_{\text{eq}}\langle(1+N_{\overline{\eta}''})N_{\overline{\eta}'}\rangle_{\text{eq}}.\end{aligned} \tag{2.29}$$

Thus, this part of (2.28) yields $w_{\zeta''\zeta'}w_{\overline{\zeta}''\overline{\zeta}'}$. The remaining part of (2.28) is a triple sum $\Sigma_{\eta'\eta''\overline{\eta}'}$ or $\Sigma_{\eta'\eta''\overline{\eta}''}$. It vanishes with respect to the first part in the thermodynamic limit $N=\Sigma N_\eta\to\infty$. This proves (2.27). The compounding of the fermion parts is simple. Since $\mathscr{B}^l_\zeta n_\zeta$ is a linear combination of n_ζ's, Theorem 1 can be applied repeatedly. We thus obtain

Theorem 3:

$$\langle M^k n_{\zeta'}\rangle_b=(\mathscr{B}^l_{\zeta'})^k n_{\zeta'}. \tag{2.30}$$

For the repeated Λ_d operator we have

$$\langle(\Lambda_d)^k\mathbf{n}_{\zeta'}\rangle_b=\sum_{\{n^k_\zeta\}}P_{\{\mathbf{n}^k_\zeta\}}\mathscr{B}^l_{\zeta'}\sum_{\{n^{k-1}_\zeta\}}P_{\{n^{k-1}_\zeta\}}\mathscr{B}^l_{\zeta'}\cdots\sum_{\{n^1_\zeta\}}P_{\{n^1_\zeta\}}\mathscr{B}^l_{\zeta'}n_{\zeta'}, \tag{2.31}$$

where $P_{\{n^i_\zeta\}}$ are the projectors $|\{n^i_\zeta\}\rangle\langle\{n^i_\zeta\}|$. Since the projectors commute with the $\mathscr{B}^l$ operators and since $P_iP_j=P^2_i\delta_{ij}=P_i\delta_{ij}$, we obtain

Theorem 4:

$$\langle(\Lambda_d)^k\mathbf{n}_{\zeta'}\rangle_b=\sum_{\{n_\zeta\}}|\{n_\zeta\}\rangle\langle\{n_\zeta\}|(\mathscr{B}^l_{\zeta'})^k n_{\zeta'}. \tag{2.32}$$

Using this result and (2.16) we find upon reconstituting the exponential in (2.16),

$$\chi^d_{BA}(i\omega)=\beta\int_0^\infty dt\,e^{-i\omega t}\sum_{\{n_\zeta\}}p_{\text{eq}}(\{n_\zeta\}) \times\sum_{\zeta'\zeta''}[-(\mathscr{B}^l_{\zeta'}n_{\zeta'})a_{\zeta'}+n_{\zeta'}\dot{a}_{\zeta'}]b_{\zeta''}e^{-t\mathscr{B}^l_{\zeta''}}n_{\zeta''}, \tag{2.33}$$

where

$$o_\zeta=(\zeta|o|\zeta) \tag{2.34}$$

for any one-body operator such as a and b. The result can also be written in terms of the resolvent[14]

$$\chi^d_{BA}(i\omega)=\beta\sum_{\zeta'\zeta''}\left\langle[-(\mathscr{B}^l_{\zeta'}n_{\zeta'})a_{\zeta'}+n_{\zeta'}\dot{a}_{\zeta'}]b_{\zeta''}\times\frac{1}{i\omega+\mathscr{B}^l_{\zeta''}}n_{\zeta''}\right\rangle_{\text{eq}}. \tag{2.35}$$

In the result for L^d_{BA} a few changes occur. For $(\dot{B}^R(t))_d$ we have

$$(\dot{B}^R(t))_d=e^{-\Lambda_d t}\sum_\zeta[-\Lambda_d\mathbf{n}_\zeta(\zeta|b|\zeta)+\mathbf{n}_\zeta(\zeta|\dot{b}|\zeta)]. \tag{2.36}$$

When the exponential is expanded we now also encounter terms with $[(-t)^k/k!](\Lambda_d)^{k+1}$. The procedure is clearly the same. We find

$$L^d_{BA}(i\omega)=\beta\int_0^\infty dt\,e^{-i\omega t}\sum_{\zeta'\zeta''}\langle[-(\mathscr{B}^l_{\zeta'}n_{\zeta'})a_{\zeta'}+n_{\zeta'}\dot{a}_{\zeta'}] \times e^{-t\mathscr{B}^l_{\zeta''}}[-(\mathscr{B}^l_{\zeta''}n_{\zeta''})b_{\zeta''}+n_{\zeta''}\dot{b}_{\zeta''}]\rangle_{\text{eq}}; \tag{2.37}$$

or in terms of the resolvent

$$L^d_{BA}(i\omega)=\beta\sum_{\zeta'\zeta''}\left\langle[-(\mathscr{B}^l_{\zeta'}n_{\zeta'})a_{\zeta'}+n_{\zeta'}\dot{a}_{\zeta'}]\frac{1}{i\omega+\mathscr{B}^l_{\zeta''}}\times[-(\mathscr{B}^l_{\zeta''}n_{\zeta''})b_{\zeta''}+n_{\zeta''}\dot{b}_{\zeta''}]\right\rangle_{\text{eq}}. \tag{2.38}$$

For the electrical conductivity likewise,

$$\sigma^d_{\mu\nu}(i\omega) = \frac{\beta q^2}{\Omega}\int_0^\infty dt\, e^{-i\omega t} \sum_{\zeta'\zeta''} \langle [-(\mathscr{B}^l_{\zeta'} n_{\zeta'})(\zeta'|r_\nu - r^{\rm eq}_\nu|\zeta') + n_{\zeta'}(\zeta'|v_\nu|\zeta')]$$
$$\times e^{-t\mathscr{B}^l_{\zeta''}}[-(\mathscr{B}^l_{\zeta''} n_{\zeta''})(\zeta''|r_\mu - r^{\rm eq}_\mu|\zeta'') + n_{\zeta''}(\zeta''|v_\mu|\zeta'')]\rangle_{\rm eq} \tag{2.39}$$

or in terms of the resolvent

$$\sigma^d_{\mu\nu}(i\omega) = \frac{\beta q^2}{\Omega}\sum_{\zeta'\zeta''}\langle [-(\mathscr{B}^l_{\zeta'} n_{\zeta'})(r_\nu - r^{\rm eq}_\nu)_{\zeta'} + n_{\zeta'} v_{\nu\zeta'}]\frac{1}{i\omega + \mathscr{B}^l_{\zeta''}} \times [-(\mathscr{B}^l_{\zeta''} n_{\zeta''})(r_\mu - r^{\rm eq}_\mu)_{\zeta''} + n_{\zeta''} v_{\mu\zeta''}]\rangle_{\rm eq}. \tag{2.40}$$

Note that in (2.37)–(2.40) the exponential $\exp(-t\mathscr{B}^l)$ or resolvent operator only operates on the particle densities to their right.

a. No collisional current. For the linear case the averages can be carried out in a grand canonical ensemble. For simplicity we first consider (2.39) in the absence of collisional current, i.e., when $(\zeta|r_\mu - r^{\rm eq}_\mu|\zeta) = 0$. Thus, with

$$p_{\rm eq}(\{n_\zeta\}) = (1/Z)e^{\alpha n - \beta\Sigma_\zeta n_\zeta \epsilon_\zeta}, \tag{2.41}$$

where α/β is the chemical potential and $Z = \Pi_\zeta(1 + e^{\alpha - \beta\epsilon_\zeta})$ is the partition sum, we must evaluate

$$\frac{1}{Z}\sum_{\zeta'\zeta''}\sum_n e^{\alpha n}\sum_{\{n_\zeta\}}{}' \prod_\zeta e^{-\beta n_\zeta \epsilon_\zeta} v_{\nu\zeta'} v_{\mu\zeta''} n_{\zeta'} e^{-t\mathscr{B}^l_{\zeta''}} n_{\zeta''}; \tag{2.42}$$

here Σ' denotes the restricted sum subject to $\Sigma_\zeta n_\zeta = n$. Combining however, Σ_n and $\Sigma'_{\{n_\zeta\}}$ to an unrestricted sum, we can interchange the Π and this sum, obtaining

$$(2.42) = \frac{1}{Z}\sum_{\zeta'\zeta''} v_{\nu\zeta'} v_{\mu\zeta''}\prod_\zeta \sum_{\{n_\zeta = 0,1\}} e^{(\alpha - \beta\epsilon_\zeta)n_\zeta} n_{\zeta'} \times \sum_{k=0}^\infty \frac{(-t)^k}{k!}(\mathscr{B}^l_{\zeta''})^k n_{\zeta''}. \tag{2.43}$$

For $k = 0$ the sum is trivial. For $k = 1$ we obtain (omitting $\Sigma_{\zeta'} v_{\nu\zeta'}$ for the time being)

$$\frac{1}{Z}\sum_{\zeta''\bar\zeta} v_{\mu\zeta''}\prod_\zeta \sum_{\{n_\zeta=0,1\}} e^{(\alpha-\beta\epsilon_\zeta)n_\zeta} n_{\zeta'}(w_{\zeta''\bar\zeta} n_{\zeta''} - w_{\bar\zeta\zeta''} n_{\bar\zeta}). \tag{2.44}$$

We split this into two sums and we interchange the summation indices ζ'' and $\bar\zeta$ in the second sum. We then find

$$(2.44) = \frac{1}{Z}\sum_{\zeta''\bar\zeta}\prod_\zeta \sum_{\{n_\zeta=0,1\}} e^{(\alpha-\beta\epsilon_\zeta)n_\zeta} n_{\zeta'} n_{\zeta''} w_{\zeta''\bar\zeta}(v_{\mu\zeta''} - v_{\mu\bar\zeta})$$
$$= \frac{1}{Z}\sum_{\zeta''}\prod_\zeta \cdots n_{\zeta'} n_{\zeta''}(\mathscr{B}^0_{\zeta''} v_{\mu\zeta''}), \tag{2.45}$$

where the operator $\mathscr{B}^0$ *acting on the matrix element* $v_{\mu\zeta''}$ is to be understood in the sense of (2.23) even though w may not be reversible as in (2.26); we signified this by the superscript zero on the Boltzmann operator. For the sum over $\{n_\zeta\}$ we first consider $\zeta'' = \zeta'$. This gives

$$\frac{1}{Z}\prod_\zeta(1 + e^{\alpha - \beta\epsilon_\zeta})\frac{e^{\alpha - \beta\epsilon_{\zeta'}}}{1 + e^{\alpha-\beta\epsilon_{\zeta'}}}\mathscr{B}^0_{\zeta'} v_{\mu\zeta'} = \langle n_{\zeta'}\rangle_{\rm eq}\mathscr{B}^0_{\zeta'} v_{\mu\zeta'}. \tag{2.46}$$

Next we consider all $\zeta'' \neq \zeta'$. The result is likewise found to be

$$\sum_{\zeta''\neq\zeta'}\langle n_{\zeta'}\rangle_{\rm eq}\langle n_{\zeta''}\rangle_{\rm eq}\mathscr{B}^0_{\zeta''} v_{\mu\zeta''} = \langle n_{\zeta'}\rangle_{\rm eq}\sum_{\zeta''}\langle n_{\zeta''}\rangle_{\rm eq}\mathscr{B}^0_{\zeta''} v_{\mu\zeta''} - \langle n'_\zeta\rangle^2_{\rm eq}\mathscr{B}^0_{\zeta'} v_{\mu\zeta'}. \tag{2.47}$$

The first term to the right is zero:

$$\sum_{\zeta''\bar\zeta}\langle n_{\zeta''}\rangle_{\rm eq} w_{\zeta''\bar\zeta}(v_{\mu\zeta''} - v_{\mu\bar\zeta}) = 0 \tag{2.48}$$

as is found from interchange of the indices ζ'', $\bar\zeta$ and detailed balance. Thus, combining (2.46) and (2.47), we obtain

$$(2.44) = \langle n_{\zeta'}\rangle_{\rm eq}(1 - \langle n_{\zeta'}\rangle_{\rm eq})\mathscr{B}^0_{\zeta'} v_{\mu\zeta'} = -(1/\beta)(\partial\langle n_{\zeta'}\rangle_{\rm eq}/\partial\epsilon_{\zeta'})\mathscr{B}^0_{\zeta'} v_{\mu\zeta'}. \tag{2.49}$$

If we now take the term for $k = 2$ of (2.43), we have, denoting by

$$\Psi \equiv \prod_\zeta \sum_{\{n_\zeta=0,1\}} e^{(\alpha-\beta\epsilon_\zeta)n_\zeta},$$

the following result

$$\frac{1}{Z}\sum_{\zeta''} v_{\mu\zeta''}\Psi n_{\zeta'}\mathscr{B}^l_{\zeta''}(\mathscr{B}^l_{\zeta''} n_{\zeta''})$$
$$= \frac{1}{Z}\sum_{\zeta''} v_{\mu\zeta''}\Psi n_{\zeta'}\sum_{\bar\zeta}(w_{\zeta''\bar\zeta}\mathscr{B}^l_{\zeta''} n_{\zeta''} - w_{\bar\zeta\zeta''}\mathscr{B}^l_{\bar\zeta} n_{\bar\zeta})$$
$$\overset{*}{=} \frac{1}{Z}\sum_{\zeta''}\Psi n_{\zeta'}(\mathscr{B}^l_{\zeta''} n_{\zeta''})\sum_{\bar\zeta} w_{\zeta''\bar\zeta}(v_{\mu\zeta''} - v_{\mu\bar\zeta'})$$
$$= \frac{1}{Z}\sum_{\zeta''}\Psi n_{\zeta'}(\mathscr{B}^l_{\zeta''} n_{\zeta''})\mathscr{B}^0_{\zeta''} v_{\mu\zeta''}$$
$$= \frac{1}{Z}\sum_{\zeta''}\Psi n_{\zeta'}(\mathscr{B}^0_{\zeta''} v_{\mu\zeta''})\sum_{\bar\zeta}(w_{\zeta''\bar\zeta} n_{\zeta''} - w_{\bar\zeta\zeta''} n_{\bar\zeta})$$
$$\overset{**}{=} \frac{1}{Z}\sum_{\zeta''}\Psi n_{\zeta'} n_{\zeta''}\sum_{\bar\zeta} w_{\zeta''\bar\zeta}(\mathscr{B}^0_{\zeta''} v_{\mu\zeta''} - \mathscr{B}^0_{\bar\zeta} v_{\mu\bar\zeta})$$
$$= \frac{1}{Z}\sum_{\zeta''}\Psi n_{\zeta'} n_{\zeta''}\mathscr{B}^0_{\zeta''}(\mathscr{B}^0_{\zeta''} v_{\mu\zeta''}), \tag{2.50}$$

where in * we interchange ζ'' and $\bar\zeta$ in one term and in ** we interchanged ζ'' and $\bar\zeta$ in one term. Likewise, we find that the term of order k in (2.43) produces the result involving $(\mathscr{B}^0_{\zeta''})^k v_{\mu\zeta''}$. The final result, valid for any linear Boltzmann collision operator, is therefore

$$\sigma^d_{\mu\nu}(i\omega) = -\left(\frac{q^2}{\Omega}\right)\sum_\zeta \frac{\partial\langle n_\zeta\rangle_{\rm eq}}{\partial\epsilon_\zeta} v_{\mu\zeta}\frac{1}{i\omega + \mathscr{B}^0_\zeta} v_{\mu\zeta}. \tag{2.51}$$

We still note that a similar, but not identical result, follows from the Boltzmann equation (5.10) of Sec. 5. We then find $\sigma_{\nu\mu}$ rather than $\sigma_{\mu\nu}$ (which are equal, however, due to the Onsager relations) and the resolvent operation of $\mathscr{B}^0$ appears in front of $(\partial\langle n_\zeta\rangle_{\rm eq}/\partial\epsilon_\zeta)v_{\mu\zeta}$. The equivalence of these results is only trivial if the collisions are elastic; then $\partial\langle n_\zeta\rangle_{\rm eq}/\partial\epsilon_\zeta$ is a collisional invariant.

At this point we also note that (2.51) shows a close correspondence with Verboven's result[15] for the original Kubo theory:

$$\sigma^{\rm Verboven}_{\mu\nu} = -\frac{1}{\Omega}\int_0^\infty dt\, e^{-i\omega t}\,{\rm tr}\left[\frac{\partial f}{\partial h} j_\nu j_\mu(t)\right] = -\frac{1}{\Omega}\lim_{\delta\to 0}\sum_\zeta \frac{\partial f_\zeta}{\partial\epsilon_\zeta} j_{\nu\zeta}\frac{1}{i(\omega - l) + \delta} j_{\mu\zeta}, \tag{2.52}$$

where f is the Fermi function, tr the one-particle trace, j the one-particle current ($= qv/\Omega$) and l the one-particle Liouville operator; clearly, the van Hove limit has brought about the change $il \to -\mathscr{B}^0$, causing convergence of the Fourier–Laplace integral and yielding and explicit result for the conductivity.

The result (2.51) can be further simplified by introducing a relaxation time τ_ζ. Let us put

$$\mathscr{B}^0_\zeta v_{\mu\zeta} = \sum_{\zeta'} w_{\zeta\zeta'}(v_{\mu\zeta} - v_{\mu\zeta'}) = \frac{1}{\tau_\zeta} v_{\mu\zeta}, \tag{2.53}$$

where τ is a c number. In addition we require $(\mathscr{B}^0_\zeta)^k v_{\mu\zeta} = (\tau_\zeta)^{-k} v_{\mu\zeta}$ for any k. This is strictly only satisfied if $1/\tau$ is an eigenvalue of $\mathscr{B}^0$, being independent of ζ. Now in all usual cases τ depends on ζ only via ϵ_ζ. Thus τ is an eigenvalue for elastic collisions, for then $\mathscr{B}^0$ decomposes into contributions $\mathscr{B}^0(\epsilon_\zeta)$ for separate energy sheets. Indeed, we have in that case

$$\begin{aligned}
(\mathscr{B}^0_\zeta)^2 v_{\mu\zeta} &= \sum_{\zeta'} w_{\zeta\zeta'}\left(\frac{v_{\mu\zeta}}{\tau(\epsilon_\zeta)} - \frac{v_{\mu\zeta'}}{\tau(\epsilon_{\zeta'})}\right) \\
&= \sum_{\zeta'} \frac{2\pi\lambda^2}{\hbar} |(\zeta'|v|\zeta)|^2 \delta(\epsilon_\zeta - \epsilon_{\zeta'}) \\
&\quad \times \left(\frac{v_{\mu\zeta}}{\tau(\epsilon_\zeta)} - \frac{v_{\mu\zeta}}{\tau(\epsilon_{\zeta'})}\right) \\
&= \frac{1}{\tau(\epsilon_\zeta)} \sum_{\zeta'} w_{\zeta\zeta'}(v_{\mu\zeta} - v_{\mu\zeta'}) = \frac{1}{[\tau(\epsilon_\zeta)]^2} v_{\mu\zeta};
\end{aligned} \tag{2.54}$$

and so on for $k = 3, 4 \cdots$. For nonelastic collisions we can only maintain the result (2.54) as an approximation in that we write $\tau(\epsilon_\zeta) \approx \tau(\epsilon_{\zeta'})$. *However, this approximation is not tantamount to the usual "relaxation time approximation"* in which one sets $\mathscr{B}^l_\zeta \langle n_\zeta \rangle_t = [\langle n_\zeta \rangle_t - \langle n_\zeta \rangle_{eq}]/\tau(\epsilon_\zeta)$; this ansatz requires that $\langle n_\zeta \rangle_{eq} \approx \langle n_{\zeta'} \rangle_{eq}$ in order to arrive at the form (2.53) and (2.56 (see below) for the relaxation time, cf., e.g., Nag.[15a] Since $\langle n_\zeta \rangle_{eq}$ depends exponentially on ϵ_ζ, while $\tau(\epsilon_\zeta)$ depends on ϵ_ζ via a low power of ϵ_ζ, the present approximation is considerably better. We therefore have for any linear Boltzmann process

$$\sigma^d_{\mu\nu}(i\omega) \approx -\frac{q^2}{\Omega} \sum_\zeta \frac{\partial \langle n_\zeta \rangle_{eq}}{\partial \epsilon_\zeta} \frac{v_{\nu\zeta} v_{\mu\zeta}}{i\omega + 1/\tau(\epsilon_\zeta)}. \tag{2.55}$$

Let μ refer to the direction of a polar axis [this direction refers to the current response, but it is easier to switch the indices ν and μ (Onsager) so that μ refers to the direction of the applied field] and let $\mathbf{v}_\zeta = (v_\zeta, \chi_\zeta, \psi_\zeta)$ and $\mathbf{v}_{\zeta'} = (v_{\zeta'}, \chi_{\zeta'}, \psi_{\zeta'})$ be the polar representations. Then

$$(v_{\mu\zeta} - v_{\mu\zeta'})/v_{\mu\zeta} = 1 - \cos\chi_{\zeta'}/\cos\chi_\zeta. \tag{2.56}$$

Thus the relaxation time is determined by

$$\frac{1}{\tau_\zeta} = \sum_{\zeta'} w_{\zeta\zeta'}\left(1 - \frac{\cos\chi_{\zeta'}}{\cos\chi_\zeta}\right). \tag{2.57}$$

The standard applications involve impurity scattering and lattice scattering. For Bloch states we have

$$\sum_{\zeta'} \to \sum_{\mathbf{k}'} = \frac{-\Omega}{8\pi^3} \int\int\int k'^2\, dk'\, d\phi\, d(\cos\theta).$$

For impurity scattering,

$$1 - \cos\chi_{k'}/\cos\chi_k = (1 - \cos\theta) - \tan\chi \sin\theta \sin\phi.$$

where θ is the scattering angle between $\mathbf{v}_{\mathbf{k}'}$ and $\mathbf{v}_\mathbf{k}$ and where ϕ is the azimuthal angle for the direction of $\mathbf{v}_{\mathbf{k}'}$ around $\mathbf{v}_\mathbf{k}$. This yields the well-known result

$$\begin{aligned}
\frac{1}{\tau(\epsilon_k)} &= \frac{\lambda^2 \pi N}{\hbar} Z(\epsilon_k) \int_{-1}^{1} d(\cos\theta) |(\mathbf{k}'|v|\mathbf{k})|^2 (1 - \cos\theta) \\
&= 2\pi(N/\Omega) v_\mathbf{k} \int_{-1}^{1} d(\cos\theta)\, \sigma(\theta)(1 - \cos\theta);
\end{aligned} \tag{2.58}$$

here $Z(\epsilon)$ is the density of states and $\sigma(\theta)$ is the cross section. The application of (2.55) and (2.57) to lattice scattering will be discussed elsewhere. For randomizing collisions, we have

$$\int_{-1}^{1} d(\cos\theta)(1 - \cos\theta) = 2$$

so that,

$$\frac{1}{\tau(\epsilon_k)} = \frac{\Omega}{2\pi^2} \int w_{\mathbf{k}\mathbf{k}'}\, dk'. \tag{2.59}$$

In this case, also, Eq. (2.55) is exact.

b. Collisional current. We consider the case that the current is due to collisional current only, such as is the case in problems involving transverse magnetic fields.[16] We must now evaluate

$$\sigma^d_{\mu\nu\,\mathrm{coll}}(i\omega) = \frac{\beta q^2}{\Omega} \int_0^\infty dt\, e^{-i\omega t} \sum_{\zeta'\zeta''} R_{\nu\zeta'} R_{\mu\zeta''} \langle (\mathscr{B}^l_{\zeta'} n_{\zeta'}) \times e^{-t\mathscr{B}^l_{\zeta''}} \mathscr{B}^l_{\zeta''} n_{\zeta''} \rangle_{eq}, \tag{2.60}$$

where $R = r - r^{eq}$. Using the grand canonical ensemble, we must evaluate

$$\frac{1}{Z} \sum_{\zeta'\zeta''} R_{\nu\zeta'} R_{\mu\zeta''} \prod_\zeta \sum_{\{n_\zeta = 0,1\}} e^{(\alpha - \beta\epsilon_\zeta) n_\zeta} \mathscr{B}^l_{\zeta'} n_{\zeta'} \times \sum_{k=0}^\infty \frac{(-t)^k}{k!} (\mathscr{B}^l_{\zeta''})^{k+1} n_{\zeta''}. \tag{2.61}$$

Consider the $k = 0$ term first. Writing out the $\mathscr{B}$ operators we easily obtain

$$\left.\begin{matrix}(2.61)\\ k=0\end{matrix}\right\} = \frac{1}{Z} \sum_{\zeta'\zeta''} \prod_\zeta \sum_{\{n_\zeta = 0,1\}} e^{(\alpha - \beta\epsilon_\zeta) n_\zeta} n_{\zeta'} n_{\zeta''} (\mathscr{B}^0_{\zeta'} R_{\nu\zeta'}) \times \mathscr{B}^0_{\zeta''} R_{\mu\zeta''}. \tag{2.62}$$

For $\zeta' = \zeta''$ this yields

$$\begin{aligned}
\left.\begin{matrix}(2.62)\\ \zeta' = \zeta''\end{matrix}\right\}
&= \frac{1}{Z} \sum_{\zeta'} \prod_\zeta (1 + e^{\alpha - \beta\epsilon_\zeta}) \frac{e^{\alpha - \beta\epsilon_{\zeta'}}}{1 + e^{\alpha - \beta\epsilon_{\zeta'}}} (\mathscr{B}^0_{\zeta'} R_{\nu\zeta'}) \mathscr{B}^0_{\zeta''} R_{\mu\zeta''} \\
&= \sum_{\zeta'} \langle n_{\zeta'} \rangle_{eq} (\mathscr{B}^0_{\zeta'} R_{\nu\zeta'}) \mathscr{B}^0_{\zeta'} R_{\mu\zeta'}.
\end{aligned} \tag{2.63}$$

For $\zeta' \neq \zeta''$, we find

$$\begin{aligned}
\left.\begin{matrix}(2.62)\\ \zeta' \neq \zeta''\end{matrix}\right\} &= \sum_{\zeta'\zeta''} \langle n_{\zeta'} \rangle_{eq} \langle n_{\zeta''} \rangle_{eq} (\mathscr{B}^0_{\zeta'} R_{\nu\zeta'}) \mathscr{B}^0_{\zeta''} R_{\mu\zeta''} \\
&\quad - \sum_{\zeta'} \langle n'^2_\zeta \rangle_{eq} (\mathscr{B}^0_{\zeta'} R_{\nu\zeta'}) \mathscr{B}^0_{\zeta'} R_{\mu\zeta'}.
\end{aligned}$$

The first term is zero by (2.48). We thus obtain

$$\left.\begin{matrix}(2.61)\\ k=0\end{matrix}\right\} = \sum_{\zeta'} \langle n_{\zeta'} \rangle_{eq} (1 - \langle n_{\zeta'} \rangle_{eq}) (\mathscr{B}^0_{\zeta'} R_{\nu\zeta'}) \mathscr{B}^0_\zeta R_{\mu\zeta'}.$$

The terms for $k = 1,2\cdots$ are treated similarly. We can easily obtain the final result

$$\sigma^d_{\mu\nu\text{coll}}(i\omega) = -\frac{q^2}{\Omega}\sum_{\zeta,\text{spin}}\int_0^\infty dt\, e^{-i\omega t}\frac{\partial\langle n_\zeta\rangle_{\text{eq}}}{\partial\epsilon_\zeta}(\mathscr{B}^0_\zeta R_{\nu\zeta})e^{-t\mathscr{B}^0_\zeta}\mathscr{B}^0_\zeta R_{\mu\zeta}$$

$$= -\frac{q^2}{\Omega}\sum_{\zeta,\text{spin}}\frac{\partial\langle n_\zeta\rangle_{\text{eq}}}{\partial\epsilon_\zeta}(\mathscr{B}^0_\zeta R_{\nu\zeta})\frac{1}{i\omega+\mathscr{B}^1_\zeta}\mathscr{B}^0_\zeta R_{\mu\zeta}. \quad (2.64)$$

In this result, we can define a new scalar

$$\frac{1}{\tau^*_\zeta} = \frac{2\pi\lambda^2 N}{\hbar}\sum_{\zeta'}X(\zeta,\zeta')\delta(\epsilon_{\zeta'}-\epsilon_\zeta)\frac{R_{\mu\zeta}-R_{\mu\zeta'}}{R_{\mu\zeta}}, \quad (2.65)$$

where $X(\zeta,\zeta') = |(\zeta|v|\zeta')|^2$. Then in (2.64) we can set $\mathscr{B}^0_\zeta \to 1/\tau^*_\zeta$.

Equation (2.64) is a generalization for all frequencies of the result by Adams and Holstein[6] for the transverse magnetoconductivity, in which case $|\zeta)$ are the Landau states $|Nk_y k_z)$. To see this we note for for $\omega = 0$,

$$\sigma^d_{XX\text{coll}}(0) = -\frac{q^2}{\Omega}\sum_{\zeta,\text{spin}}\frac{\partial\langle n_\zeta\rangle_{\text{eq}}}{\partial\epsilon_\zeta}(\mathscr{B}^0_\zeta X_\zeta)X_\zeta$$

$$= -\frac{q^2}{\Omega}\sum_{\zeta\zeta',\text{spin}}\frac{\partial\langle n_\zeta\rangle_{\text{eq}}}{\partial\epsilon_\zeta}w_{\zeta\zeta'}(X_\zeta - X_{\zeta'})X_\zeta$$

$$= -\frac{q^2}{\Omega}\sum_{\zeta\zeta'}\frac{\partial\langle n_\zeta\rangle_{\text{eq}}}{\partial\epsilon_\zeta}w_{\zeta\zeta'}(X_\zeta - X_{\zeta'})^2. \quad (2.66)$$

In the final result we interchanged the summation indices, took half the sum, and we multiplied by a factor 2 due to spin summation (noting that we need the spin factor in only one sum, since in the collision spin is generally conserved). We thus obtain the same expression as given by Adams and Holstein.

2.2. *General two-body collision operator*

We return to the general case for which $\mathscr{B}$ is the nonlinear Boltzmann collision operator. For collisions between unlike particles the operator is quadratic in $\langle n_\zeta\rangle_t$. For collisions between like particles (e.g., electron–electron interaction) the operator is quartic in $\langle n_\zeta\rangle_t$. Though we did not consider the latter case, it can be carried out in a similar way as the fermion–boson interactions considered here.

The repeated Λ_d operation can formally be carried out, but leads in practice to formidable expressions. For example, for $\Lambda^2_d\mathbf{n}_\zeta$ we find

$$\langle\Lambda^2_d\mathbf{n}_{\zeta''}\rangle_b = \sum_{\{n_\zeta\}}|\{n_\zeta\})(\{n_\zeta\}|\mathscr{B}^{(2)}_{\zeta''}n_{\zeta''}, \quad (2.67)$$

where

$$\mathscr{B}^{(2)}_{\zeta''}n_{\zeta''} = \sum_{\zeta'\zeta''}w_{\zeta'\zeta''}(1-n_{\zeta''})n_{\zeta'}\times\Big\{\sum_{\zeta'''}[w_{\zeta''\zeta'''}n_{\zeta''}(1-n_{\zeta'''}) - w_{\zeta'''\zeta''}n_{\zeta'''}(1-n_{\zeta''})]$$
$$-\sum_{\zeta'''}[w_{\zeta''\zeta'''}\overline{n}_{\zeta''}(1-\overline{n}_{\zeta'''}) - w_{\zeta'''\zeta''}\overline{n}_{\zeta'''}(1-\overline{n}_{\zeta''})]\Big\}; \quad (2.68)$$

here $\overline{n}_{\zeta''}$ and $\overline{n}_{\zeta'''}$ are given by (2.19). Carrying out the summations over the Kronecker deltas results after much algebra in

$$\mathscr{B}^{(2)}_{\zeta''}n_{\zeta''} = \mathscr{B}_{\zeta''}n_{\zeta''}\sum_{\zeta''}[w_{\zeta''\zeta''}(1-n_{\zeta''}) + w_{\zeta''\zeta''}n_{\zeta''}]$$
$$-\sum_{\zeta'\zeta''}[w_{\zeta''\zeta'}(1-n_{\zeta'})n_{\zeta''} - w_{\zeta'\zeta''}n_{\zeta'}(1-n_{\zeta''})]$$
$$\times[w_{\zeta''\zeta''}(1-n_{\zeta''}) + w_{\zeta''\zeta''}n_{\zeta''}]$$
$$+\sum_{\zeta''}[w_{\zeta''\zeta''}w_{\zeta''\zeta''}n_{\zeta''}(1-n_{\zeta''}) + w^2_{\zeta''\zeta''}n_{\zeta''}(1-n_{\zeta''})]$$
$$-\sum_{\zeta''}[w_{\zeta''\zeta''}w_{\zeta''\zeta''}n_{\zeta''}(1-n_{\zeta''}) + w^2_{\zeta''\zeta''}n_{\zeta''}(1-n_{\zeta''})]. \quad (2.69)$$

Generally we will set

$$\langle\Lambda^k_d\mathbf{n}_{\zeta'}\rangle_b = \sum_{\{n_\zeta\}}|\{n_\zeta\})(\{n_\zeta\}|\mathscr{B}^{(k)}_{\zeta'}n_{\zeta'}, \quad (2.70)$$

with $\mathscr{B}^{(1)} = \mathscr{B}$. Equation (2.10) then yields[17]

$$\chi^d_{BA}(i\omega) = \beta\int_0^\infty dt\, e^{-i\omega t}\sum_{\zeta'\zeta''}\sum_{k=0}^\infty\frac{(-t)^k}{k!}\times\langle[-(\mathscr{B}_{\zeta'}n_{\zeta'})a_{\zeta'} + n_{\zeta'}\dot{a}_{\zeta'}](\mathscr{B}^{(k)}_{\zeta''}n_{\zeta''})b_{\zeta''}\rangle_{\text{eq}}$$

or (2.71)

$$\chi^d_{BA} = -\lim_{\delta\to 0}\beta\sum_{\zeta'\zeta''}\sum_{k=0}^\infty\left(\frac{-1}{i\omega+\delta}\right)^{k+1}\times\langle[-(\mathscr{B}_{\zeta'}n_{\zeta'})a_{\zeta'} + n_{\zeta'}\dot{a}_{\zeta'}](\mathscr{B}^{(k)}_{\zeta''}n_{\zeta''})b_{\zeta''}\rangle_{\text{eq}},$$

Here,

$$\lim_{\delta\to 0}\left[\frac{-1}{i\omega+\delta}\right]^{k+1} = \mathscr{P}\left(\frac{-1}{i\omega}\right)^{k+1} - (-i)^k\pi\delta^{(k)}(\omega)/k! \quad (2.71')$$

One easily notices that for linear $\mathscr{B}$ [with $\mathscr{B}^{(k)}\to(\mathscr{B}')^k$ and for $\omega\|(\mathscr{B}')^{-1}\| < 1$] this reduces to (2.35) of the previous subsection. For practical purposes this result is not useful; however, we will need this formal result in Sec. 7.

We also give the results of L and σ:

$$L^d_{BA}(i\omega) = \beta\int_0^\infty dt\, e^{-i\omega t}\sum_{\zeta'\zeta''}\sum_{k=0}^\infty\frac{(-t)^k}{k!}\times\langle[-(\mathscr{B}_{\zeta'}n_{\zeta'})a_{\zeta'} + n_{\zeta'}\dot{a}_{\zeta'}]\times[-(\mathscr{B}^{(k+1)}_{\zeta''}n_{\zeta''})b_{\zeta''} + (\mathscr{B}^{(k)}_{\zeta''}n_{\zeta''})\dot{b}_{\zeta''}]\rangle_{\text{eq}}, \quad (2.72)$$

$$\sigma^d_{\mu\nu}(i\omega) = \frac{\beta q^2}{\Omega}\int_0^\infty dt\, e^{-i\omega t}\sum_{\zeta'\zeta''}\sum_{k=0}^\infty\frac{(-t)^k}{k!}\times\langle[-(\mathscr{B}_{\zeta'}n_{\zeta'})(r_\nu - r^{\text{eq}}_\nu)_{\zeta'} + n_{\zeta'}v_{\nu\zeta'}]\times[-(\mathscr{B}^{(k+1)}_{\zeta''}n_{\zeta''})(r_\mu - r^{\text{eq}}_\mu)_{\zeta''} + (\mathscr{B}^{(k)}_{\zeta''}n_{\zeta''})v_{\mu\zeta''}]\rangle. \quad (2.73)$$

Collisional current only; extended Adams–Holstein results. When there is collisional current only a very useful result can be obtained for $\omega = 0$. Going back to (2.5) and (2.6) we have for the many-body form

$$\sigma^d_{\mu\nu}(i\omega) = \frac{\beta q^2}{\Omega}\sum_{\zeta'\zeta''}\text{Tr}\{\rho_{\text{eq}}(\Lambda_d n_{\zeta'})(r_\nu - r^{\text{eq}}_\nu)_{\zeta'}\frac{1}{i\omega+\Lambda_d}(\Lambda_d\mathbf{n}_{\zeta''})(r_\mu - r^{\text{eq}}_\mu)_{\zeta''}\}. \quad (2.74)$$

which for $\omega \to 0$ yields

$$\sigma^d_{\mu\nu}(0) = \frac{\beta q^2}{\Omega} \sum_{\zeta'\zeta''} \mathrm{Tr}\{\rho_{eq}(\Lambda_d \mathbf{n}_{\zeta'})(r_\nu - r^{eq}_\nu)_{\zeta'} n_{\zeta'}(r_\mu - r^{eq}_\mu)_{\zeta''}\}. \tag{2.75}$$

Employing Theorem 2, we find

$$\sigma^d_{\mu\nu}(0) = \frac{\beta q^2}{\Omega} \sum_{\zeta'\zeta''} \langle (\mathscr{B}_{\zeta'} n_{\zeta'}) n_{\zeta''} \rangle_{eq} R_{\nu\zeta'} R_{\mu\zeta''}. \tag{2.76}$$

The average to be found is

$$\sum_{\zeta'} \langle (\mathscr{B}_{\zeta'} n_{\zeta'}) n_{\zeta''} \rangle_{eq} R_{\nu\zeta'}$$
$$= \sum_{\{n_\zeta\}} \sum_{\zeta'\bar\zeta} p_{eq} [w_{\zeta'\bar\zeta} n_{\zeta'}(1 - \bar n_\zeta) - w_{\bar\zeta\zeta'} n_{\bar\zeta}(1 - n_{\zeta'})] n_{\zeta''} R_{\nu\zeta'}, \tag{2.77}$$

where p_{eq} is the grand canonical distribution. In one part of this sum we interchange ζ' and $\bar\zeta$ to obtain

$$(2.77) = \sum_{\{n_\zeta\}} \sum_{\zeta'\bar\zeta} p_{eq} n_{\zeta''} n_{\zeta'}(1 - n_{\bar\zeta}) w_{\zeta'\bar\zeta}(R_{\nu\zeta'} - R_{\nu\bar\zeta}). \tag{2.78}$$

For $\bar\zeta = \zeta''$ the result is zero since $n_{\zeta''}(1 - n_{\zeta''}) = 0$. For $\zeta' = \bar\zeta$ the result is zero since $w_{\zeta'\zeta'} = 0$ (we assumed that v has no diagonal part).

For $\zeta' = \zeta''$ the result is

$$\sum_{\{n_\zeta\}} p_{eq} \sum_{\bar\zeta} n_{\zeta''}(1 - n_{\bar\zeta}) w_{\zeta''\bar\zeta}(R_{\nu\zeta''} - R_{\nu\bar\zeta})$$
$$= \sum_{\bar\zeta} \langle n_{\zeta''} \rangle_{eq}(1 - \langle n_{\bar\zeta} \rangle_{eq}) w_{\zeta''\bar\zeta}(R_{\nu\zeta''} - R_{\nu\bar\zeta}). \tag{2.79}$$

For $\zeta' \neq \zeta'' \neq \bar\zeta$ the remaining contribution is found likewise

$$\sum_{\substack{\zeta'\bar\zeta \\ \zeta'\neq\zeta''\neq\bar\zeta}} \langle n_{\zeta''} \rangle_{eq} \langle n_{\zeta'} \rangle_{eq}(1 - \langle n_{\bar\zeta} \rangle_{eq}) w_{\zeta'\bar\zeta}(R_{\nu\zeta'} - R_{\nu\bar\zeta})$$
$$= \sum_{\zeta'\bar\zeta} \langle n_{\zeta''} \rangle_{eq} \langle n_{\zeta'} \rangle_{eq}(1 - \langle n_{\bar\zeta} \rangle_{eq}) w_{\zeta'\bar\zeta}(R_{\nu\zeta'} - R_{\nu\bar\zeta})$$
$$- \sum_{\bar\zeta} \langle n_{\zeta''} \rangle^2_{eq}(1 - \langle n_{\bar\zeta} \rangle_{eq}) w_{\zeta''\bar\zeta}(R_{\nu\zeta''} - R_{\nu\bar\zeta})$$
$$- \sum_{\zeta'} \langle n_{\zeta''} \rangle_{eq} \langle n_{\zeta'} \rangle_{eq}(1 - \langle n_{\zeta''} \rangle_{eq}) w_{\zeta'\zeta''}(R_{\nu\zeta'} - R_{\nu\zeta''}). \tag{2.80}$$

The double sum of (2.80) is zero, as is found by interchanging $\zeta, \bar\zeta$ in the term with $-R_{\nu\bar\zeta}$ and applying detailed balance

$$\langle n_{\zeta'} \rangle_{eq}(1 - \langle n_{\bar\zeta} \rangle_{eq}) w_{\zeta'\bar\zeta} = \langle n_{\bar\zeta} \rangle_{eq}(1 - \langle n_{\zeta'} \rangle_{eq}) w_{\bar\zeta\zeta'}. \tag{2.81}$$

The third sum of (2.80) is written as

$$\sum_{\bar\zeta} \langle n_{\zeta''} \rangle_{eq} \langle n_{\bar\zeta} \rangle_{eq}(1 - \langle n_{\zeta''} \rangle_{eq}) w_{\bar\zeta\zeta''}(R_{\nu\bar\zeta} - R_{\nu\zeta''}). \tag{2.82}$$

Applying detailed balance, (2.82) cancels the second sum of (2.80). We are thus left with (2.79). Substituting into (2.76) we obtain (with $\bar\zeta \to \zeta$, $\zeta'' \to \zeta'$)

$$\sigma^d_{\mu\nu}(0) = \frac{\beta q^2}{\Omega} \sum_{\zeta\zeta',\text{spin}} \langle n_{\zeta'} \rangle_{eq}(1 - \langle n_\zeta \rangle_{eq}) w_{\zeta'\zeta}(R_{\nu\zeta'} - R_{\nu\zeta}) R_{\mu\zeta'}. \tag{2.83}$$

For $\mu = \nu$ this can be simplified. We interchange the indices ζ, ζ' apply the detailed balance, and add the results. We then find,

$$\sigma^d_{xx}(0) = \frac{\beta q^2}{\Omega} \sum_{\zeta\zeta'} \langle n_\zeta \rangle_{eq}(1 - \langle n_{\zeta'} \rangle_{eq}) w_{\zeta\zeta'}(X_{\zeta'} - X_\zeta)^2, \tag{2.84}$$

where we include a spin factor of two. Equation (2.84) is the extended Adams–Holstein result for processes involving inelastic binary collisions. It is also valid for collisions with quasiparticles such as phonons (see Sec. 4). For these processes this formula was first given by Argyres and Roth.[9,18]

3. The nondiagonal susceptibility and conductivity

3.1. Formulas for χ_{nd}, L_{nd}, and σ_{nd}

For the nondiagonal part of the susceptibility and conductivity we found in LRT II, Sec. 7,

$$\chi^{nd}_{BA}(i\omega) = \int_0^\infty dt\, e^{-i\omega t} \times \int_0^\beta d\beta' \,\mathrm{Tr}[\rho_{eq}(\dot A^R(-i\hbar\beta'))_{nd} B^R_{nd}(t)], \tag{3.1}$$

$$L^{nd}_{BA}(i\omega) = \int_0^\infty dt\, e^{-i\omega t} \times \int_0^\beta d\beta' \,\mathrm{tr}[\rho_{eq}(\dot A^R(-i\hbar\beta'))_{nd}(\dot B^R(t))_{nd}]; \tag{3.2}$$

for the electrical conductivity specifically,

$$\sigma^{nd}_{\mu\nu}(i\omega) = \Omega \int_0^\infty dt\, e^{-i\omega t} \int_0^\beta d\beta' \,\mathrm{Tr}[\rho_{eq} J^R_{nd\nu}(-i\hbar\beta') J^R_{nd\mu}(t)]. \tag{3.3}$$

Here

$$B^R_{nd}(t) = e^{i\mathscr{L}^0 t} B^R_{nd} = e^{iH^0 t/\hbar} B_{nd} e^{-iH^0 t/\hbar},$$
$$(\dot B^R(t))_{nd} = e^{i\mathscr{L}^0 t} \dot B^R_{nd} = e^{iH^0 t/\hbar} \dot B_{nd} e^{-iH^0 t/\hbar}, \tag{3.4}$$

with[2] $B^R_{nd} = B_{nd}$, $\dot B^R_{nd} \equiv B_{nd}$, $J^R_{nd} \equiv \Sigma q\, v_{nd}/\Omega$, there being no collisional current for the nondiagonal part; also

$$(\dot A^R(-i\hbar\beta))_{nd} = e^{\beta H^0} \dot A_{nd} e^{-\beta H^0}. \tag{3.5}$$

We proceed with (3.1). We take the trace in the representation $\{|\gamma\rangle\}$, giving

$$\chi^{nd}_{BA}(i\omega) = \int_0^\infty dt\, e^{-i\omega t} \int_0^\beta d\beta' \sum_{\gamma\bar\gamma} \{p_{eq}(\gamma) e^{\beta(\mathscr{E}_\gamma - \mathscr{E}_{\bar\gamma})} \langle\gamma|\dot A_{nd}|\bar\gamma\rangle \times e^{it(\mathscr{E}_{\bar\gamma} - \mathscr{E}_\gamma)/\hbar} \langle\bar\gamma|B_{nd}|\gamma\rangle\}, \tag{3.6}$$

where we used $H^0|\gamma\rangle = \mathscr{E}_\gamma|\gamma\rangle$. Carrying out the $d\beta'$ integration, this yields

$$\chi^{nd}_{BA}(i\omega) = \int_0^\infty dt\, e^{-i\omega t} \sum_{\gamma\bar\gamma} p_{eq}(\gamma) \frac{e^{\beta(\mathscr{E}_\gamma - \mathscr{E}_{\bar\gamma})} - 1}{\mathscr{E}_\gamma - \mathscr{E}_{\bar\gamma}} e^{it(\mathscr{E}_{\bar\gamma} - \mathscr{E}_\gamma)/\hbar} \times \langle\gamma \dot A_{nd}|\bar\gamma\rangle \langle\bar\gamma|B_{nd}|\gamma\rangle. \tag{3.7}$$

Now, in second quantization form,

$$\dot A_{nd} = \sum_{\zeta'\zeta''} c^\dagger_{\zeta'} c_{\zeta''}(\zeta'|\dot a_{nd}|\zeta'') = \sum_{\zeta'\zeta''}{}' c^\dagger_{\zeta'} c_{\zeta''}(\zeta'|\dot a|\zeta''), \tag{3.8}$$

where Σ' denotes $\zeta' \neq \zeta''$; note that by this convention the subscript "nd" on $\dot a$ can be deleted. Consider first fixed $|\gamma\rangle$ and a fixed pair ζ', ζ'' out of the sum (3.8). Since

$$c^\dagger_{\zeta'} c_{\zeta''}|\bar\gamma\rangle$$
$$\equiv c^\dagger_{\zeta'} c_{\zeta''}|\{\bar n_\zeta\},\{\bar N_\eta\}\rangle = (-1)^{\Sigma(1,\zeta'-1)}(-1)^{\Sigma(1,\zeta''-1)} \times (1 - \bar n_{\zeta'})^{1/2}(\bar n_{\zeta''})^{1/2}|\ldots,1 - n_{\zeta'},\ldots,1 - \bar n_{\zeta''},\ldots,\{\bar N_\eta\}\rangle, \tag{3.9}$$

we find that the matrix element $\langle\gamma|\dot A_{nd}|\bar\gamma\rangle$ is nonzero only if $|\bar\gamma\rangle$ is the connected state, for which

$$\bar n_{\zeta'} = 1 - n_{\zeta'}, \quad \bar n_{\zeta''} = 1 - n_{\zeta''},$$
$$\text{all other } \bar n_\zeta = \text{all other } n_\zeta,$$
$$\text{all } \bar N_\eta = \text{all } N_\eta; \tag{3.10}$$

this connected state is denoted as $|\overline{\gamma}_{\zeta'\zeta''}\rangle$. For all other terms of the series (3.8) the matrix element between $|\gamma\rangle$ and $|\overline{\gamma}_{\zeta'\zeta''}\rangle$ yields zero. Hence we arrive at

$$\langle\gamma|\dot{A}_{\rm nd}|\overline{\gamma}_{\zeta'\zeta''}\rangle = (-1)^{\Sigma(1,\zeta'-1)}(-1)^{(1,\zeta''-1)}(n_{\zeta'})^{1/2} \times(1-n_{\zeta''})^{1/2}(\zeta'|\dot{a}|\zeta'') \quad (\zeta'\neq\zeta''). \tag{3.11}$$

Likewise we need $\langle\overline{\gamma}_{\zeta'\zeta''}|B_{\rm nd}|\gamma\rangle$. Let again

$$B_{\rm nd} = \sum_{\zeta'''\zeta''''}{}' c^\dagger_{\zeta'''}c_{\zeta''''}(\zeta'''|b|\zeta''''). \tag{3.12}$$

With $|\gamma\rangle$ given, and for fixed ζ''', ζ'''', the state $|\gamma\rangle$ must be so chosen that $|\gamma\rangle$ is connected to $|\overline{\gamma}\rangle$ by

$$n_{\zeta'''} = 1-\overline{n}_{\zeta'''},\ n_{\zeta''''} = 1-\overline{n}_{\zeta''''}, \quad \text{all other } n_\zeta = \text{all other } \overline{n}_\zeta, \quad \text{all } N_\eta = \text{all } \overline{N}_\eta. \tag{3.13}$$

Now (3.13) is incompatible with (3.10) unless either $\zeta''' = \zeta'$ and $\zeta'''' = \zeta''$, or $\zeta''' = \zeta''$ and $\zeta'''' = \zeta'$. For these two cases the matrix element is, respectively,

$$\langle\overline{\gamma}_{\zeta'\zeta''}|c^\dagger_{\zeta'}c_{\zeta''}|\gamma\rangle = (-1)^{\Sigma(1,\zeta'-1)}(-1)^{\Sigma(1,\zeta''-1)} \times(n_{\zeta''})^{1/2}(1-n_{\zeta'})^{1/2}, \tag{3.14a}$$

$$\langle\overline{\gamma}_{\zeta'\zeta''}|c^\dagger_{\zeta''}c_{\zeta'}|\gamma\rangle = (-1)^{\Sigma(1,\zeta'-1)}(-1)^{\Sigma(1,\zeta''-1)} \times(n_{\zeta'})^{1/2}(1-n_{\zeta''})^{1/2}. \tag{3.14b}$$

All other terms of the series (3.12) give zero matrix elements. Moreover, when we multiply (3.11) with (3.14a) we obtain zero since

$$(n_{\zeta'})^{1/2}(1-n_{\zeta''})^{1/2}(n_{\zeta''})^{1/2}(1-n_{\zeta'})^{1/2} = 0 \quad \text{for } n_{\zeta''}, n_{\zeta'} = 0,1.$$

Hence, only (3.14b) contributes to $\langle\overline{\gamma}_{\zeta'\zeta''}|B_{\rm nd}|\gamma\rangle$, the relevant matrix element being

$$\langle\overline{\gamma}_{\zeta'\zeta''}|B_{\rm nd}|\gamma\rangle = (-1)^{\Sigma(1,\zeta'-1)}(-1)^{\Sigma(1,\zeta''-1)}(n_{\zeta'})^{1/2} \times(1-n_{\zeta''})^{1/2}(\zeta''|b|\zeta'). \tag{3.15}$$

We substitute (3.11) and (3.15) into (3.7). This gives

$$\chi^{\rm nd}_{BA}(i\omega) = \int_0^\infty dt\, e^{-i\omega t}\sum_{\{n_\zeta\}}\sum_{\{N_\eta\}}\sum_{\zeta'\zeta''}{}' p_{\rm eq}(\{n_\zeta\},\{N_\eta\}) \times\frac{e^{\beta(\mathscr{E}_\gamma-\mathscr{E}_{\overline{\gamma}})}-1}{\mathscr{E}_\gamma-\mathscr{E}_{\overline{\gamma}}}e^{it(\mathscr{E}_{\overline{\gamma}}-\mathscr{E}_\gamma)/\hbar}n_{\zeta'}(1-n_{\zeta''}) \times(\zeta'|\dot{a}|\zeta'')(\zeta''|b|\zeta'). \tag{3.16}$$

Since $|\gamma\rangle$ differs from $|\overline{\gamma}\rangle$ by the lowering of $n_{\zeta''}$ and the raising of $n_{\zeta'}$, we have with ϵ again denoting the fermion energies

$$\mathscr{E}_{\overline{\gamma}}-\mathscr{E}_\gamma = \epsilon_{\zeta''}-\epsilon_{\zeta'}. \tag{3.17}$$

So, (3.16) gives, carrying out the equilibrium averaging,

$$\chi^{\rm nd}_{BA}(i\omega) = \int_0^\infty dt\, e^{-i\omega t}\sum_{\zeta'\zeta''}{}'\langle n_{\zeta'}\rangle_{\rm eq}(1-\langle n_{\zeta''}\rangle_{\rm eq}) \times\frac{1-e^{-\beta(\epsilon_{\zeta''}-\epsilon_{\zeta'})}}{\epsilon_{\zeta''}-\epsilon_{\zeta'}}e^{it(\epsilon_{\zeta''}-\epsilon_{\zeta'})/\hbar}(\zeta'|\dot{a}|\zeta'')(\zeta''|b|\zeta'). \tag{3.18}$$

Finally, with

$$\int_0^\infty dt\, e^{i\alpha t} = 2\pi\delta_-(\alpha) = i\mathscr{P}\left(\frac{1}{\alpha}\right)+\pi\delta(\alpha), \tag{3.19}$$

where $\mathscr{P}$ denotes the principal part, we find

$$\chi^{\rm nd}_{BA}(i\omega) = \hbar\sum_{\zeta'\zeta''}{}'\langle n_{\zeta'}\rangle_{\rm eq}(1-\langle n_{\zeta''}\rangle_{\rm eq})(\zeta'|\dot{a}|\zeta'')(\zeta''|b|\zeta') \times\frac{1-e^{-\beta(\epsilon_{\zeta''}-\epsilon_{\zeta'})}}{\epsilon_{\zeta''}-\epsilon_{\zeta'}} \times\left[i\mathscr{P}\frac{1}{\epsilon_{\zeta''}-\epsilon_{\zeta'}-\hbar\omega}+\pi\delta(\epsilon_{\zeta''}-\epsilon_{\zeta'}-\hbar\omega)\right]. \tag{3.20}$$

We still note that for $\omega = 0$ (direct-current result), the delta function does not usually contribute, unless ζ' and ζ'' refer to different eigenstates with the same energy.

For L_{BA} the result is analogous, with $\dot{b}$ replacing b. For the electrical conductivity we have in particular

$$\sigma^{\rm nd}_{\mu\nu}(i\omega) = \Omega\hbar\sum_{\zeta'\zeta''}{}'\langle n_{\zeta'}\rangle_{\rm eq}(1-\langle n_{\zeta''}\rangle_{\rm eq})(\zeta'|j_\nu|\zeta'')(\zeta''|j_\mu|\zeta') \times\frac{1-e^{-\beta(\epsilon_{\zeta''}-\epsilon_{\zeta'})}}{\epsilon_{\zeta''}-\epsilon_{\zeta'}} \times\left[i\mathscr{P}\frac{1}{\epsilon_{\zeta''}-\epsilon_{\zeta'}-\hbar\omega}+\pi\delta(\epsilon_{\zeta''}-\epsilon_{\zeta'}-\hbar\omega)\right], \tag{3.21}$$

where $j = qv/\Omega$ is the one-particle current density.

3.2. *The quantum mechanical Hall effect*

For the Hall effect in strong magnetic fields, $\{|\zeta)\}$ are Landau states. It has long been realized that the diagonal matrix elements of the current yield zero, so no Hall effect results. This problem has been circumvented by some authors (see Ref. 7) by including the external electric field in the unperturbed Hamiltonian H^0. To obtain results an expansion of the one-particle von Neumann equation in powers λV is employed up to orders $(\lambda V)^2$; in the δ_+ function that is found to occur, the delta part is retained and the principal part is, unjustifiably, neglected. In our opinion it is fortuitous that the right Hall conductivity is found in this way. For that reason, we will indicate here that the quantum mechanical Hall effect stems solely from the nondiagonal part of the conductivity response formula. Since the nondiagonal part has not been considered in the past, the cause for the problems with the absence of Hall effect in earlier theories is evident.[19]

We consider the Hamiltonian

$$h^0 = (\mathbf{p}+e\mathbf{A})^2/2m, \quad \mathbf{A} = (0,Bx,0), \tag{3.22}$$

where we employed the Landau gauge, the magnetic field being in the z direction. The one-particle eigenstates (in wave mechanical form) and eigenvalues are

$$|\zeta) = \phi_N(x+\hbar k_y)e^{ik_y y}e^{ik_z z}/A^{1/2}, \tag{3.23}$$

$$\mathscr{E}_{Nk_yk_z} = (N+1/2)\hbar\omega_0 + \hbar^2k_z^2/2m, \quad N = 0,1,2,\cdots, \tag{3.24}$$

where $\omega_0 = |q|B/m$ is the cyclotron frequency and where ϕ_N represents harmonic oscillator wavefunctions. We also write $|\zeta) = |Nk_yk_z)$ and we set $x_0 = \hbar k_y/m\omega_0$. The relevant matrix elements are[7] for a solid dimensions $L_xL_yL_z$ $(A = L_yL_z)$,

$$(\zeta|x|\zeta') = x_0\delta_{NN'}\delta_{kk'} + (\hbar/2m\omega_0)^{1/2}[(N+1)^{1/2}\delta_{N'N+1} + (N)^{1/2}\delta_{N',N-1}]\delta_{kk'}, \tag{3.25}$$

$$(\zeta|y|\zeta') = (L_y/2)\delta_{NN'}\delta_{kk'} = y^{\rm eq}\delta_{NN'}\delta_{kk'}, \tag{3.26}$$

$$(\zeta|v_x|\zeta') = i(\hbar\omega_0/2m)^{1/2}[-(N+1)^{1/2}\delta_{N',N+1} + (N)^{1/2}\delta_{N',N-1}]\delta_{kk'}, \quad (3.27)$$

$$(\zeta|v_y|\zeta') = (\hbar\omega_0/2m)^{1/2}[(N+1)^{1/2}\delta_{N',N+1} + (N)^{1/2}\delta_{N',N-1}]\delta_{kk'}. \quad (3.28)$$

with $\delta_{kk'} = \delta_{k_y k_y'}\delta_{k_z k_z'}$. From the latter two equations we note that there is no diagonal ponderomotive current, neither in the x nor in the y direction. The matrix element of (3.25) indicates that in the absence of an external electric field there are stable orbits with fixed center x_0, where x_0 equals what previously we termed $x^{\rm eq}$ [see Eq. (2.6)]. Were we concerned with the conductivity σ_{xx}, as for transverse magnetoresistance, then there is a collisional contribution since by (3.25) $(\zeta|x - x^{\rm eq}|\zeta')$ is nonzero both in its diagonal and nondiagonal matrix elements. For the Hall effect, however, we need σ_{yx}; here the collisional contribution [see (2.83)] is zero, since $(\zeta|y - y^{\rm eq}|\zeta) = 0$. Consequently, for the Hall effect we have only a nondiagonal ponderomotive contribution as we stated above.

We consider the charge carriers to be electrons. Then $j = -ev/\Omega$. We obtain from (3.27) and (3.28),

$$(\zeta'|j_x|\zeta'')(\zeta''|j_y|\zeta') = (ie^2\hbar\omega_0/2m\Omega^2)\{\underset{(a)}{-(N'+1)^{1/2}(N'')^{1/2}\delta_{N'',N'+1}} + \underset{(b)}{(N')^{1/2}(N''+1)^{1/2}\delta_{N'',N'-1}}\}\delta_{k'k''}. \quad (3.29)$$

We also have for the allowed transitions

$$\epsilon_{\zeta''} - \epsilon_{\zeta'} = \hbar\omega_0 \quad [\text{term (a)}],$$
$$\epsilon_{\zeta''} - \epsilon_{\zeta'} = -\hbar\omega_0 \quad [\text{term (b)}]. \quad (3.30)$$

From (3.21) we thus find

$$\sigma_{yx}^{\rm nd}(0) = \frac{e}{2B\Omega}\sum_k \sum_{N=0,1,2,\cdots} \{(N+1)\langle n_N\rangle_{\rm eq}(1 - \langle n_{N+1}\rangle_{\rm eq})(1 - e^{-\beta\hbar\omega_0}) - N\langle n_N\rangle_{\rm eq}(1 - \langle n_{N-1}\rangle_{\rm eq})(1 - e^{\beta\hbar\omega_0})\}. \quad (3.31)$$

(In this expression we suppressed the index k, thus $n_N \equiv n_{Nk}$, etc.). In the second term we change $N \rightarrow N+1$. We then obtain the general exact result

$$\sigma_{yx}^{\rm nd}(0) = \frac{e}{2B\Omega}\sum_k \sum_{N=0,1,2,\cdots} (N+1)\{\langle n_N\rangle_{\rm eq}(1 - \langle n_{N+1}\rangle_{\rm eq})(1 - e^{-\beta\hbar\omega_0}) - \langle n_{N+1}\rangle_{\rm eq}(1 - \langle n_N\rangle_{\rm eq})(1 - e^{\beta\hbar\omega_0})\}. \quad (3.32)$$

The same result has been derived from the quantum mechanical Boltzmann equation (Sec. 6).

In the paper on applications[20] we will investigate (3.32) in detail, and derive a result for the oscillatory Hall effect. Here we consider only the steady Hall effect in nondegenerate semiconductors. We split (3.32) as follows (dropping the super nd since this is the total contribution):

$$\sigma_{yx}(0) = \frac{e}{2B\Omega}\sum_k \sum_N (N+1)\{\langle n_N\rangle_{\rm eq}(1 - \langle n_{N+1}\rangle_{\rm eq}) + \langle n_{N+1}\rangle_{\rm eq}(1 - \langle n_N\rangle_{\rm eq})e^{\beta\hbar\omega_0}\}$$
$$- \frac{e}{2B\Omega}\sum_k \sum_N (N+1)\{\langle n_N\rangle_{\rm eq}(1 - \langle n_{N+1}\rangle_{\rm eq})e^{-\beta\hbar\omega_0} + \langle n_{N+1}\rangle_{\rm eq}(1 - \langle n_N\rangle_{\rm eq})\}. \quad (3.33)$$

From Boltzmann statistics we have ($\epsilon_k = \hbar^2 k_z^2/2m, \langle n\rangle_{\rm eq} \ll 1$):

$$\langle n_N\rangle_{\rm eq} = e^{-\beta[(N+1/2)\hbar\omega_0 + \epsilon_k - \epsilon_F]}, \quad (3.34)$$

$$\langle n_{N+1}\rangle_{\rm eq} e^{\beta\hbar\omega_0} = e^{-\beta[(N+3/2)\hbar\omega_0 + \epsilon_k - \epsilon_F + \beta\hbar\omega_0]} = \langle n_N\rangle_{\rm eq}. \quad (3.35)$$

The two parts within each { } of (3.33) are found to be equal; we thus obtain

$$\sigma_{yx}(0) = \frac{e}{B\Omega}\sum_k \sum_N (N+1)\langle n_N\rangle_{\rm eq} - \frac{e}{B}\sum_k \sum_N (N+1)\langle n_{N+1}\rangle_{\rm eq}. \quad (3.36)$$

In the last term we change $N+1 \rightarrow N$. We then finally find

$$\sigma_{yx}(0) = \frac{e}{B\Omega}\sum_k \sum_N \langle n_N\rangle_{\rm eq} = \frac{e}{B}\frac{\langle n_{\rm total}\rangle}{\Omega} = \frac{en_0}{B}, \quad (3.37)$$

where n_0 is the equilibrium electron density. For high fields this gives $\rho_{yx} \simeq -1/\sigma_{yx} = -B/en_0$, the well-known result.

Note. The nondiagonal Hall effect is the only effect of this kind, as far as we presently see. If we repeat the above derivation for the nondiagonal magnetoconductance, we obtain no contribution. For, analogous to (3.32) we obtain the exact result

$$\sigma_{xx}^{\rm nd}(0) = \frac{ei}{2B\Omega}\sum_k \sum_N (N+1)\{\langle n_N\rangle_{\rm eq}(1 - \langle n_{N+1}\rangle_{\rm eq})(1 - e^{-\beta\hbar\omega_0}) + \langle n_{N+1}\rangle_{\rm eq}(1 - \langle n_N\rangle_{\rm eq})(1 - e^{\beta\hbar\omega_0}), \quad (3.38)$$

which differs from (3.32) by the sign of the two contributions and by the factor i.

Again this is split as follows:

$$\sigma_{xx}^{\rm nd}(0) = \frac{ei}{2B\Omega}\sum_k \sum_N (N+1)\{\langle n_N\rangle_{\rm eq}(1 - \langle n_{N+1}\rangle_{\rm eq}) - \langle n_{N+1}\rangle_{\rm eq}(1 - \langle n_N\rangle_{\rm eq})e^{\beta\hbar\omega_0}$$
$$- \frac{ei}{2B\Omega}\sum_k \sum_N (N+1)\{\langle n_N\rangle_{\rm eq}(1 - \langle n_{N+1}\rangle_{\rm eq})e^{-\beta\hbar\omega_0} - \langle n_{N+1}\rangle_{\rm eq}(1 - \langle n_N\rangle_{\rm eq})\}. \quad (3.39)$$

Since the two terms within each { } are found to cancel each other for nondegenerate statistics we find $\sigma_{xx}^{\rm nd} = 0$. For degenerate statistics there may be a finite imaginary result; this could contribute to the dielectric constant in metals.

4. The diagonal part for electron-phonon interaction

4.1. *The general form of the transition probabilities*

For electron-phonon interaction we have generally instead of Eq. (1.10),

$$\lambda V = i\sum_{\zeta'\zeta''}\sum_{\mathbf{q}'} F(\mathbf{q}')[c^\dagger_{\zeta''}c_{\zeta'}a_{\mathbf{q}'}(\zeta''|e^{i\mathbf{q}'\cdot\mathbf{r}}|\zeta') - c^\dagger_{\zeta''}c_{\zeta'}a^\dagger_{\mathbf{q}'}(\zeta''|e^{-i\mathbf{q}'\cdot\mathbf{r}}|\zeta')], \tag{4.1}$$

where the symbols have their usual meaning; $|\zeta)$ is a general form of one particle fermion state. In case $|\zeta) = e^{i\mathbf{k}\cdot\mathbf{r}}$, Eq. (4.1) condenses to LRT II, Eq. (8.52). The purpose of this section is to show that Theorems 1 and 2 of Sec. 2 remain valid, with the nonlinear Boltzmann operator still given by (2.13), though $w_{\zeta\zeta'}$ is differently defined.

The transition rates $w_{\gamma\bar{\gamma}}$ are calculated as in Sec. 2. Denoting the two parts of (4.1) by the superscripts "abs" and "em" (for absorption and emission of a phonon), we easily find that for given $|\gamma\rangle$ and fixed ζ', ζ'', $\mathbf{q}'$ of the series (4.1), the only connected states in the matrix element $\langle\gamma|\lambda V^{\text{abs}}|\bar{\gamma}\rangle$ are the states $|\bar{\gamma}_{\zeta'\zeta''\mathbf{q}'}\rangle$ such that

$$\begin{aligned} \bar{n}_\zeta &= n_\zeta \quad \text{for } \zeta \neq \zeta' \text{ and } \zeta \neq \zeta'', \\ \bar{n}_{\zeta'} &= 1 - n_{\zeta'}, \quad \bar{n}_{\zeta''} = 1 - n_{\zeta''}, \\ \bar{N}_\mathbf{q} &= N_\mathbf{q} - \delta_{\mathbf{q}\mathbf{q}'}. \end{aligned} \tag{4.2}$$

Likewise, the only connected states in the matrix element $\langle\gamma|\lambda V^{\text{em}}|\bar{\gamma}\rangle$ are the states $|\bar{\gamma}_{\zeta'\zeta''\mathbf{q}'}\rangle$ such that

$$\begin{aligned} \bar{n}_\zeta &= n_\zeta \quad \text{for } \zeta \neq \zeta' \text{ and } \zeta \neq \zeta'', \\ \bar{n}_{\zeta'} &= 1 - n_{\zeta'}, \quad \bar{n}_{\zeta''} = 1 - n_{\zeta''}, \\ \bar{N}_\mathbf{q} &= N_\mathbf{q} + \delta_{\mathbf{q}\mathbf{q}'}. \end{aligned} \tag{4.3}$$

With these data we find

$$W^{\text{abs}}_{\gamma\gamma_{\zeta'\zeta''\mathbf{q}'}} = Q(\zeta',\mathbf{q}'\to\zeta'')(1-n_{\zeta''})n_{\zeta'}N_{\mathbf{q}'}, \tag{4.4}$$

$$W^{\text{em}}_{\gamma\gamma_{\zeta'\zeta''\mathbf{q}'}} = Q(\zeta'\to\zeta'',\mathbf{q}')(1-n_{\zeta''})n_{\zeta'}(1+N_{\mathbf{q}'}), \tag{4.5}$$

where

$$Q(\zeta',\mathbf{q}'\to\zeta'') = (2\pi/\hbar)|F(\mathbf{q}')|^2|(\zeta''|e^{i\mathbf{q}'\cdot\mathbf{r}}|\zeta')|^2\delta(\epsilon_{\zeta'} - \epsilon_{\zeta''} + E_{\mathbf{q}'}), \tag{4.6}$$

$$Q(\zeta'\to\zeta'',\mathbf{q}') = (2\pi/\hbar)|F(\mathbf{q}')|^2|(\zeta''|e^{-i\mathbf{q}'\cdot\mathbf{r}}|\zeta')|^2\delta(\epsilon_{\zeta'} - \epsilon_{\zeta''} - E_{\mathbf{q}'}). \tag{4.7}$$

For the operator result $\langle Mn_{\zeta''}\rangle_b$ we find from (1.3) k_y by performing the boson average (we further drop the prime on $\mathbf{q}$),

$$\langle Mn_{\zeta''}\rangle_b = \sum_{\zeta'\zeta''\mathbf{q}}\{Q(\zeta',\mathbf{q}\to\zeta'')(1-n_{\zeta''})n_{\zeta'}\langle N_\mathbf{q}\rangle_{\text{eq}} + \{Q(\zeta'\to\zeta'',\mathbf{q})(1-n_{\zeta''})n_{\zeta'}(1+\langle N_\mathbf{q}\rangle_{\text{eq}})\} \times (n_{\zeta''} - \bar{n}_{\zeta''}). \tag{4.8}$$

Introducing

$$w_{\zeta'\zeta''} = \sum_\mathbf{q}\{Q(\zeta',\mathbf{q}\to\zeta'')\langle N_\mathbf{q}\rangle_{\text{eq}}) + Q(\zeta'\to\zeta'',\mathbf{q})(1+\langle N_\mathbf{q}\rangle_{\text{eq}}\} \tag{4.9}$$

we find that (4.9) takes exactly the form of (2.21). This proves the validity of Theorems 1 and 2. Also, with the definition (4.9), the property (2.15) remains intact.

4.2. *Magnetic transport phenomena*

For the special case of Landau states, Eq. (3.23) the matrix elements occurring in (4.6) and (4.7) are

$$\begin{aligned}(\zeta''|e^{i\mathbf{q}\cdot\mathbf{r}}|\zeta') &= \int d^3r\, e^{-ik''_y y}e^{-ik''_z z}\phi_{N''}\Big(x + \frac{\hbar k''_y}{m\omega_0}\Big)e^{i\mathbf{q}\cdot\mathbf{r}} \times e^{ik'_y y}e^{ik'_z z}\phi_{N'}\Big(x+\frac{\hbar k'_y}{m\omega_0}\Big)\frac{1}{L_yL_z} \\ &= \delta_{k''_y,k'_y+q_y}\delta_{k''_z,k'_z+q_z}J_{N',N''}(q_x,k''_y,k'_y),\end{aligned} \tag{4.10}$$

where, following Argyres[21] we defined

$$J_{N',N''}(q_x,k''_y,k'_y) = \int_{-\infty}^{\infty} dx\, \phi_{N''}\Big(x+\frac{\hbar k''_y}{m\omega_0}\Big)e^{iq_x x} \times \phi_{N'}\Big(x+\frac{\hbar k'_y}{m\omega_0}\Big). \tag{4.11}$$

Likewise

$$(\zeta''|e^{-i\mathbf{q}\cdot\mathbf{r}}|\zeta') = \delta_{k''_y,k'_y-q_y}\delta_{k''_z,k'_z-q_z}J_{N'N''}(-q_x,k''_y,k'_y). \tag{4.12}$$

Substituting (4.10) and (4.12) into (4.6) and (4.7), we obtain for (4.9),

$$\begin{aligned} w_{\zeta'\zeta''} &= (2\pi/\hbar)\sum_\mathbf{q}|F(\mathbf{q})|^2\{|J_{N',N''}(q_x,k''_y,k'_y)|^2\delta_{k'',k'+q}\langle N_q\rangle_{\text{eq}} \\ &\times\delta(\epsilon_{\zeta'} - \epsilon_{\zeta''} + E_\mathbf{q}) + |J_{N',N''}(-q_x,k''_y,k'_y)|^2 \\ &\times\delta_{k'',k'-q}(1+\langle N_\mathbf{q}\rangle_{\text{eq}}) \\ &\times\delta(\epsilon_{\zeta'} - \epsilon_{\zeta''} - E_\mathbf{q}\}; \end{aligned} \tag{4.13}$$

here $\delta_{k'',k'+\mathbf{q}}$ stands for $\delta_{k''_y,k'_y+q_y}\delta_{k''_z,k'_z+q_z}$ and it is to be remembered that ζ represents N, k_y, k_z.

Argyres[21] indicated that $J_{N',N''}$ depends only on $q_x^2 + (k'_y - k''_y)^2$, i.e., on $q_x^2 + q_y^2 = q_\perp^2$. Enck *et al.*[22] have calculated $J_{N',N''}$. The result is

$$|J_{N',N''}(q_\perp^2)|^2 = \frac{N''!}{N'!}\exp\Big(-\frac{\lambda^2q_\perp^2}{2}\Big)\Big(\frac{\lambda^2q_\perp^2}{2}\Big)^{N'-N''} \times\Big[\mathscr{L}^{N'-N''}_{N''}\Big(\frac{\lambda^2q_\perp^2}{2}\Big)\Big]^2, \quad N''\leqslant N', \tag{4.14}$$

where $\mathscr{L}^m_n$ is an associated Laguerre polynomial and where $\lambda^2 = \hbar/m\omega_0$. Also,

$$|(J_{N',N''}(q_\perp^2)|^2 = \frac{N'!}{N''!}\exp\Big(-\frac{\lambda^2q_\perp^2}{2}\Big)\Big(\frac{\lambda^2q_\perp^2}{2}\Big)^{N''-N'} \times\Big[\mathscr{L}^{N''-N'}_{N'}\Big(\frac{\lambda^2q_\perp^2}{2}\Big)\Big], \quad N'\leqslant N''. \tag{4.15}$$

Since $|J_{N',N''}|^2$ is the same for $\pm q_x$, Eq. (4.13) simplifies to

$$\begin{aligned} w_{\zeta'\zeta''} = \frac{2\pi}{\hbar}\sum_\mathbf{q}|F(\mathbf{q})|^2|J_{N',N''}|^2\{&\delta_{k'',k'+q}\langle N_\mathbf{q}\rangle_{\text{eq}} \\ \times\delta(\epsilon_{\zeta'} - \epsilon_{\zeta''} + E_\mathbf{q}) &+ \delta_{k'',k'-q}(1+\langle N_\mathbf{q}\rangle_{\text{eq}}) \\ \times\delta(\epsilon_{\zeta'} - \epsilon_{\zeta''} - E_\mathbf{q})\}&. \end{aligned} \tag{4.16}$$

This is the transition rate that is to be used in the generalized Adams-Holstein result of Eq. (2.84). Various applications will be discussed in a forthcoming article.[20]

B. QUANTUM MECHANICAL BOLTZMANN EQUATION

5. The diagonal Boltzmann equation

In this section we derive the quantum mechanical Boltzmann equation of LRT II Sec. 8 by a faster method than before. This method is well adapted in order to find an extension that includes the nondiagonal part, to be set forth in Sec. 6. The inhomogeneous Markovian master equation reads [see LRT II, Eqs. (4.30)]

$$\frac{\partial \rho_d^R(t)}{\partial t} + \Lambda_d \rho_d^R(t) = \beta F(t) \rho_{eq} [-\Lambda_d A_d + (\dot{A})_d]. \qquad (5.1)$$

The first moment equation of this is the Boltzmann equation. Thus, with $\langle n_\zeta \rangle_t = \mathrm{Tr}[\mathbf{n}_\zeta \rho(t)]$,

$$\frac{\partial \langle n_\zeta \rangle_t}{\partial t} + \mathrm{Tr}\{\mathbf{n}_\zeta \Lambda_d \rho_d^R(t)\} = \beta F(t) \mathrm{Tr}\{\rho_{eq} \mathbf{n}_\zeta [-\Lambda_d \mathrm{A}_d + (\dot{A})_d]\}, \qquad (5.2)$$

where we noticed that $\mathbf{n}_\zeta$ and ρ_{eq} commute. We now apply Lemma 1 of LRT II [Eq. (C1)]; for any two operators C and D we have

$$\mathrm{Tr}\,(C\Lambda_d D) = \mathrm{Tr}(D\Lambda_d C). \qquad (5.3)$$

Thus we obtain

$$\frac{\partial \langle n_\zeta \rangle_t}{\partial t} + \mathrm{Tr}\{\rho_d^R(t)\Lambda_d \mathbf{n}_\zeta\} = \beta F(t) \mathrm{Tr}\{\rho_{eq} \mathbf{n}_\zeta (-\Lambda_d A_d)\} + \beta F(t) \mathrm{Tr}\{\rho_{eq} \mathbf{n}_\zeta (\dot{A})_d\}. \qquad (5.4)$$

Using Theorem 2, Eq. (2.16), the second term to the left becomes

$$\text{2nd term lhs} = \langle \mathscr{B}_\zeta n_\zeta \rangle_t = -\left(\frac{\partial \langle n_\zeta \rangle_t}{\partial t}\right)_{coll}. \qquad (5.5)$$

For $(\dot{A})_d$ we write $\Sigma_{\zeta'} \mathbf{n}_{\zeta'} (\zeta'|\dot{a}|\zeta')$. The second term to the right then involves the average

$$\sum_{\zeta'} \langle n_\zeta n_{\zeta'} \rangle_{eq} (\zeta'|\dot{a}|\zeta') = \sum_{\zeta' \neq \zeta} \langle n_\zeta \rangle_{eq} \langle n_{\zeta'} \rangle_{eq} (\zeta'|\dot{a}|\zeta') + \langle n_\zeta^2 \rangle_{eq} (\zeta|\dot{a}|\zeta)$$
$$= \langle n_\zeta \rangle_{eq} \sum_{\zeta'} \langle n_{\zeta'} \rangle_{eq} (\zeta'|\dot{a}|\zeta') + [\langle n_\zeta^2 \rangle - \langle n_\zeta \rangle^2] (\zeta|\dot{a}|\zeta); \qquad (5.6)$$

the first sum is zero since $\langle \dot{A}_d \rangle_{eq} = 0$ [see LRT II, Eq. (6.22')]. We thus have for (5.4)

$$\text{2nd term rhs} = \beta F(t) \langle n_\zeta \rangle_{eq} (1 - \langle n_\zeta \rangle_{eq}) (\zeta|\dot{a}|\zeta). \qquad (5.7)$$

For the first term on the right we write $A_d = \sum_{\zeta'} \mathbf{n}_{\zeta'} (\zeta'|a|\zeta')$; this term involves the average

$$\sum_{\zeta'} \langle n_\zeta (-\Lambda_d \mathbf{n}_{\zeta'})_b (\zeta'|a|\zeta') \rangle_{eq} = -\sum_{\zeta'} \langle n_\zeta \mathscr{B}_{\zeta'} n_{\zeta'} \rangle_{eq} (\zeta'|a|\zeta'); \qquad (5.8)$$

this was computed in LRT II Eqs. (8.43) ff. The result is

lst term rhs

$$= -\beta F(t) \langle n_\zeta \rangle_{eq} (1 - \langle n_\zeta \rangle_{eq}) \sum_{\zeta'} \{ [(\zeta|a|\zeta) - (\zeta'|a|\zeta')] \times [w_{\zeta\zeta'} (1 - \langle n_{\zeta'} \rangle_{eq}) + w_{\zeta'\zeta} \langle n_{\zeta'} \rangle_{eq}]\}. \qquad (5.9)$$

Thus from (5.5), (5.7), and (5.9) we find the quantum Boltzmann equation

$$\frac{\partial \langle n_\zeta \rangle_t}{\partial t} - \beta F(t) \langle n_\zeta \rangle_{eq} (1 - \langle n_\zeta \rangle_{eq}) (\zeta|\dot{a}|\zeta) - \beta F(t) \langle n_\zeta \rangle_{eq} (1 - \langle n_\zeta \rangle_{eq}) \times \sum_{\zeta'} \{ [(\zeta'|a|\zeta') - (\zeta|a|\zeta)] [w_{\zeta\zeta'} (1 - \langle n_{\zeta'} \rangle_{eq}) + w_{\zeta'\zeta} \langle n_{\zeta'} \rangle_{eq}]\}$$
$$= \sum_{\zeta'} [w_{\zeta'\zeta} \langle n_{\zeta'} \rangle_t (1 - \langle n_\zeta \rangle_t) - w_{\zeta\zeta'} \langle n_\zeta \rangle_t (1 - \langle n_{\zeta'} \rangle_t)]. \qquad (5.10)$$

On the nature of the two streaming terms we commented in LRT II.

6. The full quantum Boltzmann equation

We now start from the inhomogeneous complete evolution equation, LRT II, Eq. (4.41),

$$\frac{\partial \rho^R(t)}{\partial t} + (\Lambda_d + i\mathscr{L}^0) \rho^R(t) = F(t) \rho_{eq} \int_0^\beta d\beta' e^{\hbar\beta'\mathscr{L}^0} [-\Lambda_d A_d + (\dot{A})_d + (\dot{A})_{nd}]. \qquad (6.1)$$

We seek an equation for $\partial \langle c_{\zeta_1}^\dagger c_{\zeta_2} \rangle_t / \partial t$. Thus, we multiply (6.1) by $c_{\zeta_1}^\dagger c_{\zeta_2}$ and take the trace; next we use the lemma

$$\mathrm{Tr}\,[C(\Lambda_d + i\mathscr{L}^0)D] = \mathrm{Tr}[D(\Lambda_d - i\mathscr{L}^0)C] \qquad (6.2)$$

(see LRT II, Appendix C). Then we obtain

$$\frac{\partial \langle c_{\zeta_1}^\dagger c_{\zeta_2} \rangle_t}{\partial t} + \mathrm{Tr}[\rho^R(t)(\Lambda_d - i\mathscr{L}^0) c_{\zeta_1}^\dagger c_{\zeta_2}] = \beta F(t) \mathrm{Tr}[\rho_{eq} (-\Lambda_d A_d) c_{\zeta_1}^\dagger c_{\zeta_2} + \rho_{eq} (\dot{A})_d c_{\zeta_1}^\dagger c_{\zeta_2}] + F(t) \int_0^\beta d\beta' \,\mathrm{Tr}[\rho_{eq} (e^{\hbar\beta'\mathscr{L}^0} (\dot{A})_{nd} c_{\zeta_1}^\dagger c_{\zeta_2}], \qquad (6.3)$$

where we notice that $\mathscr{L}^0$ (diagonal operator) = 0.

For the second term to the left we note that Λ_d destroys a nondiagonal operator. Thus

$$\Lambda_d c_{\zeta_1}^\dagger c_{\zeta_2} = \Lambda_d \mathbf{n}_{\zeta_1} \delta_{\zeta_1\zeta_2};$$

the result for this part is by (5.5),

$$\mathrm{Tr}[\rho^R(t) \Lambda_d c_{\zeta_1}^\dagger c_{\zeta_2}] = \langle \mathscr{B}_{\zeta_1} n_{\zeta_1} \rangle_t \delta_{\zeta_1\zeta_2}. \qquad (6.4)$$

For the other part of this term we have

$$-i\mathscr{L}^0 c_{\zeta_1}^\dagger c_{\zeta_2} = (i/\hbar)[c_{\zeta_1}^\dagger c_{\zeta_2}, H^0], \qquad (6.5)$$

so that we obtain

$$\mathrm{Tr}[\rho^R(t)(-i\mathscr{L}^0 c_{\zeta_1}^\dagger c_{\zeta_2})] = \frac{i}{\hbar} \sum_{\gamma\bar{\gamma}} \{ \langle \gamma|\rho^R(t)|\bar{\gamma}\rangle \langle \bar{\gamma}|c_{\zeta_1}^\dagger c_{\zeta_2}|\gamma\rangle \mathscr{E}_\gamma - \langle \gamma|\rho^R(t)|\bar{\gamma}\rangle \mathscr{E}_{\bar{\gamma}} \langle \bar{\gamma}|c_{\zeta_1}^\dagger c_{\zeta_2}|\gamma\rangle \}$$
$$= \frac{i}{\hbar} \sum_{\gamma\bar{\gamma}} \langle \gamma|\rho^R(t)|\bar{\gamma}\rangle \langle \bar{\gamma}|c_{\zeta_1}^\dagger c_{\zeta_2}|\gamma\rangle (\mathscr{E}_\gamma - \mathscr{E}_{\bar{\gamma}}). \qquad (6.6)$$

Now if we take $|\gamma\rangle = |\{n_\zeta\}, \{N_\eta\}\rangle$, then $|\bar{\gamma}\rangle$ can only be such that

$$\bar{n}_{\zeta_1} = 1 - n_{\zeta_1}, \quad \bar{n}_{\zeta_2} = 1 - n_{\zeta_2},$$
$$\text{all other } \bar{n}_\zeta = n_\zeta, \qquad (6.7)$$
$$\text{all } \bar{N}_\eta = N_\eta.$$

Since n_{ζ_2} is lowered and n_{ζ_1} is raised we have

$$\mathscr{E}_{\bar{\gamma}} = \mathscr{E}_\gamma + \epsilon_{\zeta_1} - \epsilon_{\zeta_2}. \tag{6.8}$$

Thus we find

$$\begin{aligned}\mathrm{Tr}[\rho^R(t)(-i\mathscr{L}^0 c^\dagger_{\zeta_1} c_{\zeta_2})] &= (i/\hbar)(\epsilon_{\zeta_2} - \epsilon_{\zeta_1}) \sum_\gamma \langle\gamma|\rho^R(t) c^\dagger_{\zeta_1} c_{\zeta_2}|\gamma\rangle \\ &= (i/\hbar)(\epsilon_{\zeta_2} - \epsilon_{\zeta_1}) \langle c^\dagger_{\zeta_1} c_{\zeta_2}\rangle_t.\end{aligned} \tag{6.9}$$

This is an off-diagonal contribution. There is no diagonal part since $\epsilon_{\zeta_2} - \epsilon_{\zeta_1} = 0$ for $\zeta_1 = \zeta_2$.

For the first term on the right of (6.3) we note that

$$\begin{aligned}\mathrm{Tr}[\rho_{\mathrm{eq}}(-\Lambda_d A_d + (\dot{A})_d) c^\dagger_{\zeta_1} c_{\zeta_2}] &= \mathrm{Tr}[\rho_{\mathrm{eq}}(-\Lambda_d A_d + (\dot{A})_d) \mathrm{n}_{\zeta_1}]\delta_{\zeta_1\zeta_2}.\end{aligned} \tag{6.10}$$

Thus this term yields [(5.7) + (5.9)] times $\delta_{\zeta_1\zeta_2}$.

Finally, the last term of (6.3) is obtained in the following manner: we substitute

$$(\dot{A})_{\mathrm{nd}} = \sum_{\zeta_3\zeta_4} c^\dagger_{\zeta_3} c_{\zeta_4} (\zeta_3|\dot{a}|\zeta_4)(1 - \delta_{\zeta_3\zeta_4}). \tag{6.11}$$

Then,

$$\begin{aligned}&\mathrm{Tr}[\rho_{\mathrm{eq}}(e^{\hbar\beta'\mathscr{L}^0}(\dot{A})_{\mathrm{nd}}) c^\dagger_{\zeta_1} c_{\zeta_2}] \\ &= \sum_{\zeta_3\zeta_4} (\zeta_3|\dot{a}|\zeta_4) \mathrm{Tr}\{\rho_{\mathrm{eq}} e^{\beta' H^0} c^\dagger_{\zeta_3} c_{\zeta_4} e^{-\beta' H^0} c^\dagger_{\zeta_1} c_{\zeta_2}\}(1 - \delta_{\zeta_3\zeta_4}) \\ &= \sum_{\zeta_3\zeta_4} (\zeta_3|\dot{a}|\zeta_4) \sum_{\gamma\bar{\gamma}} p_{\mathrm{eq}}(\gamma) e^{\beta'(\mathscr{E}_\gamma - \mathscr{E}_{\bar{\gamma}})} \langle\gamma|c^\dagger_{\zeta_3} c_{\zeta_4}|\bar{\gamma}\rangle \\ &\quad \times \langle\bar{\gamma}|c^\dagger_{\zeta_1} c_{\zeta_2}|\gamma\rangle.\end{aligned} \tag{6.12}$$

Now if $|\gamma\rangle = |\{n_\zeta\}, \{N_\eta\}\rangle$, $|\bar{\gamma}\rangle$ must satisfy the rule (6.7) to make the matrix element $\langle\bar{\gamma}|c^\dagger_{\zeta_1} c_{\zeta_2}|\gamma\rangle$ nonzero. However, in order that $\langle\gamma|c^\dagger_{\zeta_3} c_{\zeta_4}|\bar{\gamma}\rangle$ is nonzero, we must have $\zeta_3 = \zeta_2$ and $\zeta_4 = \zeta_1$ by a similar argument as in Sec. 3. Thus

$$\langle\gamma|c^\dagger_{\zeta_3} c_{\zeta_4}|\bar{\gamma}\rangle\langle\bar{\gamma}|c^\dagger_{\zeta_1} c_{\zeta_2}|\gamma\rangle = (1 - n_{\zeta_1}) n_{\zeta_2} \delta_{\zeta_3\zeta_2}\delta_{\zeta_4\zeta_1}. \tag{6.13}$$

For $\mathscr{E}_{\bar{\gamma}} - \mathscr{E}_\gamma$ we have again (6.8).

Thus (6.12) gives

$$\begin{aligned}&\mathrm{Tr}[\rho_{\mathrm{eq}}(e^{\hbar\beta'\mathscr{L}^0}(\dot{A})_{\mathrm{nd}} c^\dagger_{\zeta_1} c_{\zeta_2}] \\ &= (\zeta_2|\dot{a}|\zeta_1) \sum_\gamma [p_{\mathrm{eq}}(\gamma) e^{\beta'(\epsilon_{\zeta_2} - \epsilon_{\zeta_1})}(1 - n_{\zeta_1}) n_{\zeta_2}](1 - \delta_{\zeta_1\zeta_2}) \\ &= (\zeta_2|\dot{a}|\zeta_1) e^{\beta'(\epsilon_{\zeta_2} - \epsilon_{\zeta_1})}(1 - \langle n_{\zeta_1}\rangle_{\mathrm{eq}})\langle n_{\zeta_2}\rangle_{\mathrm{eq}}(1 - \delta_{\zeta_1\zeta_2}).\end{aligned} \tag{6.14}$$

Integration over $d\beta'$ yields for the streaming term

$$\begin{aligned}F(t)\int_0^\beta d\beta'(\text{above}) &= F(t)\frac{1 - e^{-\beta(\epsilon_{\zeta_1} - \epsilon_{\zeta_2})}}{\epsilon_{\zeta_1} - \epsilon_{\zeta_2}}(1 - \langle n_{\zeta_1}\rangle_{\mathrm{eq}} \\ &\quad \times \langle n_{\zeta_2}\rangle_{\mathrm{eq}}(\zeta_2|\dot{a}|\zeta_1)(1 - \delta_{\zeta_1\zeta_2}).\end{aligned} \tag{6.15}$$

We can *a posteriori* combine this term with the result due to $(\dot{A})_d$, for we have

$$\begin{aligned}&\beta F(t)\langle n_{\zeta_1}\rangle_{\mathrm{eq}}(1 - \langle n_{\zeta_1}\rangle_{\mathrm{eq}})(\zeta_1|a|\zeta_1) \\ &= F(t)\frac{1 - e^{-\beta(\epsilon_{\zeta_1} - \epsilon_{\zeta_2})}}{\epsilon_{\zeta_1} - \epsilon_{\zeta_2}}\langle n_{\zeta_2}\rangle_{\mathrm{eq}}(1 - \langle n_{\zeta_1}\rangle_{\mathrm{eq}})(\zeta_2|\dot{a}|\zeta_1)\delta_{\zeta_1\zeta_2}.\end{aligned} \tag{6.16}$$

The total effect of the streaming due to $(\dot{A})_d + (\dot{A})_{\mathrm{nd}}$ is thus the result (6.15) with the factor $(1 - \delta_{\zeta_1\zeta_2})$ omitted.

Collecting all terms, the full quantum Boltzmann equation becomes

$$\begin{aligned}&\frac{\partial\langle c^\dagger_{\zeta_1} c_{\zeta_2}\rangle_t}{\partial t} - F(t)\frac{1 - e^{-\beta(\epsilon_{\zeta_1} - \epsilon_{\zeta_2})}}{\epsilon_{\zeta_1} - \epsilon_{\zeta_2}}\langle n_{\zeta_2}\rangle_{\mathrm{eq}}(1 - \langle n_{\zeta_1}\rangle_{\mathrm{eq}}) \\ &\quad \times(\zeta_2|\dot{a}|\zeta_1) - \beta F(t)\langle n_{\zeta_1}\rangle_{\mathrm{eq}}(1 - \langle n_{\zeta_1}\rangle_{\mathrm{eq}}) \\ &\quad \times\sum_{\zeta'}\{[(\zeta'|a|\zeta') - (\zeta_1|a|\zeta_1)] \\ &\quad \times[w_{\zeta_1\zeta'}(1 - \langle n_{\zeta'}\rangle_{\mathrm{eq}}) + w_{\zeta'\zeta_1}\langle n_{\zeta'}\rangle_{\mathrm{eq}}]\}\delta_{\zeta_1\zeta_2} \\ &= \sum_{\zeta'}[w_{\zeta'\zeta_1}\langle n_{\zeta'}\rangle_t(1 - \langle n_{\zeta_1}\rangle_t) - w_{\zeta_1\zeta'}\langle n_{\zeta_1}\rangle_t(1 - \langle n_{\zeta'}\rangle_t)] \\ &\quad \times\delta_{\zeta_1\zeta_2} - (i/\hbar)(\epsilon_{\zeta_2} - \epsilon_{\zeta_1})\langle c^\dagger_{\zeta_1} c_{\zeta_2}\rangle_t.\end{aligned} \tag{6.17}$$

7. Equivalence with the linear response results

We will show complete equivalence of the Boltzmann procedure with the linear response procedure, by demonstrating that the response formulas also lead to the Boltzmann equation (6.17).

According to the general linear response idea, we have that

$$L_{BA}(i\omega) = \int_0^\infty dt\, e^{-i\omega t}\phi_{\dot{B}A}(t), \tag{7.1}$$

where ϕ is the response function [see LRT I, Eq. (3.9)]. Thus from (2.72) we have for the diagonal part of ϕ.

$$\begin{aligned}\phi^d_{\dot{B}A}(t) = \beta\sum_{\zeta'\zeta''}\sum_{k=0}^\infty \frac{(-t)^k}{k!}\langle[-(\mathscr{B}_{\zeta'} n_{\zeta'}) a_{\zeta'} + n_{\zeta'}\dot{a}_{\zeta'}] \\ \times[-(\mathscr{B}^{(k+1)}_{\zeta''} n_{\zeta''}) b_{\zeta''} + (\mathscr{B}^{(k)}_{\zeta''} n_{\zeta''})\dot{b}_{\zeta''}]\rangle_{\mathrm{eq}}.\end{aligned} \tag{7.2}$$

Thus for the current $J_{B,d}$ caused by the force $F(t)$ we find [see LRT I, Eq. (2.20)],

$$\begin{aligned}\langle\Delta J_{B,d}\rangle_t = \beta\int_0^t d\tau\, F(\tau)\sum_{\zeta'\zeta''}\sum_{k=0}^\infty \frac{(-t+\tau)^k}{k!} \\ \times\langle[-(\mathscr{B}_{\zeta'} n_{\zeta'}) a_{\zeta'} + n_{\zeta'}\dot{a}_{\zeta'}] \\ \times[-(\mathscr{B}^{(k+1)}_{\zeta''} n_{\zeta''}) b_{\zeta''} + (\mathscr{B}^{(k)}_{\zeta''} n_{\zeta''})\dot{b}_{\zeta''}]\rangle_{\mathrm{eq}}.\end{aligned} \tag{7.3}$$

But also generally, cf. (2.36), for any current average,

$$\langle\Delta J_{B,d}\rangle_t = \sum_{\zeta''}[-\langle\mathscr{B}_{\zeta''} n_{\zeta''}\rangle_t b_{\zeta''} + \langle n_{\zeta''}\rangle_t\dot{b}_{\zeta''}]. \tag{7.4}$$

Comparing (7.3) with (7.4) we find the following identities:

$$\begin{aligned}\langle n_\zeta\rangle_t = \beta\int_0^t d\tau\, F(\tau)\sum_{\zeta'}\sum_{k=0}^\infty \frac{(-t+\tau)^k}{k!} \\ \times\langle[-(\mathscr{B}_{\zeta'} n_{\zeta'}) a_{\zeta'} + n_{\zeta'}\dot{a}_{\zeta'}](\mathscr{B}^{(k)}_\zeta n_\zeta)\rangle_{\mathrm{eq}},\end{aligned} \tag{7.5}$$

$$\begin{aligned}\langle\mathscr{B}_\zeta n_\zeta\rangle_t = \beta\int_0^t d\tau\, F(\tau)\sum_{\zeta'}\sum_{k=0}^\infty \frac{(-t+\tau)^k}{k!} \\ \times\langle[-(\mathscr{B}_{\zeta'} n_{\zeta'}) a_{\zeta'} + n_{\zeta'}\dot{a}_{\zeta'}](\mathscr{B}^{(k+1)}_\zeta n_\zeta)\rangle_{\mathrm{eq}}.\end{aligned} \tag{7.6}$$

We now differentiate (7.5); in differentiating the integral only the term with $k = 0$ survives, and in differentiating the integrand we replace $k \to k+1$. Thus we obtain

$$\begin{aligned}\frac{\partial\langle n_\zeta\rangle_t}{\partial t} &= \beta F(t)\sum_{\zeta'}\langle[-(\mathscr{B}_{\zeta'} n_{\zeta'}) a_{\zeta'} + n_{\zeta'}\dot{a}_{\zeta'}] n_\zeta\rangle_{\mathrm{eq}} \\ &\quad - \beta\int_0^t d\tau\, F(\tau)\sum_{\zeta'}\sum_{k=0}^\infty \frac{(-t+\tau)^k}{k!} \\ &\quad \times\langle[-(\mathscr{B}_{\zeta'} n_{\zeta'}) a_{\zeta'} + n_{\zeta'}\dot{a}_{\zeta'}](\mathscr{B}^{(k+1)}_\zeta n_\zeta)\rangle_{\mathrm{eq}}.\end{aligned} \tag{7.7}$$

The last part is just $-\langle \mathscr{B}_{\zeta'} n_\zeta \rangle_t$ by (7.6). We thus have

$$\frac{\partial \langle n_\zeta \rangle_t}{\partial t} - \beta F(t) \sum_{\zeta'} [\langle n_\zeta n_{\zeta'} \rangle_{\text{eq}} \dot{a}_{\zeta'} - \langle n_\zeta \mathscr{B}_{\zeta'} n_{\zeta'} \rangle_{\text{eq}} a_{\zeta'}] = -\langle \mathscr{B}_\zeta n_\zeta \rangle_t \tag{7.8}$$

which one easily recognizes as the diagonal Boltzmann equation.

The proof has the drawback that one cannot obtain (7.5) *and* (7.6) if either one of the sets of matrix elements $\{b_\zeta\}$ or $\{\dot{b}_\zeta\}$ is zero, as if often the case. For the linear Boltzmann operator one can, however, easily deduce (7.5) from (7.6) and vice versa.

For the nondiagonal part we proceed similarly. From (7.1) and the result for L_{BA} analogous to (3.18) we have the response function

$$\phi^{\text{nd}}_{BA}(t) = \sum_{\zeta'\zeta''}{}' \langle n_{\zeta'} \rangle_{\text{eq}} (1 - \langle n_{\zeta''} \rangle_{\text{eq}}) \frac{1 - e^{-\beta(\epsilon_{\zeta''} - \epsilon_{\zeta'})}}{\epsilon_{\zeta''} - \epsilon_{\zeta'}} e^{it(\epsilon_{\zeta''} - \epsilon_{\zeta'})/\hbar} \times (\zeta'|\dot{a}|\zeta'')(\zeta''|\dot{b}|\zeta'). \tag{7.9}$$

This gives for the contribution to current due to the nondiagonal part of ρ (Σ' means $\zeta' \neq \zeta''$):

$$\langle \Delta J_{B,\text{nd}} \rangle_t = \int_0^t d\tau\, F(\tau) \sum_{\zeta'\zeta''}{}' \langle n_{\zeta'} \rangle_{\text{eq}} (1 - \langle n_{\zeta''} \rangle_{\text{eq}}) \times \frac{1 - e^{-\beta(\epsilon_{\zeta''} - \epsilon_{\zeta'})}}{\epsilon_{\zeta''} - \epsilon_{\zeta'}} e^{i(t-\tau)(\epsilon_{\zeta''} - \epsilon_{\zeta'})/\hbar} \times (\zeta'|\dot{a}|\zeta'')(\zeta''|\dot{b}|\zeta'). \tag{7.10}$$

But generally we have also

$$\langle \Delta J_{B,\text{nd}} \rangle_t = \sum_{\zeta'\zeta''}{}' \langle c^\dagger_{\zeta''} c_{\zeta'} \rangle_t (\zeta''|\dot{b}|\zeta'). \tag{7.11}$$

Comparing (7.10) and (7.11) we conclude that

$$\langle c^\dagger_{\zeta''} c_{\zeta'} \rangle_t = \int_0^t d\tau\, F(\tau) \langle n_{\zeta'} \rangle_{\text{eq}} (1 - \langle n_{\zeta''} \rangle_{\text{eq}}) \frac{1 - e^{-\beta(\epsilon_{\zeta''} - \epsilon_{\zeta'})}}{\epsilon_{\zeta''} - \epsilon_{\zeta'}} \times e^{i(t-\tau)(\epsilon_{\zeta''} - \epsilon_{\zeta'})/\hbar} (\zeta'|\dot{a}|\zeta''). \tag{7.12}$$

Differentiating we find

$$\frac{\partial \langle c^\dagger_{\zeta''} c_{\zeta'} \rangle_t}{\partial t} = F(t) \langle n_{\zeta'} \rangle_{\text{eq}} (1 - \langle n_{\zeta''} \rangle_{\text{eq}}) \frac{1 - e^{-\beta(\epsilon_{\zeta''} - \epsilon_{\zeta'})}}{\epsilon_{\zeta''} - \epsilon_{\zeta'}} (\zeta'|\dot{a}|\zeta'') - (i/\hbar)(\epsilon_{\zeta'} - \epsilon_{\zeta''}) \langle c^\dagger_{\zeta''} c_{\zeta'} \rangle_t, \tag{7.13}$$

which corresponds to the nondiagonal part ($\zeta'' \neq \zeta'$) of (6.17). We still note that one can also write the streaming term differently; from equilibrium statistics one has

$$\langle n_{\zeta'} \rangle_{\text{eq}} (1 - \langle n_{\zeta''} \rangle_{\text{eq}}) (1 - e^{-\beta(\epsilon_{\zeta''} - \epsilon_{\zeta'})}) / (\epsilon_{\zeta''} - \epsilon_{\zeta'}) = \langle n_{\zeta''} \rangle_{\text{eq}} (1 - \langle n_{\zeta'} \rangle_{\text{eq}}) (1 - e^{-\beta(\epsilon_{\zeta'} - \epsilon_{\zeta''})}) / (\epsilon_{\zeta'} - \epsilon_{\zeta''}). \tag{7.14}$$

C. INHOMOGENEOUS SYSTEMS

8. Boltzmann equation for the one particle Wigner function

It is in the nature of the quantum mechanical results that the streaming term associated with the spatial gradient is absent. This is due to the fact that we deal with c operators or occupancies of states of the set $\{|\zeta\rangle\}$ (which may for certain systems represent momentum states or Bloch functions), the specification of which is incompatible with spatial localization. A classical analog can be obtained, however, by appealing to the Wigner function, see, e.g., de Groot.[23] We will show that, curiously, *the nondiagonal parts of the full quantum Boltzmann equation, lead to the recovery of the spatial gradient term, necessary for inhomogeneous systems.*

The many particle Wigner function is defined as (h^{-3N}) times the Weyl transform of the density operator $\rho(t)$; thus we have [23]

$$\rho(p,q,t) = (1/h^{3N}) \text{Tr}\{\rho(t) \Delta(p,q)\}, \tag{8.1}$$

where

$$\Delta(p,q) = \int d^{3N}v\, e^{(i/\hbar)\sum_i \mathbf{p}_i \cdot \mathbf{v}_i} \prod_{i=1}^N |\mathbf{q}_i + \tfrac{1}{2}\mathbf{v}_i\rangle \langle \mathbf{q}_i - \tfrac{1}{2}\mathbf{v}_i|; \tag{8.2}$$

the subscript i stands for the coordinates of particle i. The antisymmetrical second quantization form is obtained by writing

$$\Delta^s(p,q) = \frac{1}{N!} \int d^{3N}q : \prod_{j=1}^N \psi^\dagger(\mathbf{q}_j) \Delta(p,q) \psi(\mathbf{q}_j): \tag{8.3}$$

where :II: denotes a normal ordered product. For one factor of the product (8.3) we find

$$\begin{aligned} &\int d^3q\, \psi^\dagger(\mathbf{q}) |\mathbf{q} + \tfrac{1}{2}\mathbf{v}\rangle \langle \mathbf{q} - \tfrac{1}{2}\mathbf{v}| \psi(\mathbf{q}) \\ &= \int d^3q\, \psi^\dagger(\mathbf{q}) e^{-(i/\hbar)\mathbf{v}\cdot\mathbf{P}} |\mathbf{q} - \tfrac{1}{2}\mathbf{v}\rangle \langle \mathbf{q} - \tfrac{1}{2}\mathbf{v}| \psi(\mathbf{q}) \\ &= \int d^3q\, \psi^\dagger(\mathbf{q}) e^{-(i/\hbar)\mathbf{v}\cdot\mathbf{P}} |\delta(\mathbf{Q} - \mathbf{q} + \tfrac{1}{2}\mathbf{v}) \psi(\mathbf{q}) \\ &= \psi^\dagger(\mathbf{Q} + \tfrac{1}{2}\mathbf{v}) e^{-(i/\hbar)\mathbf{v}\cdot\mathbf{P}} \psi(\mathbf{Q} + \tfrac{1}{2}\mathbf{v}). \end{aligned} \tag{8.4}$$

Here capitals refer to the quantum operators. Noting that $\mathbf{Q} = \mathbf{q}$ and $\mathbf{P} = (\hbar/i)\nabla$, we obtain with

$$e^{-\mathbf{v}\cdot\nabla} f(\mathbf{q}) = f(\mathbf{q} - \mathbf{v}) \tag{8.5}$$

the following form (compare Balescu[24]):

$$\rho(p,q,t) = \frac{1}{h^{3N} N!} \int d^{3N}v\, e^{(i/\hbar)pv} \times \text{Tr}\{\rho(t) : \prod_j \psi^+(\mathbf{q}_j + \tfrac{1}{2}\mathbf{v}_j) \psi(\mathbf{q}_j - \tfrac{1}{2}\mathbf{v}_j) :\}, \tag{8.6}$$

where we wrote

$$pv = \sum_{i=1}^N \mathbf{p}_i \cdot \mathbf{v}_i. \tag{8.7}$$

To change to the occupation number form used in the rest of this article we write

$$\begin{aligned} \psi^\dagger(\mathbf{q}) &= \frac{1}{\Omega^{1/2}} \sum_{\mathbf{k}'} e^{-i\mathbf{k}'\cdot\mathbf{q}} \phi^*_{\mathbf{k}'}(\mathbf{q}) c^\dagger_{\mathbf{k}'}, \\ \psi(\mathbf{q}) &= \frac{1}{\Omega^{1/2}} \sum_{\mathbf{k}} e^{i\mathbf{k}\cdot\mathbf{q}} \phi_{\mathbf{k}}(\mathbf{q}) c_{\mathbf{k}}, \end{aligned} \tag{8.8}$$

where $\phi_{\mathbf{k}}$ is a periodic function on the lattice for Bloch electrons and $\phi_{\mathbf{k}} \equiv 1$ for free particles; as usual $\Sigma_{\mathbf{k}} \to (\Omega/8\pi^3) \int d^3k$. Substitution of (8.8) into (8.6) results in

$$\rho(p,q,t) = \frac{\Omega^N}{(8\pi^3)^{2N}h^{3N}N!}$$
$$\times\int\int\int d^{3N}k'\, d^{3N}k\, d^{3N}v\, e^{iv[(p/\hbar)-(k'+k)/2]}e^{iq(k-k')}$$
$$\times\prod_j \phi^*_{k'}(\mathbf{q}_j+\tfrac{1}{2}\mathbf{v}_j)\phi_{\mathbf{k}}(\mathbf{q}_j-\tfrac{1}{2}\mathbf{v}_j)\mathrm{Tr}\{\rho(t):\prod_j c^\dagger_{k'}c_{\mathbf{k}}:\}. \tag{8.9}$$

We make the transformation

$$\left.\begin{matrix}\mathbf{k}_i-\mathbf{k}'_i=\mathbf{u}_i\\ \mathbf{k}_i+\mathbf{k}'_i=2\mathbf{k}_i\end{matrix}\right\}\quad\text{or}\quad\left.\begin{matrix}\mathbf{k}_i=\mathbf{K}_i+\tfrac{1}{2}\mathbf{u}_i\\ \mathbf{k}'_i=\mathbf{K}_i-\tfrac{1}{2}\mathbf{u}_i\end{matrix}\right\} \tag{8.10}$$

with the Jacobian being unity. We also develop $\phi_{\mathbf{k}}$ in a Fourier series on the reciprocal lattice

$$\phi_{\mathbf{k}}(\mathbf{q}+\tfrac{1}{2}\mathbf{v})=\sum_{\mathbf{g}}A'_{\mathbf{k}}(\mathbf{g})e^{i\mathbf{q}\cdot\mathbf{g}}e^{i\mathbf{v}\cdot\mathbf{g}/2},$$
$$\phi^*_{\mathbf{k}'}(\mathbf{q}-\tfrac{1}{2}\mathbf{v})=\sum_{\mathbf{g}'}A^*_{\mathbf{k}'}(\mathbf{g}')e^{-i\mathbf{q}\cdot\mathbf{g}'}e^{i\mathbf{v}\cdot\mathbf{g}'/2}. \tag{8.11}$$

We now find that the integration over $d^{3N}v$ and subsequently over $d^{3N}K$ can be carried out. The result is found to be with $\mathbf{p}=\hbar\mathbf{k}$

$$\rho(p,q,t)=\left(\frac{\Omega}{8\pi^3h^3}\right)^N\frac{1}{N!}\int d^{3N}u\sum_{\mathbf{gg}'}e^{iq(u+g-g')}$$
$$\times\prod_j A^*_{\mathbf{k}_j-(1/2)\mathbf{u}_j-(1/2)(\mathbf{g}_j+\mathbf{g}'_j)}(\mathbf{g}')A_{\mathbf{k}_j+(1/2)\mathbf{u}_j-(1/2)(\mathbf{g}_j+\mathbf{g}'_j)}(\mathbf{g})$$
$$\times\mathrm{Tr}\{\rho(t):\prod_j c^\dagger_{\mathbf{k}_j-(1/2)\mathbf{u}_j-(1/2)(\mathbf{g}_j+\mathbf{g}'_j)}c_{\mathbf{k}_j+(1/2)\mathbf{u}_j-(1/2)(\mathbf{g}_j+\mathbf{g}'_j)}:\}. \tag{8.12}$$

We make the further change of variables $\mathbf{u}+\mathbf{g}-\mathbf{g}'\rightarrow\mathbf{u}$. The subscripts on A and c now become $\mathbf{k}-\mathbf{g}+\frac{1}{2}\mathbf{u}$ and those on A^* and $c^\dagger$ become $\mathbf{k}-\mathbf{g}'-\frac{1}{2}\mathbf{u}$. Since $\mathbf{k}$ can be shifted by a reciprocal lattice vector in the extended zone scheme we can replace $\mathbf{k}-\mathbf{g}+\frac{1}{2}\mathbf{u}\rightarrow\mathbf{k}+\frac{1}{2}\mathbf{u}$ and $\mathbf{k}-\mathbf{g}'-\frac{1}{2}\mathbf{u}\rightarrow\mathbf{k}-\frac{1}{2}\mathbf{u}$. The integrand now becomes

$$e^{iqu}\sum_{\mathbf{gg}'}\prod_j A^*_{\mathbf{k}_j-(1/2)\mathbf{u}_j}(\mathbf{g}'_j)A_{\mathbf{k}_j+(1/2)\mathbf{u}_j}(\mathbf{g}_j)$$
$$=e^{iqu}\prod_j\phi^*_{\mathbf{k}_j-(1/2)\mathbf{u}_j}(0)\phi_{\mathbf{k}_j+(1/2)\mathbf{u}_j}(0), \tag{8.13}$$

where we used (8.11). Substituting into (8.12) we obtain the second quantization form sought for

$$\rho(p,q,t)$$
$$=\left(\frac{\Omega}{8\pi^3h^3}\right)^N\frac{1}{N!}\int d^{3N}u\, e^{iqu}\prod_j\phi^*_{\mathbf{k}_j-(1/2)\mathbf{u}_j}(0)\phi_{\mathbf{k}_j+(1/2)\mathbf{u}_j}(0)$$
$$\times\mathrm{Tr}\{\rho(t):\prod_j c^\dagger_{\mathbf{k}_j-(1/2)\mathbf{u}_j}c_{\mathbf{k}_j+(1/2)\mathbf{u}_j}:\}. \tag{8.14}$$

For free particles we have $\phi^*(0)=\phi(0)=1$, so we obtain the result given by Balescu.[24]

In the present section we need the one-particle Wigner function, in phase-space (μ space), denoted as $\rho_1(\mathbf{p},\mathbf{q},t)$. We have in the case of Bloch electrons

$$\rho_1(\mathbf{p},\mathbf{q},\mathrm{t})=\frac{\Omega}{8\pi^3h^3}\int d^3u e^{i\mathbf{q}\cdot\mathbf{u}}\phi^*_{\mathbf{k}-(1/2)\mathbf{u}}(0)\phi_{\mathbf{k}+(1/2)\mathbf{u}}(0)$$
$$\times\langle c^\dagger_{\mathbf{k}-(1/2)\mathbf{u}}c_{\mathbf{k}+(1/2)\mathbf{u}}\rangle_t \tag{8.15}$$

and in the case of plane waves

$$\rho_1(\mathbf{p},\mathbf{q},t)=\frac{\Omega}{8\pi^3h^3}\int d^3u\, e^{i\mathbf{q}\cdot\mathbf{u}}\langle c^\dagger_{\mathbf{k}-(1/2)\mathbf{u}}c_{\mathbf{k}+(1/2)\mathbf{u}}\rangle_t. \tag{8.16}$$

8.1. *Wigner function transport equation for free particles*

We start from the full quantum mechanical Boltzmann equation (6.17) with $\zeta_1=\mathbf{k}-\frac{1}{2}\mathbf{u}$, $\zeta_2=\mathbf{k}+\frac{1}{2}\mathbf{u}$. We multiply this equation by $(\Omega/8\pi^3h^3)e^{i\mathbf{q}\cdot\mathbf{u}}$ and integrate over $\mathbf{u}$; we thus obtain

$$\frac{\partial\rho_1(\mathbf{p},\mathbf{q},t)}{\partial t}+\frac{\Omega}{8\pi^3h^3}\int d^3u\, e^{i\mathbf{q}\cdot\mathbf{u}}\{-\mathbf{F}(t)\cdot\frac{1-e^{-\beta(\epsilon_{\mathbf{k}-(1/2)\mathbf{u}}-\epsilon_{\mathbf{k}+(1/2)\mathbf{u}})}}{\epsilon_{\mathbf{k}-(1/2)\mathbf{u}}-\epsilon_{\mathbf{k}+(1/2)\mathbf{u}}}$$
$$\times\langle n_{\mathbf{k}+(1/2)\mathbf{u}}\rangle_{\mathrm{eq}}(1-\langle n_{\mathbf{k}-(1/2)\mathbf{u}}\rangle_{\mathrm{eq}})(\mathbf{k}+\tfrac{1}{2}\mathbf{u}|\mathbf{v}|\mathbf{k}-\tfrac{1}{2}\mathbf{u})+(i/\hbar)(\epsilon_{\mathbf{k}+(1/2)\mathbf{u}}-\epsilon_{\mathbf{k}-(1/2)\mathbf{u}})\langle c^\dagger_{\mathbf{k}-(1/2)\mathbf{u}}c_{\mathbf{k}+(1/2)\mathbf{u}}\rangle_t\}$$
$$=\frac{\Omega}{8\pi^3h^3}\int d^3u\, e^{i\mathbf{q}\cdot\mathbf{u}}\sum_{\mathbf{k}'}\{w_{\mathbf{k}'\mathbf{k}}\langle n_{\mathbf{k}'}\rangle_t(1-\langle n_{\mathbf{k}}\rangle_t)-w_{\mathbf{k}\mathbf{k}'}\langle n_{\mathbf{k}}\rangle_t(1-\langle n_{\mathbf{k}'}\rangle_t)\}\delta_{\mathbf{u},0}, \tag{8.17}$$

where we noticed that $(\mathbf{k}|\mathbf{r}-\mathbf{r}^{\mathrm{eq}}|\mathbf{k})\equiv 0$ for plane wave states. Further for plane waves,

$$\epsilon_{\mathbf{k}-(1/2)\mathbf{u}}-\epsilon_{\mathbf{k}+(1/2)\mathbf{u}}=(1/2m)[(\mathbf{p}-\tfrac{1}{2}\hbar\mathbf{u})^2-(\mathbf{p}+\tfrac{1}{2}\hbar\mathbf{u})^2]$$
$$=-(\hbar/m)\mathbf{p}\cdot\mathbf{u} \tag{8.18}$$

and

$$(\mathbf{k}+\tfrac{1}{2}\mathbf{u}|\mathbf{v}|\mathbf{k}-\tfrac{1}{2}\mathbf{u})=(\hbar\mathbf{k}/m)\delta_{\mathbf{u},0}. \tag{8.19}$$

We also note the Fourier inversion of (8.16)

$$\langle c^\dagger_{\mathbf{k}-(1/2)\mathbf{u}}c_{\mathbf{k}+(1/2)\mathbf{u}}\rangle_t=\frac{h^3}{\Omega}\int d^3q\, e^{-i\mathbf{q}\cdot\mathbf{u}}\rho_1(\mathbf{p},\mathbf{q},t), \tag{8.20}$$

and *a fortiori*

$$\langle c^\dagger_{\mathbf{k}-(1/2)\mathbf{u}}c_{\mathbf{k}+(1/2)\mathbf{u}}\rangle_{\mathrm{eq}}=\frac{h^3}{\Omega}\int d^3q\, e^{-i\mathbf{q}\cdot\mathbf{u}}\rho_{1\mathrm{eq}}(\mathbf{p},\mathbf{q}). \tag{8.21}$$

We will now compute the various terms of (8.17).

For the third term on the lhs we substitute (8.18) and (8.21), to yield

$$\text{3d term lhs}=\frac{i}{8\pi^3}\int d^3u\, e^{-i\mathbf{q}\cdot\mathbf{u}}\frac{\mathbf{p}\cdot\mathbf{u}}{m}\int d^3\bar{q}\, e^{-i\bar{\mathbf{q}}\cdot\mathbf{u}}\rho_1(\mathbf{p},\bar{\mathbf{q}},t). \tag{8.22}$$

Changing the order of integration we first evaluate

$$\int d^3u\, e^{i(\mathbf{q}-\bar{\mathbf{q}})\cdot\mathbf{u}}\mathbf{p}\cdot\mathbf{u}.$$

Now since

$$\int d^3u e^{i(\mathbf{q}-\bar{\mathbf{q}})\cdot\mathbf{u}}=8\pi^3\delta(\mathbf{q}-\bar{\mathbf{q}}), \tag{8.23}$$

differentiation to $\bar{\mathbf{q}}$ gives (i.e., operate with $\nabla_{\mathbf{q}}$ on both sides):

$$-i\int d^3u\, \mathbf{u} e^{i(\mathbf{q}-\bar{\mathbf{q}})\cdot\mathbf{u}} = 8\pi^3\nabla_{\mathbf{q}}\delta(\mathbf{q}-\bar{\mathbf{q}}); \quad (8.24)$$

hence

$$\int d^3u\, e^{i(\mathbf{q}-\bar{\mathbf{q}})\cdot\mathbf{u}}\mathbf{p}\cdot\mathbf{u} = 8\pi^3 i\mathbf{p}.\nabla_{\bar{q}}\delta(\mathbf{q}-\bar{\mathbf{q}}). \quad (8.25)$$

Carrying out the remaining integration over $d^3\bar{q}$, noticing

$$\int d^3\bar{q}\, [\nabla_{\bar{q}}\delta(\mathbf{q}-\bar{\mathbf{q}})]\rho_1(\mathbf{p},\bar{\mathbf{q}},t) = -\nabla_{\mathbf{q}}\rho_1(\mathbf{p},\mathbf{q},t), \quad (8.26)$$

we obtain

$$\text{3rd term lhs} = (\mathbf{p}/m)\cdot\nabla_{\mathbf{q}}\rho_1(\mathbf{p},\mathbf{q},t), \quad (8.27)$$

which is the standard inhomogeneous streaming term of the Boltzmann equation. It is quite peculiar that this term comes from the nondiagonal part of the full quantum mechanical transport equation!

For the second term on the lhs of (8.17) we obtain, noticing (8.19)

2nd term lhs

$$= \frac{\Omega}{8\pi^3h^3}\int d^3u\, e^{i\mathbf{q}\cdot\mathbf{u}}[-\beta\mathbf{F}(t)]\cdot\langle n_{\mathbf{k}}\rangle_{\text{eq}}$$
$$\times(1-\langle n_{\mathbf{k}}\rangle_{\text{eq}})\frac{\hbar\mathbf{k}}{m}\delta_{\mathbf{u},0} = \frac{\Omega}{8\pi^3h^3}\int d^3u\, e^{i\mathbf{q}\cdot\mathbf{u}}$$
$$\times\frac{\mathbf{F}(t)}{\hbar}\frac{\partial\langle n_{\mathbf{k}}\rangle_{\text{eq}}}{\partial\epsilon_{\mathbf{k}}}.\nabla_{\mathbf{k}}\epsilon_{\mathbf{k}}\delta_{\mathbf{u},0} = \frac{\Omega}{8\pi^3h^3}$$
$$\times\int d^3u\, e^{i\mathbf{q}\cdot\mathbf{u}}\frac{\mathbf{F}(t)}{\hbar}.\nabla_{\mathbf{k}}\langle c^\dagger_{\mathbf{k}-(1/2)\mathbf{u}}c_{\mathbf{k}+(1/2)\mathbf{u}}\rangle_{\text{eq}}\delta_{\mathbf{u},0}$$
$$= \frac{\mathbf{F}(t)}{8\pi^3\hbar}.\nabla_{\mathbf{k}}\int d^3\bar{q}\,\rho_{1\text{eq}}(\mathbf{p},\bar{\mathbf{q}})\int d^3u\, e^{i\mathbf{u}\cdot(\mathbf{q}-\bar{\mathbf{q}})}\delta_{\mathbf{u},0}. \quad (8.28)$$

We must now elaborate on the meaning of the Kronecker delta $\delta_{\mathbf{u},0}$. In LRT II, Sec. IIA, we indicated that diagonal parts of many-body operators are never sharp, but are "fuzzy." We must therefore give a certain extension $|\Delta\mathbf{u}|^3$ to the volume integration in $\mathbf{u}$ space. We may do this by considering a wave packet rather than a plane wave, which reflects the fact that $\mathbf{k}$ (and so $\mathbf{u}$) is not a sharp quantum number when the system is subject to chemical or other gradients. Thus we write

$$\int d^3u e^{i\mathbf{u}\cdot(\mathbf{q}-\bar{\mathbf{q}})}\delta_{\mathbf{u},0} \approx \int_{|\Delta u|^3} d^3u\, e^{i\mathbf{u}\cdot(\mathbf{q}-\bar{\mathbf{q}})}$$
$$= \prod_{xyz}\frac{\sin[\Delta u_x(\bar{q}_x-q_x)/2]}{(\bar{q}_x-q_x)/2}, \quad (8.29)$$

where we integrated over $(-\Delta u_x/2, \Delta u_x/2)$ and similarly for the other directions. The rhs has its maximum of Δu_x for $\bar{q}_x = q_x$; the x-direction width is

$$\frac{1}{\Delta u_x}\int_{-\infty}^{\infty}\frac{\sin[\Delta u_x(\bar{q}_x-q_x)/2]}{(\bar{q}_x-q_x)/2}d\bar{q}_x = \frac{2\pi}{\Delta u_x}. \quad (8.30)$$

We may thus replace the rhs of (8.29) by a function in $\bar{\mathbf{q}}$ space which has a magnitude $|\Delta u|^3 \equiv \Pi_{xyz}\Delta u_x$ for $\bar{\mathbf{q}}$ within the rectangular box of volume $8\pi^3/|\Delta u|^3$ centered on $\mathbf{q}$, and which is zero elsewhere. Thus, carrying out the $\bar{\mathbf{q}}$ integration next, we have

$$\int d^3\bar{q}\rho_{1\text{eq}}(\mathbf{p},\bar{\mathbf{q}})\int d^3u\, e^{i\mathbf{u}\cdot(\mathbf{q}-\bar{\mathbf{q}})}\delta_{\mathbf{u},0}$$
$$\approx |\Delta u|^3\int_{8\pi^3/|\Delta u|^3} d^3\bar{q}\rho_{1\text{eq}}(\mathbf{p},\bar{\mathbf{q}}). \quad (8.31)$$

According to the uncertainty principle, the volume of a microcell in phase space is $|\Delta p|^3|\Delta q|^3 \equiv \hbar^3|\Delta u|^3\omega(\mathbf{q}) = h^3$; thus $\omega(\mathbf{q}) = 8\pi^3/|\Delta u|^3$ is the minimum accessible volume in position space centered on $\mathbf{q}$. From (8.18) and (8.31) we thus obtain

$$\text{2nd term lhs} = \frac{\mathbf{F}(t)}{\hbar}.\nabla_{\mathbf{k}}\frac{1}{\omega(\mathbf{q})}\int_{\omega(\mathbf{q})} d^3\bar{q}\rho_{1\text{eq}}(\mathbf{p},\bar{\mathbf{q}}). \quad (8.32)$$

We can treat the collision term in a similar way. The part linear in $\langle n_{\mathbf{k}}\rangle_t$ goes as before. For the quadratic part we need

$$\frac{\Omega}{8\pi^3h^3}\int d^3 u e^{i\mathbf{q}\cdot\mathbf{u}}\sum_{\mathbf{k}'} w_{\mathbf{k}'\mathbf{k}}\langle n_{\mathbf{k}'}\rangle_t\langle n_{\mathbf{k}}\rangle_t\delta_{\mathbf{u},0}. \quad (8.32')$$

For $\langle n_{\mathbf{k}'}\rangle_t$ and $\langle n_{\mathbf{k}}\rangle_t$ we write

$$\langle n_{\mathbf{k}'}\rangle_t = \frac{h^3}{\Omega}\int d^3q'\, e^{-i\mathbf{q}'\cdot\mathbf{u}}\rho_1(\mathbf{p}',\mathbf{q}',t) \quad (\text{for } \mathbf{u}\to 0).$$

$$\langle n_{\mathbf{k}}\rangle_t = \frac{h^3}{\Omega}\int d^3q''\, e^{-i\mathbf{q}''\cdot\mathbf{v}}\rho_1(\mathbf{p},\mathbf{q}'',t) \quad (\text{for } \mathbf{v}\to 0).$$

Hence

$$(8.32') = \frac{\Omega}{8\pi^3h^3}\left(\frac{h^3}{\Omega}\right)^2\int d^3q'\, e^{-i\mathbf{q}'\cdot\mathbf{u}}\rho_1(\mathbf{p}',\mathbf{q}',t)$$
$$\times\int d^3q''\, e^{-i\mathbf{q}''\cdot\mathbf{v}}\rho_1(\mathbf{p},\mathbf{q}'',t)$$
$$\times\int d^3u\, e^{i\mathbf{q}\cdot\mathbf{u}}\delta_{\mathbf{u},0}\delta_{\mathbf{v},0}. \quad (8.32'')$$

Now we multiply (8.32″) by $(\Omega/8\pi^3)\int d^3v\, e^{i\mathbf{v}\cdot\mathbf{q}}\delta_{v,0} \approx 1$. Thus we obtain

$$(8.32'') = \frac{1}{h^3}\left(\frac{1}{8\pi^3}\right)^2(h^3)^2\int d^3q'\rho_1(\mathbf{p},\mathbf{q}',t)\int e^{i\mathbf{u}\cdot(\mathbf{q}-\mathbf{q}')}d^3u\,\delta_{\mathbf{u},0}$$
$$\times\int d^3q''\rho_1(\mathbf{p},\mathbf{q}'',t)\int e^{i\mathbf{v}\cdot(\mathbf{q}-\mathbf{q}'')}d^3v\delta_{\mathbf{v},0}$$
$$\simeq\frac{1}{h^3}\left(\frac{1}{8\pi^3}\right)^2(h^3)^2|\Delta u|^3$$
$$\times\int_{8\pi^3/|\Delta u|^3} d^3q'\rho_1(\mathbf{p},\mathbf{q}',t))|\Delta v|^3\int_{8\pi^3/|\Delta v|^3} d^3q''\rho_1(\mathbf{p},\mathbf{q}'',t)$$
$$= \frac{1}{h^3}\frac{h^3}{\omega(\mathbf{q})}\int_{\omega(\mathbf{q})} d^3q'\rho_1(\mathbf{p}',\mathbf{q}',t)$$
$$\times\frac{h^3}{\omega(\mathbf{q})}\int_{\omega(\mathbf{q})} d^3q''\rho_1(\mathbf{p},\mathbf{q}'',t). \quad (8.32''')$$

Thus one finds

$$\text{coll term} = \frac{1}{h^3}\sum_{\mathbf{k}'}\Big\{w_{\mathbf{k}'\mathbf{k}}\frac{h^3}{\omega(\mathbf{q})}\int_{\omega(\mathbf{q})} d^3\bar{\mathbf{q}}\rho_1(\hbar\mathbf{k}',\bar{\mathbf{q}},t)$$
$$\times\Big[1-\frac{h^3}{\omega(\mathbf{q})}\int_{\omega(\mathbf{q})} d^3\bar{\mathbf{q}}\rho_1(\hbar\mathbf{k},\bar{\mathbf{q}},t)\Big]$$
$$- w_{\mathbf{k}\mathbf{k}'}\frac{h^3}{\omega(\mathbf{q})}\int_{\omega(\mathbf{q})} d^3\bar{\mathbf{q}}\rho_1(\hbar\mathbf{k}',\bar{\mathbf{q}},t)$$
$$\times\Big[1-\frac{h^3}{\omega(\mathbf{q})}\int_{\omega(\mathbf{q})} d^3\bar{\mathbf{q}}\rho_1(\hbar\mathbf{k}',\bar{\mathbf{q}},t)\Big]\Big\}. \quad (8.33)$$

Both (8.32) and (8.33) can be written in simpler form by introducing *coarse-grained Wigner functions* (setting $\rho_1 \equiv \rho$):

$$\tilde{\rho}(\mathbf{p},\mathbf{q},t) = \frac{1}{\omega(\mathbf{q})}\int_{\omega(\mathbf{q})} d^3\bar{q}\, \rho(\mathbf{p},\bar{\mathbf{q}},t), \quad \mathbf{q}\in\omega(\mathbf{q}), \tag{8.34}$$

$$\tilde{\rho}_{\rm eq}(\mathbf{p},\mathbf{q}) = \frac{1}{\omega(\mathbf{q})}\int_{\omega(\mathbf{q})} d^3\bar{q}\, \rho_{\rm eq}(\mathbf{p},\bar{\mathbf{q}}), \quad \mathbf{q}\in\omega(\mathbf{q}). \tag{8.35}$$

Since the integration involves a volume in phase-space (h^3) which is larger than the minimum support of the Wigner function,[23] we expect $\tilde{\rho}$ to be positive definite.[25]

Collecting terms, we find the transport equation

$$\frac{\partial\rho(\mathbf{p},\mathbf{q},t)}{\partial t} + \frac{\mathbf{p}}{m}\cdot\nabla_{\mathbf{q}}\rho(\mathbf{p},\mathbf{q},t) + \mathbf{F}(t)\cdot\nabla_{\mathbf{p}}\tilde{\rho}_{\rm eq}(\mathbf{p},\mathbf{q})$$
$$= \sum_{\mathbf{k}'}\{w_{\mathbf{k}'\mathbf{k}}\tilde{\rho}(\hbar\mathbf{k}',\mathbf{q},t)[1 - h^3\tilde{\rho}(\hbar\mathbf{k},\mathbf{q},t)] - w_{\mathbf{k}\mathbf{k}'}\tilde{\rho}(\hbar\mathbf{k},\mathbf{q},t)[1 - h^3\tilde{\rho}(\hbar\mathbf{k}',\mathbf{q},t)]\}. \tag{8.36}$$

This result is near exact.[26] If the gradient $\nabla_{\mathbf{q}}\rho$ is slowly varying over the cell volumes $\omega(\mathbf{q})$, we can also replace ρ and $\tilde{\rho}$ in the first two terms on the lhs of (8.36). We have then a transport equation for $\tilde{\rho}(\mathbf{p},\mathbf{q},t)$ alone.

The classical distribution $f(\mathbf{p},\mathbf{q},t)$ is related to the classical limit of the Wigner function, if this limit exists (cf. M. J. Groenewold[27]). We have

$$n(\mathbf{p},\mathbf{q},t) = \lim_{\hbar\to 0} h^3\tilde{\rho}(\mathbf{p},\mathbf{q},t), \tag{8.37}$$

where n has the dimension of a number. To obtain a density in phase-space, we must divide by the volume of a microcell h^3. Thus

$$f(\mathbf{p},\mathbf{q},t) = \lim_{\hbar\to 0}\tilde{\rho}(\mathbf{p},\mathbf{q},t). \tag{8.38}$$

From (8.36) we obtain

$$\frac{\partial f(\mathbf{p},\mathbf{q},t)}{\partial t} + \frac{\mathbf{p}}{m}\cdot\nabla_{\mathbf{q}} f(\mathbf{p},\mathbf{q},t) + \mathbf{F}(t)\cdot\nabla_{\mathbf{p}} f_{\rm eq}(\mathbf{p},\mathbf{q})$$
$$= \frac{\Omega}{8\pi^3}\int d^3k'\,[w_{\mathbf{k}'\mathbf{k}} f(\mathbf{p}',\mathbf{q},t) - w_{\mathbf{k}\mathbf{k}'} f(\mathbf{p},\mathbf{q},t)]. \tag{8.39}$$

The effects of the exclusion principle in the collision term have disappeared; the only quantum mechanical attribute remaining is $w_{\mathbf{k}'\mathbf{k}}$ given by the "golden rule." It is a small matter to rewrite the collision term in terms of the classical cross section.

We assume elastic scattering with N_0 heavy obstacles. Then $w_{\mathbf{k}\mathbf{k}'} = w_{\mathbf{k}'\mathbf{k}}$ and [cf. (2.25)]

$$w_{\mathbf{k}\mathbf{k}'} = (2\pi N_0\lambda^2/\hbar)|(\mathbf{k}|v|\mathbf{k}')|^2\delta(\epsilon_{\mathbf{k}} - \epsilon_{\mathbf{k}'}). \tag{8.40}$$

For the cross section we have by definition, denoting by Ω' the solid angle of the scattered vector $\mathbf{k}'$ taking the sample volume 1 cm^3,

$$\sigma(\Omega')d\Omega' = \frac{1}{v_{\mathbf{k}}}\int_{|\mathbf{k}'|\text{only}} d^3k'\frac{w_{\mathbf{k}\mathbf{k}'}}{8\pi^3N_0}. \tag{8.41}$$

Since,

$$d^3k' = \frac{d\epsilon_{\mathbf{k}'}dS_\epsilon}{|\nabla_{\mathbf{k}}\epsilon_{\mathbf{k}'}|} = \frac{d\epsilon_{\mathbf{k}'}k'^2d\Omega'}{\hbar v_{\mathbf{k}'}} \tag{8.42}$$

(where S is an energy surface) we easily obtain

$$\sigma(\Omega') = (m^2\lambda^2/\hbar^4 4\pi^2)|(\mathbf{k}|v|\mathbf{k}')|^2. \tag{8.43}$$

Now substituting (8.40), (8.42), and (8.43) in the collision term we obtain

$$\text{coll term} = N_0\int d\Omega'\frac{p}{m}\sigma(\Omega')[f(\mathbf{p}',\mathbf{q},t) - f(\mathbf{p},\mathbf{q},t)]. \tag{8.44}$$

For two-body collisions we can likewise recover the standard collision term.

8.2. *Wigner function transport equations for Bloch electrons*

The quantum mechanical Boltzmann equation is multiplied by

$$(\Omega/8\pi^3h^3)e^{i\mathbf{q}\cdot\mathbf{u}}\phi^*_{\zeta_1}(0)\phi_{\zeta_2}(0),$$
$$\zeta_1 = \mathbf{k} - \tfrac{1}{2}\mathbf{u},\ \zeta_2 = \mathbf{k} + \tfrac{1}{2}\mathbf{u} \tag{8.45}$$

and integrated over d^3u. The result is the same as (8.17), providing that all terms except $\partial\rho_1/\partial t$ are multiplied by $\phi^*_{\mathbf{k}-(1/2)\mathbf{u}}(0)\phi_{\mathbf{k}+(1/2)\mathbf{u}}(0)$. We now need the Fourier inversion of (8.15) which reads

$$\langle c^\dagger_{\mathbf{k}-(1/2)\mathbf{u}}c_{\mathbf{k}+(1/2)\mathbf{u}}\rangle_t\phi^*_{\mathbf{k}-(1/2)\mathbf{u}}(0)\phi_{\mathbf{k}+(1/2)\mathbf{u}}(0)$$
$$= \frac{h^3}{\Omega}\int d^3q\, e^{i\mathbf{q}\cdot\mathbf{u}}\rho(\mathbf{p},\mathbf{q},t). \tag{8.46}$$

In the terms with $\delta_{\mathbf{u},0}$ we use [28]

$$|\phi_{\mathbf{k}}(0)|^2 = 1. \tag{8.47}$$

The procedure is similar as in the previous subsection. The field term with $\mathbf{F}(t) = -e\mathbf{E}(t)$ survives only for $\mathbf{u} = 0$, since

$$(\mathbf{k} + \tfrac{1}{2}\mathbf{u}|\mathbf{v}|\mathbf{k} - \tfrac{1}{2}\mathbf{u}) = (1/\hbar)\nabla_{\mathbf{k}}\epsilon_{\mathbf{k}}\delta_{\mathbf{u},0}, \tag{8.48}$$

as we show in the Appendix. Noticing (8.47), we find that the term becomes the same as (8.32). The collision term also remains unchanged, i.e., we find (8.33). Some new aspects occur in the other streaming term: we have upon substituting (8.46),

$$\frac{\Omega}{8\pi^3h^3}\int d^3u\, e^{i\mathbf{q}\cdot\mathbf{u}}\frac{i}{\hbar}(\epsilon_{\mathbf{k}+(1/2)\mathbf{u}} - \epsilon_{\mathbf{k}-(1/2)\mathbf{u}})$$
$$\times\phi^*_{\mathbf{k}-(1/2)\mathbf{u}}(0)\phi_{\mathbf{k}+(1/2)\mathbf{u}}(0)\langle c^\dagger_{\mathbf{k}-(1/2)\mathbf{u}}c_{\mathbf{k}+(1/2)\mathbf{u}}\rangle_t.$$
$$= \frac{i}{8\pi^3\hbar}\int d^3u\, e^{i\mathbf{q}\cdot\mathbf{u}}(\epsilon_{\mathbf{k}+(1/2)\mathbf{u}} - \epsilon_{\mathbf{k}-(1/2)\mathbf{u}})$$
$$\times\int d^3\bar{\mathbf{q}}\, e^{-i\bar{\mathbf{q}}\cdot\mathbf{u}}\rho(\mathbf{p},\bar{\mathbf{q}},t). \tag{8.49}$$

We write

$$\epsilon_{\mathbf{k}+(1/2)\mathbf{u}} - \epsilon_{\mathbf{k}-(1/2)\mathbf{u}} = (e^{(1/2)\mathbf{u}\cdot\nabla_{\mathbf{k}}} - e^{-(1/2)\mathbf{u}\cdot\nabla_{\mathbf{k}}})\epsilon_{\mathbf{k}}. \tag{8.50}$$

For the integration over d^3u we now have

$$\int d^3u\,[e^{i\mathbf{u}\cdot(\mathbf{q}-\bar{\mathbf{q}}-(i/2)\nabla_{\mathbf{k}})} - e^{i\mathbf{u}\cdot(\mathbf{q}-\bar{\mathbf{q}}+(i/2)\nabla_{\mathbf{k}})}]\epsilon_{\mathbf{k}}$$
$$= 8\pi^3[\delta(\bar{\mathbf{q}} - \mathbf{q} + (i/2)\nabla_{\mathbf{k}}) - \delta(\bar{\mathbf{q}} - \mathbf{q}(i/2)\nabla_{\mathbf{k}})]\epsilon_{\mathbf{k}}$$
$$= 16\pi^3\sum_{n=0}^{\infty}\frac{1}{(2n+1)!}[(\nabla_{\mathbf{q}})^{2n+1}\delta(\bar{\mathbf{q}} - \mathbf{q})] * [((i/2)\nabla_{\mathbf{k}})^{2n+1}\epsilon_{\mathbf{k}}], \tag{8.51}$$

where * means a contraction over all tensor components to a scalar. Carrying out the subsequent $d^3\bar{q}$ integration we obtain

$$\text{streaming term} = -\frac{2i}{\hbar}\sum_{n=0}^{\infty}\frac{1}{(2n+1)!}(\nabla_{\mathbf{q}})^{2n+1}\rho(\mathbf{p},\mathbf{q},t)^* \times\left(\frac{i}{2}\nabla_{\mathbf{k}}\right)^{2n+1}\epsilon_{\mathbf{k}}. \quad (8.52)$$

We note that this term is real, despite the occurrence of i in its factors.

Collecting terms we obtain the following transport equation for the coarse-grained Wigner function $\tilde{\rho}$:

$$\frac{\partial\tilde{\rho}(\hbar\mathbf{k},\mathbf{q},t)}{\partial t} - \frac{e\mathbf{E}}{\hbar}\cdot\nabla_{\mathbf{k}}\tilde{\rho}_{\text{eq}}(\hbar\mathbf{k},\mathbf{q}) - \frac{2i}{\hbar} \times\sum_{n=0}^{\infty}\frac{1}{(2n+1)!}(\nabla_{\mathbf{q}})^{2n+1}\tilde{\rho}(\hbar\mathbf{k},\mathbf{q},t)^*\left(\frac{i}{2}\nabla_{\mathbf{k}}\right)^{2n+1}\epsilon_{\mathbf{k}}$$
$$= \sum_{k'}\{w_{\mathbf{k}\mathbf{k}}\tilde{\rho}(\hbar\mathbf{k},\mathbf{q},t)[1 - h^3\rho(\hbar\mathbf{k}',\mathbf{q},t)] - w_{\mathbf{k}\mathbf{k}'}\tilde{\rho}(\hbar\mathbf{k},\mathbf{q},t)[1 - h^3\tilde{\rho}(\hbar\mathbf{k}',\mathbf{q},t)]\}. \quad (8.53)$$

To obtain the classical limit, we must now state more precisely what is meant by this. If it means a $\mathbf{p},\mathbf{q}$ description, with no reference to the quantum mechanical energies $\epsilon_{\mathbf{k}}$ of the Bloch states, then we must write $\mathbf{p} = \hbar\mathbf{k}$ everywhere and

$$((i/2)\nabla_{\mathbf{k}})^{2n+1}\epsilon_{\mathbf{k}} \rightarrow \left(\frac{\hbar i}{2}\nabla_{\mathbf{p}}\right)^{2n+1}\epsilon_{\mathbf{p}}. \quad (8.54)$$

For the limit of the streaming term (8.48) we then have

$$\lim_{\hbar\to 0} -\frac{2i}{\hbar}\sum_{n=0}^{\infty}\frac{1}{(2n+1)!}(\nabla_{\mathbf{q}})^{2n+1}\tilde{\rho}(\mathbf{p},\mathbf{q},t)^*((\hbar i/2)\nabla_{\mathbf{p}})^{2n+1}\epsilon_{\mathbf{p}} = \nabla_{\mathbf{q}}\tilde{\rho}(\mathbf{p},\mathbf{q},t)\cdot\nabla_{\mathbf{p}}\epsilon_{\mathbf{p}}, \quad (8.55)$$

the higher-order terms giving zero. Likewise

$$\lim_{\hbar\to 0} -(e\mathbf{E}/\hbar)\nabla_{\mathbf{k}}\tilde{\rho}_{\text{eq}}(\mathbf{p},\mathbf{q},) = -e\mathbf{E}\nabla_{\mathbf{p}}\tilde{\rho}_{\text{eq}}(\mathbf{p},\mathbf{q}). \quad (8.56)$$

The collision term is treated the same as in the previous subsection; we thus recover the standard classical Boltzmann equation as given in (8.39).

However, it is customary in solid state physics to use a *semiclassical* $\mathbf{k},\mathbf{q}$ description with the Hamiltonian given by Wannier's theorem[29]:

$$h_{\text{Wannier}}(\mathbf{k},\mathbf{q}) = \epsilon(\mathbf{k}) + \mathscr{V}(\mathbf{q}) = \epsilon(-i\nabla_{\mathbf{q}}) + \mathscr{V}(\mathbf{q}). \quad (8.57)$$

The classical limit is now taken as

$$F(\mathbf{k},\mathbf{q},t) = \lim_{\hbar\to 0} h^3\tilde{\rho}_1(\hbar\mathbf{k},\mathbf{q},t). \quad (8.58)$$

Here F is the number of electrons "occupying $\mathbf{k}$ at time t in the neighborhood of $\mathbf{q}$" (formulation of Ziman, *op. cit.* Sec. 7.3); more specifically, F is the number of electrons occupying $\mathbf{k}$ within $|\Delta k|^3$ in the coarse graining cell $\omega(\mathbf{q})$ centered on $\mathbf{q}$ at time t. The normalization is

$$2\sum_{\text{cells}|\Delta k|^3} F(\mathbf{k},\mathbf{q},t) = 2\int\frac{d^3k}{|\Delta k|^3}F(\mathbf{k},\mathbf{q},t) = \frac{\omega(\mathbf{q})}{4\pi^3}\int d^3kF(\mathbf{k},\mathbf{q},t) = N(\mathbf{q},t). \quad (8.59\text{a})$$

Note that the density of states in $\mathbf{k}$ space, excluding spin, is now $z(\mathbf{k}) = \omega(\mathbf{q})/8\pi^3$ [where all the volumes $\omega(\mathbf{q})$ might be chosen to be of equal size ω]; $N(q,t)$ is the number of electrons in $\omega(\mathbf{q})$ at time t. The further normalization is

$$\sum_{\text{cells}\,\omega(\mathbf{q})} N(\mathbf{q},t) = \int\frac{d^3q}{\omega(\mathbf{q})}N(\mathbf{q},t) = \frac{1}{4\pi^3}\int\int d^3k\,d^3q\,F(\mathbf{k},\mathbf{q},t) = N(t). \quad (8.59\text{b})$$

Multiplying both sides of (8.53) by h^3 we obtain with (8.58),

$$\frac{\partial F(\mathbf{k},\mathbf{q},t)}{\partial t} - \frac{e\mathbf{E}}{\hbar}\cdot\nabla_{\mathbf{k}}F_{\text{eq}}(\mathbf{k},\mathbf{q}) - \frac{2i}{\hbar} \times\sum_{n=0}^{\infty}\frac{1}{(2n+1)!}(\nabla_{\mathbf{q}})^{2n+1}F(\mathbf{k},\mathbf{q},t)^*\left(\frac{i}{2}\nabla_{\mathbf{k}}\right)^{2n+1}\epsilon_{\mathbf{k}}$$
$$= \int d^3k\,z(\mathbf{k})\{w_{\mathbf{k}'\mathbf{k}}F(\mathbf{k}',\mathbf{q},t)[1 - F(\mathbf{k},\mathbf{q},t)] - w_{\mathbf{k}\mathbf{k}'}F(\mathbf{k},\mathbf{q},t)[1 - F(\mathbf{k}',\mathbf{q},t)]\}. \quad (8.60)$$

In the collision term the effects due to the exclusion principle are now retained. Equation (8.60) differs from the usual result in the occurrence of higher-order spatial derivatives of F. Only in the effective mass approximation $\epsilon_{\mathbf{k}} = \frac{1}{2}\hbar^2\mathbf{k}\mathbf{k}:M^{-1}$ these higher-order derivatives drop out.

ACKNOWLEDGMENTS

We are indebted to Dr. C. van Weert of the University of Amsterdam for drawing our attention to Eq. (8.16) for the one particle Wigner function, and we would like to thank Dr. P. H. Handel of the University of Missouri for discussions on the existence of a relaxation time. We acknowledge the support of the National Sciences and Engineering Research Council of Canada under grant A9522.

APPENDIX: MATRIX ELEMENTS OF BLOCH FUNCTIONS

The matrix element $(\mathbf{k}|\mathbf{v}|\mathbf{k}')$ for Bloch states is computed similarly as the diagonal matrix element by Reitz.[30] We start from the Schrödinger equation with $(\mathbf{r}|\mathbf{k})\equiv\psi_{\mathbf{k}}(\mathbf{r})$:

$$\nabla_{\mathbf{r}}^2\psi_{\mathbf{k}}(\mathbf{r}) = (2m/\hbar^2)[\mathscr{V}(\mathbf{r}) - \epsilon_{\mathbf{k}}]\psi_{\mathbf{k}}(\mathbf{r}). \quad (\text{A1})$$

Taking the $\mathbf{k}$-gradient of both sides we find

$$\nabla_{\mathbf{r}}^2(\nabla_{\mathbf{k}}\psi_{\mathbf{k}}(\mathbf{r})) + (2m/\hbar^2)[(\nabla_{\mathbf{k}}\epsilon_{\mathbf{k}})\psi_{\mathbf{k}}(\mathbf{r}) + (\epsilon_{\mathbf{k}} - \mathscr{V}(\mathbf{r}))\nabla_{\mathbf{k}}\psi_{\mathbf{k}}(\mathbf{r})] = 0. \quad (\text{A2})$$

Now, $\psi_{\mathbf{k}}(\mathbf{r}) = e^{i\mathbf{k}\cdot\mathbf{r}}\phi_{\mathbf{k}}(\mathbf{r})$ so that

$$\nabla_{\mathbf{k}}\psi_{\mathbf{k}}(\mathbf{r}) = i\mathbf{r}\psi_{\mathbf{k}}(\mathbf{r}) + e^{i\mathbf{k}\cdot\mathbf{r}}\nabla_{\mathbf{k}}\phi_{\mathbf{k}}(\mathbf{r}), \quad (\text{A3})$$

$$\nabla_{\mathbf{r}}^2(\nabla_{\mathbf{k}}\psi_{\mathbf{k}}(\mathbf{r})) = 2i\nabla_{\mathbf{r}}\psi_{\mathbf{k}}(\mathbf{r}) - (i2m/\hbar^2)(\epsilon_{\mathbf{k}} - \mathscr{V}(\mathbf{r}))\mathbf{r}\psi_{\mathbf{k}}(\mathbf{r}) + \nabla_{\mathbf{r}}^2(e^{i\mathbf{k}\cdot\mathbf{r}}\nabla_{\mathbf{k}}\phi_{\mathbf{k}}(\mathbf{r})), \quad (\text{A4})$$

where we used (A1) for the second term. Substituting (A4) into (A2) and using (A3) we obtain

$$2i\nabla_{\mathbf{r}}\psi_{\mathbf{k}}(\mathbf{r}) + (2m/\hbar^2)(\nabla_{\mathbf{k}}\epsilon_{\mathbf{k}})\psi_{\mathbf{k}}(\mathbf{r}) + [\nabla_{\mathbf{r}}^2 + (2m/\hbar^2)(\epsilon_{\mathbf{k}} - \mathscr{V}(\mathbf{r}))]e^{i\mathbf{k}\cdot\mathbf{r}}\nabla_{\mathbf{k}}\phi_{\mathbf{k}}(\mathbf{r}) = 0. \quad (\text{A5})$$

Multiplying by $-\frac{1}{2}\psi_{\mathbf{k}'}^*(\mathbf{r})$ and integrating over all space we get

$$(\mathbf{k}'|-i\nabla_{\mathbf{r}}|\mathbf{k}) = -i\int\psi_{\mathbf{k}'}^*(\mathbf{r})\nabla_{\mathbf{r}}\psi_{\mathbf{k}}(\mathbf{r})\,d^3r = (m/\hbar^2)\nabla_{\mathbf{k}}\epsilon_{\mathbf{k}}\int\psi_{\mathbf{k}'}^*(\mathbf{r})\psi_{\mathbf{k}}(\mathbf{r})\,d^3r$$
$$+ \frac{1}{2}\int\psi_{\mathbf{k}'}^*(\mathbf{r})\nabla_{\mathbf{r}}^2(e^{i\mathbf{k}\cdot\mathbf{r}}\nabla_{\mathbf{k}}\phi_{\mathbf{k}}(\mathbf{r}))\,d^3r$$
$$+ (m/\hbar^2)\int\psi_{\mathbf{k}'}^*(\mathbf{r})(\epsilon_{\mathbf{k}} - \mathscr{V}(\mathbf{r}))e^{i\mathbf{k}\cdot\mathbf{r}}\,\nabla_{\mathbf{k}}\phi_{\mathbf{k}}(\mathbf{r})\,d^3r. \quad (\text{A6})$$

Since ψ is normalized the first term gives $(m/\hbar^2)\nabla_{\mathbf{k}}\epsilon_{\mathbf{k}}\delta_{\mathbf{kk}'}$. For the second term we use Green's theorem; the bilinear concomitant vanishes since the integrand is periodic. We thus have for this term

$$\tfrac{1}{2}\int e^{i\mathbf{k}\cdot\mathbf{r}}\nabla_{\mathbf{k}}\phi_{\mathbf{k}}(\mathbf{r})\nabla_{\mathbf{r}}^2\psi_{\mathbf{k}'}^*(\mathbf{r})\,d^3r$$

$$= \tfrac{1}{2}\int e^{i\mathbf{k}\cdot\mathbf{r}}\nabla_{\mathbf{k}}\phi_{\mathbf{k}}(\mathbf{r})\frac{2m}{\hbar^2}[\mathscr{V}(\mathbf{r}) - \epsilon_{\mathbf{k}}]\psi_{\mathbf{k}'}^*(\mathbf{r})\,d^3r, \qquad \text{(A7)}$$

where we used the Schrödinger equation (A1). It thus cancels the third term of (A6). The result therefore is

$$(\mathbf{k}'|\mathbf{v}|\mathbf{k}) = (\hbar/m)(\mathbf{k}'| - i\nabla_{\mathbf{r}}|\mathbf{k}) = (1/\hbar)\nabla_{\mathbf{k}}\epsilon_{\mathbf{k}}\delta_{\mathbf{kk}'} \qquad \text{(A8)}$$

also

$$(\mathbf{k} + \tfrac{1}{2}\mathbf{u}|\mathbf{v}|\mathbf{k} - \tfrac{1}{2}\mathbf{u}) = (1/\hbar)\nabla_{\mathbf{k}}\epsilon_{\mathbf{k}}\delta_{\mathbf{u},0} \qquad \text{(A9)}$$

which is the result of (8.48).

[1]K. M. van Vliet, "LRT I" J. Math. Phys. **19**, 1345–1370 (1978).

[2]Unbracketed, nonsuperscripted operators are Schrödinger operators; if special emphasis is needed a superscript S is used.

[3]J. O. Vigfussen, "Time relaxation of the solutions of master equations for large systems," preprint.

[4]K. M. van Vliet, "LRT II" J. Math. Phys. **20**, 2573–2595 (1979).

[5]W. Kohn and J. M. Luttinger, Phys. Rev. **84**, 814 (1951).

[6]E. N. Adams and T. D. Holstein, J. Phys. Chem. Solids **10**, 254 (1959).

[7]A. H. Kahn and H. P. R. Frederikse, *Solid State Physics*, Vol. 9, edited by F. Seitz and D. Turnbull (Academic, New York, 1959), p. 257.

[8]P. N. Argyres, Phys. Rev. **109**, 1115 (1958).

[9]P. N. Argyres and L. M. Roth, J. Phys. Chem. Solids **12**, 89 (1959). In Sec. 3 of this paper the authors use also the many-body von Neumann equation in order to treat electron–phonon scattering.

[10]R. Kubo, S. J. Miyake, and N. Nashitsume, *Solid State Physics*, Vol. 17, edited by F. Seitz and D. Turnbull (Academic, New York, 1964), p. 269.

[11]This extra current component has also been found in less explicit form by P. N. Argyres in *Lectures in Theoretical Physics*, Vol. 7, Boulder, Colorado, edited by W. E. Britten, B. D. Downs, and J. Downs (Interscience, New York, 1966), p. 183.

[12]Throughout this paper we use ζ as denoting an arbitrary state (like in $\{n_\zeta\}$), while ζ' and ζ'' (or ζ_1 etc.) refer to specific states.

[13]The LRT II the Boltzmann operator was denoted by $\tilde{M}$.

[14]The Green's operator $(i\omega + \mathscr{B}')^{-1}$ is a Green's operator in the extended sense; the eigenvalue zero is to be omitted in the spectral decomposition.

[15]E. Verboven, Physica **26**, 1091–1116 (1960).

[15a]B. R. Nag, *Theory of Electrical Transport in Semiconductors* (Pergamon, New York, 1972), Sec. 4.1.

[16]M. Charbonneau and K. M van Vliet, Phys. Status Solidi (b) **101**, 509 (1980).

[17]In (2.70)$\mathscr{B}^{(k)}$ is the Boltzmann operator of order k. However, we could also have introduced an operator $\mathscr{B}^k n$ by (2.70); this operator is then only *defined* by the relation to $\langle A^k \mathbf{n}_\zeta\rangle_b$. The operator is not produced by repeated operation $\mathscr{B}\cdots\mathscr{B}n_\zeta$, since the nonlinear $\mathscr{B}$ can only operate on an occupation number so that $\mathscr{B}\cdots\mathscr{B}n_\zeta$ does not exist. The advantage of writing $\mathscr{B}^k n$ rather than $\mathscr{B}^{(k)}n_\zeta$ is that we can now sum the series to an exponential $\exp(-t\mathscr{B})n_\zeta$, even for nonlinear $\mathscr{B}$. Of course, such an operator expression has only formal validity.

[18]L. M. Roth and P. N. Argyres, in *Semiconductors and Semimetals*, Vol. 3, edited by R. K. Willardson and A. C. Beer (Academic, New York, 1967), p. 421.

[19]The original Kubo theory (i.e., prior to the van Hove limit) contains in essence both diagonal and nondiagonal parts (see LRT I). Therefore, in contrast to the approaches based on the density operator like Refs. 6 and 7 Kubo's theory yields correctly the Hall effect (see Ref. 10).

[20]P. Vasilopoulos, K. M. van Vliet, and M. Charbonneau, "Linear response theory revisited IV: Applications." To be published.

[21]P. N. Argyres, J. Phys. Chem. Solids **4**, 19 (1957).

[22]R. C. Enck, A. S. Saleh, and H. Y. Fan, Phys. Rev. **182**, 790 (1969).

[23]S. R. de Groot, *La transformation de Weyl et la fonction de Wigner: une forme alternative de la mécanique quantique* (Les Presses de l'Université de Montréal, Montréal, 1974).

[24]R. Balescu, *Statistical Mechanics of Charged Particles* (Interscience, New York, 1963), Chap. 14.

[25]No stringent proof is available, however, for the nonequilibrium $\bar{\rho}$ to our knowledge.

[26]The only pertinent approximation in this result stems from the linearization of the linear response result in the von Neumann equation, LRT II Eqs. (3.9)–(3.11). This approximation could have been avoided had we assumed from the beginning that the quantum states are plane waves; for general quantum states, however, it seems difficult to avoid this approach, since the commutator $[A,\rho]$ cannot be easily evaluated except via Kubo's lemma when $\rho \to \rho_{eq}$. More correctly, ρ_{eq} is the *local* equilibrium distribution.

[27]H. J. Groenewold, Physica **12**, 405 (1946); also Meddlelser (Copenhagen) **30**, no. 19 (1946).

[28]The full wave function is $\psi_{\mathbf{k}}(\mathbf{q}) = \Omega^{-1/2} e^{i\mathbf{k}\cdot\mathbf{q}}\phi_{\mathbf{k}}(\mathbf{q})$. With $\phi_{\mathbf{k}}(\mathbf{q}) = \Sigma_{\mathbf{g}} A_{\mathbf{k}}(\mathbf{g})e^{i\mathbf{q}\cdot\mathbf{g}}$, where $\mathbf{g}$ is a reciprocal lattice vector, we have $|\phi_{\mathbf{k}}(0)|^2 = \Sigma_{\mathbf{gg}'} A_{\mathbf{k}}^*(\mathbf{g}')A_{\mathbf{k}}(\mathbf{g}) \simeq \Sigma_{\mathbf{g}} |A_{\mathbf{k}}(\mathbf{g})|^2 = 1$ by Parseval's theorem.

[29]J. Ziman, *Electrons and Phonons* (Clarendon Press, Oxford, 1960).

[30]J. Reitz, in *Solid State Physics*, Vol. 1, edited by F. Seitz and D. Turnbull (Academic, New York, 1955), p.2.

Tensor fields on crystals

D. B. Litvin
Department of Physics, The Pennsylvania State University—The Berks Campus, P. O. Box 2150, Reading, Pennsylvania 19608

(Received 23 July 1981; accepted for publication 16 September 1981)

A new method is presented to determine the irreducible representations of the space group of a crystal contained in the representation whose basis functions are the components of a tensor field defined on the atoms of a crystal. This reducible representation is the direct product of a tensor representation, dependent only on the tensor, and a permutation representation dependent only on how the atoms permute under elements of the space group. The permutation representation is first separately reduced prior to the reduction of the direct product. The permutation representation is shown to be an induced representation and its reduction is facilitated using the theory of induced representations. Examples and tables of results of applying this method are given in the case of a polar vector tensor field, applicable to lattice vibrational problems, and crystals, as the diamond structure, of space group symmetry 0_h^7.

PACS numbers: 61.50.Em, 02.20. + b

I. INTRODUCTION

In many problems in solid-state physics it is often necessary to determine the irreducible representations of the space group of a crystal contained in a tensor field representation, a reducible representation of the space group whose basis functions are components of a tensor defined on the atoms of the crystal. In lattice vibrational problems[1,2] the basis functions of the tensor field representation are components of a three component tensor defined on each atom, the displacements of each atom. In classifying magnetic ordering in crystals by irreducible representations of a nonmagnetic space group,[3–5] one reduces a tensor field representation whose basis functions are the components of the atomic spins. Also, in applying the tensor-field criterion[6] in the Landau theory of continuous phase transitions, one reduces a tensor field representation, as in the case of magnetostructural phase transitions where the basis functions are components of a six-component tensor[7] defined on each atom.

The tensor field representation is the direct product of a permutation representation of the atoms of the crystal, representing how the atoms of the crystal permute under the space group elements of the crystal, and a tensor representation associated with the transformation of the tensor components defined on the atoms. In the case of lattice vibrational problems, the tensor representation is the polar vector representation, in the case of classification of magnetic ordering, it is the axial vector representation, and in the case of magnetostructural phase transitions, it is the direct product of the polar and axial vector representations.

To determine the irreducible representations contained in the tensor field representation one could use the standard group theoretical projection operator method[8] as has been done, for example, in the case of lattice vibrational problems.[1] Such a method, while of course giving the correct irreducible representations, does not take into account the common property of all tensor field representations defined on a specific crystal: The permutation representation component of the tensor field representation is the same for all tensor field representations defined on the crystal. This commonality has led to an alternate method to determine the irreducible representations contained in the tensor field representation: First determine the irreducible representations contained in the permutation representation, and then those contained in the tensor field representation.

Lulek[9] has considered the lattice vibrational problem of molecules using such a method. The irreducible representation of the point group of the molecule contained in the permutational representation, there called the positional representation, are determined using the theory of representations of permutation groups. Kuzma, Kupolowski, and Lulek[10] have applied this method to the cases of the lattice vibrations of a regular tetrahedron and cube. Birman, Kotzev, and Litvin,[11] in the context of the tensor-field criterion of the Landau theory of continuous phase transitions, have also used such a method. They have derived using the theory of color groups the $k = 0$ irreducible representations of a space group contained in the permutation representation for all possible crystals. Berenson, Kotzev, and Litvin[12] have then tabulated the $k = 0$ irreducible representations of a space group in the tensor field representation, for all possible crystals in the cases where the tensor representation is taken to be the polar vector representation, the axial vector representation, the product of the polar and axial vector representations, and the symmetrized square of the polar vector representation.

In this paper we shall consider the problem of determining all irreducible representations of the space group of a crystal contained in a tensor field representation defined on a crystal. In Sec. II we show that the tensor field representation defined on an arbitrary crystal is the direct sum of the tensor field representations defined on the arbitrary crystal's constituent simple crystals. The structure of the permutation representation of a simple crystal is derived in Sec. III. In Sec. IV, using the theory of induced representations, a general method is derived to determine all irreducible representations of the space group of a crystal contained in the permutation representation of a simple crystal. As an example, all irreducible representations contained in the permutation representations of all simple crystals of a crystal of space group symmetry O_h^7 are derived and tabulated. Finally, in Sec. V, we discuss determining all irreducible representations of the space group of a crystal contained in a tensor field representation defined on a simple crystal. As an example we consider the polar vector tensor-field representation of the diamond structure in conjunction with the lattice vibrational problem in this structure.

II. TENSOR FIELD REPRESENTATION

Consider a crystal of space group symmetry **G** and let $\mathbf{r}_i$, $i = 1,2\cdots$, denote the atomic position vectors of the atoms of the crystal. To each atom of the crystal we associate a q-component tensor $\mathscr{T}$ with components $\mathscr{T}_s$, $s = 1,2,...,q$. The q-component function $\mathscr{T}(\mathbf{r}_i)_s$, $s = 1,2,...,q$ defined on the atomic positions $\mathbf{r}_i$, $i = 1,2,\cdots$, is called a q-component tensor field on the crystal. The corresponding tensor field representaion $D_{\mathbf{G}}^{\mathrm{TF}}(\mathrm{Crys})$ of the space group **G** is that representation of **G** whose basis functions are the components $\mathscr{T}(\mathbf{r}_i)_s$, $s = 1,2,...,q$, $i = 1,2,\cdots$, of the tensor field.

The tensor field representation $D_{\mathbf{G}}^{\mathrm{TF}}(\mathrm{Crys})$ can be written as

$$D_{\mathbf{G}}^{\mathrm{TF}}(\mathrm{Crys}) = D_{\mathbf{G}}^{\mathrm{PERM}}(\mathrm{Crys}) \times D_{\mathbf{G}}^{T}, \tag{1}$$

where $D_{\mathbf{G}}^{\mathrm{PERM}}(\mathrm{Crys})$ is the permutation representation of the atoms of the crystal, representing how the atoms of the crystal permute under elements of the space group of the crystal, and $D_{\mathbf{G}}^{T}$ is the representation of **G** called the tensor representation whose basis functions are the q components of the tensor $\mathscr{T}$. It is the purpose of this paper to derive a method to determine the irreducible representations of **G** contained in a tensor field representation $D_{\mathbf{G}}^{\mathrm{TF}}(\mathrm{Crys})$ defined by Eq. (1).

A crystal of space group symmetry **G** can be partitioned into "simple crystals."[13] Each simple crystal consists of all atoms whose atomic position vectors can be obtained by applying all elements of the space group **G** to any one atomic position vector **r**, and is said to be generated by **G** from **r**. A crystal can be considered as consisting of a certain number of simple crystals, no two simple crystals have atoms in common, and the elements of **G** permute the atoms of each simple crystal among themselves.

Let the tensor field be defined on a crystal consisting of m simple crystals generated by **G** from $\mathbf{r}_j$, $j = 1,2,...,m$. Because the elements of **G** permute the atoms of each simple crystal among themselves,

$$D_{\mathbf{G}}^{\mathrm{PERM}}(\mathrm{Crys}) = D_{\mathbf{G}}^{\mathrm{PERM}}(\mathbf{r}_1) + D_{\mathbf{G}}^{\mathrm{PERM}}(\mathbf{r}_2) + \cdots + D_{\mathbf{G}}^{\mathrm{PERM}}(\mathbf{r}_m), \tag{2}$$

that is, the permutation representation of the atoms of the crystal is the direct sum of the permutation representations $D_{\mathbf{G}}^{\mathrm{PERM}}(\mathbf{r}_j)$, $j = 1,2,...,m$, of each of the simple crystals. Substituting Eq. (2) into Eq. (1), the tensor field representation is written

$$D_{\mathbf{G}}^{\mathrm{TF}}(\mathrm{Crys}) = [D_{\mathbf{G}}^{\mathrm{PERM}}(\mathbf{r}_1) + D_{\mathbf{G}}^{\mathrm{PERM}}(\mathbf{r}_2) + \cdots + D_{\mathbf{G}}^{\mathrm{PERM}}(\mathbf{r}_m)] \times D_{\mathbf{G}}^{T}, \tag{3}$$

and subsequently as

$$D_{\mathbf{G}}^{\mathrm{TF}}(\mathrm{Crys}) = D_{\mathbf{G}}^{\mathrm{TF}}(\mathbf{r}_1) + D_{\mathbf{G}}^{\mathrm{TF}}(\mathbf{r}_2) + \cdots + D_{\mathbf{G}}^{\mathrm{TF}}(\mathbf{r}_m), \tag{4}$$

where $D_{\mathbf{G}}^{\mathrm{TF}}(\mathbf{r}_j)$, the tensor field representation of the jth simple crystal, is defined by

$$D_{\mathbf{G}}^{\mathrm{TF}}(\mathbf{r}_j) = D_{\mathbf{G}}^{\mathrm{PERM}}(\mathbf{r}_j) \times D_{\mathbf{G}}^{T}. \tag{5}$$

The tensor field representation of the crystal is, by Eq (4), the direct sum of the tensor field representations associated with each simple crystal. To determine the irreducible representation of **G** contained in $D_{\mathbf{G}}^{\mathrm{TF}}(\mathrm{Crys})$ is then equivalent to determining the irreducible representations of **G** contained in each of the tensor field representations $D_{\mathbf{G}}^{\mathrm{TF}}(\mathbf{r}_j)$, $j = 1,2,...,m$, of each simple crystal. Consequently, in what follows, we shall restrict ourselves to the case of a crystal consisting of a single simple crystal. We shall consider a single simple crystal generated by **G** from the atomic position vector **r**, and the tensor field representation $D_{\mathbf{G}}^{\mathrm{TF}}(\mathbf{r})$ defined on this simple crystal:

$$D_{\mathbf{G}}^{\mathrm{TF}}(\mathbf{r}) = D_{\mathbf{G}}^{\mathrm{PERM}}(\mathbf{r}) \times D_{\mathbf{G}}^{T}. \tag{6}$$

Common to all tensor field representations $D_{\mathbf{G}}^{\mathrm{TF}}(\mathbf{r})$ defined on a specific simple crystal generated by **G** from **r**, is the permutation representation $D_{\mathbf{G}}^{\mathrm{PERM}}(\mathbf{r})$ of the atoms of the simple crystal.

III. PERMUTATION REPRESENTATION $D_{\mathbf{G}}^{\mathrm{PERM}}(\mathbf{r})$

Let $D_{\mathbf{G}}^{\mathrm{PERM}}(\mathbf{r})$ be the permutation representation of the atoms of a simple crystal generated by a space group **G** from the atom position vector **r**. The position vector **r** can be characterized by its site space group **G(r)**, the subgroup of elements G of **G** such that

$$G\mathbf{r} = \mathbf{r} + \mathbf{t}, \tag{7}$$

where **t** is a primitive translation of the space group **G**. The point group **R(r)** of **G(r)** is called the "site point group" of **r**. One can expand the space group **G** into a coset decomposition with respect to **G(r)**,

$$\mathbf{G} = \mathbf{G}(\mathbf{r}) + G_2\mathbf{G}(\mathbf{r}) + \cdots + G_n\mathbf{G}(\mathbf{r}), \tag{8}$$

and define the set of atom positions $G_i\mathbf{r}$, $i = 1,2,...,n$, where G_i is a coset representative in Eq. (8). The coordinates of this set of atom positions, for one or two of each class of space groups **G**, each **r**, and a specific choice of coset representatives, are given in the *International Tables for X-Ray Crystallography*.[14] They are called there the "coordinates of equivalent positions" and the site point group **R(r)** is called the "point symmetry" of each of the equivalent positions.

In addition, we characterize the position vector **r** from which a simple crystal is generated by **G** by the "site subgroup" **H(r)**, the subgroup of elements of the space group **G** such that

$$G\mathbf{r} = \mathbf{r}. \tag{9}$$

Elements of the site subgroup **H(r)** are, in general, of the form $(R\,|\mathbf{v}(R) + \mathbf{t}_R)$ where R is an element of the site point group **R(r)**, $\mathbf{v}(R)$ the nonprimitive translation associated with R, and $\mathbf{t}_R$ a specific primitive translation. The site subgroup **H(r)** is isomorphic to the site point group **R(r)**. However, if the choice of the origin of the space group **G** is taken to be that given in the *International Tables for X-Ray Crystallography*,[14] then the site point group **R(r)** is not necessarily a subgroup of the space group $\mathbf{G}$. As we shall show below, it is the site subgroup $\mathbf{H}(\mathbf{r})$ of the position vector $\mathbf{r}$ from which the simple crystal is generated by $\mathbf{G}$ which plays a central role in determining the irreducible representations of G contained in the permutation representation $D_{\mathbf{G}}^{\mathrm{PERM}}(\mathbf{r})$.

To determine the structure of the permutation representation $D_{\mathbf{G}}^{\mathrm{PERM}}(\mathbf{r})$ we expand the space group $\mathbf{G}$ into a coset decomposition with respect to the site subgroup $\mathbf{H}(\mathbf{r})$:

$$\mathbf{G} = \mathbf{H}(\mathbf{r}) + G_2\mathbf{H}(\mathbf{r}) + G_3\mathbf{H}(\mathbf{r}) + \cdots. \quad (10)$$

Since all elements $\mathbf{H}(\mathbf{r})$ leave $\mathbf{r}$ invariant, the atomic position vectors of the simple crystal generated by $\mathbf{G}$ from $\mathbf{r}$ are in a one-to-one correspondence with the cosets of Eq. (10). That is, the atomic position vectors $\mathbf{r}_i$, $i = 1,2,3,\cdots$, of the simple crystal are such that $\mathbf{r}_i = G_i\mathbf{r}$, $i = 1,2,3,\cdots$, where G_i is a coset representative of Eq. (10). Since the permutation representation $D_{\mathbf{G}}^{\mathrm{PERM}}(\mathbf{r})$ is the representation of $\mathbf{G}$ whose basis functions are the atomic position vectors $\mathbf{r}_i = G_i\mathbf{r}$, $i = 1,2,3,\cdots$, the (i,j)th component of the matrix $D_{\mathbf{G}}^{\mathrm{PERM}}(\mathbf{r})$ is one if $G\mathbf{r}_j = \mathbf{r}_i$ or zero if $G\mathbf{r}_j \neq \mathbf{r}_i$. Consequently, the matrices of the permutation representation $D_{\mathbf{G}}^{\mathrm{PERM}}(\mathbf{r})$ are defined by

$$D_{\mathbf{G}}^{\mathrm{PERM}}(\mathbf{r})[G]_{ij} = \begin{cases} 1 & \text{if } G_i^{-1}GG_j \in \mathbf{H}(\mathbf{r}), \\ 0 & \text{otherwise,} \end{cases} \quad (11)$$

where $i,j = 1,2,3,\cdots$, and G_i and G_j are coset representatives of Eq. (10). It follows from Eq. (11) that the permutation representation $D_{\mathbf{G}}^{\mathrm{PERM}}(\mathbf{r})$ is the representation of the space group $\mathbf{G}$ "induced" by the identity representation $D_{\mathbf{H}(\mathbf{r})}^{1}$ of the site subgroup $\mathbf{H}(\mathbf{r})$.[15] We shall write

$$D_{\mathbf{G}}^{\mathrm{PERM}}(\mathbf{r}) = D_{\mathbf{H}(\mathbf{r})}^{1}\uparrow\mathbf{G} \quad (12)$$

to denote the permutation representation as the representation of $\mathbf{G}$ induced by the identity representation of the site subgroup $\mathbf{H}(\mathbf{r})$.

IV. REDUCTION OF PERMUTATION REPRESENTATION

A. General reduction

We determine the irreducible representations of a space group $\mathbf{G}$ contained in the permutation representation $D_{\mathbf{G}}^{\mathrm{PERM}}(\mathbf{r})$: Let $D_{\mathbf{G}}^{(\mathbf{k}^*,\nu)}$ denote the $(\mathbf{k}^*,\nu)$th irreducible representation of the space group $\mathbf{G}$, and $D_{\mathbf{G}(\mathbf{k})}^{\nu}$ the νth irreducible representation of the group $\mathbf{G}(\mathbf{k})$ of the wave vector $\mathbf{k}$.[16] We have

$$D_{\mathbf{G}}^{(\mathbf{k}^*,\nu)} = D_{\mathbf{G}(\mathbf{k})}^{\nu}\uparrow\mathbf{G}, \quad (13)$$

that is, the irreducible representation $D_{\mathbf{G}}^{(\mathbf{k}^*,\nu)}$ of $\mathbf{G}$ is induced by the irreducible representation $D_{\mathbf{G}(\mathbf{k})}^{\nu}$ of $\mathbf{G}(\mathbf{k})$. We decompose the permutation representation

$$D_{\mathbf{G}}^{\mathrm{PERM}}(\mathbf{r}) = \sum_{(\mathbf{k}^*,\nu)} d(\mathbf{k}^*,\nu) D_{\mathbf{G}}^{(\mathbf{k}^*,\nu)}, \quad (14)$$

where $d(\mathbf{k}^*,\nu)$ is the number of times the irreducible representation $D_{\mathbf{G}}^{(\mathbf{k}^*,\nu)}$ of the space group $\mathbf{G}$ is contained in the permutation representation $D_{\mathbf{G}}^{\mathrm{PERM}}(\mathbf{r})$. We shall determine the coefficients $d(\mathbf{k}^*,\nu)$ of Eq. (14) using the theory of induced representations.[17,18]

The number of times the irreducible representation $D_{\mathbf{G}}^{(\mathbf{k}^*,\nu)}$ is contained in $D_{\mathbf{G}}^{\mathrm{PERM}}(\mathbf{r})$ is called the "intertwining number of $D_{\mathbf{G}}^{(\mathbf{k}^*,\nu)}$ with $D_{\mathbf{G}}^{\mathrm{PERM}}(\mathbf{r})$" and is denoted by the symbol $\mathrm{I}[D_{\mathbf{G}}^{(\mathbf{k}^*,\nu)}, D_{\mathbf{G}}^{\mathrm{PERM}}(\mathbf{r})]$. From Eq. (14) we have then that

$$d(\mathbf{k}^*,\nu) = \mathrm{I}[D_{\mathbf{G}}^{(\mathbf{k}^*,\nu)}, D_{\mathbf{G}}^{\mathrm{PERM}}(\mathbf{r})]. \quad (15)$$

Using Eqs. (12) and (13) we can rewrite this as

$$d(\mathbf{k}^*,\nu) = \mathrm{I}[D_{\mathbf{G}(\mathbf{k})}^{\nu}\uparrow\mathbf{G}, D_{\mathbf{H}(\mathbf{r})}^{1}\uparrow\mathbf{G}]. \quad (16)$$

To evaluate the intertwining number on the right-hand side of Eq. (16) using the Intertwining Number Theorem[18] requires the introduction of a double coset decomposition of $\mathbf{G}$: We expand the space group $\mathbf{G}$ into a double coset decomposition[17] with respect to the site subgroup $\mathbf{H}(\mathbf{r})$ and the group $\mathbf{G}(\mathbf{k})$ of the wavevector $\mathbf{k}$,

$$\mathbf{G} = \sum_i \mathbf{H}(\mathbf{r}) G_i \mathbf{G}(\mathbf{k}), \quad (17)$$

where the $\mathbf{G}_i$ are double coset representatives. For each double coset representative in Eq. (17) we define the group $\mathbf{L}_i$,

$$\mathbf{L}_i = \mathbf{H}(\mathbf{r}) \cap G_i \mathbf{G}(\mathbf{k}) G_i^{-1}, \quad (18)$$

and the representation D_i^{ν} of the group $G_i\mathbf{G}(\mathbf{k})G_i^{-1}$:

$$D_i^{\nu}(G_i G(\mathbf{k}) G_i^{-1}) \equiv D_{\mathbf{G}(\mathbf{k})}^{\nu}(G(\mathbf{k})). \quad (19)$$

Using the Intertwining Number Theorem,[18] Eq. (16) can be rewritten as

$$d(\mathbf{k}^*,\nu) = \sum_i \mathrm{I}[D_i^{\nu}\downarrow\mathbf{L}_i, D_{\mathbf{H}(\mathbf{r})}^{1}\downarrow\mathbf{L}_i], \quad (20)$$

where the summation is over all "i" corresponding to double coset representatives G_i of Eq. (17), with $\mathbf{L}_i$ and D_i^{ν} defined, respectively, by Eqs. (18) and (19). A symbol $D_{\mathbf{A}}^{\alpha}\downarrow\mathbf{B}$ denotes the representation of the subgroup $\mathbf{B}$ of $\mathbf{A}$ subduced onto $\mathbf{B}$ from the representation $D_{\mathbf{A}}^{\alpha}$ of $\mathbf{A}$,[15] the representation of $\mathbf{B}$ found by restricting the representation $D_{\mathbf{A}}^{\alpha}(A)$ to elements $A \in \mathbf{B}$. Equation (20) can be rewritten as

$$d(\mathbf{k}^*,\nu) = \sum_i \mathrm{I}[D_{\mathbf{L}_i}^{1}, (D_i^{\nu}\downarrow\mathbf{L}_i) \times (D_{\mathbf{H}(\mathbf{r})}^{1}\downarrow\mathbf{L}_i)], \quad (21)$$

where $D_{\mathbf{L}_i}^{1}$ is the identity representation of $\mathbf{L}_i$. Finally, since by Eq. (18), $\mathbf{L}_i$ is a subgroup of $\mathbf{H}(\mathbf{r})$, $D_{\mathbf{H}(\mathbf{r})}^{1}\downarrow\mathbf{L}_i = D_{\mathbf{L}_i}^{1}$, and

$$d(\mathbf{k}^*,\nu) = \sum_i \mathrm{I}[D_{\mathbf{L}_i}^{1}, D_i^{\nu}\downarrow\mathbf{L}_i]. \quad (22)$$

Consequently, the number $d(\mathbf{k}^*,\nu)$ of times the irreducible representation $D_{\mathbf{G}}^{(\mathbf{k}^*,\nu)}$ of the space group $\mathbf{G}$ is contained in the permutation representation $D_{\mathbf{G}}^{\mathrm{PERM}}(\mathbf{r})$ is equal to the sum, over the index i, of the number of times the identity representation of $\mathbf{L}_i$ is contained in the subduced representation $D_i^{\nu}\downarrow\mathbf{L}_i$.

Equation (22) can be reformulated in terms of the irreducible representations $D_{\mathbf{G}(\mathbf{k})}^{\nu}$ of the group $\mathbf{G}(\mathbf{k})$ of the wavevector $\mathbf{k}$: an intertwining number on the right-hand side of Eq. (22) is defined by

$$\mathrm{I}[D_{\mathbf{L}_i}^{1}, D_i^{\nu}\downarrow\mathbf{L}_i] = \frac{1}{|\mathbf{L}_i|} \sum_{\mathbf{L}_i} \chi_i^{\nu}(L_i), \quad (23)$$

where $|\mathbf{L}_i|$ is the order of the group $\mathbf{L}_i$ and $\chi_i^{\nu}(L_i)$ is the character of $D_i^{\nu}(L_i)$ defined by Eq. (19), $D_i^{\nu}(L_i) = D_i^{\nu}(G_i G(\mathbf{k}) G_i^{-1}) \equiv D_{\mathbf{G}(\mathbf{k})}^{\nu}(G(\mathbf{k}))$ for the elements $G(\mathbf{k}) = G_i^{-1} L_i G_i^{-1}$ of $\mathbf{G}(\mathbf{k})$. Since $D_i^{\nu}(L_i) = D_{\mathbf{G}(\mathbf{k})}^{\nu}(G_i^{-1} L_i G_i)$, $|\mathbf{L}_i| = |G_i^{-1}\mathbf{L}_i G_i|$, and

$G_i^{-1}\mathbf{L}_iG_i$ is a subgroup of $\mathbf{G}(\mathbf{k})$, we may rewrite Eq. (23) as

$$\mathrm{I}[D^1_{\mathbf{L}_i}, D^\nu_i \downarrow \mathbf{L}_i] = \frac{1}{|G_i^{-1}\mathbf{L}_iG_i|}\sum_{\mathbf{L}_i}\chi^\nu_{\mathbf{G}(\mathbf{k})}(G_i^{-1}L_iG_i), \quad (24)$$

and subsequently,

$$\mathrm{I}[D^1_{\mathbf{L}_i}, D^\nu_i \downarrow \mathbf{L}_i] = \mathrm{I}[D^1_{G_i^{-1}\mathbf{L}_iG_i}, D^\nu_{\mathbf{G}(\mathbf{k})}\downarrow G_i^{-1}\mathbf{L}_iG_i]. \quad (25)$$

Substituting Eq. (25) into Eq. (22), the coefficients $d(\mathbf{k}^*,\nu)$ of Eq. (14) are given in terms of the irreducible representation $D^\nu_{\mathbf{G}(\mathbf{k})}$ by

$$d(\mathbf{k}^*,\nu) = \sum_i \mathrm{I}[D^1_{G_i^{-1}\mathbf{L}_iG_i}, D^\nu_{\mathbf{G}(\mathbf{k})}\downarrow G_i^{-1}\mathbf{L}_iG_i]. \quad (26)$$

Consequently, the number $d(\mathbf{k}^*,\nu)$ of times the irreducible representation $D^{(\mathbf{k}^*,\nu)}_{\mathbf{G}}$ of the space group $\mathbf{G}$ is contained in the permutation representation $D^{\mathrm{PERM}}_{\mathbf{G}}(\mathbf{r})$ is equal to the sum, over the index i, of the number of times the identity representation of $G_i^{-1}\mathbf{L}_iG_i$, a subgroup of $\mathbf{G}(\mathbf{k})$, is contained in the representation $D^\nu_{\mathbf{G}(\mathbf{k})}$. Equation (25) provides a three-step method to determine the number $d(\mathbf{k}^*,\nu)$ of times in an irreducible representation $D^{(\mathbf{k}^*,\nu)}_{\mathbf{G}}$ is contained in the permutation representation $D^{\mathrm{PERM}}_{\mathbf{G}}(\mathbf{r})$:

(1) Determine the double coset representatives G_i of Eq. (17).

(2) Determine for each i the subgroup $G_i^{-1}\mathbf{L}_iG_i$ of $\mathbf{G}(\mathbf{k})$ using Eq.(18).

(3) Determine for each subgroup $G_i^{-1}\mathbf{L}_iG_i$ the number of times the identity representation is contained in $D^\nu_{\mathbf{G}(\mathbf{k})}\downarrow G_i^{-1}\mathbf{L}_iG_i$ using Eq. (24). The coefficient $d(\mathbf{k}^*,\nu)$ of Eq. (14), is given by Eq. (26) as the sum of the numbers determined in the above third step.

The calculation of the number of times the identity representation is contained in $D^\nu_{\mathbf{G}(\mathbf{k})}\downarrow G_i^{-1}\mathbf{L}_iG_i$, Eq. (24), can be simplified by taking into account the structure of the irreducible representations $D^\nu_{\mathbf{G}(\mathbf{k})}$ of the group $\mathbf{G}(\mathbf{k})$ of the wave vector $\mathbf{k}$.

B. k inside the Brillouin zone

Let $(R\,|\mathbf{v}(R)+\mathbf{t})$ denote an element of the group $\mathbf{L}_i$ defined by Eq. (18), $\mathbf{R}(\mathbf{L}_i)$ the point group of $\mathbf{L}_i$, and $(R_i\,|\mathbf{v}(R_i))$ the double coset representatives G_i of Eq. (17). Since $(R\,|\mathbf{v}(R)+\mathbf{t})$ is contained in $\mathbf{H}(\mathbf{r})$,

$$\mathbf{v}(R)+\mathbf{t} = \mathbf{r} - R\mathbf{r}, \quad (27)$$

and since $(R\,|\mathbf{v}(R)+\mathbf{t})$ is also contained in $G_i^{-1}\mathbf{G}(\mathbf{k})G_i$,

$$R_i^{-1}RR_i\mathbf{k} = \mathbf{k}+\mathbf{K}, \quad (28)$$

where $\mathbf{K}$ is a reciprocal lattice vector. If $\mathbf{k}$ is inside the Brillouin Zone $\mathbf{K}=0$ and the matrix of the irreducible representation $D^\nu_{\mathbf{G}(\mathbf{k})}(G_i^{-1}L_iG_i)$ can be written as[19]

$$D^\nu_{\mathbf{G}(\mathbf{k})}(G_i^{-1}L_iG_i) = \exp\{i\mathbf{k}\cdot R_i^{-1}[\mathbf{v}(R)+\mathbf{t}-\mathbf{v}(R_i) + R\mathbf{v}(R_i)]\}D^\nu_{\mathbf{R}(\mathbf{k})}(R_i^{-1}RR_i), \quad (29)$$

where $D^\nu_{\mathbf{R}(\mathbf{k})}$ is the νth irreducible representation of the point group $\mathbf{R}(\mathbf{k})$ of $\mathbf{G}(\mathbf{k})$. Using Eqs. (27) and (28) one finds that the exponential term equals one, and

$$D^\nu_{\mathbf{G}(\mathbf{k})}(G_i^{-1}L_iG_i) = D^\nu_{\mathbf{R}(\mathbf{k})}(R_i^{-1}RR_i). \quad (30)$$

Consequently, for wavevectors $\mathbf{k}$ within the Brillouin Zone, Eq. (26) becomes

$$d(\mathbf{k}^*,\nu) = \sum_i \mathrm{I}[D^1_{R_i^{-1}\mathbf{R}(\mathbf{L}_i)R_i}, D^\nu_{\mathbf{R}(\mathbf{k})}\downarrow R_i^{-1}\mathbf{R}(\mathbf{L}_i)R_i], \quad (31)$$

where $\mathbf{R}(\mathbf{L}_i)$ is the point group of $\mathbf{L}_i$, R_i the rotational part of a double coset representative, $\mathbf{R}(\mathbf{k})$ the point group of the wavevector $\mathbf{k}$, and $R_i^{-1}\mathbf{R}(\mathbf{L}_i)R_i$ a subgroup of $\mathbf{R}(\mathbf{k})$.

To determine $d(\mathbf{k}^*,\nu)$ is then a point group problem entailing three steps analogous to the three steps given in the preceding subsection:

(1) Determine the double coset representatives R_i in

$$\mathbf{R} = \sum_i \mathbf{R}(\mathbf{r})R_i\mathbf{R}(\mathbf{k}), \quad (32)$$

where $\mathbf{R}$ is the point group of the space group $\mathbf{G}$, $\mathbf{R}(\mathbf{k})$ of $\mathbf{G}(\mathbf{k})$, and $\mathbf{R}(\mathbf{r})$ is the site point group, the point group of $\mathbf{H}(\mathbf{r})$.

(2) Determine for each double coset representative R_i the subgroup $R_i^{-1}\mathbf{R}(\mathbf{L}_i)R_i$ of $\mathbf{R}(\mathbf{k})$ from

$$R_i^{-1}\mathbf{R}(\mathbf{L}_i)R_i = R_i^{-1}\mathbf{R}(\mathbf{r})R_i\cap\mathbf{R}(\mathbf{k}). \quad (33)$$

(3) Determine for each subgroup $R_i^{-1}\mathbf{R}(\mathbf{L}_i)R_i$ the number of times the identity representation is contained in $D^\nu_{\mathbf{R}(\mathbf{k})}$ subduced onto $R_i^{-1}\mathbf{R}(\mathbf{L}_i)R_i$. The coefficient $d(\mathbf{k}^*,\nu)$, Eq. (31), is the sum of the numbers calculated in step three above.

For the special case of $\mathbf{k}=0$, $\mathbf{R}(\mathbf{k})=\mathbf{R}$, there is only one double coset representative in Eq. (32), $R_1=E$, and $R_1^{-1}\mathbf{R}(\mathbf{L}_1)R_1=\mathbf{R}(\mathbf{r})$. From Eq. (31) we have

$$d(0,\nu) = \mathrm{I}[D^1_{\mathbf{R}(\mathbf{r})}, D^\nu_{\mathbf{R}}\downarrow\mathbf{R}(\mathbf{r})], \quad (34)$$

and the number $d(0,\nu)$ of times $D^{(0,\nu)}_{\mathbf{G}}$ is contained in the permutation representation $D^{\mathrm{PERM}}_{\mathbf{G}}(\mathbf{r})$ is equal to the number of times the identity representation is contained in $D^\nu_{\mathbf{R}}$ subduced onto the site point group $\mathbf{R}(\mathbf{r})$. Tables of $d(0,\nu)$ for all space groups $\mathbf{G}$ and site point groups $\mathbf{R}(\mathbf{r})$ are given by Kotzev, Litvin, and Birman.[11]

As an example we consider the space group $\mathbf{G}=0^7_h$ and the simple crystal generated by 0^7_h from $\mathbf{r}=(\frac{1}{8},\frac{1}{8},\frac{1}{8})$, Wyckoff (c) position in the notation of Ref. 14. The site point group is $\mathbf{R}(\mathbf{r})=D^{(xyz)}_{3d}$. We shall determine the number of times an irreducible representation $D^{(\mathbf{k}^*,\nu)}_{\mathbf{G}}$ of the space group $\mathbf{G}$, with $\mathbf{k}=(k_x,k_x,k_x)\equiv\Lambda$, is contained in the permutation representation $D^{\mathrm{PERM}}_{\mathbf{G}}(\mathbf{r})$.

The point group $\mathbf{R}(\mathbf{k})=C^{(xyz)}_{3v}$, and there are two double coset representatives, in this case, in Eq. (32), $R_1=E$ and $R_2=C_{2y}$. The corresponding subgroups, Eq. (33), are $R_1^{-1}\mathbf{R}(\mathbf{L}_i)R_1=C^{(xyz)}_{3v}$ and $R_2^{-1}\mathbf{R}(\mathbf{L}_2)R_2=C^{(x\bar{z})}_m$. For this wavevector $\mathbf{k}=\Lambda$, the only nonzero intertwining numbers in Eq. (31) are

$$\begin{aligned}&\mathrm{I}[D^1_{C_{3v}}, D^1_{C_{3v}}\downarrow C_{3v}]=1,\\ &\mathrm{I}[D^1_{C_m}, D^1_{C_{3v}}\downarrow C_m]=1,\\ &\mathrm{I}[D^1_{C_m}, D^3_{C_{3v}}\downarrow C_m]=1,\end{aligned} \quad (35)$$

where for the index ν of the irreducible representation $D^\nu_{\mathbf{R}(\mathbf{k})}$ we have used the conventions of Zak, Casher, Gluck, and Gur.[19] From Eqs. (32) and (35), we have that the only nonzero coefficients $d(\mathbf{k}^*,\nu)$ with $\mathbf{k}=\Lambda$ are

$$\begin{aligned}&d(\Lambda^*,1)=2,\\ &d(\Lambda^*,3)=1.\end{aligned} \quad (36)$$

TABLE I. Irreducible representations $G_G^{(k^*,\nu)}$ contained in the permutation representation $D_G^{\mathrm{PERM}}(r)$ of a simple crystal generated by $G = O_h^7$ from a point r: The points r are denoted in the Wyckoff position notation of Ref. 14: (a) = (0,0,0), (b) = $(\frac{1}{2},\frac{1}{2},\frac{1}{2})$, (c) $(\frac{1}{8},\frac{1}{8},\frac{1}{8})$, (d) = $(\frac{5}{8},\frac{5}{8},\frac{5}{8})$, (e) = (x,x,x), (f) = $(x,0,0)$, (g) = (x,x,z), (h) = $(\frac{1}{8},x,\frac{1}{4}-x)$, and (i) = (x,y,z). The number $d(k^*,\nu)$ of times $D_G^{(k^*,\nu)}$ is contained in $D_G^{\mathrm{PERM}}(r)$ is found at the intersection of the νth row of the k th subtable, and the column under the Wycokoff notation for the point r. The notation for k and indexation of ν is that of Ref. 20.

Γ	(a)	(b)	(c)	(d)	(e)	(f)	(g)	(h)	(i)
1	1	1	1	1	1	1	1	1	1
2									1
3						1	1	1	2
4			1	1	1	1	2	2	2
5							1	1	3
6								1	1
7	1	1			1	1	1		1
8						1	1	1	2
9							1	2	3
10					1	1	2	1	3
Δ									
1	1	1	1	1	2	3	4	3	6
2						1	2	3	6
3						1	3	3	6
4	1	1	1	1	2	3	3	3	6
5			1	1	2	2	6	6	12
Σ									
1	1	1	1	1	3	4	7	7	12
2			1	1	1	2	5	5	12
3			1	1	1	2	5	7	12
4	1	1	1	1	3	4	7	5	12
Λ									
1	2	2	2	2	4	4	5	4	8
2							3	4	8
3			1	1	2	4	8	8	16
Ξ									
1	1	1	2	2	4	6	12	12	24
2	1	1	2	2	4	6	12	12	24
Θ									
1	2	2	2	2	6	8	14	12	24
2			2	2	2	4	10	12	24
Υ									
1	1	1	1	1	2	2	4	3	6
2						2	2	3	6
3			1	1	1	1	3	4	6
4					1	1	3	2	6
W									
1	1		1	1	2	3	6	6	12
2		1	1	1	2	3	6	6	12

Γ	(a)	(b)	(c)	(d)	(e)	(f)	(g)	(h)	(i)
K									
1	1	1	2	2	3	4	7	7	12
2					1	2	5	5	12
3	1	1	1	1	3	4	7	5	12
4			1	1	1	2	5	7	12
L									
1	1	1	1	1	2	2	3	2	4
2							1	2	4
3				1	1	2	4	4	8
4			1	1	1	2	3	2	4
5	1	1			1		1	2	4
6			1		1	2	4	4	8
U									
1	1	1	1	1	3	4	7	5	12
2	1	1	2	2	3	4	7	7	12
3			1	1	1	2	5	7	12
4					1	2	5	5	12
Z									
1	1	1	2	2	4	6	12	12	24
Q									
1	1	1	2	2	4	6	12	12	24
2	1	1	2	2	4	6	12	12	24
S									
1	1	1	1	1	3	4	7	5	12
2	1	1	2	2	3	4	7	7	12
3			1	1	1	2	5	7	12
4					1	2	5	5	12
A									
1	1	1	2	2	4	6	12	12	24
2	1	1	2	2	4	6	12	12	24
B									
1	1	1	2	2	4	6	12	12	24
2	1	1	2	2	4	6	12	12	24
M									
1	2	2	3	3	6	8	14	12	24
2			1	1	2	4	10	12	24
N									
1	1	1	2	2	4	6	12	12	24
2	1	1	2	2	4	6	12	12	24

Consequently, the permutation representation $D_G^{\mathrm{PERM}}(\mathbf{r})$ for $\mathbf{G} = O_h^7$ and $\mathbf{r} = (\frac{1}{8}, \frac{1}{8}, \frac{1}{8})$, contains the irreducible representation $D_G^{(\Lambda^*,1)}$ twice and $D_G^{(\Lambda^*,3)}$ once, and no other irreducible representations of the space group $\mathbf{G} = O_h^7$ with the wavevector $\mathbf{k} = \mathbf{\Lambda}$. This information can be found in Table I at the intersection of the "c" column and the first and third rows of subtable $\mathbf{\Lambda}$.

C. k on the Brillouin Zone

For wavevectors **k** on the Brillouin Zone, in place of Eq. (29), one writes[19]

$$D^{\nu}_{G(\mathbf{k})}(G_i^{-1}L_iG_i) = e^{i\mathbf{k}\cdot\mathbf{t}(R_i^{-1}R(L_i)R_i)}\bar{D}^{\nu}_{\mathbf{R}(\mathbf{k})}(R_i^{-1}R(L_i)R_i), \tag{37}$$

where the primitive translation $\mathbf{t}(R_i^{-1}R(L_i)R_i)$ is determined by

$$G_i^{-1}L_iG_i = (R_i^{-1}R(L_i)R_i|\mathbf{v}(R_i^{-1}R(L_i)R_i) + \mathbf{t}(R_i^{-1}R(L_i)R_i)) \tag{38}$$

and $\bar{D}^{\nu}_{\mathbf{R}(\mathbf{k})}$ is the νth irreducible ray representation[19] of the point group $\mathbf{R}(\mathbf{k})$ of the wavevector $\mathbf{k}$.

Using Eq. (37), Eq. (26) can be rewritten for wavevectors $\mathbf{k}$ on the Brillouin Zone, as

$$d(\mathbf{k}^*,\nu) = \sum_i \mathrm{I}\left[D^{1}_{R_i^{-1}\mathbf{R}(\mathbf{L}_i)R_i}, e^{i\mathbf{k}\cdot\mathbf{t}}\bar{D}^{\nu}_{\mathbf{R}(\mathbf{k})} \downarrow R_i^{-1}\mathbf{R}(\mathbf{L}_i)R_i\right], \tag{39}$$

where $\bar{D}^{\nu}_{\mathbf{R}(\mathbf{k})}$ is the νth irreducible ray representation of $\mathbf{R}(\mathbf{k})$, and $\mathbf{t} = \mathbf{t}(R_i^{-1}R(L_i)R_i)$ is defined by Eq. (38).

The coefficients $d(\mathbf{k}^*,\nu)$ are determined again by a three-step procedure:

(1) The double coset representatives G_i are determined from Eq. (17).

(2) The subgroups $R_i^{-1}\mathbf{R}(\mathbf{L}_i)R_i$ of $\mathbf{R}(\mathbf{k})$ are determined from Eq. (33), and the translations $\mathbf{t}(R_i^{-1}\mathbf{R}(L_i)R_i)$ from Eq. (38).

(3) Determine for each subgroup $R_i\mathbf{R}(\mathbf{L}_i)R_i$ the number of times the identity representation is contained in $e^{i\mathbf{k}\cdot\mathbf{t}}\bar{D}^{\nu}_{\mathbf{R}(\mathbf{k})}$ subduced onto $R_i^{-1}\mathbf{R}(\mathbf{L}_i)R_i$. The coefficient $d(\mathbf{k}^*,\nu)$, Eq. (39), is the sum of the numbers calculated in step three above.

As an example we again consider the space group $\mathbf{G} = 0_h^7$ and the simple crystal generated by 0_h^7 from the (c) position $\mathbf{r} = (\frac{1}{8}, \frac{1}{8}, \frac{1}{8})$. We shall determine the number of times irreducible representations $D_{\mathbf{G}}^{(\mathbf{k}^*,\nu)}$ with $\mathbf{k} = (3\pi/2a, 3\pi/2a, 0) \equiv \mathbf{K}$ are contained in the permutation representation $D_{\mathbf{G}}^{\mathrm{PERM}}(\mathbf{r})$. The site subgroup $\mathbf{H}(\mathbf{r}) = (C_{3v}^{(xyz)}|0,0,0) + (\bar{1}|\frac{1}{4}, \frac{1}{4}, \frac{1}{4})(C_{3v}^{(xyz)}|0,0,0)$ and $\mathbf{G}(\mathbf{k})$ consists of the elements $(E|0,0,0)$, $(m^{(z)}|\frac{1}{4}, \frac{1}{4}, \frac{1}{4})$, $(m^{(\bar{x}y)}|0,0,0)$, $(C_2^{(xy)}|0,0,0)$ and all primitive translations of $\mathbf{G} = 0_h^7$. There are two double coset representatives, Eq. (17), $G_1 = (E|0,0,0)$ and $G_2 = (C_2^x|0,0,0)$. The corresponding subgroups of $\mathbf{R}(\mathbf{k}) = C_{2v}^{(xy,\bar{x}y,z)}$ are $R_1^{-1}\mathbf{R}(\mathbf{L}_1)R_1 = C_2^{\bar{x}y}$ with $\mathbf{t}(E) = \mathbf{t}(m^{\bar{x}y}) = 0$, and $R_2^{-1}\mathbf{R}(\mathbf{L}_2)R_2 = C_2^{xy}$ with $\mathbf{t}(E) = 0$ and $\mathbf{t}(C_2^{xy}) = (0, -\frac{1}{2}, -\frac{1}{2})$. Using Eq. (39) and the numbering of Ref. 19 for the index ν of irreducible ray representations, the nonzero coefficients $d(\mathbf{k}^*,\nu)$ for $\mathbf{k} = \mathbf{K}$, are in this example:

$$\begin{aligned} d(\mathbf{K}^*,1) &= 2, \\ d(\mathbf{K}^*,3) &= 1, \\ d(\mathbf{K}^*,4) &= 1. \end{aligned} \tag{40}$$

Consequently, the permutation representation $D_{\mathbf{G}}^{\mathrm{PERM}}(\mathbf{r})$ for $\mathbf{G} = 0_h^7$ and $\mathbf{r} = (\frac{1}{8}, \frac{1}{8}, \frac{1}{8})$ contains the irreducible representation $D_{\mathbf{G}}^{(\mathbf{K}^*,1)}$ twice, the irreducible representations $D_{\mathbf{G}}^{(\mathbf{K}^*,3)}$ and $D_{\mathbf{G}}^{(\mathbf{K}^*,4)}$ each once, and no other irreducible representations with the wavevector $\mathbf{k} = \mathbf{K}$. This information is found in Table I at the intersection of the (c) column and rows of subtable $\mathbf{K}$.

In Table I we have tabulated all irreducible representations of the space group $\mathbf{G} = 0_h^7$ contained in the permutation representations $D_{\mathbf{G}}^{\mathrm{PERM}}(\mathbf{r})$ for all simple crystals generated by $\mathbf{G} = 0_h^7$.

TABLE II. The irreducible representations $D^{\nu}_{R(k)}$ contained in the direct product $D^{\bar{\nu}}_{R(k)} \times (D^{T}_{G} \downarrow R(k))$ for $G = 0_h^7$ and the polar vector tensor representation $D^{T}_{G} = D^{V}_{G}$: The irreducible representations $D^{\nu}_{R(k)}$ contained in the direct product are listed to the right of the irreducible representation $D^{\bar{\nu}}_{R(k)}$. Irreducible representations $D^{\nu}_{R(k)}$ are denoted by k_ν in the notation and indexation of Ref. 20.

Γ_1	Γ_{10}	Θ_1	$2\Theta_1 + \Theta_2$	Z_1	$3Z_1$
Γ_2	Γ_9	Θ_2	$\Theta_1 + 2\Theta_2$		
Γ_3	$\Gamma_9 + \Gamma_{10}$			Q_1	$Q_1 + 2Q_2$
Γ_4	$\Gamma_7 + \Gamma_8 + \Gamma_9 + \Gamma_{10}$			Q_2	$2Q_1 + Q_2$
Γ_5	$\Gamma_6 + \Gamma_8 + \Gamma_9 + \Gamma_{10}$				
Γ_6	Γ_5	X_1	$X_1 + X_3 + X_4$	S_1	$S_1 + S_2 + S_3$
Γ_7	Γ_4	X_2	$X_2 + X_3 + X_4$	S_2	$S_1 + S_2 + S_4$
Γ_8	$\Gamma_4 + \Gamma_5$	X_3	$X_1 + X_2 + X_4$	S_3	$S_1 + S_3 + S_4$
Γ_9	$\Gamma_2 + \Gamma_3 + \Gamma_4 + \Gamma_5$	X_4	$X_1 + X_2 + X_3$	S_4	$S_2 + S_3 + S_4$
Γ_{10}	$\Gamma_1 + \Gamma_3 + \Gamma_4 + \Gamma_5$				
				A_1	$2A_1 + A_2$
		W_1	$W_1 + 2W_2$	A_2	$A_1 + 2A_2$
Δ_1	$\Delta_1 + \Delta_5$	W_2	$2W_1 + W_2$		
Δ_2	$\Delta_2 + \Delta_5$			B_1	$2B_1 + B_2$
Δ_3	$\Delta_3 + \Delta_5$			B_2	$B_1 + 2B_2$
Δ_4	$\Delta_4 + \Delta_5$	K_1	$K_1 + K_2 + K_3$		
Δ_5	$\Delta_1 + \Delta_2 + \Delta_3 + \Delta_4 + \Delta_5$	K_2	$K_1 + K_2 + K_4$	M_1	$M_1 + 2M_2$
		K_3	$K_1 + K_3 + K_4$	M_2	$2M_1 + M_2$
		K_4	$K_2 + K_3 + K_4$		
Σ_1	$\Sigma_1 + \Sigma_2 + \Sigma_4$			N_1	$2N_1 + N_2$
Σ_2	$\Sigma_1 + \Sigma_2 + \Sigma_3$			N_2	$N_1 + 2N_2$
Σ_3	$\Sigma_2 + \Sigma_3 + \Sigma_4$	L_1	$L_5 + L_6$		
Σ_4	$\Sigma_1 + \Sigma_3 + \Sigma_4$	L_2	$L_4 + L_6$		
		L_3	$L_4 + L_5 + 2L_6$		
		L_4	$L_2 + L_3$		
Λ_1	$\Lambda_1 + \Lambda_3$	L_5	$L_1 + L_3$		
Λ_2	$\Lambda_2 + \Lambda_3$	L_6	$L_1 + L_2 + 2L_3$		
Λ_3	$\Lambda_1 + \Lambda_2 + 2\Lambda_3$				
		U_1	$U_1 + U_2 + U_3$		
Ξ_1	$2\Xi_1 + \Xi_2$	U_2	$U_1 + U_2 + U_4$		
Ξ_2	$\Xi_1 + 2\Xi_2$	U_3	$U_1 + U_3 + U_4$		
		U_4	$U_2 + U_3 + U_4$		

V. REDUCTION OF TENSOR FIELD REPRESENTATION

The tensor field representation $D_G^{\mathrm{TF}}(\mathbf{r})$ of a simple crystal is defined by Eq. (6)

$$D_G^{\mathrm{TF}}(\mathbf{r}) = D_G^{\mathrm{PERM}}(\mathbf{r}) \times D_G^T, \tag{6}$$

where $D_G^{\mathrm{PERM}}(\mathbf{r})$ is the permutation representation of the atomic positions of the simple crystal, and D_G^T is the tensor representation. In the preceding section we have derived a method to reduce the permutation representation and here shall assume that the coefficients $d(\mathbf{k}^*,\nu)$ of Eq. (13) are known. Substituting Eq. (13) into Eq. (6) we have

$$D_G^{\mathrm{TF}}(\mathbf{r}) = \sum_{\mathbf{k}^*,\nu} d(\mathbf{k}^*,\nu)[D_G^{(\mathbf{k}^*,\nu)} \times D_G^T]. \tag{41}$$

To determine the irreducible representations in $D_G^{\mathrm{TF}}(\mathbf{r})$ one must reduce the direct product of irreducible representations $D_G^{(\mathbf{k}^*,\nu)}$ and the tensor representation D_G^T. If

$$D_G^{(\mathbf{k}^*,\nu)} \times D_G^T = \sum_{\bar{\mathbf{k}}^*,\bar{\nu}} C(\mathbf{k}^*,\nu;\bar{\mathbf{k}}^*,\bar{\nu}) D_G^{(\bar{\mathbf{k}}^*,\bar{\nu})}, \tag{42}$$

then the reduced form of the tensor field representation is

$$D_G^{\mathrm{TF}}(\mathbf{r}) = \sum_{\mathbf{k}^*,\nu} b(\mathbf{k}^*,\nu) D_G^{(\mathbf{k}^*,\nu)}, \tag{43}$$

where

$$b(\mathbf{k}^*,\nu) = \sum_{\bar{\mathbf{k}}^*,\bar{\nu}} d(\bar{\mathbf{k}}^*,\bar{\nu}) C(\bar{\mathbf{k}}^*,\bar{\nu};\mathbf{k}^*,\nu). \tag{44}$$

We shall consider here tensor representations D_G^T which are independent of the translational components of the elements of $\mathbf{G}$, that is, which are $\mathbf{k}=0$ representations of $\mathbf{G}$. Consequently, in Eq. (42), $\bar{\mathbf{k}}^* = \mathbf{k}^*$. Abbreviating $C(\mathbf{k}^*,\nu;\mathbf{k}^*,\bar{\nu})$ by $C(\mathbf{k}^*,\nu,\bar{\nu})$, we can write Eqs. (42) and (44), respectively, as

$$D_G^{(\mathbf{k}^*,\nu)} \times D_G^T = \sum_{\bar{\nu}} C(\mathbf{k}^*,\nu,\bar{\nu}) D_G^{(\mathbf{k}^*,\bar{\nu})} \tag{45}$$

and

$$b(\mathbf{k}^*,\nu) = \sum_{\bar{\nu}} d(\mathbf{k}^*,\bar{\nu}) C(\mathbf{k}^*,\bar{\nu},\nu), \tag{46}$$

where the coefficients $C(\mathbf{k}^*,\bar{\nu},\nu)$ are defined as the intertwining numbers

$$C(\mathbf{k}^*,\bar{\nu},\nu) = \mathrm{I}[D_G^{(\mathbf{k}^*,\nu)}, D_G^{(\mathbf{k}^*,\bar{\nu})} \times D_G^T]. \tag{47}$$

Using Eq. (3), this can be rewritten as

$$C(\mathbf{k}^*,\bar{\nu},\nu) = \mathrm{I}[D_{\mathbf{G}(\mathbf{k})}^{\nu}, D_{\mathbf{G}(\mathbf{k})}^{\bar{\nu}} \times (D_G^T \downarrow \mathbf{G}(\mathbf{k}))] \tag{48}$$

and since D_G^T is a $\mathbf{k}=0$ representation of $\mathbf{G}$,

$$C(\mathbf{k}^*,\bar{\nu},\nu) = \mathrm{I}[D_{\mathbf{R}(\mathbf{k})}^{\nu}, D_{\mathbf{R}(\mathbf{k})}^{\bar{\nu}} \times (D_G^T \downarrow \mathbf{R}(\mathbf{k}))], \tag{49}$$

where, if $\mathbf{k}$ is a wavevector inside the Brillouin Zone, $D_{\mathbf{R}(\mathbf{k})}^{\nu}$ and $D_{\mathbf{R}(\mathbf{k})}^{\bar{\nu}}$ are irreducible representations of the point group $\mathbf{R}(\mathbf{k})$, and if $\mathbf{k}$ is on the Brillouin Zone, $D_{\mathbf{R}(\mathbf{k})}^{\nu}$ and $D_{\mathbf{R}(\mathbf{k})}^{\bar{\nu}}$ are replaced by $\bar{D}_{\mathbf{R}(\mathbf{k})}^{\nu}$ and $\bar{D}_{\mathbf{R}(\mathbf{k})}^{\bar{\nu}}$, irreducible ray representations of $\mathbf{R}(\mathbf{k})$. For the space group $\mathbf{G} = O_h^7$ and $D_G^T = D_G^V$, the polar vector representation, the irreducible representations $D_{\mathbf{R}(\mathbf{k})}^{\nu}$ contained in $D_{\mathbf{R}(\mathbf{k})}^{\bar{\nu}} \times (D_G^T \downarrow \mathbf{R}(\mathbf{k}))$ have been calculated and are tabulated in Table II. From this table the coefficients $C(\mathbf{k}^*,\bar{\nu},\nu)$ can be found for the case $\mathbf{G} = O_h^7$ and $D_G^T = D_G^V$. For example, for $\mathbf{k} = \Lambda$ from Table II one finds the nonzero coefficients $C(\Lambda^*,\bar{\nu},\nu)$:

$$\begin{aligned} &C(\Lambda^*,1,1) = C(\Lambda^*,1,3) = 1,\\ &C(\Lambda^*,2,2) = C(\Lambda^*,2,3) = 1,\\ &C(\Lambda^*,3,1) = C(\Lambda^*,3,2) = 1,\\ &C(\Lambda^*,3,3) = 2. \end{aligned} \tag{50}$$

The number $b(\mathbf{k}^*,\nu)$, Eq. (43), of times an irreducible representation $D_G^{(\mathbf{k}^*,\nu)}$ is contained in a tensor field representation $D_G^{\mathrm{TF}}(\mathbf{r})$ is determined from Eq. (46), with the coefficients $d(\mathbf{k}^*,\bar{\nu})$ calculated from Eq. (31) and $C(\mathbf{k}^*,\bar{\nu},\nu)$ from Eq. (49). For $\mathbf{k} = \Lambda$, the nonzero coefficients $d(\Lambda^*,\bar{\nu})$ are given in Eq. (36) and the nonzero coefficients $C(\Lambda^*,\bar{\nu},\nu)$ in Eq. (50). Using Eq. (45) we have

$$\begin{aligned} &b(\Lambda^*,1) = 3,\\ &b(\Lambda^*,2) = 1,\\ &b(\Lambda^*,3) = 4. \end{aligned} \tag{51}$$

Consequently, the tensor field representation $D_G^{\mathrm{TF}}(\mathbf{r})$ for $\mathbf{G} = O_h^7$, $\mathbf{r} = (\frac{1}{8},\frac{1}{8},\frac{1}{8})$, $D_G^T = D_G^V$, and $\mathbf{k} = \Lambda$, contains the irreducible representation $D_G^{(\Lambda^*,1)}$ three times, $D_G^{(\Lambda^*,2)}$ once, and $D_G^{(\Lambda^*,3)}$ four times.

For this case, where $D_G^T = D_G^V$, is the polar vector representation, the irreducible representations contained in the tensor field representation $D_G^{\mathrm{TF}}(\mathbf{r})$, Eq. (6), are the lattice vibration irreducible representations of the simple crystal generated by $\mathbf{G}$ from $\mathbf{r}$. For the diamond structure, $\mathbf{G} = O_h^7$, $\mathbf{r} = (0,0,0)$, the (a) position according to Ref. 14, we find for $\mathbf{k} = \Lambda$, from Eq. (46) and Tables I and II, the nonzero coefficients are $b(\Lambda^*,1) = b(\Lambda^*,3) = 2$. That is, the lattice vibration decomposition for the diamond structure at $\mathbf{k} = \Lambda$ is $2D_G^{(\Lambda^*,1)} + 2D_G^{(\Lambda^*,3)}$, in agreement with Ref. 20.

ACKNOWLEDGMENTS

This author would like to thank Professor J. L. Birman of City College—CUNY, Dr. J. N. Kotzev of the University of Sofia, and Dr. R. Berenson of Nassau Community College for stimulating discussions on many aspects of group theory.

[1]A. A. Maradudin and S. H. Vosko, Rev. Mod. Phys. **40**, 1 (1958).
[2]J. L. Birman, *Theory of Crystal Space Groups and Infra-Red and Raman Lattice Processes of Insulating Crystals* (Springer, Berlin, 1974).
[3]I. E. Dzialoshinsky, Z. Eksp. Teoret. Fiz. **32**, 1547 (1957) [Sov. Phys. JETP **5**, 1259 (1957)].
[4]S. Alexander, Phys. Rev. **127**, 420 (1962).
[5]E. F. Bertaut, J. Phys. (Paris), Suppl. A **24**, 217 (1968).
[6]J. L. Birman, Phys. Rev. Lett. **17**, 1216 (1966).
[7]J. L. Birman and Y. A. Izyumov, Phys. Rev. B **21**, 1089 (1980).
[8]L. Jansen and M. Boon, *Theory of Finite Groups. Applications in Physics* (North-Holland, Amsterdam, 1967), p. 223.
[9]T. Lulek, Acta Phys. Pol. A **57**, 407 (1980).
[10]M. Kuzma, I. Kuplowski, and T. Lulek, Acta Phys. Pol. A **57**, 415 (1980).
[11]J. N. Kotzev, D. B. Litvin, and J. L. Birman, to be published.
[12]R. Berenson, J. N. Kotzev, and D. B. Litvin, to be published.
[13]W. Opechowski and R. Guccione, in *Magnetism*, edited by G. T. Rado and H. Suhl (Academic, New York, 1965), Vol. IIA, p. 105.

[14] *International Tables for X-Ray Crystallography*, edited by N. F. M. Henry and K. Lonsdale (Kynoch, Birmingham, England, 1952), Vol. I.
[15] Reference 8, Sec. II.6.
[16] Reference 8, p. 252.
[17] C. J. Bradley, J. Math. Phys. 7, 1145 (1966).
[18] A. J. Coleman, in *Group Theory and Its Applications*, edited by M. Loebl (Academic, New York, 1968).
[19] J. Zak, A. Casher, M. Gluck, and Y. Gur, *The Irreducible Representations of Space Groups* (Benjamin, New York, 1969).
[20] J. L. Warren and T. G. Worlton, *Symmetry Properties of the Lattice Dynamics of Twenty-Three Crystals*, ANL-8053 (Argonne National Laboratory, Chicago, 1973), p. 508.

Erratum: Jet bundles and path structures [J. Math. Phys. 21, 1340 (1980)]

R. A. Coleman
Department of Mathematics and Statistics, University of Regina, Regina, Saskatchewan S4S 0A2, Canada
H. Korte
Department of Philosophy, University of Regina, Saskatchewan S4S 0A2, Canada

(Received 22 October 1981; accepted for publication 30 October 1981)

PACS numbers: 04.20. − q, 99.10. + g

1. In the 6th line from the bottom of the right column of p. 1340, replace
"the space of one-directions $\mathbb{D}_p^1(M)$" by "the set of bases of $\mathbb{D}_p^1(M)$."

2. In the first paragraph of Sec. 3, replace "$\mu(s_1)Is_2$" by "$\mu(s_1) = s_2$."

3. In the last line of Eq. (3.6), delete the superfluous left bracket.

4. The last line of Eq. (4.20) should be "$+ 2\overline{X}^n_{n\rho}\,\xi^\rho_1 + \overline{X}^n_{nn})$."

5. In Eq. (5.16), replace "f^i_{jk}" by "F^i_{jk}."

6. In the second line of Eq. (5.18), "$\Xi^{f\alpha}$" should be "$\Xi^{f\alpha}_2$."

7. In the numerator of the third line of Eq. (5.19), "$2f^\alpha_n\xi^\rho_1$" should be "$2f^\alpha_{n\rho}\xi^\rho_1$."

8. In the statement of Theorem 7 on p. 1347 and on lines 1 and 5 of the left column of p. 1348, insert "the set of bases of" prior to "$\mathbb{D}_p^1(M)$."

9. On line 6 of the left column of p. 1348 change "points" to "bases."

10. Following Eq. (6.26), the three occurrences of "$\forall_\alpha$" should be "$\forall\alpha$."