# The exponential mapping in Clifford algebras

John Froelich
*Department of Mathematics, The University of Iowa, Iowa City, Iowa 52242*

Nikos Salingaros
*Division of Mathematics, Computer Science and Systems Design, The University of Texas at San Antonio, San Antonio, Texas 78285*

(Received 13 October 1983; accepted for publication 27 December 1983)

The exponential mapping takes the Lie algebra of the Lorentz group into the Lorentz group. Each element of the group is defined as a formal power series, while the product of two exponential elements usually involves an infinite sum of commutator terms, such as in the Baker–Campbell–Hausdorff formula. Because of the special arithmetic in Clifford algebras, many Baker–Campbell–Hausdorff-like formulas and identities can be calculated or summed exactly when they involve only elements from the algebra. We calculate exact identities for the Baker–Campbell–Hausdorff formula and related formulas in the quaternion and dihedral algebras. These are useful in treatments of the Lorentz group, and make possible a truly finite (as opposed to infinitesimal) description of a transformation group in physics.

PACS numbers: 02.10. + w, 02.20. + b

## I. INTRODUCTION

The Lorentz group is the basis for the description of physical fields in space-time. A general element of the Lorentz group combines a pure Lorentz transformation (boost) with a rotation in three-space. Mathematically, the boosts and spatial rotations are described on an equal footing, yet the physical processes they describe are entirely distinct. For this reason, it is convenient and often necessary to separate the spatial rotation part from the boost part of a general Lorentz transformation. The mathematical description naturally mixes them, and a complete separation is highly nontrivial. This problem has not been solved previously in any generality, because the usual treatments of the Lorentz group are almost exclusively concerned with infinitesimal transformations. For that purpose, the usual local Lie algebra characterized by the structure constants and the commutation relations is sufficient, and is generally followed in physical discussions.[1,2]

In contradistinction, the description of the Lorentz group as a finite Lie transformation group requires the exponentiation of the corresponding Lie algebra.[3,4] The exponential mapping is in principle well known, but exact forms are not available for computations. Formulas such as the Baker–Campbell–Hausdorff formula[4–11] are given as an infinite series of commutator terms which cannot in general be summed explicitly.

Recently, we have tried to formulate a physical problem directly in the Lie group,[12] and have applied the exponential mapping to describe the motion of a charged particle in an electromagnetic field.[12,13] Those results are, however, given only in the first few terms corresponding to the commutator expansion of the exponential mapping.

In this paper, we consider the exponential mapping in a Clifford algebra. Since a Clifford algebra is algebraically closed, it follows that all exponentials, and therefore the exponential map, can be written exactly in terms of elementary functions. One can in this case completely sum all of the commutator terms in the exponential.

The solution of a problem which is obtained directly in the Lie transformation group can therefore be given exactly. Such solutions have not been available in the past because of the considerable difficulty in providing an exact finite description of a Lie transformation group such as the Lorentz group. The possibility of giving an exact as opposed to an approximate description of certain physical systems follows a three-step logic. Some systems believed to possess an exact solution in fact do not, since that solution is purely local. After noting this limitation, one may obtain a more accurate though still approximate solution by adding higher-order terms derived from the exponential mapping. An example of this procedure is discussed in Refs. 12 and 13. The final step is to obtain the genuinely exact solution of the physical system by including all the terms from the exponential mapping. The physical consequences of this result are more appropriately discussed elsewhere; here we present the mathematical identities that make such exact solutions possible.

## II. THE EXPONENTIAL MAPPING IN THE QUATERNION ALGEBRA

In a description of finite automorphisms, it is frequently the case that the generators of one transformation are orthogonal to the other transformation. Examples of this are spatial rotations about two orthogonal axes, or a spatial rotation compared with a pure Lorentz transformation (boost). In these physically relevant cases, the infinitesimal generators of the transformations define a subalgebra of the full algebra of the automorphism group. In certain special cases it is possible to write down the transformation operators, which are exponentials, exactly.

The first case considered here is when the transformations define a quaternion subalgebra. The algebraic rules are given by the usual anticommuting basis elements $e_1$, $e_2$, and $e_3$, with product $\vee$ (see Table I):

$$e_i \vee e_j = -\delta_{ij} + \epsilon_{ijk} e_k, \quad i,j,k = 1,2,3. \tag{1}$$

 0022-2488/84/082347-04$02.50  

TABLE I. Multiplication table for the quaternion algebra.

| $\vee$ | $e_1$ | $e_2$ | $e_3$ |
|---|---|---|---|
| $e_1$ | $-1$ | $e_3$ | $-e_2$ |
| $e_2$ | $-e_3$ | $-1$ | $e_1$ |
| $e_3$ | $e_2$ | $-e_1$ | $-1$ |

This paper is concerned with providing an exact solution of the following concrete problem: How can we separate an exponential containing $e_1$ and $e_2$ into the product of two exponentials, each containing $e_1$ and $e_2$ separately. Because the exponential mapping generates an infinite number of commutators, these expressions will contain higher-order terms in $e_3$. For the purpose of the physical applications in mind, the correction terms in $e_3$ are grouped with $e_1$ instead of with $e_2$. The expression that is to be solved is therefore the following:

$$\exp(xe_1 + ye_2) = \exp(\beta_1 e_1 + \beta_3 e_3) \vee \exp(\gamma e_2). \tag{2}$$

Here, $x$ and $y$ are scalars; the three scalars $\beta_1, \beta_3$, and $\gamma$ are functions of $x$ and $y$ (there is no $\beta_2$ in this exercise). The above separation uniquely determines $\beta_1, \beta_3$, and $\gamma$: a related expression was given approximately in a previous paper.[12] These functions may be determined by a process of straightforward algebraic manipulation. Because this procedure and the ensuing result is unusual, some of the details are given. First, recall the direct expressions for the individual exponentials in the Clifford algebra[12]:

$$\exp(\beta_1 e_1 + \beta_3 e_3) = \cos\beta + [(\beta_1 e_1 + \beta_3 e_3)/\beta]\sin\beta,$$
$$\beta = \sqrt{\beta_1^2 + \beta_3^2}. \tag{3}$$

(In the above exponential, it does not matter which two out of the three quaternion basis elements appear in the exponent.)

Expression (2) may be expanded out, then separated into a scalar component and the coefficients of the elements $e_1$, $e_2$, and $e_3$, to obtain the following set of equations:

$$\cos z = \cos\beta\cos\gamma, \quad z = \sqrt{x^2 + y^2}, \tag{4a}$$
$$(x/z)\sin z = (\beta_1/\beta)\sin\beta\cos\gamma - (\beta_3/\beta)\sin\beta\sin\gamma, \tag{4b}$$
$$(y/z)\sin z = \cos\beta\sin\gamma, \tag{4c}$$
$$0 = (\beta_3/\beta)\sin\beta\cos\gamma + (\beta_1/\beta)\sin\beta\sin\gamma. \tag{4d}$$

Dividing (4c) by (4a) gives an expression for $\gamma$ as a function of $x$ and $y$:

$$\tan\gamma = (y/z)\tan z, \tag{5a}$$
$$\Rightarrow \gamma = \arctan(y\tan z/z). \tag{5b}$$

Equation (4a) is used to obtain $\beta$:

$$\cos\beta = \cos z/\cos\gamma \Rightarrow \beta = \arccos(\cos z/\cos\gamma). \tag{6}$$

Equation (4d) alternately gives the identity

$$\tan\gamma = -\beta_3/\beta_1, \tag{7}$$

which may be used to obtain $\cos\gamma$ and $\sin\gamma$ by elementary trigonometry:

$$\cos\gamma = \pm\beta_1/\beta, \quad \sin\gamma = \mp\beta_3/\beta. \tag{8}$$

Equations (8) trivially give $\beta_1$ and $\beta_3$ in terms of $\beta$ and $\gamma$ as

$$\beta_1 = \pm\beta\cos\gamma, \quad \beta_3 = \mp\beta\sin\gamma. \tag{9}$$

The functions $\cos\gamma$ and $\sin\gamma$ may be written in terms of $x$ and $y$ by utilizing Eq. (5a):

$$\cos\gamma = \pm z/\sqrt{x^2 + y^2\sec^2 z}, \tag{10a}$$
$$\sin\gamma = \pm y\tan z/\sqrt{x^2 + y^2\sec^2 z}. \tag{10b}$$

Finally, $\beta_1, \beta_3$, and $\gamma$ are obtained as functions of $x$ and $y$ by using (6), (9), and (10). The combination $P$ has been defined for convenience. The signs are determined by the behavior at the limit of small positive parameters:

$$\beta_1 = (\cos z/P)\arccos P,$$
$$\beta_3 = (-y\sin z/zP)\arccos P,$$
$$P = \sqrt{x^2\cos^2 z + y^2}/z, \quad z = \sqrt{x^2 + y^2},$$
$$\gamma = \arctan(y\tan z/z). \tag{11}$$

These functions give an exact result for the exponential identity (2). This is one of the few instances where an exponential identity of this type is given exactly.

We should note that an equivalent solution of Eqs. (4) can be obtained via Cramer's rule. That method has the advantage of *a priori* forcing the correct sign upon the result.

## III. THE EXPONENTIAL MAPPING IN THE DIHEDRAL ALGEBRA

Consider the exponential mapping in the dihedral algebra $N_1$.[14,15] This algebra is also known as the "generalized quaternions" and is defined by the three anticommuting elements $e_1, e_2, e_3$, where two of these have square equal to $+1$ while the third one has square equal to $-1$. The multiplication properties are given in Table II. (There are six possible definitions of $N_1$, all permutation-wise equivalent.)

The problem addressed is still the same, namely to separate the exponential (2) exactly in the dihedral algebra. Because of the change in metric, the situation is not as simple as in the case of the quaternion algebra. This comes about because the square of $xe_1 + ye_2$ has three possible branches, each necessitating a separate investigation. From the algebraic rules (Table II), the square of this element is always equal to

$$(xe_1 + ye_2) \vee (xe_1 + ye_2) = x^2 - y^2. \tag{12}$$

The three cases arise as this expression (12) is positive, negative, or zero. It is convenient to define the following positive, real quantities $\chi$ and $\psi$ as follows:

$$\text{(i)}\ \chi = \sqrt{x^2 - y^2}, \quad \text{when } |x| > |y|, \tag{13a}$$
$$\text{(ii)}\ \psi = \sqrt{y^2 - x^2}, \quad \text{when } |y| > |x|, \tag{13b}$$
$$\text{(iii)}\ y^2 - x^2 = 0, \quad \text{when } |y| = |x|. \tag{13c}$$

The exponentials in each respective case are given by the following expressions:

TABLE II. Multiplication table for the dihedral algebra.

| $\vee$ | $e_1$ | $e_2$ | $e_3$ |
|---|---|---|---|
| $e_1$ | 1 | $e_3$ | $e_2$ |
| $e_2$ | $-e_3$ | $-1$ | $e_1$ |
| $e_3$ | $-e_2$ | $-e_1$ | 1 |

(i) $\exp(xe_1 + ye_2) = \cosh\chi + [(xe_1 + ye_2)/\chi] \times \sinh\chi, \quad |x| > |y|,$ (14a)

(ii) $\exp(xe_1 + ye_2) = \cos\psi + [(xe_1 + ye_2)/\psi] \times \sin\psi, \quad |y| > |x|,$ (14b)

(iii) $\exp(xe_1 + ye_2) = 1 + x(e_1 \pm e_2), \quad |x| = |y|.$ (14c)

Case (i) is examined first. Following the same procedure as in the previous section, one obtains an analogous set of expressions after separating terms [note that expression (3) now changes to hyperbolic functions]:

$$\cosh\chi = \cosh\beta\cos\gamma, \tag{15a}$$

$$(x/\chi)\sinh\chi = (\beta_1/\beta)\sinh\beta\cos\gamma - (\beta_3/\beta)\sinh\beta\sin\gamma, \tag{15b}$$

$$(y/\chi)\sinh\chi = \cosh\beta\sin\gamma, \tag{15c}$$

$$0 = (\beta_3/\beta)\sinh\beta\cos\gamma + (\beta_1/\beta)\sinh\beta\sin\gamma. \tag{15d}$$

Proceeding as above, divide (15c) by (15a) to obtain the analogous expression to (5),

$$(y/\chi)\tanh\chi = \tan\gamma, \tag{16a}$$

$$\Rightarrow \gamma = \arctan(y\tanh\chi/\chi). \tag{16b}$$

Equation (15d) gives exactly the same relations as before: Eqs. (7)–(9). From Eq. (15a), one can solve for $\beta$ to obtain the analogous relation to (6),

$$\beta = \operatorname{arccosh}(\cosh\chi/\cos\gamma). \tag{17}$$

The expressions for $\cos\gamma$ and $\sin\gamma$ are obtained from $\tan\gamma$ (16a) as follows:

$$\cos\gamma = \pm\chi/\sqrt{x^2 - y^2\operatorname{sech}^2\chi}, \tag{18a}$$

$$\sin\gamma = \pm y\tanh\chi/\sqrt{x^2 - y^2\operatorname{sech}^2\chi}. \tag{18b}$$

Finally, the expressions for $\beta_1, \beta_3$, and $\gamma$ are obtained in this case from (17), (18), and (9):

Case (i)
$$\begin{aligned} &\beta_1 = (\cosh\chi/Q)\operatorname{arccosh} Q, \\ &\beta_3 = (-y\sinh\chi/\chi Q)\operatorname{arccosh} Q, \\ &Q = \sqrt{x^2\cosh^2\chi - y^2}/\chi, \quad \chi = \sqrt{x^2 - y^2}, \\ &\gamma = \arctan(y\tanh\chi/\chi). \end{aligned} \tag{19}$$

These equations (19) provide an exact solution for the separation (2) in the dihedral algebra, under the special condition $|x| > |y|$.

It is very easy to go through the same calculations for case (ii), when $|y| > |x|$. The results are the following:

Case (ii)
$$\begin{aligned} &\beta_1 = (\cos\psi/R)\operatorname{arccosh} R, \\ &\beta_3 = -(y\sin\psi/\psi R)\operatorname{arccosh} R, \\ &R = \sqrt{y^2 - x^2\cos^2\psi}/\psi, \quad \psi = \sqrt{y^2 - x^2}, \\ &\gamma = \arctan(y\tan\psi/\psi). \end{aligned} \tag{20}$$

Note that both functions $Q$ and $R$ are defined for all values of $x$ and $y$ in each respective case. The remaining case occurs when $|x| = |y|$. In this case, the solution is the simplest, and is given as follows (for simplicity, we have assumed $x = y > 0$):

Case (iii)
$$\begin{aligned} &\beta_1 = \operatorname{arccosh}\sqrt{1 + x^2}/\sqrt{1 + x^2}, \\ &\beta_3 = -x\operatorname{arccosh}\sqrt{1 + x^2}/\sqrt{1 + x^2}, \\ &\gamma = \arctan x. \end{aligned} \tag{21}$$

This completes the solution of the exponential identity (2) for all values of $x, y$, when the problem takes values in the dihedral algebra.

## IV. AN EXACT FORM OF THE BAKER–CAMPBELL–HAUSDORFF FORMULA

While the separation of the rotation from the boost in the Lorentz group requires the identity derived in the preceding section, the exponential mapping is usually described by means of the Baker–Campbell–Hausdorff formula.[4–11] We provide exact forms of this and a related formula when the elements take values in a Clifford algebra. The derivation is much simpler than for the formulas previously derived, and again involve only elementary functions. The exact identities which we wish to obtain are the following, with $\alpha = \Sigma_{i=1}^3 \alpha_i e_i$ and $\zeta = \Sigma_{i=1}^3 \zeta_i e_i$ to be determined:

$$\exp(xe_1) \vee \exp(ye_2) = \exp(\alpha), \tag{22a}$$

$$\exp(xe_1) \vee \exp(ye_2) \vee \exp(-xe_1) = \exp(\zeta). \tag{22b}$$

These are easy to calculate directly. As before, we treat the quaternion case first. Expanding (22a) in the quaternion algebra and separating gives the identities [here, $\alpha = (\alpha_1^2 + \alpha_2^2 + \alpha_3^2)^{1/2}$]:

$$\cos\alpha = \cos x\cos y, \tag{23a}$$

$$(\alpha_1/\alpha)\sin\alpha = \sin x\cos y, \tag{23b}$$

$$(\alpha_2/\alpha)\sin\alpha = \cos x\sin y, \tag{23c}$$

$$(\alpha_3/\alpha)\sin\alpha = \sin x\sin y. \tag{23d}$$

From (23a), the value for $\alpha$ is obtained:

$$\alpha = \arccos(\cos x\cos y). \tag{24}$$

By squaring (23b), (23c), and (23d) and adding, the coefficients are easily computed as

$$\begin{aligned} \alpha_1 &= \frac{\sin x\cos y\arccos(\cos x\cos y)}{(1 - \cos^2 x\cos^2 y)^{1/2}}, \\ \alpha_2 &= \frac{\cos x\sin y\arccos(\cos x\cos y)}{(1 - \cos^2 x\cos^2 y)^{1/2}}, \\ \alpha_3 &= \frac{\sin x\sin y\arccos(\cos x\cos y)}{(1 - \cos^2 x\cos^2 y)^{1/2}}. \end{aligned} \tag{25}$$

The coefficients $\alpha_i$ (25) substituted in (22a) give an exact form for the Baker–Campbell–Hausdorff formula in the quaternion algebra.

The second exact identity (22b) is also derived here in the same fashion. Expanding the exponentials and separating as above gives the following relations:

$$\cos\zeta = \cos y, \tag{26a}$$

$$\zeta_1 = 0, \tag{26b}$$

$$(\zeta_2/\zeta)\sin\zeta = \sin y\cos 2x, \tag{26c}$$

$$(\zeta_3/\zeta)\sin\zeta = \sin y\sin 2x. \tag{26d}$$

Equations (26a) and (26b) imply that $y = \zeta = (\zeta_2^2 + \zeta_3^2)^{1/2}$. This expansion (22b) therefore assumes the particularly simple exact form in the quaternion algebra:

$$\exp(xe_1) \vee \exp(ye_2) \vee \exp(-xe_1) = \exp[y\cos(2x)e_2 + y\sin(2x)e_3]. \tag{27}$$

It is instructive to compare this identity with the usual

result obtained from the series expansion. The following formula is given in Ref. 5, p. 372:

$$\exp(-A)\exp(B)\exp(A) = \exp\left(\sum_{n=0}^{\infty} \frac{\{BA^n\}}{n!}\right), \tag{28a}$$

$$\{BA^n\} = [\cdots[B,A],A],\cdots],A] \quad n \text{ times}. \tag{28b}$$

Setting $A = -xe_1$ and $B = ye_2$ in (28) gives a trigonometric series which is easily seen to be precisely the result obtained above (27);

$$\begin{aligned}\sum_{n=0}^{\infty} \frac{1}{n!}\{(ye_2)(-xe_1)^n\} &= y\sum_{n=0}^{\infty}(-1)^n \frac{x^n 2^n}{n!} e_2 e_1^n \\ &= y\sum_{n=0}^{\infty}\frac{(2x)^{2n}}{(2n)!}(-1)^n e_2 \\ &\quad + y\sum_{n=0}^{\infty}\frac{(2x)^{2n+1}}{(2n+1)!}(-1)^n e_3 \\ &= y\cos(2x)e_2 + y\sin(2x)e_3. \end{aligned} \tag{29}$$

This completes our derivation of the Baker–Campbell–Hausdorff and the related formulas (22a) and (22b) in the quaternion algebra.

For completeness, we include the analogous results for cases in the dihedral algebra, where $e_1^2 = +1$, $e_2^2 = -1$, and $e_1e_2 = e_3$. The formulas for (22a) and (22b) are given for appropriate values of the parameters, as follows:

$$\exp(xe_1) \vee \exp(ye_2) = \exp(\alpha_1 e_1 + \alpha_2 e_2 + \alpha_3 e_3),$$

$$\alpha_1 = \frac{\operatorname{arccosh}(\cosh x \cos y)\sinh x \cos y}{(\cosh^2 x \cos^2 y - 1)^{1/2}},$$

$$\alpha_2 = \frac{\operatorname{arccosh}(\cosh x \cos y)\cosh x \sin y}{(\cosh^2 x \cos^2 y - 1)^{1/2}},$$

$$\alpha_3 = \frac{\operatorname{arccosh}(\cosh x \cos y)\sinh x \sin y}{(\cosh^2 x \cos^2 y - 1)^{1/2}}, \tag{30a}$$

$$\exp(xe_1) \vee \exp(ye_2) \vee \exp(-xe_1) = \exp[y\cosh(2x)e_2 + y\sinh(2x)e_3]. \tag{30b}$$

These are exact formulas for the expansions (22a) and (22b) in the dihedral algebra. Corresponding formulas for regions of the domain where Eqs. (30) are not defined are also easily derived.

In conclusion, we hope to have illustrated how the exponential mapping in Clifford algebras can be expressed in closed form using only elementary functions. The identities obtained are of importance in solutions of physical problems. In addition to the examples given here for the quaternion and dihedral algebras, we have indicated that one may do this directly for any larger Clifford algebra in general.


[1] J. D. Bjorken and S. D. Drell, *Relativistic Quantum Mechanics* (McGraw-Hill, New York, 1964).
[2] J. D. Jackson, *Classical Electrodynamics* (Wiley, New York, 1975), 2nd ed.
[3] S. Helgason, *Differential Geometry, Lie Groups, and Symmetric Spaces* (Academic, New York, 1978).
[4] V. S. Varadarajan, *Lie Groups, Lie Algebras, and Their Representations* (Prentice-Hall, Englewood Cliffs, NJ, 1974).
[5] W. Magnus, A. Karrass, and D. Solitar, *Combinatorial Group Theory* (Dover, New York, 1976), 2nd ed. (Note sign mistake in [[x,y]x] term, p. 368.)
[6] W. Magnus, Commun. Pure Appl. Math. **7**, 649 (1954).
[7] D. Finkelstein, Commun. Pure Appl. Math. **8**, 245 (1955).
[8] K. Goldberg, Duke Math. J. **23**, 13 (1956).
[9] E. Eriksen, J. Math. Phys. **9**, 790 (1968).
[10] D. Z. Djokovic, Math. Z. **143**, 209 (1975).
[11] W. Witschel, J. Phys. A **8**, 143 (1975).
[12] N. Salingaros, "Relativistic motion of a charged particle, the Lorentz group, and the Thomas precession," J. Math. Phys. **25**, 706 (1984).
[13] N. Salingaros, "Particle in an external electromagnetic field," Phys. Rev. D **28**, 2473 (1983).
[14] N. Salingaros, J. Math. Phys. **22**, 226 (1981).
[15] Y. Ilamed and N. Salingaros, J. Math. Phys. **22**, 2091 (1981).

# A construction relating Clifford algebras and Cayley–Dickson algebras

G. P. Wene
*Division of Mathematics, Computer Science and Systems Design, The University of Texas at San Antonio, San Antonio, Texas 78285*

(Received 19 July 1983; accepted for publication 23 December 1983)

A review of the applications of the octonions in physics is given. A construction is presented. Both the Cayley–Dickson algebras and the Clifford algebras arise naturally under this construction from the quaternion algebras. The mathematical properties of the algebras constructed are discussed.

PACS numbers: 02.10. + w, 03.65.Fd

## I. INTRODUCTION

Clifford algebras are being widely used as mathematical tools for the descriptions of physical phenomena. The applications range from the traditional problems to candidates for unified descriptions in nature (see Refs. 1–7).

It is well known that the field $\mathbb{R}$ of real numbers, the complex numbers $\mathbb{C}$, and the quaternion division ring $\mathbb{H}$ can all be generalized as real algebras in two different ways into higher dimensions. The generalizations are the Clifford algebras and the Cayley–Dickson algebras. The octonions $\mathcal{O}$ (or Cayley numbers) are an eight-dimensional Cayley–Dickson algebra. Faulkner and Ferrar[8] emphasize that all notions of exceptionality in algebra and in geometry are manifestations of one underlying structure; that is, nonclassical Lie algebras, nonassociative alternative algebras, nonspecial Jordan algebras, and non-de Sarguesian projective planes are all related, in one way or another, to the octonions. The purpose of this paper is to demonstrate that there is essentially one construction and that both families of algebras arise naturally under this construction from $\mathbb{H}$.

Current interest in the use of octonionic structures was initiated by Gürsey,[9] who noted that specializing one of the seven nonscalar Cayley units (to play the role of the imaginary unit) automatically achieves a rationale for $\mathrm{Su}(3)_c$. That idea of Gürsey and Günaydin[10] is elaborated on in the paper of Horwitz and Biedenharn[4] where the octonionic multiplication rules are reproduced in a minimal ideal of $A^{0,7}$ [$\mathrm{Su}(3)_c$ is identified as a subset of the automorphisms of $A^{0,7}$ that leave this minimal ideal invariant]. A Hilbert space over the real Clifford algebra $A^{0,7}$ is discussed as a potential model for the unification of weak, electromagnetic, and strong interactions utilizing the exceptional Lie groups.

Truini and Biedenharn,[11] using the concept of a Jordan pair, show that two copies of the $3\times 3$ matrix algebra with entries from the complex octonions ($\mathcal{M}_3^8$) can be used to define a quantum mechanics over the complex octonion plane having $\mathcal{C}_6 \oplus \mathcal{U}(1)$ as automorphism group. The group $\mathcal{C}_6 \oplus \mathcal{U}(1)$ is large enough to accommodate a color–flavor structure which is not ruled out by experimental evidence. In this quantum mechanics, the concept of observable becomes the concept of Hermitian pairs, the standard Wigner relation in which infinitesimal symmetry generators are directly related to observables is recovered, and time reversal can be implemented. This structure is neither a projective geometry nor a lattice.

Dixon[12] constructs a symplectic algebra based on $S \oplus \mathbb{R} \oplus \mathbb{C} \oplus \mathbb{H} \oplus \mathcal{O}$ as a setting for a strong–weak–electromagnetic unification theory ($S$ is the $3\times 3$ matrix algebra with entries from $\mathbb{R}$ and each of the algebras $\mathbb{R}$, $C$, $\mathbb{H}$, and $\mathcal{O}$ can be constructed from $\mathbb{R}$ via iteration of the Cayley–Dickson process). The algebra constructed is sufficiently large to accommodate four families of quarks and leptons, neutrinos with completely chargeless right-handed components and Dirac masses, and groups $\mathrm{Su}(3)_c$ and $\mathrm{Su}(2)$.

We repeat, for emphasis, that each of the division algebras $\mathbb{C}$, $\mathbb{H}$, and $\mathcal{O}$ can be constructed from the reals by iteration of the Cayley–Dickson process. Kugo and Townsend[13] link the existence and properties of each of these division algebras with the existence and properties of supersymmetric field theories in various space-time dimensions. An association of $\mathbb{C}$ and $\mathbb{H}$ with $N = 2$ and $N = 4$ supersymmetry, respectively, in two and three space-time dimensions, is also found in the construction of supersymmetric models with spins $\leqslant \frac{1}{2}$ by stochastic methods.[14] Let $D$ denote the dimension of space-time; $D = t + s$, where $t$ is the time dimension. Lukierski and Nowicki[15] derive a connection between division algebras and supersymmetry, associating $\mathbb{R}$, $\mathbb{C}$, $\mathbb{H}$ with $D = 3, 4, 5$, respectively ($t = 1$). The paper of Kugo and Townsend shows that spinors in space-time of dimension $D$ are associated for signature $s - t = 1, 2, 4, 8$ (and if $t = 0, 1, 2$) with the algebras $\mathbb{R}$, $\mathbb{C}$, $\mathbb{H}$, $\mathcal{O}$.

We end our current literature citations with Sudbery,[16] who shows that just as the exceptional Lie algebras can be described in terms of the octonions, so can the two exceptional Lie superalgebras $G(3)$ and $F(4)$.

Earlier papers have detailed octonionic formulations of field equations and quantum mechanics. Buoncristiani[17] showed that the classical Yang–Mills field equations may be written in a simple form utilizing the split octonion algebra. The split octonion algebra is also applied in Oliveira and Maia[18] to the study of relativistic wave equations in curved space. Günaydin, Piron, and Ruegg[19] showed that the non-de Sarguesian projective octonion plane of Moufang[20] can be orthocomplemented and thus interpreted as a quantum mechanics. Nahm[21] demonstrated that for the Yang–Mills theory derived from supergravity with 11 dimensions, upon compactification of seven dimensions to a sphere, there exists a natural octonionic generalization. A more detailed survey of the octonions is given in Sorgsepp and Lôhmus.[22]

We will give the construction of a class of algebras, ob-

 0022-2488/84/082351-03$02.50  

serve that the Clifford and the Cayley–Dickson algebras are contained in this class, and develop the mathematical properties of the algebras in that class.

## II. THE CONSTRUCTION

By a generalized quaternion algebra $Q$ we will mean an $\mathbb{R}$-algebra with a basis $\{e,i,j,k\ (=ij)\}$ where

$e$ is the identity,

$$i^2 = \gamma_1 e, \tag{1a}$$

$$j^2 = \gamma_2 e, \tag{1b}$$

$$ij = -ji = k, \tag{1c}$$

$\gamma_1, \gamma_2 = \pm 1$. We will begin with this class of algebras that includes $H$. For a detailed discussion of the quaternion algebras see Ref. 23.

We define two maps on the algebra $Q$. The first, denoted by $\sigma$, is an automorphism of $Q$ of period 2. That is,

$$(x^\sigma)^\sigma = x, \tag{2a}$$

$$(xy)^\sigma = x^\sigma y^\sigma \tag{2b}$$

for each $x, y$ in $Q$. The map $\sigma$ is determined by

$$e^\sigma = e, \quad i^\sigma = -i, \quad j^\sigma = -j, \quad k^\sigma = k. \tag{3}$$

The second map, denoted by *, is given by

$$e^* = e, \quad i^* = -i, \quad j^* = -j, \quad k^* = -k, \tag{4}$$

and satisfies

$$(x^*)^* = x, \tag{5}$$

$$(x^*)^\sigma = (x^\sigma)^* \tag{6}$$

for all $x$ in $Q$.

Let $\gamma_3 = \pm 1$. We construct an algebra $Q(\gamma_3)$ of dimension 8 over $R$ having $Q$ as a subalgebra [with the identity of $Q(\gamma_3)$ in $Q$] as follows: $Q(\gamma_3)$ consists of all ordered pairs $q = (q_1,q_2)$, $q_1$ in $Q$, addition and multiplication by scalars defined componentwise, and multiplication defined by

$$\begin{aligned}(q_1,q_2)(q_3,q_4) &= (q_1q_3 + \gamma_3(\lambda q_4^* q_2 + (1-\lambda)q_2q_4^\sigma),\ \lambda(q_2q_3^* + q_4q_1) \\ &\quad + (1-\lambda)(q_1q_4 + q_2q_3^\sigma))\end{aligned} \tag{7}$$

for all $q_i$ in $Q$ and some $\lambda$ in $\mathbb{R}$. Then $e = (e,0)$ is an identity element for $Q(\gamma_3)$, $Q' = \{(q,0) | q \in Q\}$ is a subalgebra of $Q(\gamma_3)$ isomorphic to $Q$, $v = (0,e)$ is an element of $Q(\gamma_3)$ such that $v^2 = \gamma_3 e$, and $Q(\gamma_3)$ is the vector space direct sum $Q(\gamma_3) = Q' + vQ'$ of the four-dimensional vector spaces $Q'$ and $vQ'$.

*Lemma 1:* The map $\sigma'$ induced in $Q(\gamma_3)$ by $\sigma$,

$$(q_1,q_2)^{\sigma'} = (q_1^\sigma, -q_2^\sigma) \tag{8}$$

for $q_1,q_2$ in $Q_1$, is an automorphism of period 2.

*Proof:* Compute each of the products $((q_1,q_2)(q_3,q_4))^{\sigma'}$ and $(q_1,q_2)^{\sigma'}(q_3,q_4)^{\sigma'}$ and use the fact that $(q^*)^\sigma = (q^\sigma)^*$ for all $q$ in $Q$.

*Lemma 2:* The map $*'$ induced in $Q(\gamma_3)$ by *,

$$(q_1,q_2)^{*'} = (q_1^*, -q_2) \tag{9}$$

for $q_1,q_2$ in $Q$, commutes with $\sigma'$. That is, $(q^{*'})^{\sigma'} = (q^{\sigma'})^{*'}$ for each $q$ in $Q(\gamma_3)$. Furthermore,

$$(q^{*'})^{*'} = q \tag{10}$$

for each $q$ in $Q(\gamma_3)$.

*Proof:*

$$((x,y)^{\sigma'})^{*'} = (x^\sigma, -y^\infty)^{*'} = (x^{\sigma *}, y^\sigma), \tag{11}$$

$$((x,y)^{*'})^{\sigma'} = (x^*, -y)^\sigma = (x^{*\sigma}, y^\sigma) = (x^{\sigma *}, y^\sigma). \tag{12}$$

The process can be repeated. After $n-2$ steps the iterated $2^n$-dimensional algebra will be denoted by $Q(\gamma_3,\gamma_4,\ldots,\gamma_n)$.

The algebras $Q(\gamma_3,\ldots,g_n)$ are not necessarily associative. If $\lambda = 1$, we obtain the Cayley–Dickson algebras upon iteration of this process (for further discussions of the algebras arising from the Cayley–Dickson process see Refs. 24–31). For $\lambda = 0$, we derive the Clifford algebras.

## III. MATHEMATICAL PROPERTIES

We have observed that the induced map $\sigma'$ will always be an automorphism of period 2.

The map * in $Q$ is an involution in $Q$ in the sense that

$$(q^*)^* = q, \tag{13}$$

$$(qp)^* = p^*q^* \tag{14}$$

for each $p,q$ in $Q$. We determine necessary and sufficient conditions that the induced map $*'$ be an involution in $Q(\gamma_3,\ldots,\gamma_n)$.

*Lemma 3:* The induced map $*'$ is an involution in $Q(\gamma_3)$ if and only if $\lambda = 1$.

*Proof:* Compute the first component of each of the products

$$((q_1,q_2)(q_3,q_4))^{*'} \text{ and } (q_3,q_4)^{*'}(q_1,q_2)^{*'}.$$

If * is an involution, we must have

$$(1-\lambda)q_3^\sigma{}^* q_2^* = (1-\lambda)q_3q_2^\sigma. \tag{15}$$

If $\lambda \neq 1$, then

$$q_3^\sigma{}^* q_2^* = (1-\lambda)q_3q_2^\sigma. \tag{16}$$

This must be true for all $q_2,q_3$ in $Q$. Let $q_2 = e$ and $q_3 = k$. Then

$$k\sigma^* = k \tag{17}$$

which is false. Hence $\lambda = 1$.

We can rephrase Lemma 3 as

*Lemma 3':* The induced map $*'$ is an involution if and only if $Q(\gamma_3,\ldots,\gamma_n)$ is a Cayley–Dickson algebra.

There is, of course, the involution in the Clifford algebras that Chevalley[32] calls the main antiautomorphism.

There is a nondegenerate, symmetric, quadratic form $q$ associated with $Q$,

$$q(x,y) = x_0y_0 - x_1y_1\gamma_1 - x_2y_2\gamma_2 - x_3y_3\gamma_1\gamma_3, \tag{18}$$

where $x = x_0e + x_ii + x_2j + x_3k$, $y = y_0e + y_ii + y_2j + y_3k$. We can extend $q$ to $Q(\gamma_3)$ via

$$q((u,v),(x,y)) = q(u,x) + \gamma_3 q(v,y). \tag{19}$$

A first step in the systematic study of the properties of any algebra is to determine how much associativity is present. An $\mathbb{R}$-algebra $A$ is called power associative in case the subalgebra $\mathbb{R}[x]$ generated by any element $x$ of $A$ is associative. An $\mathbb{R}$-algebra $A$ is power associative if and only if, for each $x$ in $A$,

$$x^2x = xx^2 \tag{20}$$

and

$$x^2x^2 = x(xx^2). \tag{21}$$

For a proof of this see Ref. 33. All associative algebras are power associative as are the Lie algebras, Jordan algebras, and the Cayley–Dickson algebras.

If an algebra $A$ is not power associative, it may satisfy the weaker flexible rule

$$(xy)x = x(yx) \tag{22}$$

for all $x, y$ in $A$. Note that a flexible algebra satisfies the associativity of cubes in (20). For a nice discussion of algebras satisfying this weaker form of associativity see Ref. 34.

We end with

**Theorem 4**: If we begin with a quaternion algebra $Q$ and a fixed value for $\lambda$, then for $n \geqslant 3$, the following statements are equivalent:

(i) $Q(\gamma_3,...,\gamma_n)$ is flexible.

(ii) $Q(\gamma_3,...,\gamma_n)$ has associativity of cubes.

(iii) $Q(\gamma_3,...,\gamma_n)$ is power associative.

(iv) $Q(\gamma_3,...,\gamma_n)$ is either a Cayley–Dickson algebra or a Clifford algebra.

*Proof:* We show that if $x^2x = xx^2$ for each $x$ in $Q(\gamma_3,...,\gamma_n)$, then $\lambda = 0$ or $\lambda = 1$. To do this we show that if $x \in Q(\gamma_3)$ and $x^3x = xx^2$, then $\lambda = 0$ or $\lambda = 1$.

Let $x = (k,e) \in Q(\gamma_3)$. The first component of $x^2x$ is

$$(k^2 + \gamma_3)k + \gamma_3 2(1 - \lambda)k. \tag{23}$$

The first component of $xx^2$ is

$$(k^2 + \gamma_3)k + \gamma_3 2(1 - \lambda)(1 - 2\lambda)k. \tag{24}$$

Equating (23) and (24) and simplifying, we get

$$(1 - \lambda)k = (1 - \lambda)(1 - 2\lambda)k. \tag{25}$$

If $1 - \lambda \neq 0$, we get $\lambda = 0$. If $1 - \lambda = 0$, we get $\lambda = 1$. Hence the algebras that we construct are either the Cayley–Dickson algebras or the Clifford algebras.

*Corollary 5:* $Q(\gamma_3,...,\gamma_n)$ is associative if and only if it is a Clifford algebra.

## IV. CONCLUSION

Each vector space may be made into one of infinitely many algebras; these algebras may have few properties in common. With a single vector space (physical properties described mathematically) it is possible to define two or more radically different algebras with products (operations or transformations) describing different phenomena.

For a fixed value of $\lambda$, the above construction, upon iteration, gives rise to a collection of algebras of increasing dimensions. We have shown that only for $\lambda = 0$ or $\lambda = 1$ do these algebras have any "reasonable" properties, and in those cases we have either the Clifford algebras or the Cayley–Dickson algebras.

There is another construction that lets us associate with each Cayley–Dickson algebra a Clifford algebra. If $Q$ is a Cayley–Dickson algebra, there is a nondegenerate symmetric quadratic form $q$ associated with the vector space $Q$, which can be used to define a nondegenerate, symmetric, bilinear form $f$ or $Q$. It is well known that $Q$ has a basis $\{e,u_1,u_2,...,u_n\}$ where $e$ is the identity of $Q$ and

$$u_iu_j = -u_ju_i, \quad i \neq j, \quad i,j = 1,2,...,n, \tag{26}$$

$$u_iu_i = a_ie, \quad i = 1,...,n, \tag{27}$$

where $a_i$ is a nonzero scalar and $f(u_i,u_j) = \delta_{ij}a_i$, $i,j = 1,2,...,n$.

(28)

As in Lang,[35] we can construct a Clifford algebra for the vector space spanned by $u_1,u_2,...,u_n$ (identifying the identity of the resulting Clifford algebra with the identity of $A$). Since the dimension of $Q$ is $2^n$, the dimension of the corresponding Clifford algebra will be $2^{(2^n-1)}$. For $n = 3$, we see that $Q$ is an octonion algebra and the Clifford algebra is generated by the seven elements $e_1,e_2,...,e_7$. This very natural identification of $\mathcal{O}$ with a subset of the Clifford algebra $A^{0,7}$ is exploited in Horwitz and Biedenharn.[4]


[1] Y. Ilamed and N. Salingaros, J. Math. Phys. **22**, 2091–2095 (1981).
[2] F. A. Doria, Lett. Nuovo Cimento **8**, 994–996 (1973).
[3] C. N. Ktorides, J. Math. Phys. **16**, 2123–2129 (1975).
[4] L. P. Horwitz and L. C. Biedenharn, J. Math. Phys. **20**, 269–298 (1979).
[5] R.Coquereaux, Phys. Lett. B **115**, 389–395 (1982).
[6] D. Finkelstein, Int. J. Theoret. Phys. **21**, 489–503 (1982).
[7] N. Salingaros, J. Math. Phys. **23**, 1–7 (1982); **23**, 1231 (1982).
[8] J. R. Faulkner and J. C. Ferrar, Bull. London Math. Soc. **9**, 1–35 (1977).
[9] F. Gürsey, *International Symposium on Mathematical Problems in Theoretical Physics*, edited by A. Araki (Springer, New York, 1975), cf. pp. 189ff.
[10] M. Günaydin and F. Gürsey, Lett. Nuovo Cimento **6**, 401 (1973); Phys. Rev. D **9**, 3387 (1974); J. Math. Phys. **14**, 1651 (1973); F. Gürsey, *Johns Hopkins Workshop on Current Problems in High Energy Particle Theory* (John Hopkins University, Baltimore, 1974), p. 15; M. Günaydin, J. Math. Phys. **17**, 1875 (1976).
[11] P. Truini and L. C. Biedenharn, J. Math. Phys. **23**, 1327–1345 (1982).
[12] G. Dixon, Phys. Rev. D **28**, 833–843 (1983).
[13] T. Kugo and P. Townsend, Nucl. Phys. B **221**, 357–380 (1983).
[14] G. Parisi and N. Sourlas, Nucl. Phys. B **206**, 321–332 (1982); S. Cecotti and L. Girardello, Ann. Phys. N.Y. **145**, 81–99 (1983).
[15] J. Lukierski and A. Nowicki, Fortschr. Phys. **30**, 75 (1982).
[16] A. Sudbery, J. Math. Phys. **24**, 1986–1988 (1983).
[17] A. M. Buoncristiani, J. Math. Phys. **14**, 849–854 (1973).
[18] C. G. Oliveira and M. D. Maia, J. Math. Phys. **20**, 923–930 (1979).
[19] M. Günaydin, C. Piron, and H. Ruegg, Commun. Math. Phys. **61**, 69–85 (1978).
[20] R. Moufang, Abh. Math. Sem. Univ. Hamburg **9**, 207–222 (1933).
[21] W. Nahm, CERN Preprint TH. 2489, 1978.
[22] L. Sorgsepp and J. Lôhmus, Hadronic J. **2**, 1388–1459 (1979). See also Ref. 30.
[23] A. A. Albert, *Structure of Algebras*, Am. Math. Soc. Colloquium Publication XXIV (AMS, Providence, RI, 1961).
[24] R. D. Schafer, Am. J. Math. **76**, 435–446 (1954).
[25] N. Jacobson, Circ. Math., Palermo, Rendicoti, Series 2 **7**(2) 55–80 (1958).
[26] H. Braun and M. Koecher, *Jordan-Algebren* (Springer-Verlag, Berlin, 1966).
[27] R. B. Brown, Pac. J. Math. **20**, 415–422 (1967).
[28] N. Bourbaki, *Algebra I* (Addison-Wesley, Reading, MA, 1974), Chaps. 1–3.
[29] S. C. Althoen and J. F. Weidner, Am. Math. Monthly **85**(5), 368–371 (1978).
[30] L. Sorgsepp and J. Löhmus, Hadronic J. **4**, 327–353 (1981). See also Ref. 23.
[31] G. P. Wene, J. Math. Phys. **25**, 414–416 (1984).
[32] C. C. Chevalley, *The Algebraic Theory of Spinors* (Columbia U.P., Morningside Heights, NY, 1954).
[33] A. A. Albert, Sum. Bras. Math. Vol. II, Fasc 2, 21–32 (1948).
[34] J. M. Osborn, Adv. Math. **8**, 163–369 (1972).
[35] S. Lang, *Algebra* (Addison-Wesley, Reading, MA, 1971), cf. pp. 367ff.

# Partially coherent states of the real symplectic group

J. Deenen and C. Quesne[a)]
*Physique Théorique et Mathématique CP 229, Université Libre de Bruxelles, Bd du Triomphe, B 1050 Brussels, Belgium*

(Received 7 February 1984; accepted for publication 30 March 1984)

In the present paper, we introduce partially coherent states for the positive discrete series irreducible representations $\langle \lambda_d + n/2,...,\lambda_1 + n/2\rangle$ of Sp(2$d$,$R$), encountered in physical applications. These states are characterized by both continuous and discrete labels. The latter specify the row of the irreducible representation $[\lambda_1\lambda_2\cdots\lambda_d]$ of the maximal compact subgroup U($d$), while the former parametrize an element of the factor space Sp(2$d$,$R$)/$H$, where $H$ is the Sp(2$d$,$R$) subgroup leaving the $[\lambda_1\lambda_2\cdots\lambda_d]$ representation space invariant. We consider three classes of partially coherent states, respectively, generalizing the Perelomov and Barut–Girardello coherent states, as well as some recently introduced intermediate coherent states. We prove that each family of partially coherent states forms an overcomplete set in the representation space of $\langle \lambda_d + n/2,...,\lambda_1 + n/2\rangle$, and study its generating function properties. We show that it leads to a representation of the Sp(2$d$,$R$) generators in the form of differential operator matrices. Finally, we relate the latter to a boson representation, namely a generalized Dyson representation in the cases of Perelomov and Barut-Girardello partially coherent states, and a generalized Holstein–Primakoff representation in that of the intermediate partially coherent states.

PACS numbers: 02.20. + b, 21.60.Fw, 03.65.Fd

## I. INTRODUCTION

The real symplectic group Sp(2$d$,$R$), being the group of linear canonical transformations in a 2$d$-dimensional phase space,[1] plays an outstanding role in many physical problems. It is an important component of the $d$-dimensional harmonic oscillator dynamical group,[1,2] the semidirect product group $N(d) \wedge$ Sp(2$d$,$R$), where $N(d)$ is the $d$-dimensional Heisenberg–Weyl group.[3] Then the widespread use of the harmonic approximation in physical problems accounts for many of the Sp(2$d$,$R$) occurrences.[4–6] In addition, the complementarity relationship between O($n$) and Sp(2$d$,$R$)[1,7] explains the appearance of the latter in connection with the O($n$) symmetry in various fields, such as O($n$)-invariant theories,[8] and collective models.[9–18] The Sp(2$d$,$R$) irreducible representations (irreps) encountered in all those physical applications are positive discrete series,[19,20] characterized by their lowest weight $\langle \lambda_d + n/2,...,\lambda_1 + n/2\rangle$, where $[\lambda_1\lambda_2\cdots\lambda_d]$ is a partition, and $n$ is an integer greater than or equal to 2$d$. In the present paper, we shall therefore restrict ourselves to such irreps.

It is well known that Glauber's standard coherent states (CS),[21] i.e., the CS associated with the Heisenberg–Weyl group $N(1)$, can be generalized to other Lie groups in various ways. Their most commonly used generalization, due to Perelomov,[22] is applicable to any Lie group. Oppositely, the earlier generalization, due to Barut and Girardello,[23] is not applicable to compact groups. It was actually proposed by these authors only for SO(2,1), and its locally isomorphic groups SU(1,1) Sl(2,$R$), and Sp(2,$R$). In a recent paper,[17] we extended their work to the irreps $\langle(\lambda + n/2)^d\rangle$ of Sp(2$d$,$R$), and also introduced a third class of CS, intermediate between the Perelomov and Barut–Girardello ones.

The purpose of the present paper is to extend these three classes of generalized CS to the discrete series irreps $\langle \lambda_d + n/2,...,\lambda_1 + n/2\rangle$ of Sp(2$d$,$R$). In trying to generalize the Barut–Girardello CS to such irreps, we have been led to introduce the concept of partially coherent states (PCS). Contrary to the CS, which are specified by some continuous (complex) indices, the PCS are characterized by a set of continuous (complex) indices as well as by some discrete labels. The latter specify the row of the irrep $[\lambda_1\lambda_2\cdots\lambda_d]$ of the maximal compact subgroup U($d$), while the former parametrize an element of the factor space Sp(2$d$,$R$)/$H$, where $H$ is the Sp(2$d$,$R$) subgroup leaving the irrep $[\lambda_1\lambda_2\cdots\lambda_d]$ representation space invariant.

In addition to providing an appropriate generalization of Barut–Girardello CS, the PCS have some remarkable properties, which by themselves account for their study interest, and will be reviewed in the present paper. Future developments in this field will include the study of possible connections with other works,[24] detailed applications to some physical problems, and an extension of the PCS concept to other physically relevant groups.

This paper is organized as follows. In Sec. II, our notation is summarized and some properties of the Sp(2$d$,$R$) discrete series irreps listed. In Secs. III–V, three classes of PCS are defined and their main properties reviewed. They are respectively termed Perelomov, Barut–Girardello, and intermediate PCS. In Sec. VI, the PCS representations of the Sp(2$d$,$R$) generators are studied. Finally, in Sec. VII, some relations between PCS and boson representations are established.

## II. THE Sp(2$d$,$R$) IRREPS $\langle \lambda_d + n/2,...,\lambda_1 + n/2\rangle$

The positive discrete series irreps $\langle \lambda_d + n/2,...,\lambda_1 + n/2\rangle$ of Sp(2$d$,$R$) can be realized in a space of boson states built from $dn$ boson creation operators $\eta_{is}$, $i = 1,...,d$,

[a)] Maître de recherches F.N.R.S.

 0022-2488/84/082354-13$02.50  

$s = 1,...,n$, provided $n$ is greater than or equal to $2d$ (see Ref. 1). In terms of these boson creation operators and the corresponding annihilation operators $\xi_{is} = (\eta_{is})^\dagger$, the Sp($2d,R$) generators are then expressed as

$$D^\dagger_{ij} = D^\dagger_{ji} = \sum_{s=1}^{n} \eta_{is}\eta_{js}, \quad 1 \leqslant i \leqslant j \leqslant d, \tag{2.1a}$$

$$D_{ij} = D_{ji} = \sum_{s=1}^{n} \xi_{is}\xi_{js}, \quad 1 \leqslant i \leqslant j \leqslant d, \tag{2.1b}$$

and

$$E_{ij} = \frac{1}{2}\sum_{s=1}^{n} (\eta_{is}\xi_{js} + \xi_{js}\eta_{is}) = C_{ij} + \frac{n}{2}\delta_{ij}, \quad i,j = 1,...,d, \tag{2.1c}$$

where

$$C_{ij} = \sum_{s=1}^{n} \eta_{is}\xi_{js}. \tag{2.2}$$

They satisfy the following commutation relations:

$$\begin{aligned}
&[E_{ij},E_{kl}] = \delta_{jk}E_{il} - \delta_{il}E_{kj},\\
&[E_{ij},D^\dagger_{kl}] = \delta_{jk}D^\dagger_{il} + \delta_{jl}D^\dagger_{ik},\\
&[E_{ij},D_{kl}] = -\delta_{ik}D_{jl} - \delta_{il}D_{jk},\\
&[D^\dagger_{ij},D^\dagger_{kl}] = [D_{ij},D_{kl}] = 0,\\
&[D_{ij},D^\dagger_{kl}] = \delta_{ik}E_{lj} + \delta_{il}E_{kj} + \delta_{jk}E_{li} + \delta_{jl}E_{ki},
\end{aligned} \tag{2.3}$$

from which we note that the operators $C_{ij}$, satisfying the same commutation relations as the operators $E_{ij}$, generate the maximal compact subgroup U($d$) of Sp($2d,R$).

A discrete basis of the irrep $\langle \lambda_d + n/2,...,\lambda_1 + n/2\rangle$ representation space can be easily built from the lowest weight state $|(\lambda)_{\min}\rangle$. By definition, the latter satisfies the following equations:

$$D_{ij}|(\lambda)_{\min}\rangle = 0, \quad i \leqslant j, \tag{2.4a}$$

$$E_{ij}|(\lambda)_{\min}\rangle = 0, \quad i > j, \tag{2.4b}$$

$$E_{ii}|(\lambda)_{\min}\rangle = (\lambda_i + n/2)|(\lambda)_{\min}\rangle, \quad i = 1,...,d. \tag{2.4c}$$

From Eqs. (2.1c), (2.4b), and (2.4c), we note that it is the lowest weight state of an irrep $[\lambda_1\lambda_2\cdots\lambda_d]$ of U($d$), so it can be characterized by the corresponding Gel'fand[25] pattern $(\lambda)_{\min}$.

The whole representation space can be generated from $|(\lambda)_{\min}\rangle$ by applying polynomials in the $D^\dagger_{ij}$, $E_{ij}$, and $D_{ij}$ generators. By using the commutation relations(2.3), it is always possible to write such polynomials in normal form, i.e., as

$$P(D^\dagger_{ij})P'(E_{ij})P''(D_{ij}), \tag{2.5}$$

where $P$, $P'$, and $P''$ are some polynomials in the indicated operators. As a consequence of Eq. (2.4a), $P''(D_{ij})$ gives rise to a constant, so that we may neglect it in Eq. (2.5). Moreover, the action of all the polynomials $P'(E_{ij})$ upon $|(\lambda)_{\min}\rangle$ generates the representation space of the U($d$) irrep $[\lambda_1\lambda_2\cdots\lambda_d]$, whose dimension will be denoted by $\Lambda$, and whose basis states $|(\lambda)\rangle$ can be characterized by the Gel'fand patterns $(\lambda)$. We can therefore obtain any state in the representation space of $\langle \lambda_d + n/2,...,\lambda_1 + n/2\rangle$ by linearly combining the states

$$|\mathbf{N};(\lambda)\rangle = F_{\mathbf{N}}(\mathbf{D}^\dagger)|(\lambda)\rangle, \tag{2.6}$$

where $\mathbf{D}^\dagger$ denotes the $d \times d$ matrix $\|D^\dagger_{ij}\|$, $F_{\mathbf{N}}(\mathbf{D}^\dagger)$ is defined by

$$F_{\mathbf{N}}(\mathbf{D}^\dagger) = \prod_{i<j} (N_{ij}!)^{-1/2}[(1+\delta_{ij})^{-1/2}D^\dagger_{ij}]^{N_{ij}}, \tag{2.7}$$

the quantum numbers $N_{ij}$, $1 \leqslant i \leqslant j \leqslant d$, run over all non-negative integers, and $(\lambda)$ over all Gel'fand patterns of $[\lambda_1\lambda_2\cdots\lambda_d]$.

We conclude that the states $|\mathbf{N};(\lambda)\rangle$ form a discrete basis of the irrep $\langle \lambda_d + n/2,...,\lambda_1 + n/2\rangle$ representation space. Although they are characterized by a given weight in U($d$), they do not belong to a definite U($d$) irrep. In Ref. 26, we did however show how to go from the states $|\mathbf{N};(\lambda)\rangle$ to states characterized by definite irreps of U($d$) and of its subgroups U($d-1$),U($d-2$),...,U(1). Therefore, for simplicity's sake, we may here restrict ourselves to the basis states $|\mathbf{N};(\lambda)\rangle$.

The states $|\mathbf{N};(\lambda)\rangle$ do not form an orthonormal set. Let $\mathbf{M}$ be their overlap matrix, i.e., the matrix whose elements are

$$M_{\mathbf{N}'(\lambda'),\mathbf{N}(\lambda)} = \langle \mathbf{N}';(\lambda')|\mathbf{N};(\lambda)\rangle. \tag{2.8}$$

We note that since

$$M_{\mathbf{N}'(\lambda'),\mathbf{N}(\lambda)} = 0 \quad \text{if} \sum_{i<j}(N'_{ij} - N_{ij}) \neq 0, \tag{2.9}$$

the infinite-dimensional matrix $\mathbf{M}$ is block diagonal, and the submatrices on the diagonal are finite dimensional. We may therefore consider its inverse $\mathbf{M}^{-1}$, and define the dual basis states (for which we use a round bracket instead of an angular one) by the usual relation

$$|\mathbf{N};(\lambda)) = \sum_{\mathbf{N}'(\lambda')} |\mathbf{N}';(\lambda')\rangle(\mathbf{M}^{-1})_{\mathbf{N}'(\lambda'),\mathbf{N}(\lambda)}. \tag{2.10}$$

The basis states $|\mathbf{N};(\lambda)\rangle$ and their dual ones $|\mathbf{N};(\lambda))$ form a biorthogonal system, i.e.,

$$(\mathbf{N}';(\lambda')|\mathbf{N};(\lambda)\rangle = \delta_{\mathbf{N}',\mathbf{N}}\delta_{(\lambda'),(\lambda)}, \tag{2.11}$$

where

$$\delta_{\mathbf{N}',\mathbf{N}} = \prod_{i<j} \delta_{N'_{ij},N_{ij}}. \tag{2.12}$$

They give rise to the following unity resolution relation:

$$\sum_{\mathbf{N}(\lambda)} |\mathbf{N};(\lambda))\langle \mathbf{N};(\lambda)| = I, \tag{2.13}$$

where $I$ denotes the unit operator in the representation space of $\langle \lambda_d + n/2,...,\lambda_1 + n/2\rangle$.

In addition to the two discrete, nonorthogonal basis $|\mathbf{N};(\lambda)\rangle$ and $|\mathbf{N};(\lambda))$, it is also possible to consider a discrete intermediate orthogonal basis, obtained from either of the former by the standard orthonormalization procedure. Its elements (for which we use a curly bracket instead of an angular one) are defined by

$$\begin{aligned}
|\mathbf{N};(\lambda)\} &= \sum_{\mathbf{N}'(\lambda')} |\mathbf{N}';(\lambda')\rangle(\mathbf{M}^{-1/2})_{\mathbf{N}'(\lambda'),\mathbf{N}(\lambda)}\\
&= \sum_{\mathbf{N}'(\lambda')} |\mathbf{N}';(\lambda'))(\mathbf{M}^{1/2})_{\mathbf{N}'(\lambda'),\mathbf{N}(\lambda)},
\end{aligned} \tag{2.14}$$

where $\mathbf{M}^{1/2}$ denotes the square root of the Hermitian, positive definite matrix $\mathbf{M}$, and $\mathbf{M}^{-1/2}$ the inverse of $\mathbf{M}^{1/2}$. The

states $|\mathbf{N};(\lambda)\}$ satisfy the two following relations:

$$\{\mathbf{N}';(\lambda')|\mathbf{N};(\lambda)\} = \delta_{\mathbf{N}',\mathbf{N}}\delta_{(\lambda'),(\lambda)}, \tag{2.15}$$

and

$$\sum_{\mathbf{N}(\lambda)} |\mathbf{N};(\lambda)\}\{\mathbf{N};(\lambda)| = I. \tag{2.16}$$

For subsequent purposes, it is convenient to introduce a matrix notation to denote either of the three discrete basis. If we enumerate the Gel'fand patterns $(\lambda)$, associated with $[\lambda_1\lambda_2\cdots\lambda_d]$, in a given order $(\lambda)_1,(\lambda)_2,\ldots,(\lambda)_\Lambda$, then the basis states corresponding to given values of $\mathbf{N}$ can be arranged in a row vector, and the corresponding bras in a column vector. For the basis $|\mathbf{N};(\lambda)\rangle$, for instance, we obtain in this way

$$\|\mathbf{N}\rangle\rangle = (|\mathbf{N};(\lambda)_1\rangle|\mathbf{N};(\lambda)_2\rangle\cdots|\mathbf{N};(\lambda)_\Lambda\rangle), \tag{2.17}$$

and

$$\langle\langle\mathbf{N}\| = \begin{pmatrix} \langle\mathbf{N};(\lambda)_1| \\ \langle\mathbf{N};(\lambda)_2| \\ \vdots \\ \langle\mathbf{N};(\lambda)_\Lambda| \end{pmatrix}. \tag{2.18}$$

Row vectors $\|\mathbf{N})\rangle$, $\|\mathbf{N}\}\}$, and column vectors $(\langle\mathbf{N}\|$, $\{\{\mathbf{N}\|$ are similarly defined for the basis $|\mathbf{N};(\lambda)\rangle$ and $|\mathbf{N};(\lambda)\}$, respectively. In this matrix notation, Eqs. (2.11), (2.13), (2.15), and (2.16) can be rewritten as

$$(\langle\mathbf{N}'\|\mathbf{N}\rangle\rangle = \{\{\mathbf{N}'\|\mathbf{N}\}\} = \delta_{\mathbf{N}',\mathbf{N}}\mathbb{I}, \tag{2.19}$$

and

$$\sum_{\mathbf{N}} \|\mathbf{N})\rangle\langle\langle\mathbf{N}\| = \sum_{\mathbf{N}} \|\mathbf{N}\}\}\{\{\mathbf{N}\| = I, \tag{2.20}$$

where $\mathbb{I}$ denotes the $\Lambda\times\Lambda$ unit matrix.

In the next three sections, we proceed to introduce three classes of PCS, which provide us with three alternative bases for the representation space of $\langle\lambda_d + n/2,\ldots,\lambda_1 + n/2\rangle$, and can each be associated with one of the discrete bases defined in the present section. We start with Perelomov PCS in the next section.

## III. PERELOMOV PARTIALLY COHERENT STATES

Following Perelomov,[22] generalized CS can be defined for the irrep $\langle\lambda_d + n/2,\ldots,\lambda_1 + n/2\rangle$ of Sp(2$d$,$R$) by acting with the operator representing an arbitrary group element upon a fixed vector $|\psi_0\rangle$ of the representation space. They are determined by the points of the coset space Sp(2$d$,$R$)/$H$, where $H$ is the stationary subgroup of $|\psi_0\rangle$. For the reference state $|\psi_0\rangle$, it is interesting to choose the irrep lowest weight state $|(\lambda)_{\min}\rangle$. When $\lambda_1,\lambda_2,\ldots,\lambda_d$ are all different, it is clear from Eq. (2.4) that the stationary subgroup $H$ of $|(\lambda)_{\min}\rangle$ is generated by the operators $D_{ij}$, $E_{ii}$, and $E_{ij}$ $(i>j)$. Perelomov CS can then be written as

$$\exp\left\{\sum_{i<j}(1+\delta_{ij})^{-1}u_{ij}^* D_{ij}^\dagger + \sum_{i<j}\bar{u}_{ij}E_{ij}\right\}|(\lambda)_{\min}\rangle, \tag{3.1a}$$

in terms of the complex variables $u_{ij}$ $(i\leqslant j)$, and $\bar{u}_{ij}$ $(i<j)$, which completely specify them. When some of the $\lambda_i$'s are equal, the algebra of $H$ also includes some raising generators $E_{ij}$ $(i<j)$. In the extreme case where all $\lambda_i$'s are equal, it contains all the generators $D_{ij}$ and $E_{ij}$, so that Perelomov CS can then be expressed as[17]

$$|\mathbf{u}\rangle = \exp\left\{\sum_{i<j}(1+\delta_{ij})^{-1}u_{ij}^* D_{ij}^\dagger\right\}|(\lambda)_{\min}\rangle$$

$$= \exp(\tfrac{1}{2}\operatorname{tr}\mathbf{u}^*\mathbf{D}^\dagger)|(\lambda)_{\min}\rangle, \tag{3.1b}$$

where $\mathbf{u}$ denotes the $d\times d$ symmetrical complex matrix, whose elements are $u_{ij} = u_{ji}$.

Let us now introduce the concept of PCS by slightly modifying definition (3.1). Instead of a single reference state $|(\lambda)_{\min}\rangle$, we use a whole subspace of the irrep $\langle\lambda_d + n/2,\ldots,\lambda_1 + n/2\rangle$ representation space, namely the representation space of the U($d$) subgroup irrep $[\lambda_1\lambda_2\cdots\lambda_d]$. In close analogy with Eq. (3.1), we then act upon the basis states $|(\lambda)\rangle$ of this subspace with the elements of the coset space Sp(2$d$,$R$)/$H$, where $H$ is now the Sp(2$d$,$R$) subgroup leaving the subspace invariant. Since $H$ is generated by the operators $D_{ij}$ and $E_{ij}$ (whatever the values of $\lambda_1,\lambda_2,\ldots,\lambda_d$ may be), the PCS are defined by the following relation:

$$|\mathbf{u};(\lambda)\rangle = \exp(\tfrac{1}{2}\operatorname{tr}\mathbf{u}^*\mathbf{D}^\dagger)|(\lambda)\rangle, \tag{3.2}$$

where $\mathbf{u}$ is again a $d\times d$ symmetrical complex matrix. In the following, we shall refer to the states $|\mathbf{u};(\lambda)\rangle$ as Perelomov PCS (PPCS).

From definition (3.2), it is clear that the PPCS are both labeled by the continuous indices $u_{ij} = u_{ji}$, and the discrete indices $(\lambda)$. The latter can only take a finite number of values, equal to the dimension $\Lambda$ of the U($d$) irrep $[\lambda_1\lambda_2\cdots\lambda_d]$. In analogy with Eqs. (2.17) and (2.18), the PPCS can be arranged in a row vector

$$\|\mathbf{u}\rangle\rangle = (|\mathbf{u};(\lambda)_1\rangle|\mathbf{u};(\lambda)_2\rangle\cdots|\mathbf{u};(\lambda)_\Lambda\rangle), \tag{3.3}$$

and the corresponding bras in a column vector

$$\langle\langle\mathbf{u}\| = \begin{pmatrix} \langle\mathbf{u};(\lambda)_1| \\ \langle\mathbf{u};(\lambda)_2| \\ \vdots \\ \langle\mathbf{u};(\lambda)_\Lambda| \end{pmatrix}. \tag{3.4}$$

We note that when $\lambda_1 = \lambda_2 = \cdots = \lambda_d$, $\Lambda$ is equal to one, so that the PPCS then reduce to the Perelomov CS. For arbitrary $\lambda_i$ values, the PPCS can therefore be considered as a generalization of the Perelomov CS corresponding to all $\lambda_i$'s equal.

Let us now review some properties of the PPCS. From their definition, it is clear that they are generating functions for the discrete basis states $|\mathbf{N};(\lambda)\rangle$, defined in Eq. (2.6). By expanding the exponential in Eq. (3.2), we indeed obtain

$$|\mathbf{u};(\lambda)\rangle = \sum_{\mathbf{N}} F_{\mathbf{N}}(\mathbf{u}^*)|\mathbf{N};(\lambda)\rangle, \tag{3.5a}$$

or

$$\|\mathbf{u}\rangle\rangle = \sum_{\mathbf{N}} F_{\mathbf{N}}(\mathbf{u}^*)\|\mathbf{N}\rangle\rangle, \tag{3.5b}$$

where the summation over $N_{11},N_{12},\ldots,N_{dd}$ runs over all nonnegative integers, and $F_{\mathbf{N}}(\mathbf{u}^*)$ is defined by Eq. (2.7) with $\mathbf{D}^\dagger$ replaced by $\mathbf{u}^*$. This property largely accounts for the usefulness of PPCS for practical purposes. It should be noted that, in contrast, no such simple property is available for the Perelomov CS, except in the case where $\lambda_1 = \lambda_2 = \cdots = \lambda_d$.

Let us next show that the PPCS form a nonorthogonal family of states. The overlap of two PPCS,

$$\langle \mathbf{u}';(\lambda')|\mathbf{u};(\lambda)\rangle$$
$$= \langle(\lambda')|\exp(\tfrac{1}{2}\operatorname{tr}\mathbf{u}'\mathbf{D})\exp(\tfrac{1}{2}\operatorname{tr}\mathbf{u}^*\mathbf{D}^\dagger)|(\lambda)\rangle, \tag{3.6}$$

can be calculated by rewriting the operator on the right-hand side in normally ordered form,

$$\exp(\tfrac{1}{2}\operatorname{tr}\mathbf{u}'\mathbf{D})\exp(\tfrac{1}{2}\operatorname{tr}\mathbf{u}^*\mathbf{D}^\dagger)$$
$$= \exp(\tfrac{1}{2}\operatorname{tr}\mathbf{a}\mathbf{D}^\dagger)\exp(\operatorname{tr}\mathbf{b}\mathbf{E})\exp(\tfrac{1}{2}\operatorname{tr}\mathbf{c}\mathbf{D}), \tag{3.7}$$

and by using Eq. (2.4a) with $|(\lambda)_{\min}\rangle$ replaced by $|(\lambda)\rangle$. In Appendix A, it is shown that the matrices **a**, **b**, and **c**, appearing on the right-hand side of Eq. (3.7), are given in terms of $\mathbf{u}'$ and $\mathbf{u}^*$ by the following relations:

$$\mathbf{a} = \mathbf{u}^*(\mathbf{I} - \mathbf{u}'\mathbf{u}^*)^{-1} = (\mathbf{I} - \mathbf{u}^*\mathbf{u}')^{-1}\mathbf{u}^*, \tag{3.8a}$$
$$\exp\mathbf{b} = (\mathbf{I} - \mathbf{u}'\mathbf{u}^*)^{-1}, \tag{3.8b}$$

and

$$\mathbf{c} = (\mathbf{I} - \mathbf{u}'\mathbf{u}^*)^{-1}\mathbf{u}' = \mathbf{u}'(\mathbf{I} - \mathbf{u}^*\mathbf{u}')^{-1}. \tag{3.8c}$$

Equation (3.6) is transformed into

$$\langle \mathbf{u}';(\lambda')|\mathbf{u};(\lambda)\rangle$$
$$= \langle(\lambda')|\exp(\operatorname{tr}\mathbf{b}\mathbf{E})|(\lambda)\rangle$$
$$= \exp((n/2)\operatorname{tr}\mathbf{b})\langle(\lambda')|\exp(\operatorname{tr}\mathbf{b}\mathbf{C})|(\lambda)\rangle, \tag{3.9}$$

where in the last step we used Eq. (2.1c).

Since the operator $\exp(\operatorname{tr}\mathbf{b}\mathbf{C})$ belongs to the $\mathrm{Gl}(d,C)$ subgroup of $\mathrm{Sp}(2d,C)$, and the Gel'fand states $|(\lambda)\rangle$ and $|(\lambda')\rangle$, associated with the irrep $[\lambda_1\lambda_2\cdots\lambda_d]$ of $\mathrm{U}(d)$, transform under the same irrep of $\mathrm{Gl}(d,C)$, the matrix element $\langle(\lambda')|\exp(\operatorname{tr}\mathbf{b}\mathbf{C})|(\lambda)\rangle$ is an element of a $\mathrm{Gl}(d,C)$ representation matrix corresponding to $[\lambda_1\lambda_2\cdots\lambda_d]$

$$\langle(\lambda')|\exp(\operatorname{tr}\mathbf{b}\mathbf{C})|(\lambda)\rangle = D^{[\lambda_1\cdots\lambda_d]}_{(\lambda')(\lambda)}(\exp\tilde{\mathbf{b}}). \tag{3.10}$$

Here $\tilde{\mathbf{b}}$ stands for the transpose of **b**. From Eq. (3.10) and the relation

$$\exp(\operatorname{tr}\mathbf{b}) = \det(\exp\mathbf{b}) = \det(\exp\tilde{\mathbf{b}}), \tag{3.11}$$

it results that the overlap (3.9) only depends upon the matrix $\exp\tilde{\mathbf{b}}$, or equivalently, using Eq. (3.8b), upon the matrix $\mathbf{I} - \mathbf{u}^*\mathbf{u}'$. Its final form is given by

$$\langle \mathbf{u}';(\lambda')|\mathbf{u};(\lambda)\rangle = [\det(\mathbf{I} - \mathbf{u}^*\mathbf{u}')]^{-n/2}$$
$$\times D^{[\lambda_1\cdots\lambda_d]}_{(\lambda')(\lambda)}([\mathbf{I} - \mathbf{u}^*\mathbf{u}']^{-1}). \tag{3.12a}$$

We note that, as expected, the overlap (3.12a) reduces for $\lambda_1 = \lambda_2 = \cdots = \lambda_d$ to the overlap of two Perelomov CS.[17] In matrix notation, Eq. (3.12a) can be written as

$$\hat{\mathbf{K}}(\mathbf{u}',\mathbf{u}^*) = \langle\langle \mathbf{u}'\|\mathbf{u}^*\rangle\rangle$$
$$= [\det(\mathbf{I} - \mathbf{u}^*\mathbf{u}')]^{-n/2}\mathbb{D}^{[\lambda_1\cdots\lambda_d]}([\mathbf{I} - \mathbf{u}^*\mathbf{u}']^{-1}), \tag{3.12b}$$

where

$$\mathbb{D}^{[\lambda_1\cdots\lambda_d]}(g) = \|D^{[\lambda_1\cdots\lambda_d]}_{(\lambda')(\lambda)}(g)\|. \tag{3.13}$$

The nonorthogonal family of PPCS can be used as a basis in the representation space of the irrep $\langle\lambda_d + n/2,\ldots,\lambda_1 + n/2\rangle$. To prove this property, it remains to show that the set is complete (as a matter of fact, it is overcomplete as proved hereafter). For such a purpose, we have to demonstrate that in the irrep $\langle\lambda_d + n/2,\ldots,\lambda_1 + n/2\rangle$ representation space it gives rise to a unity resolution relation. The latter takes a rather unusual form

$$\sum_{(\lambda')(\lambda')} \int |\mathbf{u};(\lambda')\rangle d\hat{\sigma}_{(\lambda')(\lambda)}(\mathbf{u})\langle\mathbf{u};(\lambda)| = I, \tag{3.14a}$$

or

$$\int \|\mathbf{u}\rangle\rangle \mathbf{d}\hat{\sigma}(\mathbf{u})\langle\langle\mathbf{u}\| = I, \tag{3.14b}$$

in the sense that it contains a $\Lambda \times \Lambda$ matrix measure

$$\mathbf{d}\hat{\sigma}(\mathbf{u}) = \|d\hat{\sigma}_{(\lambda')(\lambda)}(\mathbf{u})\|. \tag{3.15}$$

This measure is given by

$$\mathbf{d}\hat{\sigma}(\mathbf{u}) = \hat{\mathbf{f}}(\mathbf{u},\mathbf{u}^*)d\mathbf{u}\,d\mathbf{u}^*, \tag{3.16}$$

where

$$d\mathbf{u}\,d\mathbf{u}^* = \prod_{i<j} d\operatorname{Re}u_{ij}\,d\operatorname{Im}u_{ij}, \tag{3.17}$$

and

$$\hat{\mathbf{f}}(\mathbf{u},\mathbf{u}^*) = \hat{A}\,[\det(\mathbf{I} - \mathbf{u}^*\mathbf{u})]^{-d-1}\hat{\mathbb{K}}^{-1}(\mathbf{u},\mathbf{u}^*), \tag{3.18a}$$

or

$$\hat{f}_{(\lambda')(\lambda)}(\mathbf{u},\mathbf{u}^*) = \hat{A}\,[\det(\mathbf{I} - \mathbf{u}^*\mathbf{u})]^{n/2-d-1}$$
$$\times D^{[\lambda_1\cdots\lambda_d]}_{(\lambda')(\lambda)}(\mathbf{I} - \mathbf{u}^*\mathbf{u}). \tag{3.18b}$$

In Eq. (3.14), the integration takes place over the origin-centered unit ball. In Eq. (3.18), $\hat{A}$ is a normalization constant, determined by the condition

$$\Lambda^{-1}\int \operatorname{tr}\mathbf{d}\hat{\sigma}(\mathbf{u}) = 1, \tag{3.19}$$

i.e.,

$$\hat{A}^{-1} = \Lambda^{-1}\int d\mathbf{u}\,d\mathbf{u}^*[\det(\mathbf{I} - \mathbf{u}^*\mathbf{u})]^{n/2-d-1}$$
$$\times\chi^{[\lambda_1\cdots\lambda_d]}(\mathbf{I} - \mathbf{u}^*\mathbf{u}), \tag{3.20}$$

where $\chi^{[\lambda_1\cdots\lambda_d]}(\mathbf{I} - \mathbf{u}^*\mathbf{u}')$ is the character of $\mathbf{I} - \mathbf{u}^*\mathbf{u}'$ in the irrep $[\lambda_1\cdots\lambda_d]$ of $\mathrm{Gl}(d,C)$.

The proof of Eq. (3.14) is detailed in Appendix B. It consists in showing that the operator on the left-hand side of this equation commutes with all the $\mathrm{Sp}(2d,R)$ generators when $\mathbf{d}\hat{\sigma}(\mathbf{u})$ is given by Eqs. (3.16)–(3.18). From the irreducibility of the representation $\langle\lambda_d + n/2,\ldots,\lambda_1 + n/2\rangle$, and Schur's lemma, it then follows that the operator is a multiple of the unit operator in the representation space of $\langle\lambda_d + n/2,\ldots,\lambda_1 + n/2\rangle$. The multiplicative constant can be set to 1 by adjusting the normalization constant $\hat{A}$ in accordance with Eq. (3.19).

In the PPCS representation, any state $|\psi\rangle$ in the representation space of $\langle\lambda_d + n/2,\ldots,\lambda_1 + n/2\rangle$ is represented by a column vector

$$\langle\langle\mathbf{u}\|\psi\rangle = \begin{pmatrix} \langle\mathbf{u};(\lambda)_1|\psi\rangle \\ \langle\mathbf{u};(\lambda)_2|\psi\rangle \\ \vdots \\ \langle\mathbf{u};(\lambda)_\Lambda|\psi\rangle \end{pmatrix}. \tag{3.21}$$

From Eq. (3.14), its expansion in terms of PPCS is given by

$$|\psi\rangle = \int \|\mathbf{u}\rangle\rangle \mathbf{d}\hat{\sigma}(\mathbf{u})\langle\langle\mathbf{u}\|\psi\rangle. \tag{3.22}$$

The PPCS overlap matrix $\hat{\mathbb{K}}(\mathbf{u}',\mathbf{u}^*)$ therefore acts as a matrix reproducing kernel since

$$\langle\langle \mathbf{u}||\psi\rangle = \int \hat{\mathbb{K}}(\mathbf{u},\mathbf{u}'^*)\mathrm{d}\hat{\sigma}(\mathbf{u}')\langle\langle \mathbf{u}'||\psi\rangle. \tag{3.23}$$

It can also be easily seen that the PPCS are not linearly independent, hence they form an overcomplete set, as from Eq. (3.22)

$$\|\mathbf{u}\rangle\rangle = \int \|\mathbf{u}'\rangle\rangle \mathrm{d}\hat{\sigma}(\mathbf{u}')\hat{\mathbb{K}}(\mathbf{u}',\mathbf{u}^*). \tag{3.24}$$

It should therefore be possible to extract complete subsets from the set of PPCS.

In the next section, we shall introduce a second class of PCS, generalizing the Barut–Girardello CS,[17,23] and associated with the discrete dual basis states $|\mathbf{N};(\lambda)\rangle$.

## IV. BARUT–GIRARDELLO PARTIALLY COHERENT STATES

Barut–Girardello CS for the irrep $\langle(\lambda + n/2)^d\rangle$ of Sp($2d,R$) are defined as the common eigenstates $|\mathbf{w}\rangle$ of the set of commuting, non-Hermitian operators $D_{ij}$, $1\leqslant i\leqslant j\leqslant d$, corresponding to the complex eigenvalues $w_{ij}^*$,

$$D_{ij}|\mathbf{w}\rangle = w_{ij}^*|\mathbf{w}\rangle, \quad 1\leqslant i\leqslant j\leqslant d. \tag{4.1}$$

Here $\mathbf{w}$ denotes the $d\times d$ symmetrical complex matrix whose elements are $w_{ij} = w_{ji}$. As shown in Ref. 17, such states do exist for any complex values of $w_{ij}$, and are unique up to a normalization factor.

Let us now turn to an arbitrary irrep $\langle\lambda_d + n/2,\ldots,\lambda_1 + n/2\rangle$, and ask whether the system of equations (4.1) has a solution in this case. To answer this question, let us consider the representation of both sides of Eq. (4.1) in the PPCS basis $|\mathbf{u};(\lambda')\rangle$,

$$\langle\mathbf{u};(\lambda')|D_{ij}|\mathbf{w}\rangle = w_{ij}^*\langle\mathbf{u};(\lambda')|\mathbf{w}\rangle, \quad 1\leqslant i\leqslant j\leqslant d. \tag{4.2}$$

From definition (3.2) of PPCS, the left-hand side of Eq. (4.2) can be transformed as follows:

$$\begin{aligned}\langle\mathbf{u};(\lambda')|D_{ij}|\mathbf{w}\rangle &= \langle(\lambda')|\exp(\tfrac{1}{2}\operatorname{tr}\mathbf{uD})D_{ij}|\mathbf{w}\rangle \\ &= \Delta_{u_{ij}}\langle(\lambda')|\exp(\tfrac{1}{2}\operatorname{tr}\mathbf{uD})|\mathbf{w}\rangle,\end{aligned} \tag{4.3}$$

where, in the last step, we have introduced the differential operator

$$\Delta_{u_{ij}} = (1+\delta_{ij})\frac{\partial}{\partial u_{ij}}. \tag{4.4}$$

Equation (4.2) is therefore equivalent to the following system of $\nu = d(d+1)/2$ independent first-order partial differential equations for $\langle\mathbf{u};(\lambda')|\mathbf{w}\rangle$,

$$\Delta_{u_{ij}}\langle\mathbf{u};(\lambda')|\mathbf{w}\rangle = \mathbf{w}_{ij}^*\langle\mathbf{u};(\lambda')|\mathbf{w}\rangle, \quad 1\leqslant i\leqslant j\leqslant d. \tag{4.5}$$

In addition to the trivial solution

$$\langle\mathbf{u};(\lambda')|\mathbf{w}\rangle = 0, \tag{4.6}$$

this system also admits the solution

$$\langle\mathbf{u};(\lambda')|\mathbf{w}\rangle = G_{(\lambda')}(\mathbf{w}^*)\exp(\tfrac{1}{2}\operatorname{tr}\mathbf{uw}^*), \tag{4.7}$$

where $G_{(\lambda')}(\mathbf{w}^*)$ is an arbitrary, nonidentically zero function in $\mathbf{w}^*$.

When $(\lambda')$ runs over the set $(\lambda)_1,(\lambda)_2,\ldots,(\lambda)_\Lambda$, we obtain $\Lambda$ systems (4.5), whose collection is equivalent to Eq. (4.1). Each one of them admits both solutions (4.6) and (4.7). If we choose the nontrivial solution for one system, and the trivial solution for the remaining ones, then we get $\Lambda$ independent solutions for the collection of $\Lambda$ systems, according to which of the latter gives rise to the nontrivial solution. These $\Lambda$ independent solutions may be specified by the Gel'fand pattern $(\lambda)$ associated with the nontrivial solution, and denoted by $|\mathbf{w};(\lambda)\rangle$. Equation (4.1) then becomes

$$D_{ij}|\mathbf{w};(\lambda)\rangle = w_{ij}^*|\mathbf{w};(\lambda)\rangle, \quad 1\leqslant i\leqslant j\leqslant d, \tag{4.8}$$

where

$$\langle\mathbf{u};(\lambda')|\mathbf{w};(\lambda)\rangle = \delta_{(\lambda'),(\lambda)}G_{(\lambda)}(\mathbf{w}^*)\exp(\tfrac{1}{2}\operatorname{tr}\mathbf{uw}^*). \tag{4.9}$$

It remains to choose the normalization of $|\mathbf{w};(\lambda)\rangle$, and thereby the value of the so far arbitrary function $G_{(\lambda)}(\mathbf{w}^*)$. Let us impose the following condition:

$$\langle(\lambda')|\mathbf{w};(\lambda)\rangle = \delta_{(\lambda'),(\lambda)}. \tag{4.10}$$

Since the PPCS $|\mathbf{0};(\lambda')\rangle$, whose parameters $u_{ij}$ are all equal to zero, reduce to the Gel'fand states $|(\lambda')\rangle$ of the irrep $[\lambda_1\lambda_2\cdots\lambda_d]$, Eq. (4.10) imposes that for any $(\lambda)$,

$$G_{(\lambda)}(\mathbf{w}^*) = 1, \tag{4.11}$$

so that Eq. (4.9) becomes

$$\langle\mathbf{u};(\lambda')|\mathbf{w};(\lambda)\rangle = \delta_{(\lambda'),(\lambda)}\exp(\tfrac{1}{2}\operatorname{tr}\mathbf{uw}^*). \tag{4.12}$$

We have therefore proved that for any complex values of the set of parameters $w_{ij} = w_{ji}$, the system of equations (4.1) has [up to some normalization factors we choose in accordance with Eq. (4.10)] $\Lambda$ independent solutions, which can be written in the PPCS basis as

$$\begin{aligned}|\mathbf{w};(\lambda)\rangle = \sum_{(\lambda')}\int &\exp(\tfrac{1}{2}\operatorname{tr}\mathbf{uw}^*) \\ &\times|\mathbf{u};(\lambda')\rangle d\hat{\sigma}_{(\lambda')(\lambda)}(\mathbf{u}),\end{aligned} \tag{4.13}$$

by making use of Eq. (3.14a). The states $|\mathbf{w};(\lambda)\rangle$, being both labeled by the continuous indices $w_{ij} = w_{ji}$, and the discrete indices $(\lambda)$, are PCS. In the following, they will be referred to as Barut–Girardello PCS (BGPCS).

Before proceeding, it is worth mentioning that, in contrast with Perelomov CS, Barut–Girardello CS cannot be defined for an irrep $\langle\lambda_d + n/2,\ldots,\lambda_1 + n/2\rangle$ for which all $\lambda_i$'s are not equal. To completely specify them, we should have indeed to diagonalize some extra Sp($2d,R$) generators in addition to the set of operators $D_{ij}$. However no generator $E_{ij}$ or $D_{ij}^\dagger$ commutes with the whole set of operators $D_{ij}$. The only way to get rid of this difficulty is to go from CS to PCS, since in the latter case we only need the set of operators $D_{ij}$.

For the BGPCS, we may introduce a matrix notation similar to that used for the PPCS. Let us denote by $\|\mathbf{w})$ the row vector

$$\|\mathbf{w}) = (|\mathbf{w};(\lambda)_1\rangle|\mathbf{w};(\lambda)_2\rangle\cdots|\mathbf{w};(\lambda)_\Lambda\rangle), \tag{4.14}$$

and by $(\mathbf{w}\|$ the Hermitian conjugate column vector

$$(\mathbf{w}\| = [\|\mathbf{w})]^\dagger. \tag{4.15}$$

From Eq. (4.12), we obtain that the overlap matrix of the PPCS with the BGPCS,

$$\langle\langle\mathbf{u}\|\mathbf{w})\rangle = \exp(\tfrac{1}{2}\operatorname{tr}\mathbf{uw}^*)\mathbb{I}, \tag{4.12'}$$

is a multiple of the $\Lambda\times\Lambda$ unit matrix $\mathbb{I}$. The multiplicative factor $\exp(\frac{1}{2}\operatorname{tr}\mathbf{uw}^*)$ is the reproducing kernel of a Bargmann Hilbert space[27] of analytic functions in the $\nu$ complex variables, $(1+\delta_{ij})^{-1/2}u_{ij}$, $1\leqslant i\leqslant j\leqslant d$, since it can be rewritten as

$$\exp(\tfrac{1}{2}\,\mathrm{tr}\,\mathbf{u}\mathbf{w}^*) = \exp\left\{\sum_{i<j}\left[(1+\delta_{ij})^{-1/2}u_{ij}\right]\times\left[(1+\delta_{ij})^{-1/2}w_{ij}^*\right]\right\}. \tag{4.16}$$

We also note that, in matrix notation, Eq. (4.13) assumes the following form:

$$\|\mathbf{w})) = \int \exp(\tfrac{1}{2}\,\mathrm{tr}\,\mathbf{u}\mathbf{w}^*)\|\mathbf{u}\rangle\rangle\,\mathbf{d}\hat{\sigma}(\mathbf{u}). \tag{4.13'}$$

The generating function properties of the BGPCS can be easily established by starting from the relation

$$\|\mathbf{w})) = \sum_{\mathbf{N}} \|\mathbf{N}))(\langle\langle\mathbf{N}\|\mathbf{w})), \tag{4.17}$$

resulting from Eq. (2.20). On the right-hand side of Eq. (4.17), $\langle\langle\mathbf{N}\|\mathbf{w})\rangle$ denotes the $\Lambda\times\Lambda$ matrix, whose elements are given by

$$\langle\mathbf{N};(\lambda')|\mathbf{w};(\lambda)) = \langle(\lambda')|F_{\mathbf{N}}(\mathbf{D})|\mathbf{w};(\lambda)). \tag{4.18}$$

By using Eqs. (4.8) and (4.10), it can be written as

$$\langle\langle\mathbf{N}\|\mathbf{w})\rangle = \mathrm{F}_{\mathbf{N}}(\mathbf{w}^*)\mathbb{I}, \tag{4.19}$$

where $F_{\mathbf{N}}(\mathbf{w}^*)$ is defined by Eq. (2.7) with $\mathbf{D}^\dagger$ replaced by $\mathbf{w}^*$. Equation (4.17) therefore assumes the form

$$\|\mathbf{w})) = \sum_{\mathbf{N}} F_{\mathbf{N}}(\mathbf{w}^*)\|\mathbf{N})), \tag{4.20}$$

from which it results that the BGPCS are generating functions for the discrete dual basis states $|\mathbf{N};(\lambda))$, with the same expansion coefficients as those appearing in the expansion (3.5) of the PPCS in terms of the discrete basis states $|\mathbf{N};(\lambda)\rangle$.

Equations (4.13) and (4.20), respectively, express the BGPCS as an expansion in terms of PPCS or discrete dual basis states. We may also ask for an expansion of the BGPCS in terms of the discrete basis states $|\mathbf{N};(\lambda)\rangle$, since this would give us their explicit form in terms of the $D_{ij}^\dagger$ generators, i.e., the analog of Eq. (3.2) for the PPCS. In Ref. 17, an equivalent question was considered for the Barut–Girardello CS associated with the irrep $\langle(\lambda+n/2)^d\rangle$. Since its solution was rather tedious, we shall omit it in the present case, and restrict ourselves to pointing out an important relationship between the BGPCS explicit form and their overlap matrix.

From the discussion following Eq. (2.5), it results that the BGPCS can be written as

$$|\mathbf{w};(\lambda)) = \sum_{(\lambda')} K_{(\lambda')(\lambda)}(\mathbf{D}^\dagger,\mathbf{w}^*)|(\lambda')\rangle, \tag{4.21}$$

where $K_{(\lambda')(\lambda)}(\mathbf{D}^\dagger,\mathbf{w}^*)$ is some analytic function in the $D_{ij}^\dagger$ generators. The overlap of two BGPCS is then given by

$$(\mathbf{w}';(\lambda')|\mathbf{w};(\lambda)) = \sum_{(\lambda'')}(\mathbf{w}';(\lambda')|K_{(\lambda'')(\lambda)}(\mathbf{D}^\dagger,\mathbf{w}^*)|(\lambda'')\rangle. \tag{4.22}$$

By using Eqs. (4.8) and (4.10), it can be transformed as follows:

$$(\mathbf{w}';(\lambda')|\mathbf{w};(\lambda)) = K_{(\lambda')(\lambda)}(\mathbf{w}',\mathbf{w}^*). \tag{4.23a}$$

We conclude that the same matrix $\mathbb{K} = \|K_{(\lambda')(\lambda)}\|$ gives us both the BGPCS explicit form (4.21) and their overlap matrix

$$((\mathbf{w}'\|\mathbf{w})) = \mathbb{K}(\mathbf{w}',\mathbf{w}^*). \tag{4.23b}$$

The family of BGPCS is a complete set in the representation space of the irrep $\langle\lambda_d+n/2,\ldots,\lambda_1+n/2\rangle$, and may therefore be used as a basis in this space. As usual, the proof of this property is based upon that of a unity resolution relation. In Appendix C, we demonstrate that the latter does exist and has a form similar to Eq. (3.14). In other words, it is possible to find a $\Lambda\times\Lambda$ matrix measure

$$\mathbf{d}\sigma(\mathbf{w}) = \|d\sigma_{(\lambda')(\lambda)}(\mathbf{w})\| = \mathbf{f}(\mathbf{w},\mathbf{w}^*)d\mathbf{w}\,d\mathbf{w}^*, \tag{4.24}$$

such that

$$\int \|\mathbf{w}))\mathbf{d}\sigma|\mathbf{w})((\mathbf{w}\| = I, \tag{4.25}$$

where the integration over each variable $w_{ij}$, $1\leqslant i\leqslant j\leqslant d$, takes place over the whole complex plane.

To prove Eq. (4.25) in Appendix C, we realize the representation space of $\langle\lambda_d+n/2,\ldots,\lambda_1+n/2\rangle$ as a subspace of a Bargmann Hilbert space of analytic functions.[27] The measure $\mathbf{d}\sigma(\mathbf{w})$ then directly derives from Bargmann measure by carrying out some appropriate integrations. Although this procedure enables us to show the existence of $\mathbf{d}\sigma(\mathbf{w})$, it is not suited for deriving its explicit form. In Ref. 17, we presented an alternative method to calculate the measure for the Barut–Girardello CS associated with the irrep $\langle(\lambda+n/2)^d\rangle$. Its extension to the case of BGPCS would enable us to calculate the weight functions $f_{(\lambda')(\lambda)}(\mathbf{w},\mathbf{w}^*)$. It will however not be considered in the present paper. The lack of explicit forms for the BGPCS and for their measure is indeed irrelevant to the most interesting application of BGPCS, namely the discrete basis state representation, which we shall now proceed to discuss.

As in the PPCS representation, any state $|\psi\rangle$ in the representation space of $\langle\lambda_d+n/2,\ldots,\lambda_1+n/2\rangle$ is represented by a column vector $((\mathbf{w}\|\psi\rangle$ in the BGPCS representation. Its expansion in terms of BGPCS is given by a relation similar to Eq. (3.22),

$$|\psi\rangle = \int \|\mathbf{w}))\mathbf{d}\sigma(\mathbf{w})((\mathbf{w}\|\psi\rangle. \tag{4.26}$$

When $|\psi\rangle$ is a discrete basis state $|\mathbf{N};(\lambda)\rangle$, it results from Eqs. (4.8) and (4.10) that

$$(\mathbf{w};(\lambda')|\mathbf{N};(\lambda)\rangle = \delta_{(\lambda')(\lambda)}F_{\mathbf{N}}(\mathbf{w}). \tag{4.27}$$

The BGPCS representation of $|\mathbf{N};(\lambda)\rangle$ is therefore a column vector with zeros everywhere, except in the row labeled by $(\lambda)$ where we have the function $F_{\mathbf{N}}(\mathbf{w})$,

$$((\mathbf{w}\|\mathbf{N};(\lambda)\rangle = \begin{pmatrix} 0 \\ \vdots \\ F_{\mathbf{N}}(\mathbf{w}) \\ 0 \\ \vdots \\ 0 \end{pmatrix}. \tag{4.28}$$

This simple result should be contrasted with the corresponding relation for $\langle\langle\mathbf{u}\|\mathbf{N};(\lambda)\rangle$, which makes use of the unknown discrete basis overlap matrix $\mathbf{M}$.

We conclude this section by pointing out the existence of relations similar to Eqs. (3.23) and (3.24), namely

$$(\{\mathbf{w}\|\psi\rangle = \int \mathbb{K}(\mathbf{w},\mathbf{w}'^*)\mathrm{d}\sigma(\mathbf{w}')(\{\mathbf{w}'\|\psi\rangle, \tag{4.29}$$

and

$$\|\mathbf{w}\}) = \int \|\mathbf{w}'\})\mathrm{d}\sigma(\mathbf{w}')\mathbb{K}(\mathbf{w}',\mathbf{w}^*). \tag{4.30}$$

They mean that the overlap matrix $\mathbb{K}(\mathbf{w}',\mathbf{w}^*)$, defined in Eq. (4.23b), acts as a reproducing kernel, and that the BGPCS are linearly dependent, hence form an overcomplete set.

In the next section, we proceed to introduce a third class of PCS, associated with the discrete intermediate orthogonal basis states $|\mathbf{N};(\lambda)\}$.

## V. INTERMEDIATE PARTIALLY COHERENT STATES

In Eqs. (3.5) and (4.20), we saw that the PPCS and the BGPCS are, respectively, generating functions for the discrete basis states $|\mathbf{N};(\lambda)\rangle$ and their dual states $|\mathbf{N};(\lambda)|$, with the same expansion coefficients $F_{\mathbf{N}}$. If we now consider the same type of expansions with $|\mathbf{N};(\lambda)\rangle$ or $|\mathbf{N};(\lambda)|$ replaced by the intermediate orthogonal basis states $|\mathbf{N};(\lambda)\}$, defined in Eq. (2.14), we obtain new states

$$|\mathbf{v};(\lambda)\} = \sum_{\mathbf{N}} F_{\mathbf{N}}(\mathbf{v}^*)|\mathbf{N};(\lambda)\}, \tag{5.1a}$$

still characterized by a $d \times d$ symmetrical complex matrix $\mathbf{v} = \|v_{ij}\|$, and a Gel'fand pattern $(\lambda)$ of the irrep $[\lambda_1\lambda_2\cdots\lambda_d]$. Hence they form a third class of PCS, that we shall call intermediate PCS (IPCS). As the other PCS, they can be arranged in a row vector, that we shall denote by $\|\mathbf{v}\}\}$. Definition (5.1a) can therefore be rewritten as

$$\|\mathbf{v}\}\} = \sum_{\mathbf{N}} F_{\mathbf{N}}(\mathbf{v}^*)\|\mathbf{N}\}\}. \tag{5.1b}$$

We now proceed to review some properties of the IPCS.

From definition (5.1), it results that the IPCS overlap matrix,

$$\check{\mathbb{K}}(\mathbf{v}',\mathbf{v}^*) = \{\{\mathbf{v}'\|\mathbf{v}\}\}, \tag{5.2}$$

is given by

$$\begin{aligned}\check{\mathbb{K}}(\mathbf{v}',\mathbf{v}^*) &= \left[\sum_{\mathbf{N}} F_{\mathbf{N}}(\mathbf{v}')F_{\mathbf{N}}(\mathbf{v}^*)\right]\mathbb{I} \\ &= \exp\left\{\sum_{i<j}\left[(1+\delta_{ij})^{-1/2}v'_{ij}\right]\right. \\ &\quad \left.\times\left[(1+\delta_{ij})^{-1/2}v^*_{ij}\right]\right\}\mathbb{I},\end{aligned} \tag{5.3}$$

or, using Eq. (4.16), by

$$\check{\mathbb{K}}(\mathbf{v}',\mathbf{v}^*) = \exp(\tfrac{1}{2}\,\mathrm{tr}\,\mathbf{v}'\mathbf{v}^*)\mathbb{I}. \tag{5.4}$$

It is therefore a multiple of the $\Lambda \times \Lambda$ unit matrix, the multiplicative factor being the reproducing kernel of a Bargmann Hilbert space of analytic functions in the $\nu$ complex variables $(1+\delta_{ij})^{-1/2}v_{ij}$, $1 \leqslant i \leqslant j \leqslant d$. By comparing Eq. (5.4) with Eq. (4.12'), we also obtain the following relation:

$$\{\{\mathbf{v}'\|\mathbf{v}\}\} = \langle(\mathbf{v}'\|\mathbf{v})\rangle. \tag{5.5}$$

It is now trivial to show that the set of states $|\mathbf{v};(\lambda)\}$ is a basis for the representation space of the irrep $\langle\lambda_d + n/2,\ldots,\lambda_1 + n/2\rangle$. For such purpose, we shall prove that they give rise to the following unity resolution relation:

$$\sum_{(\lambda)}\int |\mathbf{v};(\lambda)\}d\mu(\mathbf{v})\{\mathbf{v};(\lambda)| = I, \tag{5.6a}$$

or

$$\int \|\mathbf{v}\}\}\mathrm{d}\mu(\mathbf{v})\{\{\mathbf{v}\| = I, \tag{5.6b}$$

where

$$\mathrm{d}\mu(\mathbf{v}) = d\mu(\mathbf{v})\mathbb{I}, \quad d\mu(\mathbf{v}) = \prod_{i<j} d\mu\left[(1+\delta_{ij})^{-1/2}v_{ij}\right], \tag{5.7}$$

and $d\mu\left[(1+\delta_{ij})^{-1/2}v_{ij}\right]$ is the standard Bargmann measure[28]

$$d\mu(z) = \pi^{-1}\exp(-zz^*)d\,\mathrm{Re}\,z\,d\,\mathrm{Im}\,z, \tag{5.8}$$

corresponding to $z = (1+\delta_{ij})^{-1/2}v_{ij}$. In Eq. (5.6), the integration over each variable $(1+\delta_{ij})^{-1/2}v_{ij}$ takes place over the whole complex plane. Let us multiply both sides of Eq. (5.6b) from the left by the column vector $\{\{\mathbf{N}'\|$, and from the right by the row vector $\|\mathbf{N}\}\}$. By using the relation

$$\{\{\mathbf{v}\|\mathbf{N}\}\} = F_{\mathbf{N}}(\mathbf{v})\mathbb{I}, \tag{5.9}$$

resulting from Eqs. (5.1b) and (2.19), we obtain the equation

$$\left[\int F_{\mathbf{N}'}(\mathbf{v}^*)d\mu(\mathbf{v})F_{\mathbf{N}}(\mathbf{v})\right]\mathbb{I} = \delta_{\mathbf{N}',\mathbf{N}}\mathbb{I}, \tag{5.10}$$

whose validity for any $\mathbf{N}$ and $\mathbf{N}'$ will now be proved, thereby showing that of Eq. (5.6).

To this end, let us consider the boson states[16]

$$|\mathbf{N}] = \prod_{i<j}(N_{ij}!)^{-1/2}(a^\dagger_{ij})^{N_{ij}}|0] = F_{\mathbf{N}}(\bar{\mathbf{a}}^\dagger)|0], \tag{5.11}$$

built from $\nu$ independent boson creation operators $a^\dagger_{ij} = a^\dagger_{ji}$, $i,j = 1,\ldots,d$, acting upon the vacuum state $|0]$. The boson creation operators satisfy the following commutation relations with their corresponding annihilation operators $a_{ij} = a_{ji}$,

$$[a_{ij},a^\dagger_{kl}] = (1+\delta_{ij})^{-1}(\delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk}). \tag{5.12}$$

In Eq. (5.11), we have also introduced non-normalized boson creation operators $\bar{a}^\dagger_{ij} = \bar{a}^\dagger_{ji} = (1+\delta_{ij})^{1/2}a^\dagger_{ij}$, whose commutation relations with their corresponding annihilation operators $\bar{a}_{ij} = \bar{a}_{ji}$ assume the simpler form

$$[\bar{a}_{ij},\bar{a}^\dagger_{kl}] = \delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk}. \tag{5.13}$$

In Bargmann representation, the operators $\bar{a}^\dagger_{ij}$ and $\bar{a}_{ij}$ are, respectively, represented by the complex variables $v_{ij}$ and the corresponding differential operators

$$\Delta_{v_{ij}} = (1+\delta_{ij})\frac{\partial}{\partial v_{ij}}, \tag{5.14}$$

which satisfy the same commutation relations

$$[\Delta_{v_{ij}},v_{kl}] = \delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk} \tag{5.15}$$

as the operators themselves. The function $F_{\mathbf{N}}(\mathbf{v})$ is therefore the Bargmann representation of the boson state $|\mathbf{N}]$. Hence, when disregarding the unit matrix, Eq. (5.10) just expresses the orthonormality of boson states in Bargmann representation, and is therefore valid for any $\mathbf{N}$ and $\mathbf{N}'$. This completes the proof of Eq. (5.6).

As in the other PCS representations, any state $|\psi\rangle$ in the representation space of $\langle\lambda_d + n/2,\ldots,\lambda_1 + n/2\rangle$ is represented by a column vector $\{\{\mathbf{v}\|\psi\rangle$ in the IPCS representa-

tion. Relations similar to Eqs. (3.22)–(3.24), and (4.26), (4.29), and (4.30), can be easily established:

$$|\psi\rangle = \int d\mu(\mathbf{v})\|\mathbf{v}\}\{\{\mathbf{v}\|\psi\rangle, \tag{5.16}$$

$$\{\{\mathbf{v}\|\psi\rangle = \int d\mu(\mathbf{v}')\exp(\tfrac{1}{2}\,\mathrm{tr}\,\mathbf{v}\mathbf{v}'^{*})\{\{\mathbf{v}'\|\psi\rangle, \tag{5.17}$$

$$\|\mathbf{v}\}\} = \int d\mu(\mathbf{v}')\exp(\tfrac{1}{2}\,\mathrm{tr}\,\mathbf{v}'\mathbf{v}^{*})\|\mathbf{v}'\}\}. \tag{5.18}$$

Note the simple form of these equations, coming from the fact that in the present case both the overlap and measure matrices are multiples of the unit matrix.

In conclusion, we have shown in this section that the IPCS representation of Sp(2$d$,$R$) shares many properties with Bargmann representation of boson states. Before discussing this point in more detail, we proceed to study the PCS representations of the Sp(2$d$,$R$) generators in the next section.

## VI. PARTIALLY COHERENT STATE REPRESENTATIONS OF THE Sp(2$d$,$R$) GENERATORS

Let $X$ denote any operator acting in the representation space of the irrep $\langle\lambda_d + n/2,\ldots,\lambda_1 + n/2\rangle$. When it is applied to an arbitrary vector $|\psi\rangle$ of this space, it gives rise to a new vector $X|\psi\rangle$, whose representation is a column vector in any of the PCS representations discussed in the three preceding sections. In the PPCS representation, for instance, $X|\psi\rangle$ is represented by $\langle\langle\mathbf{u}\|X|\psi\rangle$. This column vector can be expressed in terms of the column vector $\langle\langle\mathbf{u}\|\psi\rangle$, representing $|\psi\rangle$, by acting upon the latter with a $\Lambda\times\Lambda$ matrix $\hat{\mathbf{X}}$ of differential operators $\hat{X}_{(\lambda),(\lambda')}$, i.e.,

$$\langle\langle\mathbf{u}\|X|\psi\rangle = \hat{\mathbf{X}}\langle\langle\mathbf{u}\|\psi\rangle, \tag{6.1a}$$

or

$$\langle\mathbf{u};(\lambda)|X|\psi\rangle = \sum_{(\lambda')}\hat{X}_{(\lambda),(\lambda')}\langle\mathbf{u};(\lambda')|\psi\rangle. \tag{6.1b}$$

The matrix $\hat{\mathbf{X}}$ is therefore the representation of $X$ in the PPCS basis. In the same way, let us denote by $\mathbf{X}$ and $\check{\mathbf{X}}$ the matrix representation of $X$ in the BGPCS and IPCS basis, respectively. They satisfy the following relations:

$$((\mathbf{w}\|X|\psi\rangle = \mathbb{X}((\mathbf{w}\|\psi\rangle, \tag{6.2}$$

and

$$\{\{\mathbf{v}\|X|\psi\rangle = \check{\mathbf{X}}\{\{\mathbf{v}\|\psi\rangle. \tag{6.3}$$

We now proceed to derive the matrix representations of the Sp(2$d$,$R$) generators in the PPCS and BGPCS basis. Starting with the former, we first note that the representation of $D_{ij}$ can be directly obtained from Eq. (4.3) by replacing there $|\mathbf{w}\rangle$ by $|\psi\rangle$. It is given by

$$\hat{\mathbb{D}}_{ij} = \Delta_{u_{ij}}\mathbb{I}, \tag{6.4}$$

where $\Delta_{u_{ij}}$ has been defined in Eq. (4.4). To obtain the representation of $D^{\dagger}_{ij}$ and $E_{ij}$, we proceed as in Ref. 17, and start from the identity

$$\begin{aligned}\langle\mathbf{u};(\lambda)|X|\psi\rangle &= \langle(\lambda)|[\exp(\tfrac{1}{2}\,\mathrm{tr}\,\mathbf{u}\mathbf{D})X\exp(-\tfrac{1}{2}\,\mathrm{tr}\,\mathbf{u}\mathbf{D})]\\ &\quad\times\exp(\tfrac{1}{2}\,\mathrm{tr}\,\mathbf{u}\mathbf{D})|\psi\rangle,\end{aligned} \tag{6.5}$$

where on the right-hand side we apply Baker–Campbell–Hausdorff formula

$$\exp(Y)X\exp(-Y) = X + \sum_{m=1}^{\infty}(m!)^{-1}[Y,[Y,\ldots,[Y,X]\cdots]]_m, \tag{6.6}$$

for $Y = \frac{1}{2}\,\mathrm{tr}\,\mathbf{u}\mathbf{D}$.

For the representation of $E_{ij}$, we get in this way the following relation:

$$\begin{aligned}\langle\mathbf{u};(\lambda)|E_{ij}|\psi\rangle &= \langle(\lambda)|[C_{ij} + (\mathbf{u}\mathbf{D} + (n/2)\mathbf{I})_{ij}]\\ &\quad\times\exp(\tfrac{1}{2}\,\mathrm{tr}\,\mathbf{u}\mathbf{D})|\psi\rangle\\ &= \langle(\lambda)|[C_{ij} + (\mathbf{u}\boldsymbol{\Delta}_u + (n/2)\mathbf{I})_{ij}]\\ &\quad\times\exp(\tfrac{1}{2}\,\mathrm{tr}\,\mathbf{u}\mathbf{D})|\psi\rangle,\end{aligned} \tag{6.7}$$

where in the last step we used Eq. (4.3) again. The U($d$) generator $C_{ij}$ transforms the bra $\langle(\lambda)|$ into a linear combination of bras $\langle(\lambda')|$, while the operator $(\mathbf{u}\boldsymbol{\Delta}_u + (n/2)\mathbf{I})_{ij}$ leaves it invariant. Equation (6.7) therefore becomes

$$\begin{aligned}\langle\mathbf{u};(\lambda)|E_{ij}|\psi\rangle &= \sum_{(\lambda')}[\langle(\lambda)|C_{ij}|(\lambda')\rangle\\ &\quad + (\mathbf{u}\boldsymbol{\Delta}_u + (n/2)\mathbf{I})_{ij}\delta_{(\lambda),(\lambda')}]\langle\mathbf{u};(\lambda')|\psi\rangle.\end{aligned} \tag{6.8}$$

Let us denote by $\mathring{\mathbb{C}}_{ij}$ the $\Lambda\times\Lambda$ matrix representing $C_{ij}$ in the U($d$) irrep $[\lambda_1\lambda_2\ldots\lambda_d]$, i.e.,

$$\mathring{\mathbb{C}}_{ij} = \|\langle(\lambda)|C_{ij}|(\lambda')\rangle\|. \tag{6.9}$$

This is a scalar matrix, entirely determined by the chosen Sp(2$d$,$R$) irrep. By comparing Eq. (6.8) with Eq. (6.1), we conclude that the matrix $\hat{\mathbb{E}}_{ij}$ representing $E_{ij}$ in the PPCS basis is given by

$$\hat{\mathbb{E}}_{ij} = \hat{\mathbb{C}}_{ij} + (n/2)\delta_{ij}\mathbb{I}, \tag{6.4'}$$

in terms of the matrix

$$\hat{\mathbb{C}}_{ij} = \mathring{\mathbb{C}}_{ij} + (\mathbf{u}\boldsymbol{\Delta}_u)_{ij}\mathbb{I}, \tag{6.10}$$

representing $C_{ij}$. For the representation of $D^{\dagger}_{ij}$, a similar calculation, detailed in Appendix D, leads to the following result:

$$\hat{\mathbb{D}}^{\dagger}_{ij} = \sum_k(\mathring{\mathbb{C}}_{ik}u_{kj} + u_{ik}\mathring{\mathbb{C}}_{jk}) + [(\mathbf{u}\boldsymbol{\Delta}_u + n - d - 1)\mathbf{u}]_{ij}\mathbb{I}. \tag{6.4''}$$

Equations (6.4), (6.4'), and (6.4'') can be rewritten in a more compact form by introducing the following notations. Let $\hat{\mathbb{D}}^{\dagger}$, $\hat{\mathbb{D}}$, $\hat{\mathbb{E}}$, $\hat{\mathbb{C}}$, and $\mathring{\mathbb{C}}$ denote the $d\times d$ matrices, whose elements are the $\Lambda\times\Lambda$ matrices $\hat{\mathbb{D}}^{\dagger}_{ij}$, $\hat{\mathbb{D}}_{ij}$, $\hat{\mathbb{E}}_{ij}$, $\hat{\mathbb{C}}_{ij}$, and $\mathring{\mathbb{C}}_{ij}$, respectively (hence they are $d\Lambda\times d\Lambda$ matrices), and let $\mathring{\bar{\mathbb{C}}}$ be the transpose of $\mathring{\mathbb{C}}$ when considered as a $d\times d$ matrix, i.e., $\mathring{\bar{\mathbb{C}}}_{ij} = \mathring{\mathbb{C}}_{ji}$. Then we obtain the following result:

$$\hat{\mathbb{D}}^{\dagger} = \mathring{\mathbb{C}}\mathbf{u} + \mathbf{u}\mathring{\bar{\mathbb{C}}} + (\mathbf{u}\boldsymbol{\Delta}_u + n - d - 1)\mathbf{u}\mathbb{I}, \tag{6.11a}$$

$$\hat{\mathbb{D}} = \boldsymbol{\Delta}_u\mathbb{I}, \tag{6.11b}$$

$$\hat{\mathbb{E}} = \hat{\mathbb{C}} + (n/2)\mathbf{I}\mathbb{I}, \quad \hat{\mathbb{C}} = \mathring{\mathbb{C}} + \mathbf{u}\boldsymbol{\Delta}_u\mathbb{I}, \tag{6.11c}$$

where matrices such as $\boldsymbol{\Delta}_u\mathbb{I}$ must be interpreted as the Kronecker product of a $d\times d$ matrix $\boldsymbol{\Delta}$ and a $\Lambda\times\Lambda$ matrix $\mathbb{I}$. Note that Eq. (6.11a) can be put into the more symmetrical form

$$\begin{aligned}\hat{\mathbb{D}}^{\dagger} &= \{\mathring{\mathbb{C}} + \tfrac{1}{2}[\mathbf{u}\boldsymbol{\Delta}_u + (n-d-1)\mathbf{I}]\mathbb{I}\}\,\mathbf{u}\\ &\quad + \mathbf{u}\{\mathring{\mathbb{C}} + \tfrac{1}{2}(\mathbf{u}\boldsymbol{\Delta}_u + n\mathbf{I})\mathbb{I}\},\end{aligned} \tag{6.12}$$

by using the relation

$$\sum_j u_{jl}\Delta_{u_{ik}} = \sum_l \Delta_{u_{kl}} u_{lj} - (d+1)\delta_{jk}, \tag{6.13}$$

coming from the commutation relation

$$[\Delta_{u_{ij}}, u_{kl}] = \delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk}. \tag{6.14}$$

We now turn to the matrix representation of the Sp(2$d$,$R$) generators in the BGPCS basis. From the definition (4.8) of the latter, and the Hermiticity properties of the Sp(2$d$,$R$) generators, we immediately obtain the representation of $D_{ij}^{\dagger}$ in the form

$$\mathbb{D}_{ij}^{\dagger} = w_{ij}\mathbb{I}. \tag{6.15}$$

To calculate the representation of the remaining generators, we take advantage of the facts that in Eq. (6.2) we may take a PPCS for $|\psi\rangle$, and that we already know the representation of the generators in the PPCS basis.

For the representation of $E_{ij}$, for instance, we start from the relation

$$(\mathbf{w};(\lambda)|E_{ij}|\mathbf{u};(\lambda')\rangle = \langle \mathbf{u};(\lambda')|E_{ji}|\mathbf{w};(\lambda)\rangle^*, \tag{6.16}$$

and use Eqs. (6.8) and (4.12) to obtain

$$\begin{aligned}&(\mathbf{w};(\lambda)|E_{ij}|\mathbf{u};(\lambda')\rangle \\ &\quad = [\langle(\lambda')|C_{ji}|(\lambda)\rangle^* \\ &\qquad + (\mathbf{u}^*\Delta_{u^*} + (n/2)\mathbf{I})_{ji}\delta_{(\lambda'),(\lambda)}]\exp(\tfrac{1}{2}\operatorname{tr}\mathbf{u}^*\mathbf{w}).\end{aligned} \tag{6.17}$$

By noting that

$$(\mathbf{u}^*\Delta_{u^*})_{ji}\exp(\tfrac{1}{2}\operatorname{tr}\mathbf{u}^*\mathbf{w}) = (\mathbf{w}\Delta_w)_{ij}\exp(\tfrac{1}{2}\operatorname{tr}\mathbf{u}^*\mathbf{w}), \tag{6.18}$$

where

$$\Delta_{w_{ij}} = (1+\delta_{ij})\frac{\partial}{\partial w_{ij}}, \tag{6.19}$$

Eq. (6.17) can be transformed into the following relation:

$$\begin{aligned}&(\mathbf{w};(\lambda)|E_{ij}|\mathbf{u};(\lambda')\rangle \\ &\quad = \sum_{(\lambda'')}[\langle(\lambda)|C_{ij}|(\lambda'')\rangle \\ &\qquad + (\mathbf{w}\Delta_w + (n/2)\mathbf{I})_{ij}\delta_{(\lambda),(\lambda'')}](\mathbf{w};(\lambda'')|\mathbf{u};(\lambda')\rangle.\end{aligned} \tag{6.20}$$

Hence the representation of $E_{ij}$ in the BGPCS basis is given by

$$\mathbb{E}_{ij} = \mathbb{C}_{ij} + (n/2)\delta_{ij}\mathbb{I}, \quad \mathbb{C}_{ij} = \mathring{\mathbb{C}}_{ij} + (\mathbf{w}\Delta_w)_{ij}\mathbb{I}. \tag{6.15'}$$

A similar calculation, detailed in Appendix D, leads to the following representation of $D_{ij}$:

$$\begin{aligned}\mathbb{D}_{ij} = &\sum_k (\Delta_{w_{ik}}\mathring{\mathbb{C}}_{kj} + \mathring{\mathbb{C}}_{ki}\Delta_{w_{kj}}) \\ &+ [\Delta_w(\mathbf{w}\Delta_w + n - d - 1)]_{ij}\mathbb{I}.\end{aligned} \tag{6.15''}$$

In a compact form, Eqs. (6.15), (6.15'), and (6.15'') can be rewritten as

$$\mathbb{D}^{\dagger} = \mathbf{w}\mathbb{I}, \tag{6.21a}$$

$$\mathbb{D} = \Delta_w\mathring{\mathbb{C}} + \mathring{\mathbb{C}}\Delta_w + \Delta_w(\mathbf{w}\Delta_w + n - d - 1)\mathbb{I}, \tag{6.21b}$$

$$\mathbb{E} = \mathbb{C} + (n/2)\mathbb{I}\mathbb{I}, \quad \mathbb{C} = \mathring{\mathbb{C}} + \mathbf{w}\Delta_w\mathbb{I}. \tag{6.21c}$$

Equation (6.21b) can be put into the more symmetrical form

$$\begin{aligned}\mathbb{D} = &\Delta_w\{\mathring{\mathbb{C}} + \tfrac{1}{2}[\mathbf{w}\Delta_w + (n-d-1)\mathbf{I}]\mathbb{I}\} \\ &+ \{\mathring{\mathbb{C}} + \tfrac{1}{2}(\mathbf{w}\Delta_w + n\mathbf{I})\mathbb{I}\}\Delta_w.\end{aligned} \tag{6.22}$$

There are striking similarities between Eqs. (6.11) and (6.21). One can go, for instance, from $\hat{\mathbb{D}}^{\dagger}$ in the PPCS representation to $\mathbb{D}$ in the BGPCS representation by transposing $\mathring{\mathbb{C}}$, and by replacing $\mathbf{u}$ and $\Delta_u$ by $\Delta_w$ and $\mathbf{w}$, respectively. The matrix elements of $\mathbb{D}$ are however second-order partial differential operators, whereas those of $\hat{\mathbb{D}}^{\dagger}$ are first-order ones. The PPCS representation of the Sp(2$d$,$R$) generators is therefore much more convenient for practical purposes.

To conclude this section, we would like to mention that the determination of the IPCS representation of the Sp(2$d$,$R$) generators is a very hard problem, which so far has remained unsolved. The reasons for these difficulties will become clearer in the next section, where we shall relate the PCS representations of Sp(2$d$,$R$) with its boson representations.

## VII. RELATIONS BETWEEN PARTIALLY COHERENT STATE AND BOSON REPRESENTATIONS OF Sp(2$d$,$R$)

In Refs. 16 and 17, we studied the relations between the CS representations of Sp(2$d$,$R$) associated with an irrep $\langle(\lambda + n/2)^d\rangle$, and its corresponding boson representations, i.e., both generalized Dyson[28] and Holstein–Primakoff[29] representations. The purpose of the present section is to extend such results to the PCS representations associated with the irrep $\langle\lambda_d + n/2,\ldots,\lambda_1 + n/2\rangle$.

Boson representations of Sp(2$d$,$R$) for the irrep $\langle(\lambda + n/2)^d\rangle$ were based upon a one-to-one mapping between discrete basis states of the irrep representation space and the boson states $|\mathbf{N}]$, defined in Eq. (5.11). In the case of the irrep $\langle\lambda_d + n/2,\ldots,\lambda_1 + n/2\rangle$, we note that all the discrete basis states $|\mathbf{N};(\lambda)\}$ [or $|\mathbf{N};(\lambda)\rangle$, or $|\mathbf{N};(\lambda)]$], corresponding to a given set of quantum numbers $\mathbf{N}$, and to all possible Gel'fand patterns $(\lambda)$ of the irrep $[\lambda_1\lambda_2\cdots\lambda_d]$, are mapped onto the same boson state $|\mathbf{N}]$. Whenever $\lambda_1,\lambda_2,\ldots,\lambda_d$ are not all equal, this mapping is therefore $\Lambda$-to-one, instead of one-to-one.

To recover the bijectiveness in the correspondence between discrete basis states and boson ones, we can use a procedure similar to that employed for the same purpose in studying the representation in quantum mechanics of nonbijective canonical transformations.[30,31] Let us enlarge the boson space by considering its Kronecker product with a $\Lambda$-dimensional space. The basis states of the latter are assumed to be labeled by a Gel'fand pattern $(\lambda)$ of the U($d$) irrep $[\lambda_1\lambda_2\cdots\lambda_d]$, and are denoted by $\chi_{(\lambda)}$. Their representation is a column vector with 1 in the row labeled by $(\lambda)$, and zeros elsewhere. The extra label $(\lambda)$ plays the same role as the ambiguity spin[30,31] in nonbijective canonical transformations.[32] The extended boson states can be written as

$$|\mathbf{N};(\lambda)] = |\mathbf{N}]\chi_{(\lambda)}, \tag{7.1}$$

or

$$|\mathbf{N};(\lambda)] = F_{\mathbf{N}}(\bar{\mathbf{a}}^{\dagger})|(\lambda)], \tag{7.2}$$

where $|(\lambda)]$ is defined by

$$|(\lambda)] = |0]\chi_{(\lambda)}. \tag{7.3}$$

They are clearly in one-to-one correspondence with the discrete basis states of the irrep $\langle\lambda_d + n/2,\ldots,\lambda_1 + n/2\rangle$ representation space, e.g.,

$$|\mathbf{N};(\lambda)\} \leftrightarrow |\mathbf{N};(\lambda)]. \tag{7.4}$$

It results from Eqs. (5.1) and (7.2) that, in the mapping

(7.4), the IPCS $|\mathbf{v};(\lambda)\}$ are mapped onto the states

$$|\mathbf{v};(\lambda)] = \sum_{\mathbf{N}} F_{\mathbf{N}}(\mathbf{v}^*)|\mathbf{N};(\lambda)] = \exp(\tfrac{1}{2}\,\mathrm{tr}\,\mathbf{v}^*\bar{\mathbf{a}}^\dagger)|(\lambda)], \tag{7.5}$$

which are the product of a Glauber standard CS[21] and an ambiguity spin state. The states (7.5), corresponding to $(\lambda) = (\lambda_1),(\lambda_2),\ldots,(\lambda)_A$, can be arranged in a row vector, that we denote by $\|\mathbf{v}]]$. It is obvious that in the mapping

$$\|\mathbf{v}\}\} \leftrightarrow \|\mathbf{v}]], \tag{7.6}$$

the form of Eqs. (5.4) and (5.6) is preserved, i.e.,

$$[[\mathbf{v}'\|\mathbf{v}]] = \exp(\tfrac{1}{2}\,\mathrm{tr}\,\mathbf{v}'\mathbf{v}^*)\mathbb{I}, \tag{7.7}$$

and

$$\int \|\mathbf{v}]]\,\mathrm{d}\mu(\mathbf{v})[\mathbf{v}\| = I, \tag{7.8}$$

where $\mathrm{d}\mu(\mathbf{v})$ is defined in Eq. (5.7). The set of IPCS therefore behaves in the same way as the set of standard CS with spin.

In basis (7.5), the extended boson states are represented by the column vectors

$$[[\mathbf{v}\|\mathbf{N};(\lambda)]] = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ F_{\mathbf{N}}(\mathbf{v}) \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \tag{7.9}$$

and the boson creation and annihilation operators $\bar{a}^\dagger_{ij}$ and $\bar{a}_{ij}$ by the matrices $v_{ij}\mathbb{I}$ and $\Delta_{v_{ij}}\mathbb{I}$, respectively. By comparing Eq. (5.9) with Eq. (7.9), it is clear that the right-hand side of the latter may be considered either as the representation of $|\mathbf{N};(\lambda)\}$ in the IPCS basis or as that of $|\mathbf{N};(\lambda)]$ in basis (7.5). In the same way, the IPCS representation of the Sp(2$d$,$R$) generators, $\breve{\mathbb{D}}^\dagger$, $\breve{\mathbb{D}}$, and $\breve{\mathbb{E}}$, may also be viewed as the representation of boson operators in basis (7.5). By replacing $v_{ij}$ and $\Delta_{v_{ij}}$ by $\bar{a}^\dagger_{ij}$ and $\bar{a}_{ij}$ in $\breve{\mathbb{D}}^\dagger$, $\breve{\mathbb{D}}$, and $\breve{\mathbb{E}}$, we therefore obtain a boson representation of Sp(2$d$,$R$). The latter is a generalized Holstein–Primakoff representation[29] since the mapping defined in Eq. (7.4) or Eq. (7.6), being unitary, preserves the Hermiticity properties of the Sp(2$d$,$R$) generators. The equivalence between the IPCS and Holstein–Primakoff representations accounts for the difficulties encountered in the derivation of an explicit form for $\breve{\mathbb{D}}^\dagger$, $\breve{\mathbb{D}}$, and $\breve{\mathbb{E}}$.

We can also establish a one-to-one mapping between the discrete basis states $|\mathbf{N};(\lambda)\rangle$ or their dual ones $|\mathbf{N};(\lambda)\}$, and the extended boson states $|\mathbf{N};(\lambda)]$. The PPCS or BGPCS are then mapped onto the extended Glauber CS. Hence by substituting $\bar{a}^\dagger_{ij}$ and $\bar{a}_{ij}$ for $u_{ij}$ and $\Delta_{u_{ij}}$ in $\hat{\mathbb{D}}^\dagger$, $\hat{\mathbb{D}}$, and $\hat{\mathbb{E}}$, or for $w_{ij}$ and $\Delta_{w_{ij}}$ in $\mathbb{D}^\dagger$, $\mathbb{D}$, and $\mathbb{E}$, we obtain two additional boson representations of Sp(2$d$,$R$). Both of them are generalized Dyson representations[28] since the Hermiticity properties of the generators are not preserved under the corresponding nonunitary mappings.


## ACKNOWLEDGMENT

We would like to thank M. Moshinsky for stimulating discussions on the Sp(2$d$,$R$) boson representations.


## APPENDIX A: NORMALLY ORDERED FORM OF THE OPERATOR $\exp(\frac{1}{2}\,\mathrm{tr}\,\mathbf{u}'\mathbf{D})\exp(\frac{1}{2}\,\mathrm{tr}\,\mathbf{u}^*\mathbf{D}^\dagger)$

The purpose of this appendix is to rewrite the operator $\exp(\frac{1}{2}\,\mathrm{tr}\,\mathbf{u}'\mathbf{D})\exp(\frac{1}{2}\,\mathrm{tr}\,\mathbf{u}^*\mathbf{D}^\dagger)$ in normally ordered form, as given in Eqs. (3.7) and (3.8). Since both $\exp(\frac{1}{2}\,\mathrm{tr}\,\mathbf{u}'\mathbf{D})$ and $\exp(\frac{1}{2}\,\mathrm{tr}\,\mathbf{u}^*\,\mathbf{D}^\dagger)$ are elements of the complex extension Sp(2$d$,$C$) of Sp(2$d$,$R$), this calculation can be performed in any faithful representation of Sp(2$d$,$C$).[33] We shall use here the defining $2d \times 2d$ matrix representation, given by[33]

$$\begin{aligned} D^\dagger_{ij} &= -E^{(2d)}_{i,d+j} - E^{(2d)}_{j,d+i}, \\ D_{ij} &= E^{(2d)}_{d+i,j} + E^{(2d)}_{d+j,i}, \\ E_{ij} &= E^{(2d)}_{ij} - E^{(2d)}_{d+j,d+i}. \end{aligned} \tag{A1}$$

The matrices $E^{(2d)}_{\alpha\beta}$, appearing on the right-hand side of Eq. (A1), are $2d \times 2d$ matrices with 1 in the $\alpha$th row and the $\beta$th column and zeros elsewhere.

In this representation, the operators $\frac{1}{2}\,\mathrm{tr}\,\mathbf{u}'\mathbf{D}$ and $\frac{1}{2}\,\mathrm{tr}\,\mathbf{u}^*\mathbf{D}^\dagger$ are represented by the matrices

$$\tfrac{1}{2}\,\mathrm{tr}\,\mathbf{u}'\mathbf{D} = \begin{pmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{u}' & \mathbf{0} \end{pmatrix}, \quad \tfrac{1}{2}\,\mathrm{tr}\,\mathbf{u}^*\mathbf{D}^\dagger = \begin{pmatrix} \mathbf{0} & -\mathbf{u}^* \\ \mathbf{0} & \mathbf{0} \end{pmatrix}, \tag{A2}$$

hence

$$\begin{aligned} \exp(\tfrac{1}{2}\,\mathrm{tr}\,\mathbf{u}'\,\mathbf{D}) &= \begin{pmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{u}' & \mathbf{I} \end{pmatrix}, \\ \exp(\tfrac{1}{2}\,\mathrm{tr}\,\mathbf{u}^*\mathbf{D}^\dagger) &= \begin{pmatrix} \mathbf{I} & -\mathbf{u}^* \\ \mathbf{0} & \mathbf{I} \end{pmatrix}. \end{aligned} \tag{A3}$$

To prove Eqs. (3.7) and (3.8), we also need the matrix representation of $\exp(\mathrm{tr}\,\mathbf{bE})$. The latter is given by

$$\exp(\mathrm{tr}\,\mathbf{bE}) = \begin{pmatrix} \exp\tilde{\mathbf{b}} & \mathbf{0} \\ \mathbf{0} & \exp(-\mathbf{b}) \end{pmatrix}, \tag{A4}$$

since from Eq. (A1) it follows that

$$\mathrm{tr}\,\mathbf{bE} = \begin{pmatrix} \tilde{\mathbf{b}} & \mathbf{0} \\ \mathbf{0} & -\mathbf{b} \end{pmatrix}, \tag{A5}$$

By using Eqs. (A3) and (A5), Eq. (3.7) is transformed into the relation

$$\begin{pmatrix} \mathbf{I} & -\mathbf{u}^* \\ \mathbf{u}' & \mathbf{I} - \mathbf{u}'\mathbf{u}^* \end{pmatrix} = \begin{pmatrix} \exp\tilde{\mathbf{b}} - \mathbf{a}\exp(-\mathbf{b})\mathbf{c} & -\mathbf{a}\exp(-\mathbf{b}) \\ \exp(-\mathbf{b})\mathbf{c} & \exp(-\mathbf{b}) \end{pmatrix}, \tag{A6}$$

equivalent to the following system of matrix equations:

$$\exp\tilde{\mathbf{b}} - \mathbf{a}\exp(-\mathbf{b})\mathbf{c} = \mathbf{I}, \tag{A7a}$$

$$\mathbf{a}\exp(-\mathbf{b}) = \mathbf{u}^*, \tag{A7b}$$

$$\exp(-\mathbf{b})\mathbf{c} = \mathbf{u}', \tag{A7c}$$

$$\exp(-\mathbf{b}) = \mathbf{I} - \mathbf{u}'\mathbf{u}^*. \tag{A7d}$$

The solution of Eqs. (A7b), (A7c), and (A7d) is given by Eq. (3.8). By introducing the latter into Eq. (A7a), we obtain an identity, thus completing the proof of Eqs. (3.7) and (3.8).

## APPENDIX B: UNITY RESOLUTION FOR THE SET OF PERELOMOV PARTIALLY COHERENT STATES

In this appendix, we wish to show that the operator

$$O = \int d\mathbf{u}\,d\mathbf{u}^*\|\mathbf{u}\rangle\rangle\,\hat{\mathrm{f}}(\mathbf{u},\mathbf{u}^*)\langle\langle\mathbf{u}\| \tag{B1}$$

commutes with the Sp(2$d$,$R$) generators, when the weight matrix $\hat{\mathbf{f}}(\mathbf{u},\mathbf{u}^*)$ is chosen as

$$\hat{\mathbf{f}}(\mathbf{u},\mathbf{u}^*) = \hat{A}\,(\det \mathbf{U})^{-d-1}\hat{\mathbb{K}}^{-1}(\mathbf{U}), \tag{B2}$$

where $\mathbf{U}$ is defined by

$$\mathbf{U} = \mathbf{I} - \mathbf{u}^*\mathbf{u}, \tag{B3}$$

and

$$\hat{\mathbb{K}}^{-1}(\mathbf{U}) = \hat{\mathbb{K}}^{-1}(\mathbf{u},\mathbf{u}^*) \tag{B4}$$

is the inverse of the overlap matrix $\hat{\mathbb{K}}(\mathbf{u},\mathbf{u}^*)$. To this end, we shall first establish the conditions imposed on $\hat{\mathbf{f}}(\mathbf{u},\mathbf{u}^*)$ by the equations

$$[E_{ij},O] = 0, \tag{B5a}$$

$$[D^{\dagger}_{ij},O] = 0, \tag{B5b}$$

and next prove that they are satisfied by the right-hand side of Eq. (B2). Equation (B5b) will then imply that $O$ also commutes with $D_{ij}$, since $\mathbf{U}$ is a Hermitian matrix.

The action of the Sp(2$d$,$R$) generators on $\|\mathbf{u}\rangle\rangle$, and $\langle\langle\mathbf{u}\|$ can be deduced from the results of Sec. VI since it does not depend on whether the PPCS form or do not form a basis. From Eqs. (6.1) and (6.11), and the Hermiticity property $E_{ij} = E^{\dagger}_{ji}$, Eq. (B5a) is transformed into the following equation:

$$\begin{aligned}\int d\mathbf{u}\,d\mathbf{u}^* \Big\{ &\Big[\|\mathbf{u}\rangle\rangle\mathring{\mathbf{C}}_{ij} + \Big(\mathbf{u}^*\boldsymbol{\Delta}_{u^*} + \frac{n}{2}\mathbf{I}\Big)_{ji}\|\mathbf{u}\rangle\rangle\Big] \\ &\times \hat{\mathbf{f}}(\mathbf{u},\mathbf{u}^*)\langle\langle\mathbf{u}\| - \|\mathbf{u}\rangle\rangle\hat{\mathbf{f}}(\mathbf{u},\mathbf{u}^*) \\ &\times\Big[\mathring{\mathbf{C}}_{ij} + \Big(\mathbf{u}\boldsymbol{\Delta}_u + \frac{n}{2}\mathbf{I}\Big)_{ij}\mathbb{I}\Big]\langle\langle\mathbf{u}\|\Big\} = 0,\end{aligned} \tag{B6}$$

where the differential operators $\boldsymbol{\Delta}_u(\boldsymbol{\Delta}_{u^*})$ can now be transferred from $\langle\langle\mathbf{u}\|\,(\|\mathbf{u}\rangle\rangle)$ to the weight matrix $\hat{\mathbf{f}}(\mathbf{u},\mathbf{u}^*)$ by integrating by parts. This calculation leads to the following partial differential equation for $\hat{\mathbf{f}}(\mathbf{u},\mathbf{u}^*)$:

$$[(\mathbf{u}\boldsymbol{\Delta}_u)_{ij} - (\mathbf{u}^*\boldsymbol{\Delta}_{u^*})_{ji}]\,\hat{\mathbf{f}}(\mathbf{u},\mathbf{u}^*) = [\hat{\mathbf{f}}(\mathbf{u},\mathbf{u}^*),\mathring{\mathbf{C}}_{ij}]. \tag{B7}$$

The same procedure applied to Eq. (B5b) gives rise to the additional equation

$$\begin{aligned}&\Big\{\sum_k u_{ik}(\mathbf{u}\boldsymbol{\Delta}_u)_{jk} - (n-2d-2)u_{ij} - \Delta_{u^*_{ij}}\Big\}\hat{\mathbf{f}}(\mathbf{u},\mathbf{u}^*) \\ &\quad = \hat{\mathbf{f}}(\mathbf{u},\mathbf{u}^*)(\mathring{\mathbf{C}}\mathbf{u} + \mathbf{u}\widetilde{\mathring{\mathbf{C}}})_{ij}.\end{aligned} \tag{B8}$$

In deriving this relation, we used Eq. (6.13) to let the differential operators $\boldsymbol{\Delta}_u$ directly act upon $\hat{\mathbf{f}}(\mathbf{u},\mathbf{u}^*)$.

By introducing Eq. (B2) into Eqs. (B7) and (B8), the latter are transformed into the following two equations:

$$\begin{aligned}&-(d+1)\{[(\mathbf{u}\boldsymbol{\Delta}_u)_{ij} - (\mathbf{u}^*\boldsymbol{\Delta}_{u^*})_{ji}]\det\mathbf{U}\}\hat{\mathbb{K}}^{-1}(\mathbf{U}) \\ &\quad = \det\mathbf{U}\{-[(\mathbf{u}\boldsymbol{\Delta}_u)_{ij} - (\mathbf{u}^*\boldsymbol{\Delta}_{u^*})_{ji}]\hat{\mathbb{K}}^{-1}(\mathbf{U}) \\ &\qquad + [\hat{\mathbb{K}}^{-1}(\mathbf{U}),\mathring{\mathbf{C}}_{ij}]\},\end{aligned} \tag{B9}$$

and

$$\begin{aligned}&-(d+1)\Big\{\Big[\sum_k u_{ik}(\mathbf{u}\boldsymbol{\Delta}_u)_{jk} - 2u_{ij} - \Delta_{u^*_{ij}}\Big]\det\mathbf{U}\Big\}\hat{\mathbb{K}}^{-1}(\mathbf{U}) \\ &\quad = \det\mathbf{U}\Big\{-\Big[\sum_k u_{ik}(\mathbf{u}\boldsymbol{\Delta}_u)_{jk} - nu_{ij} - \Delta_{u^*_{ij}}\Big] \\ &\qquad \times\hat{\mathbb{K}}^{-1}(\mathbf{U}) + \hat{\mathbb{K}}^{-1}(\mathbf{U})(\mathring{\mathbf{C}}\mathbf{u} + \mathbf{u}\widetilde{\mathring{\mathbf{C}}})_{ij}\Big\},\end{aligned} \tag{B10}$$

which remain to be demonstrated to complete the proof.

For such purposes, we shall use the partial differential equations satisfied by the overlap matrix $\hat{\mathbb{K}}(\mathbf{U})$. The latter can be established as follows. By using the Hermiticity property $E_{ij} = E^{\dagger}_{ji}$, the matrix $\langle\langle\mathbf{u}\|E_{ij}\|\mathbf{u}\rangle\rangle$ can be put into either of the two forms:

$$\begin{aligned}&\langle\langle\mathbf{u}\|E_{ij}\|\mathbf{u}\rangle\rangle \\ &\quad = [\mathring{\mathbf{C}}_{ij} + (\mathbf{u}\boldsymbol{\Delta}_u + (n/2)\mathbf{I})_{ij}\mathbb{I}]\langle\langle\mathbf{u}\|\mathbf{u}\rangle\rangle \\ &\quad = \langle\langle\mathbf{u}\|[\|\mathbf{u}\rangle\rangle\mathring{\mathbf{C}}_{ij} + (\mathbf{u}^*\boldsymbol{\Delta}_{u^*} + (n/2)\mathbf{I})_{ji}\|\mathbf{u}\rangle\rangle],\end{aligned} \tag{B11}$$

hence leading to the equation

$$[(\mathbf{u}\boldsymbol{\Delta}_u)_{ij} - (\mathbf{u}^*\boldsymbol{\Delta}_{u^*})_{ji}]\hat{\mathbb{K}}(\mathbf{U}) = [\hat{\mathbb{K}}(\mathbf{U}),\mathring{\mathbf{C}}_{ij}]. \tag{B12}$$

The same procedure applied to $\langle\langle\mathbf{u}\|D^{\dagger}_{ij}\|\mathbf{u}\rangle\rangle$ gives the additional equation

$$\begin{aligned}&\Big[\sum_k u_{ik}(\mathbf{u}\boldsymbol{\Delta}_u)_{jk} + nu_{ij} - \Delta_{u^*_{ij}}\Big]\hat{\mathbb{K}}(\mathbf{U}) \\ &\quad = -(\mathring{\mathbf{C}}\mathbf{u} + \mathbf{u}\widetilde{\mathring{\mathbf{C}}})_{ij}\hat{\mathbb{K}}(\mathbf{U}),\end{aligned} \tag{B13}$$

where we used Eq. (6.13) again.

From the identity $\hat{\mathbb{K}}^{-1}(\mathbf{U})\hat{\mathbb{K}}(\mathbf{U}) = \mathbb{I}$, it follows that for any first-order partial differential operator $\mathscr{D}$

$$\mathscr{D}\hat{\mathbb{K}}^{-1}(\mathbf{U}) = -\hat{\mathbb{K}}^{-1}(\mathbf{U})[\mathscr{D}\hat{\mathbb{K}}(\mathbf{U})]\hat{\mathbb{K}}^{-1}(\mathbf{U}). \tag{B14}$$

By combining Eq. (B14) with Eqs. (B12) and (B13), we obtain the following relations:

$$[(\mathbf{u}\boldsymbol{\Delta}_u)_{ij} - (\mathbf{u}^*\boldsymbol{\Delta}_{u^*})_{ji}]\hat{\mathbb{K}}^{-1}(\mathbf{U}) = [\hat{\mathbb{K}}^{-1}(\mathbf{U}),\mathring{\mathbf{C}}_{ij}], \tag{B15}$$

and

$$\begin{aligned}&\Big[\sum_k u_{ik}(\mathbf{u}\boldsymbol{\Delta}_u)_{jk} - nu_{ij} - \Delta_{u^*_{ij}}\Big]\hat{\mathbb{K}}^{-1}(\mathbf{U}) \\ &\quad = \hat{\mathbb{K}}^{-1}(\mathbf{U})(\mathring{\mathbf{C}}\mathbf{u} + \mathbf{u}\widetilde{\mathring{\mathbf{C}}})_{ij},\end{aligned} \tag{B16}$$

which imply that the right-hand sides of Eqs. (B9) and (B10) are both equal to zero.

Finally, we note that when $\lambda_1 = \lambda_2 = \cdots = \lambda_d = 0$, the overlap matrix reduces to

$$\hat{\mathbb{K}}(\mathbf{U}) = \hat{A}\,(\det\mathbf{U})^{-n/2}\mathbb{I}, \tag{B17}$$

and $\mathring{\mathbf{C}}_{ij}$ is equal to zero. Equations (B12) and (B13) then, respectively, become

$$[(\mathbf{u}\boldsymbol{\Delta}_u)_{ij} - (\mathbf{u}^*\boldsymbol{\Delta}_{u^*})_{ji}]\det\mathbf{U} = 0, \tag{B18}$$

and

$$\Big[\sum_k u_{ik}(\mathbf{u}\boldsymbol{\Delta}_u)_{jk} - 2u_{ij} - \Delta_{u^*_{ij}}\Big]\det\mathbf{U} = 0, \tag{B19}$$

thereby showing that the left-hand sides of Eqs. (B9) and (B10) do also vanish, which completes the proof.

## APPENDIX C: UNITY RESOLUTION FOR THE SET OF BARUT–GIRARDELLO PARTIALLY COHERENT STATES

The purpose of this appendix is to prove the existence of a matrix measure $\mathbf{d}\sigma(\mathbf{w})$ satisfying the unity resolution relation (4.25).

We start by noting that the Sp(2$d$,$R$) group can be embedded into the larger group Sp(2$dn$,$R$),[1] whose generators are given by

$$\begin{aligned}&D^{\dagger}_{is,\,jt} = D^{\dagger}_{jt,is} = \eta_{is}\eta_{jt}, \quad (is)\leqslant(jt), \\ &D_{is,\,jt} = D_{jt,is} = \xi_{is}\xi_{jt}, \quad (is)\leqslant(jt),\end{aligned} \tag{C1}$$

and

$$E_{is,jt} = \tfrac{1}{2}(\eta_{is}\xi_{jt} + \xi_{jt}\eta_{is}),$$

where $i,j = 1,...,d$, and $s,t = 1,...,n$. As a matter of fact, Sp($2dn,R$) contains the direct product of Sp($2d,R$) with the O($n$) group, generated by the operators

$$\Lambda_{st} = -\Lambda_{ts} = -i\sum_{i=1}^{d}(E_{is,it} - E_{it,is}), \quad 1 \leqslant s \leqslant t \leqslant n. \tag{C2}$$

We have therefore the group chain

$$\begin{array}{ccccc} \mathrm{Sp}(2dn,R) & \supset & \mathrm{Sp}(2d,R) & \times & \mathrm{O}(n), \\ \langle (1/2)^{dn}\rangle & & \langle \lambda_d + n/2,...,\lambda_1 + n/2\rangle & & (\lambda_1...\lambda_d) \\ \text{or } \langle (1/2)^{dn-1}3/2\rangle, & & & & \end{array} \tag{C3}$$

where below each group we have indicated the quantum numbers characterizing its irreps.

All the boson states belong to one of two irreps of Sp($2dn,R$), $\langle (1/2)^{dn}\rangle$ or $\langle (1/2)^{dn-1}3/2\rangle$, according to whether the total number of bosons is even or odd. Within each one of them, the Sp($2d,R$) and O($n$) groups are complementary[7] in the sense that there is a one-to-one correspondence between all their irreps, respectively, characterized by $\langle \lambda_d + n/2,...,\lambda_1 + n/2\rangle$ and $(\lambda_1...\lambda_d)$. The representation space of $\langle \lambda_d + n/2,...,\lambda_1 + n/2\rangle$ can therefore be realized by selecting all the boson states transforming under the irrep $(\lambda_1...\lambda_d)$ of O($n$), and belonging to a definite row of the latter, e.g., their highest weight states. The discrete basis states $|\mathbf{N};(\lambda)\rangle$ of such a representation space are obtained by choosing as basis states $|(\lambda)\rangle$ of the U($d$) irrep $[\lambda_1...\lambda_d]$ in Eq. (2.6) those boson states which are at the same time of highest weight with respect to O($n$).

In the Bargmann representation,[27] the boson creation and annihilation operators, $\eta_{is}$ and $\xi_{is}$, are, respectively, represented by some complex variables $z_{is}$ and the corresponding differential operators $\partial/\partial z_{is}$. The boson states are then represented by analytic functions in $z_{is}$, $i = 1,...,d$, $s = 1,...,n$. In particular, for the discrete basis states $|\mathbf{N};(\lambda)\rangle$, the latter are given by

$$\phi_{\mathbf{N},(\lambda)}(z_{is}) = \prod_{i\leqslant j}(N_{ij}!)^{-1/2} \times\left[(1+\delta_{ij})^{-1/2}\sum_{s} z_{is}z_{js}\right]^{N_{ij}}\phi_{(\lambda)}(z_{is}), \tag{C4}$$

and span a subspace of this Bargmann Hilbert space of analytic functions. In Eq. (C4), we note the appearance of the combinations $\Sigma_s z_{is}z_{js}$, which are actually the scalars with respect to $O(n)$ that can be built from the $z_{is}$ variables.

In Ref. 26, a transformation from the $z_{is}$ variables to new variables, including the combinations $\Sigma_s z_{is}z_{js}$, was proposed. These new variables were defined by

$$w_{ij} = \sum_{s=1}^{n} z_{is}z_{js}, \tag{C5a}$$

$$x_{i\alpha} = 2^{-1/2}(z_{i,2\alpha-1} - iz_{i,2\alpha}), \quad \alpha = 1,...,l, \tag{C5b}$$

$$y_{i\rho} = \begin{cases} 2^{-1/2}(z_{i,2\rho-1} + iz_{i,2\rho}), & \rho = 1,...,l-i, \\ z_{in}, & \rho = l-i+1 \text{ (only when } n = 2l+1), \end{cases} \tag{C5c}$$

where $l = [n/2]$. When this transformation is performed on $\phi_{(\lambda)}(z_{is})$, it turns out that it only depends upon the $x_{i\alpha}$ variables, so that Eq. (C4) becomes

$$\phi_{\mathbf{N},(\lambda)}(w_{ij},x_{i\alpha}) = F_{\mathbf{N}}(\mathbf{w})\phi_{(\lambda)}(x_{i\alpha}). \tag{C6}$$

By comparing Eq. (C6) with Eq. (4.19), we conclude that the former can be rewritten as

$$\phi_{\mathbf{N},(\lambda)}(w_{ij},x_{i\alpha}) = \sum_{(\lambda')}\phi_{(\lambda')}(x_{i\alpha})(\mathbf{w};(\lambda')|\mathbf{N};(\lambda)\rangle, \tag{C7}$$

provided $w_{ij}$ in $(\mathbf{w};(\lambda')|\mathbf{N};(\lambda)\rangle$ is interpreted as $\Sigma_s z_{is}z_{js}$ in accordance with Eq. (C5a). This interpretation is actually not restrictive since for any set of complex variables $w_{ij}$, $1\leqslant i\leqslant j\leqslant d$, it is always possible to find some complex variables $z_{is}$, $i = 1,...,d$, $s = 1,...,n$ ($\geqslant 2d$), such that Eq. (C5a) is fulfilled.

Let us now calculate in Bargmann space the overlap of two discrete basis states

$$\langle \mathbf{N}'(\lambda')|\mathbf{N};(\lambda)\rangle = \int\left[\prod_{is} d\mu(z_{is})\right][\phi_{\mathbf{N}',(\lambda')}(w_{ij},x_{i\alpha})]^* \times\phi_{\mathbf{N},(\lambda)}(w_{ij},x_{i\alpha}), \tag{C8}$$

where $d\mu(z)$ is the standard Bargmann measure, defined in Eq. (5.8). From Eq. (C7), it can be written as

$$\langle \mathbf{N}';(\lambda')|\mathbf{N};(\lambda)\rangle = \sum_{(\bar\lambda)(\bar\lambda')}\int\left[\prod_{is} d\mu(z_{is})\right]\langle \mathbf{N}';(\lambda')|\mathbf{w};(\bar\lambda')) \times[\phi_{(\bar\lambda')}(x_{i\alpha})]^*\phi_{(\bar\lambda)}(x_{i\alpha})(\mathbf{w};(\bar\lambda)|\mathbf{N};(\lambda)\rangle. \tag{C9}$$

When performing the change of variables (C5) in the integral on the right-hand side, Eq. (C9) is transformed into the relation

$$\langle \mathbf{N}';(\lambda')|\mathbf{N};(\lambda)\rangle = \sum_{(\bar\lambda)(\bar\lambda')}\int d\mathbf{w}\, d\mathbf{w}^*\langle \mathbf{N}';(\lambda')|\mathbf{w};(\bar\lambda')) \times f_{(\bar\lambda')(\bar\lambda)}(\mathbf{w},\mathbf{w}^*)(\mathbf{w};(\bar\lambda)|\mathbf{N};(\lambda)\rangle, \tag{C10}$$

where the function $f_{(\bar\lambda')(\bar\lambda)}(\mathbf{w},\mathbf{w}^*)$ can, at least in principle, be obtained by integrating the product of $[\phi_{(\bar\lambda')}(x_{i\alpha})]^*\phi_{(\bar\lambda)}(x_{i\alpha})$ with the Jacobian of the transformation over the $x_{i\alpha}$ and $y_{i\rho}$ variables.

Since the states $|\mathbf{N};(\lambda)\rangle$ form a basis of the representation space, and Eq. (C10) is valid for any values of $\mathbf{N}$, $\mathbf{N}'$, $(\lambda)$, and $(\lambda')$, it imposes that in this space there exists a unity resolution relation for the BGPCS, given by

$$\sum_{(\bar\lambda)(\bar\lambda')}\int d\mathbf{w}\, d\mathbf{w}^*|\mathbf{w};(\bar\lambda'))f_{(\bar\lambda')(\bar\lambda)}(\mathbf{w},\mathbf{w}^*)(\mathbf{w};(\bar\lambda)| = I. \tag{C11}$$

Equation (C11) coincides with Eq. (4.25), whose existence is therefore proved. In passing, we have also demonstrated that

if we interpret $\mathbf{w}$ in accordance with Eq. (C5a), the BGPCS $|\mathbf{w};(\lambda)\rangle$ span a subspace of a Bargmann–Hilbert space of analytic functions. This subspace, characterized by definite transformation properties with respect to O($n$), might be called a Barut–Girardello space.

## APPENDIX D: CALCULATION OF $\hat{\mathbb{D}}^{\dagger}_{ij}$ AND $\mathbb{D}_{ij}$

The purpose of this appendix is to prove Eqs. (6.4″) and (6.15″). Starting with the former, we note that from Eq. (6.5) we may write

$$\begin{aligned}\langle \mathbf{u};(\lambda)|D^{\dagger}_{ij}|\psi\rangle &= \langle(\lambda)|\left[D^{\dagger}_{ij}+(\mathbf{Cu})_{ij}+(\mathbf{Cu})_{ji}+nu_{ij}+(\mathbf{uDu})_{ij}\right]\\ &\quad\times\exp(\tfrac{1}{2}\,\mathrm{tr}\,\mathbf{uD})|\psi\rangle. \end{aligned}\tag{D1}$$

When acting upon the bra $\langle(\lambda)|$, $D^{\dagger}_{ij}$ gives zero, $C_{kl}$ transforms it into a linear combination of bras $\langle(\lambda')|$, and $(\mathbf{uDu})_{ij}$ may be replaced by

$$\sum_{kl} u_{ik}u_{lj}\Delta_{u_{kl}} = \{[\mathbf{u}\Delta_u-(d+1)]\mathbf{u}\}_{ij},\tag{D2}$$

where use has been made of Eq. (6.13). Equation (D1) therefore becomes

$$\begin{aligned}\langle \mathbf{u};(\lambda)|D^{\dagger}_{ij}|\psi\rangle = \sum_{(\lambda')}\Big\{&\sum_k\left[\langle(\lambda)|C_{ik}|(\lambda')\rangle u_{kj}\right.\\ &\left.+\langle(\lambda)|C_{jk}|(\lambda')\rangle u_{ki}\right]\\ &+[(\mathbf{u}\Delta_u+n-d-1)\mathbf{u}]_{ij}\delta_{(\lambda),(\lambda')}\Big\}\\ &\times\langle\mathbf{u};(\lambda')|\psi\rangle,\end{aligned}\tag{D3}$$

and leads to Eq. (6.4″) when definition (6.9) is taken into account.

We next prove Eq. (6.15″). For such purpose, we start from the relation

$$(\mathbf{w};(\lambda)|D_{ij}|\mathbf{u};(\lambda')\rangle = \langle\mathbf{u};(\lambda')|D^{\dagger}_{ij}|\mathbf{w};(\lambda)\rangle^*,\tag{D4}$$

and use Eqs. (D3) and (4.12) to obtain

$$\begin{aligned}&(\mathbf{w};(\lambda)|D_{ij}|\mathbf{u};(\lambda')\rangle\\ &\quad=\Big\{\sum_k\left[\langle(\lambda')|C_{ik}|(\lambda)\rangle^* u^*_{kj}\right.\\ &\qquad\left.+\langle(\lambda')|C_{jk}|(\lambda)\rangle^* u^*_{ki}\right]\\ &\qquad+[(\mathbf{u}^*\Delta_{u^*}+n-d-1)\mathbf{u}^*]_{ij}\delta_{(\lambda'),(\lambda)}\Big\}\\ &\qquad\times\exp(\tfrac{1}{2}\,\mathrm{tr}\,\mathbf{u}^*\mathbf{w}).\end{aligned}\tag{D5}$$

The last term on the right-hand side of Eq. (D5) may be transformed as follows:

$$\begin{aligned}&[(\mathbf{u}^*\Delta_{u^*}+n-d-1)\mathbf{u}^*]_{ij}\exp(\tfrac{1}{2}\,\mathrm{tr}\,\mathbf{u}^*\mathbf{w})\\ &\quad=\left[\sum_{kl}u^*_{ik}u^*_{lj}\Delta_{u^*_{kl}}+nu^*_{ij}\right]\exp(\tfrac{1}{2}\,\mathrm{tr}\,\mathbf{u}^*\mathbf{w})\\ &\quad=\left[\sum_{kl}w_{kl}\Delta_{w_{ik}}\Delta_{w_{lj}}+n\Delta_{w_{ij}}\right]\exp(\tfrac{1}{2}\,\mathrm{tr}\,\mathbf{u}^*\mathbf{w})\\ &\quad=[\Delta_w(\mathbf{w}\Delta_w+n-d-1)]_{ij}\exp(\tfrac{1}{2}\,\mathrm{tr}\,\mathbf{u}^*\mathbf{w}),\end{aligned}\tag{D6}$$

by applying Eq. (6.13) in the first and last steps. The remaining terms in Eq. (D5) may also be converted into differential operators with respect to the $w_{ij}$'s by replacing $u^*_{kj}$ and $u^*_{ki}$ by $\Delta_{w_{kj}}$ and $\Delta_{w_{ki}}$, respectively. Finally, by using Eq. (4.12) again, as well as the Hermiticity property $C^{\dagger}_{ij}=C_{ji}$, we obtain the result

$$\begin{aligned}&(\mathbf{w};(\lambda)|D_{ij}|\mathbf{u};(\lambda')\rangle\\ &\quad=\sum_{(\lambda'')}\Big\{\sum_k\left[\langle(\lambda)|C_{ki}|(\lambda'')\rangle\Delta_{w_{kj}}\right.\\ &\qquad\left.+\langle(\lambda)|C_{kj}|(\lambda'')\rangle\Delta_{w_{ki}}\right]\\ &\qquad+[\Delta_w(\mathbf{w}\Delta_w+n-d-1)]_{ij}\delta_{(\lambda),(\lambda'')}\Big\}\\ &\qquad\times(\mathbf{w};(\lambda'')|\mathbf{u};(\lambda')\rangle,\end{aligned}\tag{D7}$$

which proves Eq. (6.15″) valid.


[1] M. Moshinsky and C. Quesne, J. Math. Phys. **12**, 1772 (1971).
[2] R. C. Hwa and J. Nuyts, Phys. Rev. **145**, 1188 (1966).
[3] B. G. Wybourne, *Classical Groups for Physicists* (Wiley, New York, 1974), p. 286.
[4] C. Quesne and M. Moshinsky, J. Math. Phys. **12**, 1780 (1971).
[5] E. Chacón, D. Levi, and M. Moshinsky, J. Math. Phys. **17**, 1919 (1976).
[6] N. J. Gunther, J. Math. Phys. **20**, 1539 (1979).
[7] M. Moshinsky and C. Quesne, J. Math. Phys. **11**, 1631 (1970).
[8] L. D. Mlodinow and N. Papanicolaou, Ann. Phys. (N.Y.) **128**, 314 (1980); **131**, 1 (1981).
[9] L. C. Biedenharn and J. D. Louck, Ann. Phys. (N.Y.) **63**, 459 (1971).
[10] G. Rosensteel and D. J. Rowe, Ann. Phys. (N.Y.) **126**, 343 (1980).
[11] G. F. Filippov, V. I. Ovcharenko, and Yu. F. Smirnov, *Microscopic Theory of Collective Excitations in Nuclei* (in Russian) (Naukova Dumka, Kiev, 1981).
[12] V. Vanagas, *The Microscopic Nuclear Theory Within the Framework of the Restricted Dynamics*, in *Lecture Notes in Physics* (University of Toronto, Toronto, 1977); V. Vanagas, *The Microscopic Theory of the Collective Motion in Nuclei*, in *Group Theory and Its Applications in Physics–1980*, edited by T. H. Seligman, AIP Conference Proceedings No. 71 (A.I.P., New York, 1981).
[13] O. Castaños, A. Frank, E.Chacón, P. Hess, and M. Moshinsky, J. Math. Phys. **23**, 2537 (1982).
[14] P. Kramer, Ann. Phys. (N.Y.) **141**, 254,269 (1982).
[15] J. Deenen and C. Quesne, J. Math. Phys. **23**, 878 (1982).
[16] J. Deenen and C. Quesne, J. Math. Phys. **23**, 2004 (1982).
[17] J. Deenen and C. Quesne J. Math. Phys. **25**, 1638 (1984).
[18] J. Deenen and C. Quesne, in *Group Theoretical Methods in Physics, Proceedings, Istanbul, Turkey, 1982,* edited by M. Serdaroğlu and E. Inönü, Lecture Notes in Physics No. 180 (Springer, Berlin, 1983), p. 444.
[19] G. Rosensteel and D. J. Rowe, Int. J. Theor. Phys. **16**, 63 (1977).
[20] A. U. Klimyk, J. Math. Phys. **24**, 224 (1983).
[21] R. J. Glauber, Phys. Rev. **130**, 2529 (1963); **131**, 2766 (1963).
[22] A. M. Perelomov, Commun. Math. Phys. **26**, 222 (1972); Usp. Fiz. Nauk **123**, 23 (1977) [Sov. Phys. Usp. **20**, 703 (1977)].
[23] A. O. Barut and L. Girardello, Commun. Math. Phys. **21**, 41 (1971).
[24] V. S. Vasilevskii, Yu. F. Smirnov, and G. F. Filippov, Yad. Fiz. **32**, 987 (1980) [Sov. J. Nucl. Phys. **32**, 510 (1980)].
[25] I. M. Gel'fand and M. L. Tseitlin, Dokl. Akad. Nauk SSSR **71**, 825 (1950).
[26] J. Deenen and C. Quesne, J. Phys. A: Math. Gen **16**, 2095 (1983).
[27] V. Bargmann, Commun. Pure Appl. Math. **14**, 187 (1961).
[28] F. J. Dyson, Phys. Rev. **102**, 1217 (1956).
[29] T. Holstein and H. Primakoff, Phys. Rev. **58**, 1098 (1940).
[30] P. Kramer, M. Moshinsky, and T. H. Seligman, J. Math. Phys. **19**, 683 (1978).
[31] M. Moshinsky and T. H. Seligman, Ann. Phys. (N.Y.) **114**, 243 (1978); **120**, 402 (1979); J. Deenen, M. Moshinsky, and T. H. Seligman, Ann. Phys. (N.Y.) **127**, 458 (1980).
[32] Note however that in nonbijective canonical transformations, the ambiguity spin characterizes the irreps of the ambiguity group, itself connected with the classical phase space structure. Since, in the present case, we are discussing a quantum-mechanical problem with no simple classical analog, the corresponding concept of an ambiguity group is unfortunately missing.
[33] R. Gilmore, *Lie Groups, Lie Algebras, and Some of Their Applications* (Wiley, New York, 1974).

# The Kostant partition function for simple Lie algebras

Jeffrey R. Schmidt and Adam M. Bincer
*Physics Department, University of Wisconsin–Madison, Madison, Wisconsin 53706*

(Received 12 December 1983; accepted for publication 24 February 1984)

A simple algorithm is developed for evaluating the Kostant partition function for any simple Lie algebra. The algorithm may also be used to express the partition function for one Lie algebra in terms of the partition function for another, with the latter algebra not necessarily a subalgebra of the former. A special role in the algorithm is played by the sum of the simple roots. Explicit, closed-form expressions are given for the partition function for a variety of special cases.

PACS numbers: 02.20.Qs

## I. INTRODUCTION

We present a series of results dealing with the evaluation of Kostant's partition function. This function is used in the theory of Lie algebras to determine the (inner) multiplicity of a weight in an irreducible representation and the (outer) multiplicity in tensor products of irreducible representations.

The paper is organized as follows. In Sec. II, we define the partition function in very general terms and give an algorithm for its evaluation by introducing a set $S$ and its subset $T$. In Sec. III, we obtain a general formula for Kostant's partition function for any simple Lie algebra $L$ by taking for $S$ the set of positive roots of $L$ and for $T$ the set of simple roots together with a special root called $\mathbf{1}$. The inclusion of this root $\mathbf{1}$, which equals the sum of all simple roots and is present in every simple Lie algebra, is crucial to our arguments. As examples we give explicit formulas for the algebras $A_2$ and $A_3$ and point out a minor error in the results of Tarski on this subject.

In Sec. IV, we derive a variety of recursion formulas by means of which the Kostant partition function for the algebra $L$ can be gotten from the Kostant partition function for the algebra $L'$. It is interesting to note here that the algorithm does not require that $L'$ be a subalgebra of $L$ but only that the positive roots of $L'$, with respect to some basis of simple roots, be a subset of the positive roots of $L$. As applications of this approach we give the partition functions for $B_l$, $C_l$, and $A_{l+1}$ in the form of multiple sums over the partition function for $A_l$ even though $A_l \not\subset B_l$, $A_l \not\subset C_l$.

In view of the above we concentrate on $P_{A_l}(\mathbf{k})$, the Kostant partition function for $A_l$. To solve the recursion relations for $P_{A_l}(\mathbf{k})$ requires evaluating $(l-1)$-fold sums whose limits have a complicated dependence on $\mathbf{k}$. For arbitrary $\mathbf{k}$ and $l$ this is a formidable task which we hope to tackle at some later date. We are, however, able to solve these recursion relations when $\mathbf{k}$ is restricted to be monotonic (by which we mean that the successive components of $\mathbf{k}$ are nondecreasing or nonincreasing) and this is done in Sec. V. Even more explicit results can be obtained by imposing additional restrictions on $\mathbf{k}$. Thus, for example, we give in Sec. V closed-form expressions for $P_{A_l}(\mathbf{k})$, for $l$ up to 7, when the components of $\mathbf{k}$ are restricted to be all equal.

Under the extreme restriction, $\mathbf{k}=\mathbf{1}$, we are able to show that

$$P_L(\mathbf{1}) = 2^{l-1}$$

for *any* simple Lie algebra $L$ of rank $l$. This result is proved in Sec. VI.

## II. PARTITION FUNCTION

Let $N$ be the set of non-negative integers and consider the $l$-tuples

$$\mathbf{k} = (k_1, k_2, \ldots, k_l), \quad k_i \in N. \tag{2.1}$$

Let $S$ be a finite subset of $N^l$, the set of $l$-tuples whose components are chosen from $N$. The partition function $P_S(\mathbf{k})$ is the function whose value on $\mathbf{k}$ is the number of ways of forming $\mathbf{k}$ as linear combinations of elements of $S$ with non-negative coefficients.

To evaluate $P_S(\mathbf{k})$ we introduce $T$, a subset of $S$, and its complement $S \setminus T$,

$$S = T \cup S \setminus T \tag{2.2}$$

and write $\mathbf{k}$ as

$$\mathbf{k} = \mathbf{m} + \mathbf{n}, \tag{2.3}$$

where $\mathbf{m}$ and $\mathbf{n}$ are some linear combinations of elements from $T$ and $S \setminus T$, respectively, and by definition

$$m_i \geqslant 0, \quad n_i \geqslant 0 \quad \text{for all } i. \tag{2.4}$$

Specifically suppose that

$$\mathbf{n} = \sum_{r=1}^{x} j_r \mathbf{p}_r \equiv \mathbf{n}(j), \quad j_r \in N, \ \mathbf{p}_r \in S \setminus T, \tag{2.5}$$

where $x$ denotes the number of elements in $S \setminus T$. It follows that the number of ways of writing $\mathbf{k}$ as

$$\mathbf{k} = \mathbf{m} + \mathbf{n}(j) \tag{2.6}$$

for a particular fixed configuration of the $j_1, j_2, \ldots, j_x$ is given by

$$P_T(\mathbf{m}) = P_T[\mathbf{k} - \mathbf{n}(j)], \tag{2.7}$$

and therefore

$$P_S(\mathbf{k}) = \sum P_T[\mathbf{k} - \mathbf{n}(j)]. \tag{2.8}$$

Here $\Sigma$ stands for summation over all allowed configurations of the $j_r$. In view of Eqs. (2.4) and (2.6) we must have $k_i - n_i(j) \geqslant 0$ for all $i$, which we shall write as

$$\mathbf{n}(j) \leqslant \mathbf{k} \tag{2.9}$$

and the allowed $j_r$ must satisfy this constraint. There may be additional restrictions on the $j_r$ depending on the specific nature of the sets $S$ and $T$. We next obtain a number of results

 0022-2488/84/082367-07$02.50  

for Kostant's partition function by using Eq. (2.8) for various choices of the sets $S$ and $T$.

## III. KOSTANT'S PARTITION FUNCTION

Let $L$ be a simple Lie algebra of rank $l$ with simple roots $\boldsymbol{\alpha}_i$, which may be conveniently written as unit $l$-tuples

$$\begin{aligned}&\boldsymbol{\alpha}_1 = (1, 0, ..., 0),\\ &\boldsymbol{\alpha}_2 = (0, 1, 0, ..., 0),\\ &\vdots\\ &\boldsymbol{\alpha}_l = (0, 0, ..., 0, 1).\end{aligned} \tag{3.1}$$

Kostant's partition function[1-4] for this Lie algebra $L$, to be denoted by $P_L(\mathbf{k})$, is obtained by taking for $S$ the set $\Phi_L^+$, which is the set of positive roots of $L$

$$P_L(\mathbf{k}) \equiv P_{\Phi_L^+}(\mathbf{k}). \tag{3.2}$$

In other words, $P_L(\mathbf{k})$ is the number of ways of writing

$$\mathbf{k} = k_1\boldsymbol{\alpha}_1 + k_2\boldsymbol{\alpha}_2 + \cdots + k_l\boldsymbol{\alpha}_l \tag{3.3}$$

as linear combinations of the positive roots with non-negative coefficients. The Cartan matrix determines the positive roots, their number is finite and any $\boldsymbol{\alpha}\in\Phi_L^+$ may be written as $\boldsymbol{\alpha} = \Sigma_{i=1}^{l} s_i\boldsymbol{\alpha}_i$, $s_i\in N$, $\boldsymbol{\alpha}_i$ the simple roots.

We next observe that every simple Lie algebra possesses the positive root

$$\mathbf{1} \equiv (1, 1, ..., 1) = \sum_{i=1}^{l} \boldsymbol{\alpha}_i \tag{3.4}$$

and choose for $T$ the following subset of $\Phi_L^+$:

$$T = \{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, ..., \boldsymbol{\alpha}_l, \mathbf{1}\}. \tag{3.5}$$

To evaluate the partition function $P_T(\mathbf{m})$ we must count the number of ways of writing $\mathbf{m}$ as non-negative linear combinations of the elements of $T$, Eq. (3.5)

$$\begin{aligned}\mathbf{m} &= m_1(1, 0, ..., 0) + m_2(0, 1, ..., 0)\\ &\quad + \cdots + m_l(0, 0, ..., 0, 1)\\ &= (1, 1, ..., 1) + (m_1 - 1)(1, 0, ..., 0)\\ &\quad + (m_2 - 1)(0, 1, 0, ..., 0)\\ &\quad + \cdots + (m_l - 1)(0, 0, ..., 0, 1)\\ &= 2(1, 1, ..., 1) + (m_1 - 2)(1, 0, ..., 0)\\ &\quad + (m_2 - 2)(0, 1, 0, ..., 0)\\ &\quad + \cdots + (m_l - 2)(0, 0, ..., 0, 1)\\ &\quad \vdots\\ &= \mu(1, 1, ..., 1) + (m_1 - \mu)(1, 0, ..., 0)\\ &\quad + (m_2 - \mu)(0, 1, 0, ..., 0)\\ &\quad + \cdots + (m_l - \mu)(0, 0, ..., 0, 1),\end{aligned} \tag{3.6}$$

where

$$\mu = \min[\mathbf{m}] \tag{3.7}$$

and by $\min[\mathbf{m}]$ we mean the smallest component $m_i$. Therefore,

$$P_T(\mathbf{m}) = 1 + \min[\mathbf{m}]. \tag{3.8}$$

Combining this with the results of Sec. II yields the following formula for Kostant's partition function:

$$P_L(\mathbf{k}) = \sum \{1 + \min[\mathbf{k} - \mathbf{n}_L(j)]\}, \tag{3.9}$$

where the sum is over all $j_1, j_2, ..., j_x \in N$ such that

$$\mathbf{n}_L(j) \leqslant \mathbf{k}, \tag{3.10}$$

and where

$$\mathbf{n}_L(j) = \sum_{r=1}^{x_L} j_r\mathbf{p}_r, \quad j_r\in N, \ \mathbf{p}_r\in L\setminus T, \tag{3.11}$$

with $x_L$ denoting the number of elements of $L\setminus T \equiv \Phi_L^+\setminus T$, i.e., the number of positive roots of $L$ left after omitting $\mathbf{1}$ and the simple roots.

As a simple application of the result Eq. (3.9) suppose that $L = A_2$. Since

$$\Phi_{A_2}^+ = \{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_1 + \boldsymbol{\alpha}_2 \equiv \mathbf{1}\} \tag{3.12}$$

the set $A_2\setminus T$ is empty and we have the well-known answer

$$P_{A_2}(k_1, k_2) = 1 + \min(k_1, k_2). \tag{3.13}$$

As another example take $L = A_3$. Since

$$\Phi_{A_3}^+ = \{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3, \boldsymbol{\alpha}_1 + \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_2 + \boldsymbol{\alpha}_3, \mathbf{1}\} \tag{3.14}$$

the set $A_3\setminus T$ contains the two elements $\mathbf{p}_1 = \boldsymbol{\alpha}_1 + \boldsymbol{\alpha}_2$ and $\mathbf{p}_2 = \boldsymbol{\alpha}_2 + \boldsymbol{\alpha}_3$ so that

$$\mathbf{n}_{A_3}(j) = j_1\mathbf{p}_1 + j_2\mathbf{p}_2 = (j_1, j_1 + j_2, j_2). \tag{3.15}$$

Hence,

$$P_{A_3}(\mathbf{k}) = \sum_{j_1, j_2} \{1 + \min(k_1 - j_1, k_2 - j_1 - j_2, k_3 - j_2)\} \tag{3.16}$$

with the constraints on the summation

$$0 \leqslant j_1 \leqslant k_1, \quad 0 \leqslant j_1 + j_2 \leqslant k_2, \quad 0 \leqslant j_2 \leqslant k_3. \tag{3.17}$$

Although the sums in Eq. (3.16) are straightforward, the answer depends on the relative size of $k_1$, $k_2$, $k_3$, and $k_1 + k_3$. The results are summarized in Table I where the various cases are labeled as in Tarski[1] to facilitate comparison. We agree in all cases except (d); we note that Tarski's expression for that case is not positive.

It should be remarked that the literature on explicit

TABLE I. The partition function $P_{A_3}(k)$. Three more cases are obtained from (b), (c), and (d) by exchanging $k_1$ and $k_3$.

| Case | $P_{A_3}(k_1, k_2, k_3)$ |
|---|---|
| (a) $k_2 \leqslant k_1, k_3$ | $\frac{1}{6}(k_2 + 1)(k_2 + 2)(k_2 + 3)$ |
| (b) $k_1 \leqslant k_2 \leqslant k_3$ | $\frac{1}{6}(k_1 + 1)(k_1 + 2)(3k_2 - 2k_1 + 3)$ |
| (c) $k_1 \leqslant k_3 \leqslant k_1 + k_3 \leqslant k_2$ | $\frac{1}{6}(k_1 + 1)(k_1 + 2)(3k_3 - k_1 + 3)$ |
| (d) $k_1 \leqslant k_3 \leqslant k_2 \leqslant k_1 + k_3$ | $\frac{1}{6}(k_1 + 1)(k_1 + 2)(3k_3 - k_1 + 3) - \binom{(k_3 + k_1 - k_2 + 2)}{3}$ |

evaluations of Kostant's partition function is extremely scant. We are able to find but one reference, the paper by Tarski,[2] dealing with $A_2$, $B_2$, $G_2$, and $A_3$.

With increasing rank $l$ direct use of Eq. (3.9) becomes more and more cumbersome. In the next section, we present a recursive approach which is often more practical.

## IV. RECURSION FORMULAS

Let $L$ and $L'$ be two simple Lie algebras of rank $l$ and $l'$, respectively, with $l \geqslant l'$, and let $\Phi_L^+ \supset \Phi_{L'}^+$. Note that we do not require $L'$ to be a subalgebra of $L$ but only that its positive roots $\Phi_{L'}^+$ be a subset of $\Phi_L^+$. We require only that the coefficients of root $\xi$ in $L$, with basis $\{\alpha_1, \alpha_2, ..., \alpha_l\}_L$ be the same as the coefficients of the root $\xi$ in $L'$ with basis $\{\alpha_1', ..., \alpha_{l'}'\}_{L'}$, or $\xi = c_1\alpha_1 + \cdots + c_l\alpha_l$ in $L$ and $\xi = c_1\alpha_1' + \cdots + c_{l'}\alpha_{l'}'$ in $L'$. Direction and length of $\xi$ in $L$ and $L'$ need not be the same. We obtain Kostant's partition function $P_L$ in terms of $P_{L'}$ from Eq. (2.8) by choosing the set $S$ to be $\Phi_L^+$ and the subset $T$ to be $\Phi_{L'}^+$

$$P_L(\mathbf{k}) = \sum P_{L'}[\mathbf{k} - \mathbf{n}_{LL'}(j)], \tag{4.1}$$

where $\Sigma$ means summation over allowed configurations of the $j_r$, and where

$$\mathbf{n}_{LL'}(j) = \sum_{r=1}^{x_{LL'}} j_r \mathbf{p}_r, \quad j_r \in N, \ \mathbf{p}_r \in L \setminus L', \tag{4.2}$$

with $x_{LL'}$ the number of elements in $L \setminus L' \equiv \Phi_L^+ \setminus \Phi_{L'}^+$, i.e., the number of positive roots of $L$ that are not roots of $L'$.

A word about notation. By definition $\mathbf{k}$ is an $l$-tuple, the elements of $\Phi_L^+$ are $l$-tuples and the elements of $\Phi_{L'}^+$ are $l'$-tuples. However, when $\Phi_{L'}^+$ is viewed as a subset of $\Phi_L^+$ its elements are $l$-tuples. A moment's reflection shows that these elements of $\Phi_{L'}^+$ are $l$-tuples which all have the same $l - l'$ components identically zero. If we agree to label these $l - l'$ components as 1, 2, ..., $l - l'$, then the condition for allowed configurations $j_r$ may be stated as

$$\mathbf{n}_{LL'}(j) \leqslant \mathbf{k}, \tag{4.3}$$

which is our standard constraint, and

$$[\mathbf{n}_{LL'}(j)]_i = k_i \quad \text{for } 1 \leqslant i \leqslant l - l', \tag{4.4}$$

which ensures that $\mathbf{k} - \mathbf{n}_{LL'}(j)$ is an $l'$-tuple.

As a first application of Eq. (4.1) we consider $L = B_l$ and $L' = A_l$. Indeed, the $l^2$ positive roots of $B_l$ may be decomposed into

$$\Phi_{B_l}^+ = \Phi_{A_l}^+ \cup B_l \setminus A_l, \tag{4.5}$$

where the $l(l+1)/2$ positive roots of $A_l$ are

$$\Phi_{Al}^+ = \left\{ \sum_{m=s}^{t-1} \alpha_m, \quad 1 \leqslant s < t \leqslant l+1 \right\} \tag{4.6}$$

and

$$B_l \setminus A_l = \left\{ \sum_{m=s}^{t-1} \alpha_m + 2\sum_{m=t}^{l} \alpha_m, \quad 1 \leqslant s < t \leqslant l \right\}. \tag{4.7}$$

Therefore,

$$P_{B_l}(\mathbf{k}) = \sum P_{A_l}[\mathbf{k} - \mathbf{n}_{B_lA_l}(j)], \tag{4.8}$$

with

$$\mathbf{n}_{B_lA_l}(j) = \sum_{r=1+x_{A_l}}^{x_{B_l}} j_r \mathbf{p}_r, \quad j_r \in N, \ \mathbf{p}_r \in B_l \setminus A_l \tag{4.9}$$

and, since $l = l'$ in this case, the allowed $j_r$ obey

$$\mathbf{n}_{B_lA_l}(j) \leqslant \mathbf{k}. \tag{4.10}$$

Since the elements of $B_l \setminus A_l$ consist of the positive roots of $B_l$ that are not roots of $A_l$ and since both $\Phi_{A_l}^+$ and $\Phi_{B_l}^+$ contain as a subset the $T$ of Eq. (3.5), we can use in Eq. (4.9) $x_{A_l}$ and $x_{B_l}$ as defined by Eq. (3.11). Note that it also follows from Eqs. (3.11) and (4.9) that

$$\mathbf{n}_{B_lA_l}(j) = \mathbf{n}_{B_l}(j) - \mathbf{n}_{A_l}(j). \tag{4.11}$$

Quite explicitly suppose $l = 3$. Then

$$B_3 \setminus A_3 = \{\alpha_1 + 2\alpha_2 + 2\alpha_3, \alpha_1 + \alpha_2 + 2\alpha_3, \alpha_2 + 2\alpha_3\}, \tag{4.12}$$

$$\begin{aligned} \mathbf{n}_{B_3A_3}(j) &= j_1(\alpha_1 + 2\alpha_2 + 2\alpha_3) \\ &\quad + j_2(\alpha_1 + \alpha_2 + 2\alpha_3) + j_3(\alpha_2 + 2\alpha_3) \\ &= (j_1 + j_2, 2j_1 + j_2 + j_3, 2j_1 + 2j_2 + 2j_3), \end{aligned} \tag{4.13}$$

and therefore,

$$\begin{aligned} P_{B_3}(\mathbf{k}) = \sum_{j_1, j_2, j_3} P_{A_3}(&k_1 - j_1 - j_2, k_2 - 2j_1 - j_2 \\ &- j_3, k_3 - 2j_1 - 2j_2 - 2j_3), \end{aligned} \tag{4.14}$$

with $j_1, j_2, j_3 \in N$ subject to the constraints

$$j_1 + j_2 \leqslant k_1, \quad 2j_1 + j_2 + j_3 \leqslant k_2, \quad 2j_1 + 2j_2 + 2j_3 \leqslant k_3. \tag{4.15}$$

As a second application of Eq. (4.1) take $L = C_l$ and $L' = A_l$. It is again true that the $l^2$ positive roots of $C_l$ may be decomposed into

$$\Phi_{C_l}^+ = \Phi_{A_l}^+ \cup C_l \setminus A_l, \tag{4.16}$$

where

$$C_l \setminus A_l = \left\{ \alpha_l + \sum_{m=s}^{t-1} \alpha_m + 2\sum_{m=t}^{l-1} \alpha_m, \quad 1 \leqslant s \leqslant t \leqslant l-1 \right\} \tag{4.17}$$

and so the analog of Eqs. (4.8)–(4.10), with $B_l$ replaced by $C_l$ everywhere, follows.

As a last application of Eq. (4.1) we take $L = A_l$, $L' = A_{l-1}$, so that $l > l'$ and the additional constraint Eq. (4.4) becomes operative. It follows from Eq. (4.6) that

$$A_l \setminus A_{l-1} = \left\{ \sum_{m=1}^{t-1} \alpha_m, \ 1 < t \leqslant l+1 \right\}, \tag{4.18}$$

$$\begin{aligned} \mathbf{n}_{A_lA_{l-1}}(j) &= j_1\alpha_1 + j_2(\alpha_1 + \alpha_2) + \cdots + j_l \sum_{m=1}^{l} \alpha_m \\ &= \left( \sum_{s=1}^{l} j_s, \sum_{s=2}^{l} j_s, ..., j_{l-1} + j_l, j_l \right). \end{aligned} \tag{4.19}$$

Clearly the additional constraint, ensuring that $\mathbf{k} - \mathbf{n}_{A_lA_{l-1}}(j)$ is an $(l-1)$-tuple, reads

$$\sum_{s=1}^{l} j_s = k_1 \tag{4.20}$$

and may be used to eliminate $j_1$. Therefore,

$$P_{A_l}(k_1, k_2, ..., k_l) = \sum_{j_2, j_3, ..., j_l} P_{A_{l-1}}\left(k_2 - \sum_{s=2}^{l} j_s,\ k_3 - \sum_{s=3}^{l} j_s, ..., k_l - j_l\right) \tag{4.21}$$

with the summation constrained by

$$\sum_{s=i}^{l} j_s \leqslant k_i \quad \text{for } 2 \leqslant i \leqslant l \tag{4.22}$$

and

$$\sum_{s=2}^{l} j_s \leqslant k_1. \tag{4.23}$$

Equivalently we may write

$$P_{A_l}(k_1, k_2, ..., k_l) = \sum_{j_l=0}^{r_l} \sum_{j_{l-1}=0}^{r_{l-1}} \cdots \sum_{j_2=0}^{r_2} P_{A_{l-1}}\left(k_2 - \sum_{s=2}^{l} j_s,\ k_3 - \sum_{s=3}^{l} j_s, ..., k_l - j_l\right), \tag{4.24}$$

where

$$r_i = \min\left(k_1 - \sum_{s=i+1}^{l} j_s, k_2 - \sum_{s=i+1}^{l} j_s, ..., k_i - \sum_{s=i+1}^{l} j_s\right), \quad 2 \leqslant i \leqslant l. \tag{4.25}$$

These results are still too general to provide explicit formulas. In the next two sections we are able to obtain more explicit answers by imposing various restrictions on $\mathbf{k}$.

## V. THE PARTITION FUNCTION $P_{A_l}(\mathbf{k})$ FOR MONOTONIC $\mathbf{k}$

Even in the relatively simple case of $P_{A_3}(\mathbf{k})$ the explicit formulas had to be classified into seven cases depending on the relative size of the components of $\mathbf{k}$. For higher rank $l$ the number of such cases increases rapidly. If $\mathbf{k}$ is specialized to be monotonic, by which we mean

$$k_i \leqslant k_j \quad \text{for } i < j, \tag{5.1}$$

or

$$k_i \geqslant k_j \quad \text{for } i < j, \tag{5.2}$$

we are able to obtain results for arbitrary rank $l$.

We assume from now on without loss of generality that Eq. (5.1) holds and denote the partition function for such $k$ by $\hat{P}_{A_l}(\mathbf{k})$. For such $\mathbf{k}$ Eqs. (4.24) and (4.25) yield

$$\hat{P}_{A_l}(\mathbf{k}) = \sum_{j_l=0}^{r_l} \sum_{j_{l-1}=0}^{r_{l-1}} \cdots \sum_{j_2=0}^{r_2} P_{A_{l-1}}\left(k_2 - \sum_{s=2}^{l} j_s,\ k_3 - \sum_{s=3}^{l} j_s, ..., k_l - j_l\right) \tag{5.3}$$

with

$$r_t = k_1 - \sum_{s=t+1}^{l} j_s, \quad 2 \leqslant t \leqslant l \tag{5.4}$$

(it being understood that $\Sigma_{s=l+1}^{l} j_s = 0$). By relabeling the summation indices according to

$$j_t' = \sum_{s=t}^{l} j_s, \quad 2 \leqslant t,\ s \leqslant l \tag{5.5}$$

(and then dropping the prime) and reversing the order of the sums we convert Eq. (5.3) into

$$\hat{P}_{A_l}(\mathbf{k}) = \sum_{j_2=0}^{k_1} \sum_{j_3=0}^{j_2} \cdots \sum_{j_l=0}^{j_{l-1}} \hat{P}_{A_{l-1}}(k_2 - j_2, k_3 - j_3, ..., k_l - j_l). \tag{5.6}$$

We now assert that for $l \geqslant 3$

$$\hat{P}_{A_l}(k_1, k_2, ..., k_l) = \tfrac{1}{3}(3k_{l-1} - 2k_{l-2} + 3) \times S_l(k_1, k_2, ..., k_{l-2}), \tag{5.7}$$

i.e., that $\hat{P}_{A_l}(\mathbf{k})$ is independent of $k_l$, depends on $k_{l-1}$ through the factor $(3k_{l-1} - 2k_{l-2} + 3)$ only, and depends on the remaining components of $\mathbf{k}$ through the function $S_l$ defined by the relation (5.7). To prove Eq. (5.7) we proceed by induction on the rank $l$. First, we have from Table I that

$$\hat{P}_{A_3}(k_1, k_2, k_3) = \tfrac{1}{6}(3k_2 - 2k_1 + 3)(k_1 + 1)(k_1 + 2),$$

which is of the indicated form with

$$S_3(k_1) = \binom{k_1 + 2}{2}. \tag{5.8}$$

Next assume the validity of Eq. (5.7) for rank $l - 1$ and insert it on the right-hand side of Eq. (5.6) to obtain

$$\begin{aligned}\hat{P}_{A_l}(\mathbf{k}) &= \sum_{j_2=0}^{k_1} \sum_{j_3=0}^{j_2} \cdots \sum_{j_{l-2}=0}^{j_{l-3}} S_{l-1}(k_2 - j_2, k_3 - j_3, ..., k_{l-2} - j_{l-2}) \\ &\quad \times \sum_{j_{l-1}=0}^{j_{l-2}} \sum_{j_l=0}^{j_{l-1}} \frac{1}{3}(3k_{l-1} - 3j_{l-1} - 2k_{l-2} + 2j_{l-2} + 3) \\ &= \frac{1}{3}(3k_{l-1} - 2k_{l-2} + 3) \sum_{j_2=0}^{k_1} \sum_{j_3=0}^{j_2} \cdots \sum_{j_{l-2}=0}^{j_{l-3}} \binom{j_{l-2} + 2}{2} S_{l-1}(k_2 - j_2, ..., k_{l-2} - j_{l-2}).\end{aligned} \tag{5.9}$$

This completes the proof of Eq. (5.7) by induction and also yields the recursion for $S_l$ for $l > 3$:

$$S_l(k_1, k_2, ..., k_{l-2}) = \sum_{j_2=0}^{k_1} \sum_{j_3=0}^{j_2} \cdots \sum_{j_{l-2}=0}^{j_{l-3}} \binom{j_{l-2} + 2}{2} S_{l-1}(k_2 - j_2, ..., k_{l-2} - j_{l-2}). \tag{5.10}$$

The solution to this recursion is

$$S_l(k_1, k_2, ..., k_{l-2}) = \mathscr{L}_1\mathscr{L}_2\cdots\mathscr{L}_{l-3}\binom{k_1 + m_{l-3}}{k_1}\prod_{t=1}^{l-3}\binom{k_{l-1-t} - k_{l-2-t} + m_{l-3,t} - l + t + 2}{m_{l-3,t}}, \tag{5.11}$$

where

$$\mathscr{L}_s \equiv \sum_{m_{s,1}=0}^{m_{s-1,1}} \sum_{m_{s,2}=0}^{m_{s-1,2}} \cdots \sum_{m_{s,s-1}=0}^{m_{s-1,s-1}} \sum_{m_{s,s}=0}^{m_{s-1}} \tag{5.12}$$

and

$$m_s = \binom{s+3}{2} - 1 - \sum_{i=1}^{s} m_{s,i}, \tag{5.13}$$

and we omit the straightforward but tedious proof by induction on the rank $l$.

As one application of this result we observe that for

$$k_{l-1-t} - k_{l-2-t} = l - t - 3 \tag{5.14}$$

all the binomials in the product over $t$ vanish, except if $m_{l-3,t} = 0$ when the binomials equal unity. Consequently

$$S_l(k, k, k+1, k+3, ..., k + (l-4)(l-3)/2) = \binom{k + \frac{l(l-1)}{2} - 1}{k} f_l, \tag{5.15}$$

where

$$f_l = \mathscr{L}_1\mathscr{L}_2\cdots\mathscr{L}_{l-4}1. \tag{5.16}$$

Specifically

$$f_4 = 1, \tag{5.17}$$

$$f_5 = \mathscr{L}_1 1 = \sum_{m_{1,1}=0}^{m_0} 1 = 3, \tag{5.18}$$

$$f_6 = \mathscr{L}_1\mathscr{L}_2 1 = \sum_{m_{1,1}=0}^{m_0} \sum_{m_{2,1}=0}^{m_{1,1}} \sum_{m_{2,2}=0}^{m_1} 1 = 28, \tag{5.19}$$

$$f_7 = \sum_{m_{1,1}=0}^{m_0} \sum_{m_{2,1}=0}^{m_{1,1}} \sum_{m_{2,2}=0}^{m_1} \sum_{m_{3,1}=0}^{m_{2,1}} \sum_{m_{3,2}=0}^{m_{2,2}} \sum_{m_{3,3}=0}^{m_2} 1 = 840, \tag{5.20}$$

etc.

As another application we offer explicit formulas for $A_4$, $A_5$, $A_6$, and $A_7$ in the case when all the components of $\mathbf{k}$ are equal:

$$S_4(k, k) = \binom{k+5}{5}, \tag{5.21}$$

$$S_5(k, k, k) = \frac{1}{3}(k+3)\binom{k+8}{8}, \tag{5.22}$$

$$S_6(k, k, k, k) = \frac{1}{3}(k+3)\frac{1}{26}(k^2 + 12k + 26)\binom{k+11}{11}, \tag{5.23}$$

$$S_7(k, k, k, k, k) = (k+3)\left[3\binom{k+15}{15} - 27\binom{k+16}{16} + \frac{244}{3}\binom{k+17}{17} - 99\binom{k+18}{18} + 42\binom{k+19}{19}\right], \tag{5.24}$$

and, of course,

$$P_{A_l}(k, k, ..., k) = \tfrac{1}{3}(k+3)S_l(k, k, ..., k). \tag{5.25}$$

Finally, consider the extreme specialization of $\mathbf{k}$ when all its components are unity

$$\mathbf{k} = \mathbf{1}. \tag{5.26}$$

For such $\mathbf{k}$ Eq. (5.6) becomes

$$P_{A_l}(\mathbf{1}) = \sum_{j_2=0}^{1} \sum_{j_3=0}^{j_2} \cdots \sum_{j_l=0}^{j_{l-1}} P_{A_{l-1}}(1 - j_2, 1 - j_3, ..., 1 - j_l). \tag{5.27}$$

The sum over $j_2$ consists of two terms, $j_2 = 0$ and $j_2 = 1$, each of which contributes precisely $P_{A_{l-1}}(\mathbf{1})$ so that we have

$$P_{A_l}(\mathbf{1}) = 2P_{A_{l-1}}(\mathbf{1}), \tag{5.28}$$

which combined with $P_{A_1} = 1$ yields

$$P_{A_l}(\mathbf{1}) = 2^{l-1}. \tag{5.29}$$

It turns out that the partition of $\mathbf{1}$ equals $2^{l-1}$ for *any* simple Lie algebra, not just $A_l$. A general algorithm, proving this statement, is given in the next section.

## VI. THE PARTITION FUNCTION $P_L(\mathbf{1})$

In this section we prove the proposition

$$P_L(\mathbf{1}) = 2^{l-1}, \tag{6.1}$$

where $L$ is *any* simple Lie algebra of rank $l$.

For the sake of convenience we shall call any linear combination of positive roots totaling $\mathbf{1}$ a partition of $\mathbf{1}$. Clearly partitions of $\mathbf{1}$ can involve as summands only roots containing single or zero multiples of the simple roots. Let $O_l$ be the set of $l$-tuples satisfying

$$O_l = \{(x_1, x_2, ..., x_l) | x_i = 0 \text{ or } 1\} = \left\{\sum_{i=1}^{l} x_i\alpha_i \,\middle|\, x_i = 0 \text{ or } 1\right\}. \tag{6.2}$$

Then

$$P_L(\mathbf{1}) = P_{\theta_L}(\mathbf{1}), \tag{6.3}$$

where

$$\theta_L \equiv O_l \cap \Phi_L^+. \tag{6.4}$$

The cardinality of the set $\theta_L$ is less than or equal to that of $\Phi_L^+$. We suggest the following method for determining $\theta_L$.

*Lemma*: Let $D_L$ be the Dynkin diagram for $L$. Then $\theta_L$ is the set of $l$-tuples obtained by taking the root $\mathbf{1}$ from each subalgebra $M$ of $L$, where the allowed subalgebras $M$ have Dynkin diagrams $D_M$ consisting of connected fragments of $D_L$.

*Proof*: Consider a connected fragment $D_M$ which consists of roots $\alpha_i, \alpha_{i+1}, ..., \alpha_{i+j}$ of $L$. Then the root $\mathbf{1}$ for $M$ is $\alpha_i + \alpha_{i+1} + \cdots + \alpha_{i+j}$. Since $M \subset L$, the root system of $M$ is entirely contained in the root system of $L$. Hence the root $\mathbf{1}$ of $M$ is a root of $L$. Since $D_M$ was any connected fragment

this proves

$$\Lambda \subset \theta_L, \tag{6.5}$$

where $\Lambda$ is the set of all roots **1** for all choices of $M$.

Next let $\beta$ be any element of $\theta_L$ and consider how the roots of $L$ are obtained from the simple roots $\alpha_1, \ldots, \alpha_l$. Recall that $\alpha_i + \alpha_j$ is a root provided $\langle \alpha_i, \alpha_j \rangle < 0$, where $\langle \alpha_i, \alpha_j \rangle$ is the Cartan integer in the $i$th row and $j$th column of the Cartan matrix. This means that $\alpha_i$ and $\alpha_j$ must be connected by at least a single bound in the Dynkin diagram of $L$. Continuing to build up $\beta$ by adding in more simple roots we see that $\beta = \alpha_i + \alpha_j + \cdots + \alpha_k$ is a root only if each $\alpha_m$ in $\beta$ is connected to at least one other $\alpha_p$ in $\beta$. Thus we have shown that any element in $\theta_L$ must be a root in a Lie algebra corresponding to a connected portion of $D_L$, or, by an appropriate choice of the fragment $D_M$, that

$$\Lambda \supset \theta_L. \tag{6.6}$$

Combining Eqs. (6.5) and (6.6) yields

$$\Lambda = \theta_L, \tag{6.7}$$

and this proves the lemma.

We now use this lemma to prove our proposition Eq. (6.1) by considering what is usually a much smaller set of roots than all of $\Phi_L^+$.

We proceed by induction with the following strategy. The Eq. (6.1) is obviously true for $L = A_1 = B_1 = C_1$ by elementary considerations. Next we find the largest subalgebra $M$ of $L$, generated by a connected fragment $D_M$ of $D_L$, and determine the roots in the set

$$O_l \cap L \setminus M. \tag{6.8}$$

(Recall that $L \setminus M$ is the set of positive roots of $L$ with the positive roots of $M$ left out.) Let $\gamma$ be in this set and call $M_\gamma$ that subalgebra of $L$ for which this $\gamma$ is the root **1**. Let the complementary fragment of $D_L$, obtained upon removal of $D_{M_\gamma}$, be called $D^c_{M_\gamma}$. Then the partitions of **1** in $L$ generated by $\gamma$ are all of the form

$$\gamma + \gamma^c = \mathbf{1}, \tag{6.9}$$

where $\gamma^c$ is a linear combination of roots in $M^c_\gamma$, the subalgebra of $L$ complementary to $M_\gamma$. By summing over all such linear combinations we obtain $P_{M^c_\gamma}(\mathbf{1})$, hence

$$P_L(\mathbf{1}) = \sum_{\{\gamma\}} P_{M^c_\gamma}(\mathbf{1}). \tag{6.10}$$

TABLE II. The roots $\gamma$ in $O_l \cap A_l \setminus A_{l-1}$ and the corresponding complementary fragments $D^c_{M_\gamma}$.

| $\gamma$ | $D^c_{M_\gamma}$ |
|---|---|
| $\alpha_l$ | $D_{A_{l-1}}$ |
| $\alpha_l + \alpha_{l-1}$ | $D_{A_{l-2}}$ |
| $\vdots$ | $\vdots$ |
| $\alpha_l + \alpha_{l-1} + \cdots + \alpha_2$ | $D_{A_1}$ |
| $\alpha_l + \alpha_{l-1} + \cdots + \alpha_2 + \alpha_1$ | * |

TABLE III. The roots $\gamma$ in $O_l \cap B_l \setminus B_{l-1}$ and the corresponding complementary fragments $D^c_{M_\gamma}$.

| $\gamma$ | $D^c_{M_\gamma}$ |
|---|---|
| $\alpha_1$ | $D_{B_{l-1}}$ |
| $\alpha_1 + \alpha_2$ | $D_{B_{l-2}}$ |
| $\vdots$ | $\vdots$ |
| $\alpha_1 + \cdots + \alpha_{l-1}$ | $D_{B_1}$ |
| $\alpha_1 + \cdots + \alpha_{l-1} + \alpha_l$ | * |

### A. The $A_l$ algebra

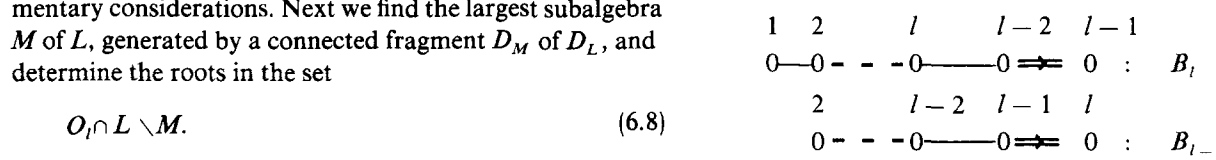
The roots $\gamma$ in $O_l \cap A_l \setminus A_{l-1}$ and the corresponding complementary fragments $D^c_{M_\gamma}$ are given in Table II. Note that the last line is itself a partition of **1** and contributes 1 on the rhs of Eq. (6.10).

By the induction hypothesis $P_{A_m}(\mathbf{1}) = 2^{m-1}$, $m < l$, and by Eq. (6.10)

$$P_{A_l}(\mathbf{1}) = 1 + \sum_{i=1}^{l-1} P_{A_i}(\mathbf{1}) = 1 + \sum_{i=1}^{l-1} 2^{i-1} = 2^{l-1}. \tag{6.11}$$

### B. The $B_l$ algebra

Using the labeling




we have the roots $\gamma$ in $O_l \cap B_l \setminus B_{l-1}$ and the corresponding complementary fragments $D^c_{M_\gamma}$ given in Table III.

By the induction hypothesis $P_{B_m}(\mathbf{1}) = 2^{m-1}$, $m < l$, and by Eq. (6.10)

$$P_{B_l}(\mathbf{1}) = 1 + \sum_{i=1}^{l-1} P_{B_i}(\mathbf{1}) = 1 + \sum_{i=1}^{l-1} 2^{i-1} = 2^{l-1}. \tag{6.12}$$

### C. The $C_l$ algebra

The proof for $C_l$ is identical to that for $B_l$ since we can use the same choice of roots $\gamma$ and the same series $D^c_{M_\gamma}$ with $B$ replaced by $C$.

TABLE IV. The roots $\gamma$ in $O_l \cap D_l \setminus D_{l-1}$ and the corresponding complementary fragments $D^c_{M_\gamma}$.

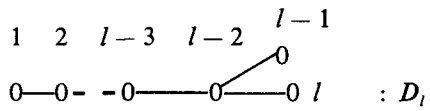
| $\gamma$ | $D^c_{M_\gamma}$ |
|---|---|
| $\alpha_1$ | $D_{D_{l-1}}$ |
| $\alpha_1 + \alpha_2$ | $D_{D_{l-2}}$ |
| $\vdots$ | $\vdots$ |
| $\alpha_1 + \cdots + \alpha_{l-3}$ | $D_{D_3}$ |
| $\alpha_1 + \cdots + \alpha_{l-2}$ | $D_{(A_1 + A_1)}$ |
| $\alpha_1 + \cdots + \alpha_{l-2} + \alpha_l$ | $D_{A_1}$ |
| $\alpha_1 + \cdots + \alpha_{l-2} + \alpha_{l-1}$ | $D_{A_1}$ |
| $\alpha_1 + \cdots + \alpha_{l-2} + \alpha_{l-1} + \alpha_l$ | * |

TABLE V. The roots $\gamma$ in $O_l \cap E_l \setminus E_{l-1}$ and the corresponding complementary fragments $D^c_{M_\gamma}$.

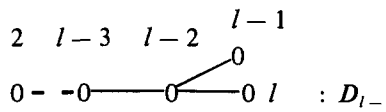
| $\gamma$ | $D^c_{M_\gamma}$ |
|---|---|
| $\alpha_l$ | $D_{E_{l-1}}$ |
| $\alpha_l + \alpha_{l-1}$ | $D_{E_{l-2}}$ |
| ⋮ | ⋮ |
| $\alpha_l + \cdots + \alpha_6$ | $D_{E_5}$ |
| $\alpha_l + \cdots + \alpha_5$ | $D_{A_4}$ |
| $\alpha_l + \cdots + \alpha_4$ | $D_{(A_2 + A_1)}$ |
| $\alpha_l + \cdots + \alpha_4 + \alpha_2$ | $D_{A_2}$ |
| $\alpha_l + \cdots + \alpha_4 + \alpha_3$ | $D_{(A_1 + A_1)}$ |
| $\alpha_l + \cdots + \alpha_4 + \alpha_3 + \alpha_2$ | $D_{A_1}$ |
| $\alpha_l + \cdots + \alpha_4 + \alpha_3 + \alpha_1$ | $D_{A_1}$ |
| $\alpha_l + \cdots + \alpha_4 + \alpha_3 + \alpha_2 + \alpha_1$ | * |

### D. The $D_l$ algebra, $l \geqslant 3$

Since $D_3 = A_3$ the proposition is true for $l = 3$. Using the labeling




: $D_l$




: $D_{l-1}$

we have the roots $\gamma$ in $O_l \cap D_l \setminus D_{l-1}$ and the corresponding complementary fragments $D^c_{M_\gamma}$ given in Table IV. Hence by the induction hypothesis and Eq. (6.10)

$$P_{D_l}(\mathbf{1}) = 1 + 1 + 1 + 1 + \sum_{i=3}^{l-1} 2^{i-1} = 2^{l-1}. \tag{6.13}$$

### E. The $G_2$ and $F_4$ algebras

We note that

$$O_l \cap \Phi^+_{G_2} = \Phi^+_{A_2}, \tag{6.14}$$

$$O_l \cap \Phi^+_{F_4} = \Phi^+_{A_4}. \tag{6.15}$$

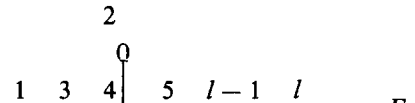
Hence the sets $O_l \cap G_2 \setminus A_2$ and $O_l \cap F_4 \setminus A_4$ are empty and

$$P_{G_2}(\mathbf{1}) = P_{A_2}(\mathbf{1}) = 2^{2-1}, \tag{6.16}$$

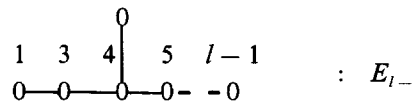
$$P_{F_4}(\mathbf{1}) = P_{A_4}(\mathbf{1}) = 2^{4-1}. \tag{6.17}$$

### F. The $E_l$ algebras, $l = 6, 7, 8$

Since $E_5 = D_5$ the proposition is true for $l = 5$. Using the labeling




: $E_l$




: $E_{l-1}$

we have the roots $\gamma$ in $O_l \cap E_l \setminus E_{l-1}$ and the corresponding complementary fragments $D^c_{M_\gamma}$ given in Table V. Hence by the induction hypothesis and Eq. (6.10)

$$P_{E_l}(\mathbf{1}) = 1 + 1 + 1 + 1 + 2 + 2 + \sum_{i=4}^{l-1} 2^{i-1} = 2^{l-1}. \tag{6.18}$$

We remark that in obtaining Eqs. (6.17) and (6.18) we also used

$$P_{L_1 + L_2}(\mathbf{k}_1 + \mathbf{k}_2) = P_{L_1}(\mathbf{k}_1) P_{L_2}(\mathbf{k}_2), \tag{6.19}$$

which expresses the obvious connection between the partition function for a semisimple Lie algebra $L = L_1 + L_2$ and the partition functions for the component algebras $L_1$ and $L_2$.


## ACKNOWLEDGMENTS

We wish to express our thanks to Dr. Georgia Benkart for many helpful comments. This research was supported in part by the Department of Energy under contract DE-AC02-76ER00881.



[1] B. Kostant, Trans. Am. Math. Soc. **93**, 53–73 (1959).
[2] J. Tarski, J. Math. Phys. **4**, 569 (1963).
[3] B. Gruber and H. J. Weber, Proc. R. Ir. Acad. Sec. A **66**, 31 (1968).
[4] J. P. Antoine and D. Speiser, J. Math. Phys. **5**, 1226 (1964).

# Shape functions for separable solutions to cross-field diffusion problems

C. D. Luning
*Department of Mathematics, Sam Houston State University, Huntsville, Texas 77340*
W. L. Perry
*Department of Mathematics, Texas A&M University, College Station, Texas 77843*

(Received 13 September 1983; accepted for publication 13 January 1984)

The shape function $S(x)$, which arises in the study of nonlinear diffusion for cross-field diffusion in plasmas, satisfies the equation $S''(x)+\lambda a(x)S^{\alpha}(x)=0$, $0<x<1$, $S(0)=S(1)=0$, $\alpha>0$. In the cases of physical interest $a(x)$ possesses an integrable singularity at some point in (0,1) but is otherwise continuous. Existence of a positive solution to this problem is established.

PACS numbers: 02.30. – f, 02.30.Hq, 52.25.Dg

## I. INTRODUCTION

The nonlinear diffusion equation $(d/dx)[D(n)(dn/dx)]=a(x)(dn/dt)$, $0\leqslant x\leqslant 1$ has been used[1–3] in the study of cross-field diffusion in toroidal multipole plasmas. In the equation, $n$ is the particle density, $a(x)$ is a positive function determined by the poloidal field and $D(n)$, the diffusion coefficient, is a nonlinear function of the density. By seeking separable solutions Berryman[1] was led to consider the nonlinear eigenvalue problem

$$S''(x)+\lambda a(x)S^{\alpha}(x)=0,\quad 0<x<1,\quad \alpha>0,$$
$$S(0)=S(1)=0 \tag{1}$$

for the shape function $S$. The function $S$ is called a shape function because in the separable solution $S$ is the factor containing the spatial dependence. The parameter $\lambda$ is the separation constant, and $a(x)$ is a positive integrable function, singular at $x=x_s$ and continuous for $x\neq x_s$.

The purpose of this paper is to prove existence of a positive solution and at the same time develop a computational method for the solution of (1). The monotone iterative methods originally developed by Keller and Cohen[4] and placed in the context of ordered Banach spaces by Amann[5] cannot be used for problem (1) since $a(x)$ is singular on (0,1). The previous results of the authors on this problem as well as the related results[6–9] on generalized Emden–Fowler equations require $a(x)$ to be continuous on [0,1]. However, these latter methods can be extended to the present case by adopting the Lebesgue integral and allowing (1) to hold almost everywhere on (0,1); that is what will be shown in this paper.

## II. EXISTENCE OF SOLUTION: AN ITERATIVE METHOD

Let $G(x,\xi)$ be the Green's function for $u''=0$, $u(0)=u(1)=0$ and let $S_0(x)=x$. Define the sequences $\{S_k(x)\}_{k=1}^{\infty}$, $\{\lambda_k\}_{k=1}^{\infty}$ by

$$\lambda_k=1\Big/\int_0^1(1-\xi)a(\xi)S_{k-1}^{\alpha}(\xi)\,d\xi,$$
$$S_k(x)=\lambda_k\int_0^1 G(x,\xi)a(\xi)S_{k-1}^{\alpha}(\xi)\,d\xi, \tag{2}$$

$k=1,2,3,\ldots$, where the integrals are taken in the Lebesgue sense.

We first show that this sequence is well defined almost everywhere. The Green's function is

$$G(x,\xi)=\begin{cases}x(1-\xi), & x\leqslant\xi,\\ \xi(1-x), & x>\xi,\end{cases}$$

so we have

$$S_k(x)=\lambda_k\Big((1-x)\int_0^x\xi a(\xi)S_{k-1}^{\alpha}(\xi)\,d\xi + x\int_x^1(1-\xi)a(\xi)S_{k-1}^{\alpha}(\xi)\,d\xi\Big), \tag{3}$$

$k=1,2,3,\ldots$. $S_1(x)$ is absolutely continuous since the indefinite integral of an integrable function is absolutely continuous. Inductive use of (3) shows that the functions $S_k$, $k=1,2,\ldots$ are defined and absolutely continuous on [0,1], and the sequence $\{\lambda_k\}_{k=1}^{\infty}$ is well defined.

We will show that the sequences $\{S_k\}_{k=0}^{\infty}$, $\{\lambda_k\}_{k=1}^{\infty}$ converge to a continuous function S(x) and a real number $\lambda$ with $S_k\searrow S$ and $\lambda_k\nearrow\lambda$ as $k\to+\infty$. We will also show that $(\lambda,S)$ is a solution pair for (1) in the sense that S'(x) is absolutely continuous and Eq. (1) holds almost everywhere.

In the next three lemmas we prove properties of the sequences $\{S_k(x)\}_{k=1}^{\infty}$ and $\{\lambda_k\}_{k=1}^{\infty}$ which will be used to establish the existence of a solution to (1).

*Lemma 1*: For $k=1,2,3,\ldots$, $S_k'(0)=1$.

*Proof*: From (3) and the usual result on calculating indefinite integrals of integrable functions we have

$$S_k'=\lambda_k\Big\{\int_x^1 a(\xi)S_{k-1}^{\alpha}(\xi)\,d\xi-\int_0^1\xi a(\xi)S_{k-1}^{\alpha}(\xi)\,d\xi\Big\} \tag{4}$$

almost everywhere. Since the right-hand side of (4) defines an absolutely continuous function, the discontinuities of $S'(x)$ are removable. Therefore, as $x\to 0$ we have $S_k'(x)\to\lambda_k\int_0^1(1-\xi)a(\xi)S_{k-1}^{\alpha}(\xi)\,d\xi=1$. We note that $S_k''=-\lambda_k a(x)S_{k-1}^{\alpha}(x)$ except at the singularity $x=x_s$.

*Lemma 2*: For $k\geqslant 1$, $0<S_k(x)<S_{k-1}(x)$, $0<x<1$ and $0<\lambda_{k-1}<\lambda_k$.

*Proof*: Consider first $v_0(x)=(S_0-S_1)(x)$. We will show that $v_0(x)>0$, $0<x<1$, and $f_0(x)=M^{1/\alpha}S_0(x)-S_1(x)$ has at most one zero on (0,1) for any $M$, $0<M<1$. The latter result is needed for the induction argument to follow.

 0022-2488/84/082374-03$02.50  

Note that $v_0''(x) = \lambda_1 a(x) S_0^\alpha(x) > 0$ except at $x = x_s$ (the singularity), where it is undefined; also $v_0(0) = 0$, $v_0'(0) = 0$. Therefore, $v_0(x) > 0$, $0 < x < 1$. Considering $f_0$, $f_0''(x) = \lambda_1 a(x) S_0^\alpha(x) > 0$ except at $x = x_s$, where it is undefined and $f_0(0) = 0, f_0'(0) = M^{1/\alpha} - 1 < 0$. Therefore, $f_0(x)$ is zero at most once on $(0,1)$.

For the induction hypothesis suppose $v_{k-1}(x) = (S_{k-1} - S_k)(x) > 0$, $0 < x < 1$ and that $f_{k-1}(x) = M^{1/\alpha} S_{k-1}(x) - S_k(x)$ has at most one zero on $(0,1)$ for any $M \in (0,1)$. Since $S_k(x) < S_{k-1}(x)$, (2) implies that $\lambda_{k+1} > \lambda_k$. Now let $v_k(x) = (S_k - S_{k+1})(x)$ and $f_k = M^{1/\alpha} S_k(x) - S_{k+1}(x)$. We have $v_k(0) = v_k'(0) = 0$ and

$$v_k''(k) = (S_k - S_{k+1})''(x) = a(x)[\lambda_{k+1} S_k^\alpha(x) - \lambda_k S_{k-1}^\alpha(x)]. \tag{5}$$

The last term in (5) has a zero on $(0,1)$ if and only if

$$(\lambda_k/\lambda_{k+1})^{1/\alpha} S_{k-1}(x) - S_k(x) \tag{6}$$

has a zero on $(0,1)$. Since $\lambda_k/\lambda_{k+1} < 1$ the induction hypothesis shows that $v_k''(x)$ has at most one zero on $(0,1)$. Moreover (5) shows that $v_k''(x)$ is positive near $x = 0$. Thus we have $v_k(0) = v_k'(0) = v_k(1) = 0$, $v_k''(x) > 0$ for $x$ near zero and $v_k''(x)$ changes sign at most, once on $(0,1)$. Therefore, $v_k(x) > 0$, $0 < x < 1$, and so $S_k(x) > S_{k+1}(x)$, $0 < x < 1$.

To show $f_k$ has at most one zero on $(0,1)$ we can proceed similarly. In this case we have

$$f_k''(x) = a(x)[\lambda_{k+1} S_k^\alpha - M^{1/\alpha} \lambda_k S_{k-1}^\alpha(x)],$$

which has a zero on $(0,1)$ if and only if

$$(M^{1/\alpha} \lambda_k/\lambda_{k+1})^{1/\alpha} S_{k-1}(x) - S_k(x)$$

does. Again the induction hypothesis shows that this term has at most one zero on $(0,1)$. Now $f_k(0) = f_k(1) = 0$, $f_k'(0) < 0, f_k''(x) > 0$ for $x$ near zero, so $f_k(x)$ has at most one zero on $(0,1)$.

*Lemma 3*: There exists a nonnegative continuous function $w(x)$ defined on $[0,1]$, positive on $(0,\epsilon)$, with $S_k(x) \geqslant w(x)$ for all $x \in [0,1]$, and all $k$.

*Proof*: Let

$$T(x) = \begin{cases} x, & 0 \leqslant x \leqslant \frac{1}{2}, \\ 1 - x, & \frac{1}{2} \leqslant x \leqslant 1. \end{cases} \tag{7}$$

Since for $k \geqslant 1$ $S_k(0) = S_k(1) = 0$, $S_k(x) \geqslant 0$, $S_k''(x) < 0$ except at $x_s$, where it is undefined, it follows that $S_k(x) \geqslant \|S_k\|_\infty T(x)$, where $\|S_k\|_\infty = \sup_{0<x<1} |S_k(x)|$. From (2),

$$S_k(x) \geqslant \lambda_k \|S_{k-1}\|_\infty^\alpha \int_0^1 G(x,\xi) a(\xi) T^\alpha(\xi) d\xi,$$

and since $0 < S_k(x) < S_0(x) < 1$, $0 < x < 1$, it follows that

$$1 > \lambda_k \|S_{k-1}\|_\infty^\alpha \int_0^1 G(x,\xi) a(\xi) T^\alpha(\xi) d\xi,$$

and thus

$$\|S_{k-1}\|_\infty^\alpha < \lambda_k^{-1} \left\| \int_0^1 G(x,\xi) a(\xi) T^\alpha(\xi) d\xi \right\|_\infty^{-1}.$$

Hence,

$$-S_k''(x) = \lambda_k a(x) S_{k-1}^\alpha(x) < a(x) \left\| \int_0^1 G(x,\xi) a(\xi) T^\alpha(\xi) d\xi \right\|_\infty^{-1} \tag{8}$$

except at $x = x_s$, where both sides are undefined. That is, there is a $C > 0$ such that $S_k''(x) > -Ca(x)$, $0 < x < 1$, $x \neq x_s$. Using $S_k(0) = 0$, $S_k(1) = 0$, $S_k'(0) = 1$, and integrating we have

$$S_k(x) \geqslant x - C \int_0^x \int_0^t a(\tau) d\tau\, dt.$$

Now let $x \leqslant x_0 < x_s$. Since $a(\tau)$ is continuous on $[0,x_0]$, it possesses a maximum on $[0,x_0]$, say $M$. Therefore, on $[0,x_0]$, $S_k(x) \geqslant x - CMx^2/2$. This establishes the existence of $w(x)$.

At this point we have shown that the sequence $\{S_k(x)\}_{k=1}^\infty$ is well defined, decreasing and bounded below by a positive function $w(x)$. Thus from (2)

$$\lambda_k \leqslant \left( \int_0^1 (1 - \xi) a(\xi) w^\alpha(\xi) d\xi \right)^{-1}$$

so that $\{\lambda_k\}_{k=1}^\infty$ is a monotone increasing sequence bounded above; therefore, it is convergent to some number $\lambda$. Since the $S_k(x)$ are absolutely continuous,

$$|S_k(x_1) - S_k(x_2)| = \left| \int_{x_1}^{x_2} S_k'(t) dt \right| \leqslant \int_{x_1}^{x_2} |S_k'(t)| dt.$$

From (4) and the fact that $0 < \lambda_k < \lambda$ we have $|S_k'(t)| \leqslant \lambda J$, where $J$ does not depend on $k$. Therefore,

$$|S_k(x_1) - S_k(x_2)| \leqslant \lambda J |x_1 - x_2|$$

and the sequence $\{S_k(x)\}_{k=1}^\infty$ is seen to be equicontinuous. Thus since $0 < S_k(x) \leqslant 1$, $k = 1,2,\ldots$, we have a uniformly bounded equicontinuous set of functions; Arzela's theorem[10] applies and we can assert the existence of a subsequence of $\{S_k(x)\}$ which converges uniformly to a continuous function $S(x) (\geqslant w(x))$ on $[0,1]$. Since the sequence $\{S_k(x)\}$ is monotone, the entire sequence converges uniformly to $S(x)$.

Now we want to show that this pair $(\lambda, S)$ solves (1). Since the functions $\lambda_k G(x,\xi) a(\xi) S_k^\alpha(\xi)$ are bounded above by an integrable function and converge a.e. (as $k \to +\infty$) to $\lambda G(x,\xi) a(\xi) S^\alpha(\xi)$ we may apply the Lebesgue convergence theorem to assert that

$$S(x) = \lambda \int_0^1 G(x,\xi) a(\xi) S^\alpha(\xi) d\xi.$$

Therefore, using (3) with $S_k$ and $S_{k-1}$ replaced by $S$ we find that $S'(x)$ is defined and absolutely continuous and that

$$S''(x) = -\lambda a(x) S^\alpha(x) \quad \text{a.e. on } (0,1),$$
$$S(0) = S(1) = 0.$$

In fact this last equation holds for all $x \neq x_s$. So we have proved

**Theorem 1:** The eigenvalue problem

$$S''(x) + \lambda a(x) S^\alpha(x) = 0, \quad 0 < x < 1, \tag{9}$$
$$S(0) = S(1) = 0,$$

where $\alpha > 0$, $a(x) > 0$ on $(0,1)$, and $a(x)$ is continuous except for an integrable singularity at $x = x_s$, has a positive solution pair in the sense described earlier. Moreover, it is obtainable as the uniform limit of the sequence $\{(\lambda_k, S_k)\}$ defined by

$$S_0(x) = x,$$
$$S_k(x) = \lambda_k \int_0^1 G(x,\xi) a(\xi) S_{k-1}^\alpha(\xi) d\xi,$$

$$\lambda_k = \left( \int_0^1 (1-\xi) a(\xi) S_{k-1}^{\alpha}(\xi) d\xi \right)^{-1},$$

$k = 1,2,...$, and $S_k \searrow S$, $\lambda_k \nearrow \lambda$, as $k \to \infty$.

*Notes*: 1. If $(\lambda, S)$ is a solution pair to (1), then so is $(\lambda c^{1-\alpha}, cS)$; thus the choice of $\lambda_k$ at each step which forced $S_k'(0) = 1$ is simply a normalization.

2. If $a(x)$ has a finite number of integrable singularities on (0,1) the theorem still holds.

3. If $a(x) \equiv 0$ on $(0, x_s)$ then the proof of Lemma 3 fails. In this case, the result of Lemma 3 could be established by considering the equivalent differential equation formulation of (2),

$$S_k''(x) = \begin{cases} 0, & 0 < x < x_s, \\ -\lambda_k a(x) S_{k-1}^{\alpha}(x), & x_s < x < 1, \end{cases}$$

$$S_k(0) = S_k(1) = 0, \quad S_k'(0) = 1.$$

In this way it is seen that $S_k(x) = x$, $0 < x < x_s$, for all $k$. From the fact that $S_k''(x) \leqslant 0$, $x_s < x < 1$, we see that a choice for $w(x)$ is

$$w(x) = \begin{cases} x, & 0 < x < x_s, \\ \dfrac{x_s}{1 - x_s}(1 - x), & x_s < x < 1. \end{cases}$$

In fact, $(\lambda = 1/(1 - x_s) x_s^{\alpha}, w(x))$ is the solution to

$$S'' + \lambda \delta(x - x_s) S = 0,$$

$$S(0) = S(1) = 0,$$

$$S'(0) = 1.$$


[1] J. G. Berryman, J. Math. Phys. **18**, 2108 (1977).
[2] J. G. Berryman and C. J. Holland, J. Math. Phys. **19**, 247 (1978).
[3] J. G. Berryman, J. Math. Phys. **21**, 1326 (1980).
[4] H. B. Keller and D. S. Cohen, J. Math. Mech. **16**, 1361 (1967).
[5] H. Amann, SIAM Rev. **18**, 620 (1976).
[6] C. D. Luning and W. L. Perry, J. Math. Phys. **22**, 1591 (1981).
[7] C. D. Luning and W. L. Perry, J. Diff. Eq. **33**, 359 (1979).
[8] C. D. Luning and W. L. Perry, J. Comput. Phys. **38**, 396 (1980).
[9] C. D. Luning and W. L. Perry, SIAM J. Math. Anal. **12**, 874 (1981).
[10] R. Courant and D. Hilbert, *Methods of Mathematical Physics* (Wiley, New York, 1953), Vol. 1, p. 59.

# A new method for summation of divergent power series

Gustavo A. Arteca, Francisco M. Fernández, and Eduardo A. Castro
*Instituto de Investigaciones Fisicoquímicas Teóricas y Aplicadas (INIFTA), División Química Teórica, Sucursal 4, Casilla de Correo 16, La Plata 1900, Argentina*

(Received 19 April 1983; accepted for publication 22 December 1983)

A new method for summation of divergent power series is developed. It only requires the knowledge of the form of both the small and large $\lambda$-power expansion ($\lambda$ being the perturbation parameter) and few coefficients of one of them to yield excellent results. Convergence is proved for a simple two-level model, and reasonable arguments are given for more complex and interesting models. The method is quite general and contains some resummation techniques reported previously as particular cases. The anharmonic, mean square, displacement function, the ground-state eigenvalue of the quantum-mechanical anharmonic oscillator, and the ground-state energy of the hydrogen atom in a magnetic field calculated in this way are shown to be of striking accuracy in the whole range of the perturbation parameter.

PACS numbers: 02.30.Lt, 03.65.Db

## I. INTRODUCTION

Perturbation theory is one of the most powerful and widely extended mathematical techniques in theoretical physics. Since the majority of physically interesting models which can be treated in this way lead to divergent series, much effort was directed to understand the reasons for this divergence. Furthermore, several summation procedures were developed and discussed during the last years, the most commonly employed ones being Padé approximants[1,2] and the Borel transform.[3]

Frequently, the problem under study leads to two different power series expansions in the small and large $\lambda$ regime ($\lambda$ being the perturbation parameter), respectively. The functions obtained through Padé approximants or Borel transform do not fulfill both series except in a few particular cases. Several renormalized perturbation techniques were proposed recently[4–16] to overcome this difficulty. In general, these procedures which allow both series to be fitted improve convergence largely. Notwithstanding, there remains much to be done in the field of divergent power series, especially due to the universality of perturbation theory approaches.

During the last years, the search for approximate eigenvalues of parameter-dependent Hamiltonians led to a new procedure based on semiclassical and variational methods.[17–21] It consists of approaching the eigenvalues by means of a variational functional that obeys a differential equation similar to that satisfied by the exact eigenvalues. The differential equation is obtained from the virial and Hellmann–Feynman theorem.[17,18]

In this paper we propose a new way to handle those problems that lead to two power series expansions, one for the small and the other one for the large $\lambda$ regime. It consists in building a suitable functional which satisfies both series. The procedure, which will be termed the functional method (FM) from now on, only requires knowing the form of both series and a few coefficients of one of them to yield highly accurate results in the whole range of $\lambda$ values. Although the form of our functional was suggested by the aforementioned variational functional procedure,[17–21] the method we propose here is wholly original, for the point of view is thoroughly different, as will be shown later.

Those series which our method can be applied to are also Padé and Borel summable, but the present procedure offers several advantages. It is easier to handle and gives us the proper analytical structure (i.e., the correct power series in the small as well as large $\lambda$ regime) of the unknown function that is approximated. Since the functional is built from the power series directly, the FM can be suitable for problems which are not associated with a Hamiltonian. This fact allows the FM to be useful in many branches of physics other than quantum mechanics. Furthermore, when summing one of the series, the convergence can be largely improved by properly adding the main term of the other one, provided it is known.

In Sec. II we set down the method and show how to build the functional that approximates some rather general power series expansions. In Sec. III, some previously reported resummation techniques are shown to be particular cases of (or simply related to) the present method and also to relate each other closely.

Convergence properties of the resulting functional series are discussed at length in Sec. IV by way of a two-level model that exhibits two level crossings in the complex plane. It is shown that the FM changes the finite convergence radius of the Rayleigh–Schrödinger series into an infinite one.

In Sec. V the method is applied to more complex problems that have real physical interest such as the anharmonic, mean-square, displacement function, the quantum-mechanical anharmonic oscillator, and Zeeman effect for the hydrogen atom. In all cases results are in excellent agreement with the exact (numerical) ones and far better than those obtained through the original power series. The FM is also more accurate than other resummation techniques.

Even though we are not able to prove rigorously in which cases the FM series converge, we will reasonably show that our procedure is very suitable for obtaining accurate results from divergent power series.

Finally, some other problems which can be handled in this way are briefly discussed in Sec. VI.

 0022-2488/84/082377-13$02.50  

## II. THE METHOD

Let us suppose that we are interested in calculating, as exactly as possible, the values of an unknown real function $E(\lambda)$, which can be expanded as

$$E(\lambda) = \sum_{i=0}^{\infty} E_i \lambda^i, \tag{2.1a}$$

$$E(\lambda) = \lambda^{\beta} e(\lambda) = \lambda^{\beta} \sum_{i=0}^{\infty} e_i \lambda^{\alpha i}, \quad \alpha < 0, \tag{2.1b}$$

when $\lambda$ is small and large enough, respectively ($\alpha$ and $\beta$ are real numbers). In general, the convergence radii of these series are not known, and they may be finite, infinite, or zero. We do not make any assumption about this. Even when the convergence radius of one of these series is infinite, results are not good enough in the whole range of $\lambda$ values, for the series is to be truncated up to the last known term.

In this paper we will show that excellent values for $E(\lambda)$ can be obtained by means of the functional

$$F = A/q^2 + q^s + \lambda q^t B(\lambda, q), \tag{2.2}$$

where $A$ is a real constant, $q$ is a new variable, and $B(\lambda, q)$ is a function that will be determined later on.

The form of the functional (2.2) is suggested by previous works on a variational-semiclassical approximation to the eigenvalues of anharmonic and coupled oscillators.[17-21] In this scheme, $A/q^2$ plays the role of the kinetic energy and the remaining of the rhs in (2.2) is a sort of expectation value of the potential energy. But this interpretation is nonsense here because we do not suppose that $E(\lambda)$ is an eigenvalue of a given Hamiltonian operator.

The change of variables $q = \lambda^{-1/(t+2)} w$ allows us to write (2.2) as

$$F = \lambda^{2/(t+2)} \{A/w^2 + \lambda^{-(s+2)/(t+2)} w^s + w^t B(\lambda, q)\}. \tag{2.3}$$

If $F$ is to fulfill (2.1), then it is necessary that

$$s = -(2/\beta)(\alpha + \beta), \quad t = (2/\beta)(1 - \beta), \tag{2.4}$$

and, besides, $B(\lambda, q)$ has to be able to be expanded as $\lambda^{\alpha}$ power series. To achieve this, let us consider the equation

$$q^{s+2} + \lambda b q^{t+2} = 1, \tag{2.5}$$

where $b$ is a real parameter that allows convergence to be improved, as will be seen later on. When $\lambda b > 0$, there is only one real root $q(\lambda)$ of (2.5), which is a continuous function of $\lambda$ for all $\lambda$ values. This equation plays the role of the stationary point condition in the variational functional method.[17-21]

If (2.5) is written in terms of $w$,

$$\lambda^{-(s+2)/(t+2)} w^{s+2} + b w^{t+2} = 1, \tag{2.6}$$

we see that the two variables

$$u = \lambda q^{t+2}, \quad v = q^{s+2} \tag{2.7}$$

are bounded:

$$u(\lambda = 0) = 0, \quad u(\lambda \to \infty) = 1/b,$$
$$v(\lambda = 0) = 1, \quad v(\lambda \to \infty) = 0,$$

and, furthermore, they can be expanded in $\lambda$- and $\lambda^{\alpha}$-power series when $\lambda$ is small and large enough, respectively. Then, good choices for $B(\lambda, q)$ are

$$B = \sum_{i=0}^{\infty} b_i u^i, \tag{2.8a}$$

$$B = \sum_{i=0}^{\infty} b'_i v^i, \tag{2.8b}$$

where the coefficients $b_i$ and $b'_i$ are adjusted in order to fulfill (2.1).

Our method consists in approaching $E(\lambda)$ by means of Eqs. (2.2) and (2.5) (the last equation fixing the proper $q$ value). The coefficients $b_i$ are best adjusted through Eq. (2.1a) because the term $b_i u^i$ only yields $\lambda^j$ terms with $j \geqslant i$. In this way, the addition of the $M$ th term to (2.8a) does not modify the coefficients $b_i$ ($i < M - 1$) computed previously. By the same reasoning, Eq. (2.8b) is found to be better in order to allow $F$ to fulfill (2.1b).

When $\beta$ and $\alpha$ are not integer numbers, and negative $\lambda$ values are allowed, the proper form of (2.1b) should be

$$E(\lambda) = |\lambda|^{\beta} \sum_{i=0}^{\infty} e_i |\lambda|^{\alpha i}, \tag{2.9}$$

but this fact does not require changing the form of Eqs. (2.2) and (2.5).

In order to obtain a general expression for the constant $A$ and the coefficients $b_i$ in terms of the original power series coefficients $E_i$, let us write $\lambda$ and $F$ as functions of $u$,

$$\lambda = u(1 - bu)^{1/\alpha}, \tag{2.10}$$

$$(1 - bu)^{-\beta/\alpha} F = A + 1 - bu + uB(u). \tag{2.11}$$

If we introduce (2.1a) in (2.11) (with $F = E$) and expand each term of the resulting equation in powers of $u$, we obtain, after equating the coefficients of $u^i$ on both sides of (2.11), the following recursion relationship:

$$\sum_{j=0}^{i} \sum_{k=0}^{j} (-b)^{i+k-j} \binom{(j-k)/\alpha}{k} \binom{-\beta/\alpha}{i-j} E_{j-k} = (A+1)\delta_{i,0} + (b_{i-1} - b\delta_{i-1,0})(1 - \delta_{i,0}), \tag{2.12}$$

where $\delta_{i,j}$ is the Kronecker delta and $\binom{a}{i}$ is the coefficient of $u^i$ in the Taylor expansion of $(1+u)^a$. The zeroth-order equation [i.e., $i = 0$ in (2.12)] gives us an $A$ value that is $(\alpha, \beta)$-independent,

$$A = E_0 - 1. \tag{2.13}$$

Coefficients $b'_i$ are determined in the same way. In this case, we have

$$\lambda = (1 - v) v^{1/\alpha} / b, \tag{2.14}$$

$$v^{-\beta/\alpha} F = (1 - v)^{\beta} e / b^{\beta} = A + v + (1 - v) B(v)/b. \tag{2.15}$$

After expanding each term in (2.15) as a $v$-power series, we obtain another general recursion relationship:

$$\sum_{j=0}^{i} \sum_{k=0}^{j} (-1)^{i+k-j} \binom{\beta}{i-j} \binom{(j-k)\alpha}{k} e_{j-k} b^{(k-j)\alpha - \beta} = A\delta_{i,0} + \delta_{i,1} + \{b'_i - b'_{i-1}(1 - \delta_{i,0})\}/b. \tag{2.16}$$

The zeroth-order equation links $A$, $b'_0$, and $e_0$ together:

$$b'_0 = b^{1-\beta} e_0 - Ab. \tag{2.17}$$

Therefore, if we place (2.13) into (2.17), $b'_0$ is thoroughly fixed in terms of $e_0$, $b$, and $E_0$. The proper $b$ value is chosen in order to achieve highest convergence rates. This fact will be discussed at length in the subsequent sections.

In all nontrivial problems, we only get a finite number of terms in the series (2.1a) and/or (2.1b). So, we have to truncate the expansions (2.8) up to the highest-order coefficient we can calculate from (2.12) and (2.16). For this reason, we will have, in a general case, two different approximate expansions for $B$,

$$B_M = \sum_{i=0}^{M} b_i u^i, \tag{2.18}$$

$$B'_M = \sum_{i=0}^{M} b'_i v^i. \tag{2.19}$$

In problems of actual physical interest, it occurs frequently that one knows $e_0$ in addition to the first $M+1$ coefficients $E_i$. In such cases, we can fix $b_0, b_1, \ldots, b_M$ through (2.12) and then $b_{M+1}$ by allowing the resulting functional to fulfill $(\lambda^{-\beta} F)(\lambda \to \infty) = e_0$; i.e.,

$$b_{M+1} = b^{M+1}\left\{b^{1-\beta}(e_0 - Ab^{\beta}) - \sum_{i=0}^{M} b_i b^{-i}\right\}. \tag{2.20}$$

As we reasoned before, in practice it is better to use (2.18) and (2.19) to deal with (2.1a) and (2.1b), respectively. But, since both $B_M$ and $B'_M$ are able to be written as $\lambda$- or $\lambda^{\alpha}$-power series alternatively, then both expansions (2.1) can be taken into account by means of either (2.18) or (2.19).

Were we to know $e_1$ in the example that led to (2.20), we would be able to take this coefficient into consideration too by adding a new term to (2.18). But this addition would modify the $b_{M+1}$ value, for $b_{M+1}$ and $b_{M+2}$ would be coupled in two equations. In general, this way does not lead to better results than those obtained by taking into account $e_0$ only and so, we will not follow it in this paper.

It is clear that the functional (2.2) and Eq. (2.5) are thoroughly determined by means of the real numbers $\alpha$ and $\beta$ [that is to say, by means of the series (2.1)] without reference to any sort of Hamiltonian operator. Thus, the method is so general that it can be applied to a wide variety of problems in many branches of physics whenever they lead to power series expansions of the form (2.1). This will be shown in the following sections.

Another very important feature of the FM is that it makes clear a relation between the variational functional method[17–21] and several rearranged and renormalized peturbation series.[4–16]

## III. RELATION TO OTHER TECHNIQUES OF RESUMMATION OF POWER SERIES

Since divergent power series appear very frequently in many fields of physics, several procedures were developed during the last years to sum them. Most of these methods were applied to the quantum-mechanical anharmonic oscillator and the $\phi^{2M}$-model field theory in order to test their performance. Good reviews of some interesting divergent series and several rigorous mathematical results about them can be found in Refs. 22 and 23. For example, Padé approximants[1,2] and the Padé–Borel method[3] were applied to the Rayleigh–Schrödinger (RS) perturbation series for the quantum-mechanical anharmonic oscillator which is known to be divergent for all $\lambda$ values.[24] This one-dimensional model gives rise to two power series like (2.1) with $\beta = \frac{1}{3}$ and $\alpha = -\frac{2}{3}$.

But it is obvious that neither the Padé approximants nor the Padé–Borel method can give us, in the general case, the correct analytical $\lambda$ dependence in the large $\lambda$ regime [see (2.1b)] when they are applied directly to the $\lambda$-power series (2.1a). The only case they lead to the proper behavior is $\alpha = -1$ and $\beta$ an integer.

Due to this difficulty, several alternative techniques were developed to treat some particular problems. For example, the variationally rearranged perturbation theory,[4–6] the geometric approximation (GA),[7,8] partition techniques[6,25] and their renormalization,[9,10] generalized Wick ordering of Hamiltonians,[11] order-dependent mappings,[12] and rearranged Hamiltonians written in second-quantization formalism[13–16] were proved to be very useful. All these procedures have succeeded in summing some particular divergent series or, at least, in diminishing their divergence rate largely. Convergence was rigorously proved for very few cases only.[11,12]

Though the procedure described in Sec. II is reminiscent of the variational functional method,[17–21] it is only based on the analytical structure of the series (2.1), and it is wholly original as far as we know. Then, it might be profitable to know if, under certain conditions, the FM can be related to other existing techniques such as those summarized before. We will try to find some of these connections in the remaining of this section.

### A. Geometric approximation

Let us suppose that a Hamiltonian $H$ can be written as

$$H = H_0 + \lambda V, \tag{3.1}$$

where the eigenvalues and eigenfunctions of $H_0$ are known,

$$H_0 \Psi_0 = E_0 \Psi_0. \tag{3.2}$$

In the usual Rayleigh–Schrödinger perturbation theory (RSPT), the eigenvalues $E$ and eigenfunctions $\Psi$ of $H$ are expanded as

$$E = \sum_{i=0}^{\infty} E_i \lambda^i, \quad \Psi = \sum_{i=0}^{\infty} \Psi_i \lambda^i. \tag{3.3}$$

The convergence of the eigenvalue-power series can be largely improved by way of the GA that consists of rewriting $H$ as

$$H = H'_0 + V', \quad H'_0 = \mu H_0, \quad V' = \lambda V + (1-\mu) H_0. \tag{3.4}$$

Since the eigenvalues and eigenfunctions of $H'_0$ are known, then the RSPT can be applied to this new problem without difficulty and the first perturbation corrections $E'_i$ are

$$E'_0 = \mu E_0, \quad E'_1 = E_1 + (1-\mu) E_0/\lambda, \quad E'_2 = E_2/\mu,$$
$$E'_3 = \{E_3 + (\mu - 1) E_2/\lambda\}/\mu^2. \tag{3.5}$$

It is customary to choose $\mu$ in order to make $E'_3$ null (i.e., according to Wigner's rule)[7,26]:

$$\mu = 1 - \lambda E_3/E_2. \tag{3.6}$$

With this $\mu$ value, the rearranged power series (up to third-order) equals the (2/1) Padé approximant built from (3.3):

$$E \simeq E_0' + E_1' \lambda + E_2' \lambda^2$$
$$= E_0 + \lambda \{E_1 + \lambda E_2/(1 - \lambda E_3/E_2)\}. \quad (3.7)$$

If the integral $S_1 = \langle \Psi_0 | \Psi_1 \rangle$ is thought to be a variable parameter, then all perturbation terms $E_i$ $(i > 1)$ will depend on it. Therefore,

$$E^{(I)} = \sum_{i=0}^{I} E_i(S_1) \lambda^i. \quad (3.8)$$

Since the energy eigenvalue $E$ does not depend on $S_1$, then a good choice for it will be

$$\frac{\partial E^{(I)}}{\partial S_1} = 0. \quad (3.9)$$

When $I = 3$, this equation leads to

$$S_1 = -E_3/E_2 = (\mu - 1)/\lambda. \quad (3.10)$$

This is a very interesting result, which poses a connection between a variational procedure and a partition technique. Other choices for $S_1$ correspond to different partitions of the Hamiltonian $H$.[8]

Equations (3.7) and (3.10) show that the stationary-point condition (3.9) (with $I = 3$) gives rise to the change of variables $\lambda \to \lambda/(1 + \lambda S_1)$, where $S_1$ is given by (3.10). On the other hand, the variable we proposed in Sec. II was $u = \lambda/(q^{s-t} + \lambda b)$, which clearly suggests that our method should equal the GA when $s = t$; i.e., when $\alpha = -1$ [see (2.4)]. Besides, our parameter $b$ seems to play the role of the integral $S_1$. In order to verify this, let us consider Eqs. (2.2), (2.4), (2.5), and (2.18) with $\beta = -\alpha = 1$ and $M = 2$, which lead to the following functional:

$$F = A/q^2 + 1 + \lambda b_0 + \lambda^2 b_1 q^2 + \lambda^3 b_2 q^4, \quad q^2 = 1/(1 + \lambda b), \quad (3.11)$$

where

$$A = E_0 - 1, \quad b_0 = E_1 - Ab, \quad b_1 = E_2, \quad b_2 = E_3 + bE_2. \quad (3.12)$$

Since the exact function $E(\lambda)$ is $b$-independent, then it seems reasonable to determine the best $b$ $(= b^*)$ value according to

$$\frac{\partial F}{\partial b}(\lambda, b = b^*) = 0. \quad (3.13)$$

Thus, we obtain

$$b^* = -E_3/E_2, \quad (3.14)$$

which together with (3.11) and (3.12) gives us (3.7) exactly. It is interesting to note that the stabilization condition (3.13) leads to $b_2 = 0$ and, besides, that $b^*$ is $\lambda$-independent. This last result seems to hold in a general case, as we will see later on.

Previous discussion shows that the GA can be thought of as a particular case of our FM when $\beta = -\alpha = 1$. Furthermore, since the functional (2.2)–(2.5) was built in order to fulfill both power series exactly, then we can conclude that the GA can only give us the correct $\lambda$ dependence for $E$ when $\beta = -\alpha = 1$. For example, it is very useful to deal with Coulomb interaction potentials.[7,27] Also, Euler's method,[28] which is often used to continuate Taylor expansions analytically, is a particular case of the FM when $\beta = -\alpha = 1$.

Recently, Bhattacharyya[29] has proposed a generalized Euler's method to deal with power series expansions, but the change of parameter he has discussed is quite different from ours. Besides, his procedure does not take into consideration the small and large power series in those cases that admit them.

Before going forward, it might be profitable to discuss how to determine the best $b$ value $(b = b^*)$. If the series rearranged through the FM converged for a given $\lambda$ value, then their plots vs $b$ ($\lambda$ fixed) should probably possess a plateau, the extension of which should increase as $M$ in Eq. (2.18) increases. Furthermore, the smaller the $\lambda$ value, the larger this plateau. Thus we are forced to find a suitable plateau criterion.

Equation (3.13) exhibits a very attracting feature: the $b^*$ value seems to be $\lambda$-independent. This fact leads to a surprising conclusion. Let us suppose that $M - 1$ coefficients $b_i$ [see Eq. (2.18)] were determined according to (2.12). Therefore, the functional can be expanded in $\lambda$-power series as

$$F = \sum_{i=0}^{M} E_i \lambda^i + \sum_{i=M+1}^{\infty} F_i \lambda^i. \quad (3.15)$$

If Eq. (3.13) held for all $\lambda$ values, then all $F_i$'s $(i \geqslant M + 1)$ should be stable for the same $b^*$ value,

$$\frac{\partial F_i}{\partial b}(b = b^*) = 0, \quad i \geqslant M + 1. \quad (3.16)$$

For example, if we expand (3.11) up to $\mathcal{O}(\lambda^4)$, we obtain $F_4 = -(E_2 b^2 + 2E_3 b)$, which is stable when $b = b^* = -E_3/E_2$. Equation (3.16) is more suitable than (3.13) to compute $b^*$ because $F_i$ is always a polynomial of degree $i$.

Usually, $F$ has no stationary points for some $M$ values, and the plateau has to be determined by another way. A good plateau criterion is always necessary in order to find the best mapping parameter,[11,12] and we can obtain it, for example, by minimizing the absolute value of the last coefficient in the expansion. Also, we can fix $b^*$ through an inflexion point condition

$$\frac{\partial^2 F}{\partial b^2}(\lambda, b = b^*) = 0, \quad (3.17)$$

but in this case $b^*$ is found to be $\lambda$-dependent. Throughout this paper, $b^*$ will be determined according to the two following rules:

*(i) Stationary point*: We will choose the stationary point with the smallest absolute value of the second derivative. The $b^*$ value obtained in this way is $\lambda$-independent.

*(ii) Inflexion point*: If there were no stationary points, we would choose the inflexion point with the smallest absolute value of the first derivative. The $b^*$ value obtained in this way changes within a finite (and always small) interval when $\lambda$ runs from 0 to $\infty$. These rules can be summarized in the following mathematical expression:

$$G(b^*) = \min_{\{b_i^*\}} G(b),$$
$$G(b) = \left|\frac{\partial F}{\partial b}\right| + |b| \left|\frac{\partial^2 F}{\partial b^2}\right|, \quad (3.18)$$

where $\{b_i^*\}$ means the set of all stationary or inflexion points. The coefficient $|b|$ of the second derivative was intro-

duced to avoid too large $b^*$ values, for they do not lead to accurate enough results.

Though we are unable to rigorously prove that the stationary points (3.13) are $\lambda$-independent, this assertion is supported by a careful numerical search on several quite different examples, some of which will be discussed later on in this paper.

### B. Scaling-variational method

Let us suppose that we are interested in the ground-state eigenvalue of the Hamiltonian,

$$H = H_0 + \lambda |x|^t, \quad H_0 = p^2 + |x|^s, \quad p = -i\frac{d}{dx}, \tag{3.19a}$$

and that we know a normalized eigenfunction $\phi(x)$ of $H_0$ with eigenvalue $E_0$,

$$H_0\phi = E_0\phi. \tag{3.19b}$$

Then, a good approach to the lowest eigenvalue $E$ of $H$ can be obtained with the upper bound

$$E(a) = \langle \phi_a | H\phi_a \rangle \geqslant E, \quad \phi_a = a^{1/2}\phi(ax). \tag{3.20}$$

The virial theorem for (3.19) states that

$$\langle \phi | |x|^s\phi \rangle = 2E_0/(s+2) = 2\langle \phi | p^2\phi \rangle / s. \tag{3.21}$$

Therefore,

$$E(a) = 2a^2E_0/(s+2) + sa^{-s}E_0/(s+2) + \lambda E_1 a^{-t},$$
$$E_1 = \langle \phi | |x|^t\phi \rangle. \tag{3.22}$$

In this simple approximation scheme, the most accurate approximation to $E$ is given by the lowest $E(a)$ value. The minimum condition $\partial E(a)/\partial a = 0$ leads at once to the best $a$ value,

$$a^{-s-2} + \lambda(s+2)tE_1 a^{-t-2}/(2sE_0) = 1. \tag{3.23}$$

The scaling-variational method just described here is also suitable for estimating excited-state eigenvalues because it gives us the proper $(\lambda,n)$ dependence of them (see Refs. 30 and 31 and references therein.)

On the other hand, the FM leads to

$$F = (E_0 - 1)/q^2 + q^s + \lambda b_0 q^t, \tag{3.24a}$$

$$q^{s+2} + \lambda b q^{t+2} = 1, \tag{3.24b}$$

in the first-order approximation [i.e., $M = 0$ in (2.18)]. From (2.12) it follows that

$$b_0 = E_1 + (s + 2 - 2E_0)\, b/(s+2). \tag{3.25}$$

Thus, the three first Taylor expansion terms for $F$ are

$$F = E_0 + E_1\lambda + F_2\lambda^2 + \cdots,$$
$$F_2 = sE_0b^2/(s+2)^2 - tE_1b/(s+2). \tag{3.26}$$

Since it is rather cumbersome to obtain a root of $(\partial F/\partial b)(b = b^*) = 0$, we will obtain $b^*$ by way of $(\partial F_2/\partial b)(b = b^*) = 0$, which leads to

$$b^* = t(s+2)E_1/(2sE_0). \tag{3.27}$$

When this result is replaced in (3.24b), we obtain (3.23) exactly with $q = 1/a$. So, we have proved that the FM also contains the scaling-variational method[30,31] as a particular case. It is found that (3.27) is also a stationary point of $F$ for all $\lambda$ values, according to what was argued before.

### C. Other partition techniques

The eigenvalues of the $2k$-anharmonic oscillator

$$H = p_x^2 + m^2x^2 + \lambda x^{2k}, \quad p_x = -i\frac{d}{dx}, \tag{3.28}$$

can be expanded in power series like (2.1) with $\alpha = -2/(k+1)$ and $\beta = 1/(k+1)$. Equation (2.10) gives rise to the following change of perturbation parameter:

$$\lambda = u(1 - bu)^{-(k+1)/2}. \tag{3.29}$$

This is exactly the same transformation proposed by Dmitrieva and Plindov,[5,6] who determined $b$ according to the virial theorem.

The change of variables $x = (1 - bu)^{1/4}y$ allows us to write (3.28) as[5,6]

$$\bar{H} = (1 - bu)^{1/2}H = p_y^2 + y^2 + u(y^{2k} - by^2). \tag{3.30}$$

Since the eigenvalues of $\bar{H}_0 = p_y^2 + y^2$ are known, then the RSPT can be applied to (3.30) without further difficulty, and we obtain the eigenvalues of $H$ as $u$-power series

$$E = (1 - bu)^{-1/2}\sum_{i=0}^{\infty}\bar{E}_i(b)\,u^i, \tag{3.31}$$

where $\bar{E}_i$ are the perturbation corrections to the eigenvalue $\bar{E}$ of $\bar{H}$.

Dmitrieva and Plindov's method[5,6] is suitable to deal with any Hamiltonian operator of the general form $H = T + V_1 + V_2$, where $V_1$ and $V_2$ are homogeneous functions of the coordinates, provided the eigenvalues and eigenfunctions of $T + V_1$ are known.

A careful examination of the partition techniques discussed by Pascual,[4] Killingbeck,[9] and Austin and Killingbeck[10] shows that they are closely related to the method of Refs. 5 and 6 and that they also lead to a mapping parameter like (3.29), though these authors did not make it explicit.

The FM applies to all problems that can be treated by means of the aforesaid procedures and to other ones that are not related to a Hamiltonian operator. Therefore, the perturbation techniques described in Refs. 4–6, 9, and 10 may be clearly thought to be particular cases of the FM.

### D. Generalized Wick-ordering method

In a very interesting paper, Caswell[11] extended Wick-order concepts that are well known in field theory and developed an approach to the eigenvalues of anharmonic oscillators and double well potential problems related to them. Caswell's method will be briefly discussed here, using the quartic anharmonic oscillator [$k = 2$ in (3.28)] as an example, in order to show that it is a particular case of our FM.

The Wick-ordered Hamiltonian $:H:$ with all creation operators to the left of all anihilation operators is related to $H$ [(3.28)] by

$$H = :H: + M - 3\lambda/(2M^2), \tag{3.32}$$

where the renormalized mass $M$ is connected with the original oscillator mass $m$ through

$$M^2 = m^2 + 3\lambda/M. \tag{3.33}$$

The RS perturbation series

$$E = m \sum_{i=0}^{\infty} E_i \left(\frac{\lambda}{m^3}\right)^i \tag{3.34}$$

can be rearranged as

$$E = M \sum_{i=0}^{\infty} E_i^M \left(\frac{\lambda}{M^3}\right)^i, \tag{3.35}$$

where the new expansion coefficients $E_i^M$ are determined in such a way that (3.35) leads to (3.34) when $M$ is expanded in $\lambda$-power series. Because of (3.33), the variable $\lambda/M^3$ remains bounded for all $\lambda$ values.

It is not difficult to prove that Caswell's perturbation series (3.34) and those obtained by Pascual[4] and Dmitrieva and Plindov[5,6] are exactly the same when $E$ is the ground-state eigenvalue of the quartic anharmonic oscillator.

Caswell[11] generalized the Wick-ordering method just described by replacing (3.33) with

$$M^2 = m^2 + l\lambda/M, \tag{3.36}$$

where the "mass shift" $l$ is determined in order to improve convergence. Caswell[11] proposed the stationary point condition

$$\frac{\partial}{\partial l}\left\{\lim_{\lambda\to\infty} \lambda^{-1/3} E^{(I)}(M)\right\}(l = l^*) = 0, \tag{3.37}$$

where $E^{(I)}(M)$ is the sum (3.35) up to the $I$ th term, in order to obtain the best $l(=l^*)$ value. Frequently, Eq. (3.37) leads to complex roots only. In such cases it is necessarily a good criterion to choose $l^*$.[11]

The Wick ordering does not lead to series like (3.35) when it is directly applied to any anharmonic oscillator with $k > 2$, but Caswell could generate them by resorting to dimensional arguments.

We will now prove that our FM contains Caswell's method as a particular case. The change of variables $z = q^{s+2}$ transforms (2.2) and (2.5) in

$$z + \lambda b z^{-1/\alpha} = 1, \tag{3.38}$$

$$F = z^{\beta/\alpha} \sum_{i=0}^{\infty} d_i (\lambda z^{-1/\alpha})^i, \tag{3.39}$$

where $d_0 = A + 1$, $d_1 = b_0 - b$, and $d_i = b_{i-1}$ $(i > 1)$. For the quartic anharmonic oscillator model discussed before ($\alpha = -2/3$), Eqs. (3.38) and (3.39) equal (3.33) and (3.35), respectively, when

$$z = m^2/M^2, \quad b = l/m^3, \quad E_i^M = m^{3i-1} d_i. \tag{3.40}$$

Therefore, the FM and Caswell's procedure yield identical results whenever the parameters $l$ and $b$ are determined according to the same criterion. In Sec. V we will compute the ground-state eigenvalue of the quartic anharmonic oscillator for different $\lambda$ values using Caswell's method and the FM with (3.37) and (3.18), respectively. Results will show the influence of the plateau criterion.

When considering the anharmonic oscillator problem, the zeroth-order Hamiltonain can be defined by several different arrangements of the creation and anihilation operators[13–16] unlike the Wick-ordering one. It was recently shown[32] that one of these partitions[15] leads to an RS series with infinite convergence radius.

### E. Order-dependent mappings

Seznec and Zinn-Justin[12] made a numerical test of a summation method based on an order-dependent mapping which poses a generalization of the techniques that are usually employed to analytically continuate power series outside their convergence radii.[28] The procedure was applied to some models of interest in field theory and to other ones closely related to the quartic anharmonic oscillator.[12] It consists of changing the perturbation parameter $\lambda$ into another one $g$ through a mapping

$$\lambda = \rho g (1 - g)^{-s}, \tag{3.41}$$

where $\rho$ is a variational parameter introduced to minimize the error of the truncated series. By way of a partition of a Hamiltonian operator like those discussed in subsections III A and III B, Seznec and Zinn-Justin[12] found that $s = 3/2$ is a good choice, but they considered other $s$ values too.

This technique is closely related to the FM for the mapping (3.41) coincides with (2.10) exactly when $\rho = 1/b$, $g = bu$, and $s = -1/\alpha$. Besides, the FM requires $\alpha = -2/3$ for the quartic anharmonic oscillator. Also, the alternative mapping (2.14) we proposed in Sec. II equals (3.41) when $\rho = 1/b$, $g = 1/(1 - v)$, and $s = -1/\alpha$. Therefore, we see that the rearranged power series proposed by Seznec and Zinn-Justin[12] and by the FM are one and the same whenever $s$ is properly chosen and $\rho$ and $b$ are determined through identical criteria.

Another redefinition of the perturbation parameter was recently proposed by Vaĭnberg *et al.*[33] and Popov and Weinberg[34] together with Padé and Borel–Padé methods, respectively, in order to improve the convergence of the perturbation series for screened Coulomb potentials.

To make the discussion clear, let us consider the funnel-like confining potential:

$$H(Z, \lambda) = p^2 - Z/r + \lambda r, \tag{3.42}$$

which is of great interest in quarkonium physics.[35] From the well-known Symanzik's theorem[2]

$$H(Z, \lambda) = \lambda^{2/3} H(Z\lambda^{-1/3}, 1), \tag{3.43}$$

Vaĭnberg *et al.*[33] obtained the following mapping:

$$\lambda' = \lambda(1 + \gamma\lambda)^{-1/3}, \tag{3.44}$$

which transforms the original RS series into a $\lambda'$-power expansion.

The change of the perturbation parameter (3.44) is not as good as those discussed in previous subsections for $\lambda'$ is an unbounded parameter. This fact makes it necessary to resort to $(N + 1/N)$ Padé approximants in order to obtain reasonable results when $\lambda$ is large enough. Furthermore, though these approximants behave properly as $\lambda^{2/3}$ in the large $\lambda$ regime, they do not give rise to the correct $\lambda^{-1/3}$-power series [see (3.43)] because $\lambda'$ cannot be expanded in such a series.

On the other hand, Popov and Weinberg[34] applied the Borel–Padé method to the $\lambda'$ expansion and chose $\gamma$ in order to remove the nearest singularity of the Borel transform to infinity. Although this procedure yields better results than the RS series itself, the approximate function obtained in this

manner does not lead to the proper $\lambda^{-1/3}$ expansion in the large $\lambda$ regime.

The FM shows that in this case the correct mapping should be [cf. (2.10)]

$$\lambda = u(1-bu)^{-3}. \tag{3.45}$$

## IV. APPLICATION TO A TWO-LEVEL MODEL

At present we cannot prove to what extent the FM improves convergence of a general power series, but in this section we will try to give convincing heuristic arguments that the method is useful to enlarge convergence radii. To do this, let us consider a simple two-level model that was previously studied by other authors[8,36,37] because it exhibits a level crossing which is of great importance in atomic and molecular physics.[36,37]

The model is posed by the symmetric matrix

$$H(\lambda) = \begin{pmatrix} 3\lambda/2 & 2\lambda \\ 2\lambda & 4-3\lambda/2 \end{pmatrix}, \tag{4.1}$$

where $\lambda$ is a real parameter. The two eigenvalues of $H(\lambda)$,

$$E_{\pm}(\lambda) = 2 \pm \tfrac{1}{2}(25\lambda^2 - 24\lambda + 16)^{1/2}, \tag{4.2}$$

are analytic functions of $\lambda$ within a finite domain determined by the branching-point singularities $\lambda_b = 12/25 \pm 16i/25$. The matrix elements were chosen to obtain a strong coupling between the two eigenvalues that is given by a level crossing in the complex plane. The convergence radius of the Taylor series for the lower level $E_-(\lambda) = E(\lambda)$ [like (2.1a)] is $R = |\lambda_b| = 4/5$. (We will not consider here the other eigenvalue explicitly for the same qualitative conclusions apply to it.)

The Borel method consists in calculating $E(\lambda)$ by means of the transformation[3,28]

$$E(\lambda) = \int_0^{\infty} g(\lambda z)\, e^{-z}\, dz, \tag{4.3}$$

where $g(z)$ obeys the following Taylor expansion:

$$g(z) = \sum_{i=0}^{\infty} \frac{E_i\, z^i}{i!}. \tag{4.4}$$

In the present case, the region of Borel summability (Borel polygon) is $\lambda\,\langle R_B = |\lambda_b|^2/\mathrm{Re}(\lambda_b) = 4/3$.[37]

The appropriate functional to deal with the $\lambda$-power expansion for (4.2) is ($\beta = -\alpha = 1$; $s = t = 0$)

$$F(\lambda) = A/q^2 + 1 + \lambda B_M(q), \tag{4.5a}$$

where

$$B_M(q) = \sum_{i=0}^{M} b_i\, u^i, \quad u = \lambda q^2 = \lambda/(1+\lambda b). \tag{4.5b}$$

The branching-point singularities of $E(\lambda(u))$ in the $u$-complex plane are found to be

$$u_b = u(\lambda_b) = (12 + 16b \pm 16i)/(16b^2 + 24b + 25). \tag{4.6}$$

Since $E(\lambda \to \infty) \to \infty$, then the mapping $\lambda(u)$ introduces a pole $u_p = 1/b$ into $E(\lambda)$. The two branching points lie on the circle of radius $R'$,

$$R' = |u_b| = (400 + 384b + 256b^2)^{1/2}/(25 + 24b + 16b^2), \tag{4.7}$$

in the $u$-complex plane. Therefore, the $u$-power series will have, for real $u$ values, the following convergence intervals:

(i) $-R' < u < R'$ or $-R'/(1+bR') < \lambda < R'/(1-bR')$ if $R' < 1/|b|$;

(ii) $-1/|b| < u < 1/|b|$ or $-\infty < \lambda < 1/|2b|$ if $R' \geq 1/|b|$.

When $b = -25/24$, the equality $R' = 1/|b|$ holds and at this point the transition from real finite convergence intervals ($b > -25/24$) to infinite ones ($b < -25/24$) takes place.

The $\lambda$ and $b$ values which allow the $u$-power series to converge are shown in Fig. 1 (region marked by hatching). It is clear that if $b$ is chosen properly, the functional $F$ should give us excellent results for $\lambda < 0$. We have studied several functions with a couple of conjugate branching points $\lambda_b = \lambda_1 \pm i\lambda_2$ and in all cases the $b$ values that fulfilled $|u_b| = 1/|b|$ gave rise to $u$-power series which converged in $(-\infty, \lambda_1)$ and $(\lambda_1, \infty)$ when $\lambda_1 > 0$ and $\lambda_1 < 0$, respectively. Numerical calculation for more complicated problems (such as those discussed in Sec. V), suggests that the convergence radii of the $u$-power series are much larger than those of the $\lambda$-power series.

The two-level model just considered is suitable to compare the performance of the FM with those of the Padé and Borel–Padé methods for the $(N+1, N)$ approximants behave properly in the large $\lambda$ regime (this is due to the fact that $\beta = -\alpha = 1$).

The fifth-order $\lambda$-power series for the ground-state level

$$E(\lambda) = 3\lambda/2 - \lambda^2 - 3\lambda^3/4 - 5\lambda^4/16 + 9\lambda^5/64 + \cdots \tag{4.8}$$

can be rearranged in a $u$-power series with the following coefficients:

$$A = -1, \quad b_0 = \tfrac{3}{2} + b, \quad b_1 = -1, \quad b_2 = -b - \tfrac{3}{4},$$
$$b_3 = -b^2 - 3b/2 - 5/16,$$
$$b_4 = -b^3 - 9b^2/4 - 15b/16 + 9/64. \tag{4.9}$$

When orders up to the third one are only included, the FM gives

$$F(\lambda, b) = \{\lambda^3(3b^2/2 - 2b - 3/4) + \lambda^2(3b-1) + 3\lambda/2\}/\{1 + 2b\lambda + \lambda^2 b^2\}, \tag{4.10}$$

which has a stationary point $b^* = -3/4$. For this particular value, the functional (4.10) leads to the (2/1) Padé approximant

$$F(\lambda) = (3\lambda/2)(1 - 17\lambda/12)/(1 - 3\lambda/4), \tag{4.11}$$

according to the discussion in subsection III A (it should be noted that $b^* = -E_3/E_2$). It is a very striking fact that this $b^*$ value leads to the largest $|u_b|$ value [$|u_b|(b = -3/4) = 1$] and that $u_b$ has no real part [$u_b(b = -3/4) = \pm i$].

Figure 2 shows the plots $F(\lambda = -1/2, b)$ vs $b$ where the functional values were computed by means of Eqs. (4.5) and (4.9) with $M = 2,3,4$. Results clearly show that the plateau becomes larger as $M$ increases. When $M = 4$, the functional is almost a constant for all $b$ values within the interval $(-2, -1/2)$.

We have calculated $E(\lambda)$ approximately for different $\lambda$ values by means of the functional (4.5) with $M = 2,3,4$. Re-

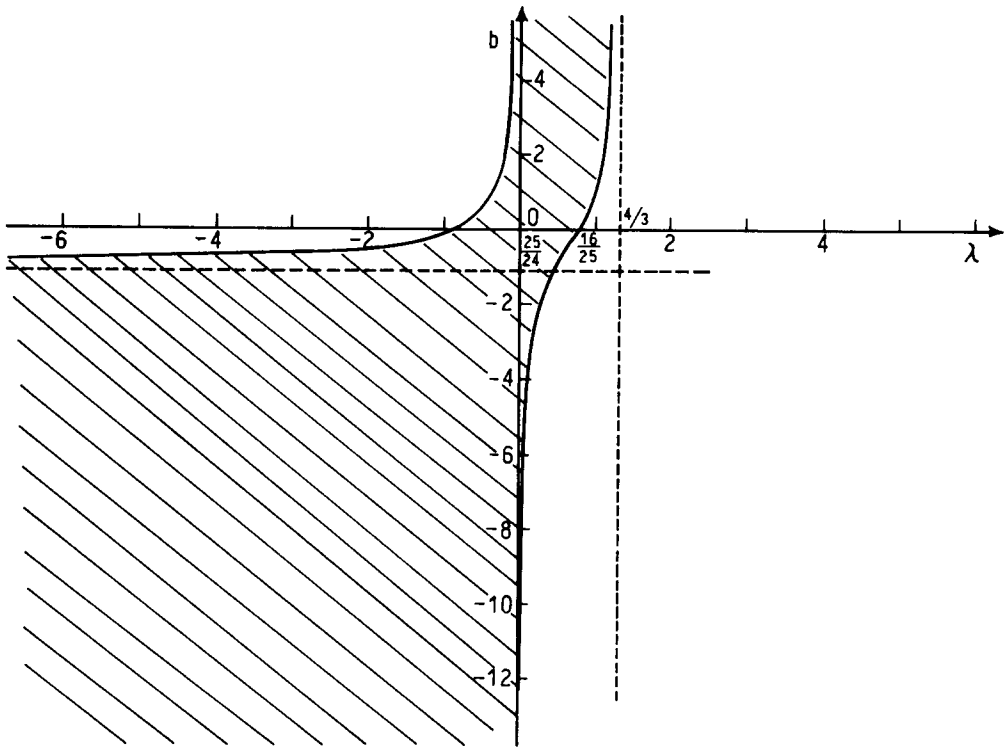



FIG. 1. Convergence region in the $\lambda$–$b$ plane for the ground state of the two-level model (4.1).

sults are shown in Table I together with those obtained through the fifth-order perturbation series and the exact values. Even the third-order functional ($M = 2$) is much more accurate than the fifth-order Taylor expansion in almost the whole range of $\lambda$ values.

The fifth-order functional ($M = 4$) exhibits two stationary points: $b^* = -3/4$ and $b^* = -1.61$, the last one being in accord with the plateau criterion discussed in Sec. III. Both $b^*$ values lead to very good results which are compared with those obtained through Padé and Borel–Padé methods [(3/2) approximants in each case] in Table II. Our functional with $b^* = -1.61$ is as accurate as the (3/2) Padé approximant, and both sets of FM results are better than Borel–Padé ones.

In general, the FM is more exact and much easier to apply than the Borel–Padé method because the integral (4.3) has to be calculated numerically. [It must be kept in mind that in this approximation $g(\lambda z)$ is written as a Padé approxi-

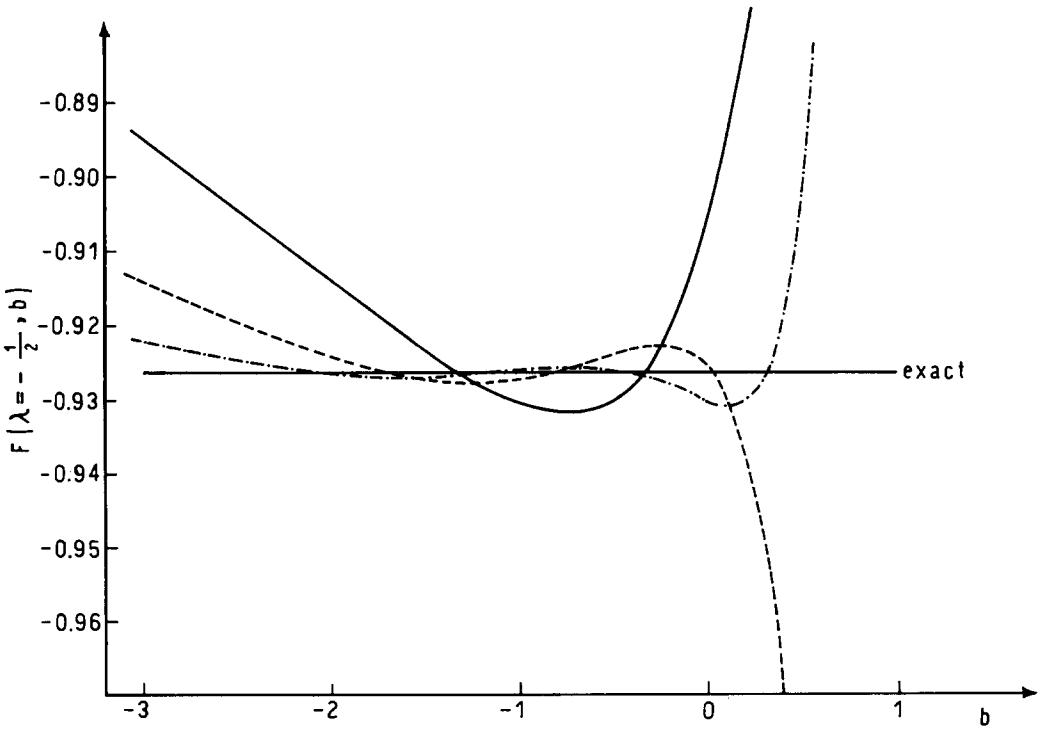



FIG. 2. Ground-state of the two-level model(4.1). $F(\lambda = -\frac{1}{2}, b)$ vs $b$ for different degrees of approximation: (——) third-order; (- - - - - -) fourth-order; (- · - · - ·) fifth-order.

TABLE I. Ground state of the two-level model (4.1) for several $\lambda$ values.

| $-\lambda$ | $-E^{a}$ | $-E^{b}$ | $-E^{c}$ | $-E^{d}$ | $-E^{\text{exact}}$ |
|---|---|---|---|---|---|
| 0.5 | 0.9318 | 0.9275 | 0.9267 | 0.9302 | 0.9262 |
| 1.0 | 2.0714 | 2.0432 | 2.0367 | 2.2031 | 2.0311 |
| 1.5 | 3.3088 | 3.2368 | 3.2195 | 4.6186 | 3.2022 |
| 2.0 | 4.6000 | 4.4694 | 4.4375 | 10.5000 | 4.4031 |
| 2.5 | 5.9239 | 5.7243 | 5.6753 | 24.2212 | 5.6199 |

[a] Third-order functional calculation, $b^{*} = -3/4$.
[b] Fourth-order functional calculation, $b^{*} = -5/4$.
[c] Fifth-order functional calculation, $b^{*} = -1.61$.
[d] Fifth-order Taylor expansion.

mant.] Furthermore, when $\beta$ is not an integer or $\alpha \neq -1$, Padé approximants do not behave properly in the large $\lambda$ regime unless a sort of renormalized power series is used. As we saw before, the techniques employed to obtain these expansions are not general enough to handle any power series like (2.1), and, besides, they are particular cases of the FM. Therefore, the FM seems to be the most useful way to deal with this sort of power series.

The FM has another very important advantage: We can introduce a new coefficient $b_5$ in order to satisfy $(|\lambda|^{-1}F)(\lambda \to \infty) = e_0 = -5/2$. This condition leads to [cf. (2.20)]

$$b_5 = b^{5}\left\{ -\frac{5\lambda}{2|\lambda|} + b - \sum_{i=0}^{4} b_i\, b^{-i} \right\}. \tag{4.12}$$

The functional obtained in this way gives us far better results as shown in Table II. In general, the addition of a coefficient to $B_M(q)$ that allows the functional to behave properly in the large $\lambda$ regime improves convergence markedly.

## V. APPLICATION TO SIMPLE MODELS OF PHYSICAL INTEREST

### A. The anharmonic, mean square, displacement function

We said before that by this time we are not able to rigorously prove if our FM leads to convergent series or not. However, in Sec. IV we gave reasonable arguments that the FM may be a powerful tool to increase the convergence radius of a given power series. In this section we will study three physically interesting problems, the first two ones being known to lead to RS series with zero convergence radii. Our goal is to show more clearly that the FM is far more general than the procedures described briefly in Sec. III by applying it to quite different problems.

TABLE II. Ground state of the two-level model (4.1) computed through different approximation methods.

| $-\lambda$ | $-E^{a}$ | $-E^{b}$ | $-E^{c}$ | $-E^{d}$ | $-E^{\text{exact}}$ |
|---|---|---|---|---|---|
| 0.5 | 0.9258<br>0.9267 | 0.9257 | 0.9260 | 0.9263 | 0.9262 |
| 1.0 | 2.0248<br>2.0367 | 2.0216 | 2.0283 | 2.0326 | 2.0311 |
| 1.5 | 3.1769<br>3.2195 | 3.1604 | 3.1915 | 3.2059 | 3.2022 |
| 2.0 | 4.3440<br>4.4375 | 4.2978 | 4.3793 | 4.4097 | 4.4031 |
| 2.5 | 5.5130<br>5.6753 | 5.4174 | 5.5783 | 5.6293 | 5.6199 |

[a] Fifth-order functional. First and second value for each $\lambda$ correspond to $b^{*} = -3/4$ and $b^{*} = -1.61$, respectively.
[b] (3/2) Padé approximant.
[c] Borel–Padé method with the (3/2) Padé approximant for $g(\lambda z)$ [see Eq. (4.3)].
[d] Fifth-order functional with Eq. (4.12).

Great interest has been recently shown in the anharmonic, mean-square, displacement function[38–42]

$$\langle x^2 \rangle = \int_0^{\infty} x^2 \exp\{-\beta' V(x)\}\, dx \times \left( \int_0^{\infty} \exp\{-\beta' V(x)\}\, dx \right)^{-1}, \tag{5.1a}$$

where

$$\beta' = 1/(kT), \quad V(x) = k_2 x^2 + k_4 x^4, \tag{5.1b}$$

for it has been proved to be useful to interpret a wide variety of experimental data[39,41,42] such as the dynamic response of the order–disorder and displacive ferroelectrics above the Curie temperature[41] or the temperature dependence of the isochoric dielectric polarizability.[39,42]

Morita and Frood[39] have shown that the function

$$E(\lambda) = 2\beta' k_2 \langle x^2 \rangle, \quad \lambda = k_4/(\beta' k_2^2) \tag{5.2}$$

can be expressed in the form of a continued fraction

$$E(\lambda) = 1/1 + 3\lambda/1 + 5\lambda/1 + 7\lambda/1 + \cdots. \tag{5.3}$$

This problem is quite appropriate for our purposes because the $\lambda$-power series

$$E(\lambda) = \sum_{i=0}^{\infty} E_i\, \lambda^{i} = 1 - 3\lambda + 24\lambda^{2} - 297\lambda^{3} + 4896\lambda^{4} - 100{,}278\lambda^{5} + \cdots \tag{5.4}$$

is strongly divergent for all $\lambda$ values. It is very easy to convince oneself that this is true by noticing that both integrals in (5.1) diverge so long as $\lambda$ takes negative values no matter how small $|\lambda|$ is. Besides, though the continued fraction (5.3) converges, it is not useful to calculate $E(\lambda)$ in the large $\lambda$ regime for too many terms are required.

The function $E(\lambda)$ can also be expanded in powers of $\lambda^{-1/2}$ [40],

$$\begin{aligned} E(\lambda) &= \lambda^{-1/2} \sum_{i=0}^{\infty} e_i\, \lambda^{-i/2} \\ &= \lambda^{-1/2}\{0.675\,978\,2399 - 0.271\,526\,7096\lambda^{-1/2} \\ &\quad + 0.077\,221\,4863\lambda^{-1} \\ &\quad - 0.015\,566\,6445\lambda^{-3/2} \\ &\quad + 0.001\,780\,083\,54\lambda^{-2} - \cdots\}. \end{aligned} \tag{5.5}$$

Taking into account the leading term of this series (i.e., $e_0 \lambda^{-1/2}$), Booth[40] suggested approximations of the form

$$E(\lambda) \simeq \sum_{i=0}^{n} a_i (\lambda + \alpha_i^2)^{-1/2}, \tag{5.6}$$

where the $2n$ parameters $a_i$ and $\alpha_i$ are determined according to

$$(i!)^{-1} \frac{\partial^{i} E}{\partial \lambda^{i}}(\lambda = 0) = E_i, \quad i = 0,1,\ldots,2n-1, \tag{5.7a}$$

$$\lim_{\lambda \to \infty} \lambda^{1/2} E(\lambda) = e_0. \tag{5.7b}$$

In this way, Booth[40] could obtain quite good results for small and large enough $\lambda$ values. However, his formula fails in the

TABLE III. Calculation of the anharmonic, mean square, displacement function with different approximation levels.

| $\lambda$ | $E^{a}$ | $E^{b}$ | $E^{c}$ | $E^{exact}$ | $E^{d}$ | $E^{e}$ | $E^{f}$ |
|---|---|---|---|---|---|---|---|
| $10^{-4}$ | 0.999 5820 | 0.999 7020 | 0.999 7015 | 0.999 7002 | 0.999 7002 | 0.999 7002 | 0.999 7002 |
| $10^{-3}$ | 0.995 914 | 0.997 041 | 0.997 036 | 0.997 024 | 0.997 024 | 0.997 024 | 0.997 024 |
| $10^{-2}$ | 0.965 671 | 0.972 257 | 0.972 222 | 0.972 144 | 0.972 140 | 0.972 144 | 0.972 144 |
| $10^{-1}$ | 0.812 166 | 0.817 663 | 0.817 621 | 0.817 561 | 0.817 279 | 0.817 480 | 0.817 552 |
| 1 | 0.467 546 | 0.467 923 | 0.467 921 | 0.467 919 | 0.467 333 | 0.467 657 | 0.467 829 |
| 10 | 0.188 895 | 0.188 902 | 0.188 902 | 0.188 902 | 0.188 736 | 0.188 818 | 0.188 864 |
| $10^{2}$ | 0.064 958 | 0.064 958 | 0.064 958 | 0.064 958 | 0.064 935 | 0.064 946 | 0.064 952 |
| $10^{3}$ | 0.021 1070 | 0.021 1072 | 0.021 1072 | 0.021 1072 | 0.021 105 | 0.021 106 | 0.021 106 |

[a,b,c] $v$-form of the functional with $M = 1,2,3$, respectively.
[d,e,f] $u$-form of the functional with $M = 2,3,4$, respectively.

intermediate regime, giving a very poor approach to $E(\lambda)$. This failure is obviously due to the fact that the approximate function (5.6) does not obey the $\lambda^{-1/2}$-power expansion which is required by the exact one.

Our FM, on the contrary, can satisfy both expansions. Using the results of Sec. II we obtain

$$F = q^{-4} + \lambda q^{-6} B(q), \quad \lambda > 0, \tag{5.8a}$$

where

$$q^{-2} + \lambda b q^{-4} = 1, \quad u = \lambda q^{-4}, \quad v = q^{-2}. \tag{5.8b}$$

It is possible to compute $E(\lambda)$ approximately by two ways; i.e., using the $u$ or the $v$ form of $F$. In the first case we have $(A = 0)$

$$b_0 = E_1 + 2b, \tag{5.9a}$$

$$b_1 = E_2 + 3bE_1 + b^2, \tag{5.9b}$$

$$b_2 = E_3 + 5bE_2 + 6b^2E_1 + b^3, \tag{5.9c}$$

$$b_3 = E_4 + 7bE_3 + 15b^2E_2 + 10b^3E_1 + b^4. \tag{5.9d}$$

As was said in Sec. II, the first $M$ coefficients $b_i$ in $B(q)$ determined with (2.12) allow the functional to reproduce the $\lambda$-power series up to the $(M+1)$th order and $b_{M+1}$ given by (2.20) takes into consideration the behavior of $E(\lambda)$ in the very large $\lambda$ regime.

In order to show how the results are improved by adding terms to $B_M$, we have computed $F$ with $M = 1,2,3$. In the first two cases there exist stationary points $b^* = 9.266$ and $b^* = 7.930$, respectively, which are $\lambda$-independent. In the third case there are no stationary points but an inflexion point that changes approximately from 9.0 to 10.7 as $\lambda$ increases from 0 to $\infty$. The best $b^*$ value was determined according to (3.18). Table III shows that our results are in an excellent agreement with the exact ones (obtained by Romberg integration) in the whole range of $\lambda$ values. Furthermore, our functional is much more accurate than Booth's formula (5.6)[40] for all $\lambda$ values.

The coefficients in $B'_M$ [(2.19)] can be easily calculated through (2.16), the first five ones being

$$b'_0 = b^{3/2} e_0, \tag{5.10a}$$

$$b'_1 = b^2 e_1 + \tfrac{3}{2} b^{3/2} e_0 - b, \tag{5.10b}$$

$$b'_2 = b^{5/2} e_2 + 2b^2 e_1 + \tfrac{15}{8} b^{3/2} e_0 - b, \tag{5.10c}$$

$$b'_3 = b^3 e_3 + \tfrac{5}{2} b'_2 - 2b'_1 + \tfrac{1}{2} b_0 - \tfrac{1}{2} b, \tag{5.10d}$$

$$b'_4 = b^{7/2} e_4 + \tfrac{5}{4} b'_1 - \tfrac{25}{8} b'_2 + 3b'_3 - \tfrac{15}{128} b'_0 + \tfrac{1}{8} b. \tag{5.10e}$$

In this case, it is not necessary to add an extra term to fit $E(\lambda)$ in the small $\lambda$ regime for $A = 0$ takes $E_0$ into account. We have calculated $F$ with $M = 2,3,4$ and the results are shown in Table III. The best $b^*$ values are stationary points placed in 15.582, 6.057, and 6.360, respectively. Also in this case, our results are much more accurate than Booth's.[40]

The FM might be useful in another very important physical area. The denominator in (5.1) which represents the classical partition function of the anharmonic oscillator is also closely related to the integral

$$Z(\lambda) = \pi^{-1/2} \int_{-\infty}^{\infty} \exp(-x^2 - \lambda x^4)\, dx, \tag{5.11}$$

representing a zero-dimensional $\phi^4$-field theory.[12,43–46] It is very easy to show that $Z(\lambda)$ obeys

$$Z(\lambda) = \sum_{i=0}^{\infty} Z_i \lambda^i = \lambda^{-1/4} \sum_{i=0}^{\infty} z_i \lambda^{-i/2}, \tag{5.12a}$$

where

$$Z_i = (-1)^i (4i)!/\{2^{4i} i! (2i)!\}, \tag{5.12b}$$

$$z_i = (-1)^i \Gamma(\{2i+1\}/4)/(2\pi^{1/2} i!). \tag{5.12c}$$

Since this problem does not differ too much from the anharmonic, mean-square, displacement function considered previously, we do not discuss it here in detail. The application of the FM is straightforward, and it might probably yield results as good as those shown in Table III.

### B. The quartic anharmonic oscillator

Since the appearance of the fundamental papers by Bender and Wu,[24] Simon,[2] and Hioe *et al.*,[47] much is known about the analytical structure of the eigenvalues of the $2k$-anharmonic oscillators (3.28). For example, even though the RS perturbation series has zero convergence radius, it was proved that the Padé approximants converge to the exact eigenvalue.[2] Furthermore, some renormalized perturbation series[4–6,9] also seem to be convergent (or at least slowly divergent).

We will consider here the quartic anharmonic oscillator only,

$$H = p^2 + x^2 + \lambda x^4, \quad p = -i \frac{d}{dx}, \tag{5.13}$$

for which the FM gives us

$$F = A/q^2 + q^2 + \lambda q^4 B(q), \tag{5.14a}$$

$$v = q^4, u = \lambda q^6, \quad q^4 + \lambda b q^6 = 1. \tag{5.14b}$$

The Cardan solution for this last equation,

$$q^2 = (1/3\lambda b)\{2\cos[(\pi - \phi)/3] - 1\}, \tag{5.15a}$$

TABLE IV. Lowest eigenvalue of the quartic anharmonic oscillator calculated by means of several approximation procedures.

| $\lambda$ | $E^a$ | $E^b$ | $E^c$ | $E^d$ | $E^e$ | $E^f$ | $E^g$ | $E^{exact48}$ |
|---|---|---|---|---|---|---|---|---|
| $10^{-3}$ | 1.000 748 69 | 1.000 748 69 | 1.000 7489 | | 1.000 748 69 | 1.000 748 70 | 1.000 748 69 | 1.000 748 69 |
| $10^{-2}$ | 1.007 373 62 | 1.007 373 61 | 1.007 3907 | | 1.007 373 32 | 1.007 3740 | 1.007 373 65 | 1.007 373 67 |
| $10^{-1}$ | 1.065 2860 | 1.065 2725 | 1.066 2036 | | 1.065 1651 | 1.065 268 | 1.065 2807 | 1.065 2855 |
| 1 | 1.392 951 | 1.392 2297 | 1.403 3233 | 1.529 2635 | 1.388 9615 | 1.392 108 | 1.392 3160 | 1.392 3516 |
| 10 | 2.452 664 | 2.449 0098 | 2.488 624 | 2.851 407 | 2.433 9912 | 2.448 912 | 2.449 1392 | 2.449 1741 |
| $10^2$ | 5.008 898 | 4.999 2298 | 5.095 161 | | 4.960 7686 | 4.999 270 | 4.999 3984 | 4.999 4176 |
| $10^3$ | 10.660 996 | 10.639 4824 | 10.851 511 | | 10.553 4402 | 10.639 717 | 10.639 7794 | 10.639 7887 |
| $10^4$ | 22.9077 | 22.860 9968 | 23.320 287 | | 22.674 157 | 22.861 574 | 22.861 6044 | 22.861 6089 |

[a] Second-order $u$-functional [$b^*$ according to (3.18)].
[b] Second-order Caswell's results[11] ($b^* = 4$).
[c] First-order $u$-functional (scaling variational method).[19,20]
[d] Second-order Killingbeck's partition scheme.[9]
[e] Second-order Pascual's[4] and Dmitrieva and Plindov's[5,6] renormalized series.
[f] First-order functional with $e_0$.
[g] Second-order functional with $e_0$.

$$\phi = \arccos(1 - \tfrac{27}{2}\lambda^2 b^2), \quad \lambda b < \tfrac{2}{3}3^{-1/2}, \tag{5.15b}$$

$$q^2 = -(1/3\lambda b)\{1 + 2/\sin(2\chi)\}, \tag{5.15c}$$

$$\chi = \arctan\{\tan^{1/3}(\psi/2)\}, \tag{5.15d}$$

$$\psi = \arcsin(1/\cos\phi), \quad \lambda b > \tfrac{2}{3}3^{-1/2}, \tag{5.15e}$$

allows us to compute $\partial(q^2)/\partial b$ exactly, and so we can calculate the stationary points very accurately.

When only the first three coefficients $b_i$ of the $u$ expansion (2.18) are included in the functional, and, taking into account that $E_0 = 1$ ($A = 0$) for the lowest eigenvalue (the only we consider here), we obtain

$$F = q^2 + \lambda b_0 q^4 + \lambda^2 b_1 q^{10} + \lambda^3 b_2 q^{16}, \tag{5.16}$$

where

$$b_0 = 3/4 + b/2, \tag{5.17a}$$

$$b_1 = -b^2/8 + 3b/4 - 21/16. \tag{5.17b}$$

When $b_i = 0$ ($i \geqslant 2$), the functional $F$ has no stationary points but an inflexion point changing from $b^* \simeq 4.00$ to $b^* \simeq 4.52$ as $\lambda$ runs from 0 to $\infty$. Taking $b^* = 4$, we obtain second-order Caswell's results[11] exactly. The inflexion point lies within a plateau that becomes larger as $\lambda$ decreases. We have computed the lowest eigenvalue of the quartic anharmonic oscillator for different $\lambda$ values using the plateau criterion of Sec. III to determine $b^*$. Results are compared in Table IV with the exact (numerical) ones[48] and with those obtained by means of other approximation methods: (i) second-order Caswell's results ($b^* = 4$)[11]; (ii) first-order FM calculation (scaling-variational method[30,31]); (iii) second-order Killingbeck's partition scheme[9]; (iv) second-order Pascual's[4] and Dmitrieva and Plindov's[5,6] renormalized series. The best results are obtained through Caswell's generalized Wick-ordering method that is a particular case of the FM. The great success of this calculation is due to a fortunate choice for $b^*$.

To show the advantage of the FM we will now discuss how to improve previous results by taking into account the large $\lambda$ behavior of the ground-state eigenvalue $E(\lambda)$. When $\lambda \to \infty$, the quantum-mechanical problem (5.13) tends to a quartic oscillator that is a very useful model to study the vibrations of some molecular rings.[49] Furthermore, it can also be related to a zero mass $\phi^4$ field theory. The main coefficient in the large $\lambda$ expansion is $e_0 = 1.060\,362\,077$, which can be introduced in $F$ through $b_{M+1}$ [see (2.18) and (2.20)]. When $M = 0$,

$$b_1 = e_0 b^{5/3} - b^2/2 - 3b/4, \tag{5.18}$$

and $F$ shows an inflexion point $b^*(\lambda = 0) \simeq 3.90$, $b^*(\lambda \to \infty) \simeq 4.35$. On the contrary, when $M = 1$,

$$b_2 = e_0 b^{8/3} - 3b^3/8 - 3b^2/2 + 21b/16, \tag{5.19}$$

and $F$ has a stationary point $b^* = 4.963$ for all $\lambda$ values. Since $F$ was forced to behave properly when $\lambda$ is small or large, the highest errors should appear in the intermediate $\lambda$ regime (boundary layer[47]). Table IV shows that these errors are not higher than $3 \times 10^{-5}$ in the second case ($M = 1$). Clearly, the FM yields excellent results in the whole range of $\lambda$ values which are far better than those obtained through other approximation procedures.

In closing this subsection, we want to discuss another very striking success of the FM. When $e_0$ is taken into account, $F$ can be expanded as

$$F \simeq e_0 \lambda^{1/3} + e_1' \lambda^{-1/3} + \cdots, \tag{5.20a}$$

where

$$e_1' = b^{-1/3} - \tfrac{2}{3}b^{-4/3}(b_0 + b_1 b^{-1} + b_2 b^{-2}) - b^{-2/3}(b_1 b^{-5/3} + 2b_2 b^{-8/3}). \tag{5.20b}$$

When using the proper $b^*$, $b_0$, $b_1$, and $b_2$ values, we obtain $e_1' = 0.361\,929$ which is an excellent estimation for the exact coefficient $e_1 = 0.362\,022$.[47] Clearly, this fact explains why our approximate eigenvalues are so good even for moderately large $\lambda$ values. In all cases studied, we have found that whenever the functional was forced to behave like $e_0 \lambda^\beta$ in the large $\lambda$ regime, then it provided a very good estimation for $e_1$. This fact suggests that the functional (2.2) has the proper form in order to approximate (2.1).

Recently, we have shown[20,21] that the eigenvalues of the anharmonic oscillators can be accurately computed by way of an improved variational functional method. But this procedure is not so general nor so accurate as the present one.

Almost all techniques for summation of series developed during the last years have been applied to the anharmonic oscillator model, for it is the simplest quantum-me-

chanical nontrivial problem that leads to an RS series with zero convergence radius. On the contrary, only a few of these procedures (or none at all) have been applied to more complex problems. Since the FM does not take into account the Hamiltonian of the problem explicitly, then it is suitable to handle even nonseparable Schrödinger equations provided the power series expansions are known. This fact will be made clear in the next subsection.

### C. The Zeeman effect in hydrogen

The hydrogen atom in a uniform magnetic field of arbitrary intensity[50] poses a quantum-mechanical problem of great interest in physics and astrophysics. The Hamiltonian describing the interaction between the electron of a hydrogenlike atom with nuclear charge $Z$ and the magnetic field $\mathscr{B}$ supposed to be aligned to the $z$ axis is (atomic units are used throughout)

$$H(Z,\lambda) = -\frac{1}{2}\left\{\frac{\partial^2}{\partial\rho^2}+\frac{1}{\rho}\frac{\partial}{\partial\rho}+\frac{\partial^2}{\partial z^2}\right\} + \frac{m^2}{(2\rho^2)}+\frac{\lambda\rho^2}{8}-\frac{Z}{(\rho^2+z^2)^{1/2}}, \tag{5.21}$$

where $\rho^2 = x^2+y^2$, $m$ is the magnetic quantum number, and $\lambda = \mathscr{B}\cdot\mathscr{B}$. The paramagnetic and spin terms were taken off because they are only trivial additive constants.[50]

For small enough $\lambda$ values, the eigenvalues of $H$ can be accurately computed through the RS series. Although it is not rigorously proved yet, most researchers believe that the $\lambda$-power expansion for this problem has zero convergence radius.

For very high magnetic field intensities ($\lambda \to \infty$), the system behaves like a two-dimensional harmonic oscillator with the electron unbounded along the $z$ direction. The scaling law $H(Z,\lambda) = \lambda^{1/2}H(Z\lambda^{-1/4},1)$ suggests that at least part of the eigenvalue $E(\lambda)$ could be expanded in powers of $\lambda^{-1/4}$:

$$E(\lambda) - \Delta E(\lambda) = \lambda^{1/2}\sum_{i=0}^{\infty} e_i \lambda^{-i/4}. \tag{5.22}$$

The other part $\Delta E(\lambda)$ exhibits a logarithmic contribution that cannot be taken into consideration by our formulation[51]

$$\Delta E(\lambda) = -\tfrac{1}{2}\ln^2\lambda + (\ln\lambda)\ln(\ln\lambda) - 2\{\ln(\ln\lambda)\}^2 + \cdots. \tag{5.23}$$

Notwithstanding, we can obtain $s$ and $t$ by considering $\alpha = -1/4$ and $\beta = 1/2$ (see Sec. II). Therefore, the functional should be ($Z = 1$)

$$F = A/q^2 + 1/q + \lambda q^2 B(q), \tag{5.24a}$$

$$q + \lambda b q^4 = 1. \tag{5.24b}$$

Since our $\lambda^{-1/4}$-power expansion is not complete, we write $B$ in terms of $u(=\lambda q^4)$ only [see Eq. (2.18)]. This example reveals the semiclassical origin of the FM for Eqs. (5.24a) and (5.24b) are reminiscent of the classical energy and trajectory, respectively.[52]

In this article we only make a brief exposition of the application of the FM to the Zeeman effect for the hydrogen atom, and a detailed discussion and calculation will be presented in a forthcoming paper.

For the hydrogen atom ($Z = 1$), the coefficients of the functional are

$$A = -1 - 1/(2n^2), \quad n = 1,2,3,..., \tag{5.25a}$$

$$b_0 = E_1 - 2bE_0 + b, \tag{5.25b}$$

$$b_1 = E_2 + b^2E_0 + 2bE_1, \tag{5.25c}$$

$$b_2 = E_3 + 6bE_2 + 3b^2E_1, \tag{5.25d}$$

$$b_3 = E_4 + 10bE_3 + 21b^2E_2 + 4b^3E_1. \tag{5.25e}$$

We consider here the ground state only for which several coefficients $E_i$ have been exactly computed,[23,53]

$$E_1 = 1/4, \quad E_2 = -53/192, \quad E_3 = 5581/4608, \quad E_4 = -21\,577\,397/2\,211\,840. \tag{5.26}$$

Furthermore, the large-field behavior of the eigenvalues can be easily taken into account for $e_0$ is also well known,

$$e_0 = \tfrac{1}{2}(N+1), \tag{5.27}$$

where $N$ is the Landau quantum number that labels the energy levels of the aforesaid two-dimensional harmonic oscillator. As the field increases, the ground state of the hydrogen atom ($n = 1$) tends to the lowest energy level of the oscillator ($N = 0$).[54,55]

Also in this example, the stationary points can be accurately determined by resorting to the analytical solution of the quartic equation (5.24b).[56] We have made two different calculations of the ground-state energy of the hydrogen atom in a magnetic field: the first one by using (5.25b) for $b_0$ and Eq. (2.20) for $b_1$. This case shows a stationary point $b^* = 1$. In the second calculation we have computed $b_0$ and $b_1$ through Eqs. (5.25b), and (5.25c) and $b_2$ by way of (2.20). Table V shows that both sets of results are in an excellent agreement with the exact (numerical) eigenvalues[57] in the whole range of field intensities. As far as we know, there is not an analytical expression in the current literature for the eigenvalues of this problem as simple and accurate as those we have presented here.

Though being an important physical problem, the Zeeman effect for a hydrogen atom is considered here as an example. More accurate calculation (taking more perturbation terms into consideration) will be presented in a forthcoming paper.[58]

## VI. FURTHER COMMENTS AND CONCLUSIONS

We hope that the arguments given in this article are reasonable enough to convince one that the FM is a very useful tool for improving convergence of a wide variety of

TABLE V. Ground-state eignevalue of the hydrogen atom in a uniform magnetic field of intensity $\mathscr{B} = \lambda^{1/2}$ (a.u.).

| $\mathscr{B}$ | $E^a$ | $E^b$ | $E^{exact57}$ |
|---|---|---|---|
| 0.03 | −0.499 775 202 | −0.499 775 223 | −0.499 775 22 |
| 0.10 | −0.497 524 28 | −0.497 526 70 | −0.497 526 48 |
| 0.40 | −0.464 519 | −0.464 816 | −0.464 6054 |
| 1.00 | −0.332 625 | −0.335 132 | −0.3312 |
| 3.00 | 0.311 02 | 0.298 53 | 0.3355 |
| 10.00 | 3.1010 | 3.0611 | 3.2522 |
| 100.0 | 44.686 | 44.496 | 45.27 |

[a] First-order $u$-form of the functional with $e_0$.
[b] Second-order $u$-form of the functional with $e_0$.

power series expansions. Besides the eigenvalues of the two-level model discussed in Sec. IV, our method increases the convergence radii of the Taylor expansions of functions like $1/(1+\lambda)$, $(1+\lambda)^{1/2}$, and several others we have studied.

The examples discussed in the previous sections show, beyond any doubt, that the FM is very useful in calculating physical properties that can be expanded in power series, even though few terms are taken into consideration.

Some problems like the confining potential models

$$H = p^2 - 1/r + \lambda r^k, \quad k = 1,2, \tag{6.1}$$

are specially suitable for being handled in this way because the eigenvalues and eigenfunctions are well known in both limits ($\lambda = 0$ and $\lambda \to \infty$). Therefore, in these cases we can compute analytical coefficients for both power expansions which are of the form (2.1) with $\beta = 2/(k+2)$ and $\alpha = -1/(k+2)$.

In Sec. III our method was shown to be more general than several other ones proposed before. To make this statement even more convincing, let us consider the Stark effect of a planar polar rigid rotator. The Hamiltonian (in appropriate units) is

$$H = -\frac{1}{2}\frac{d^2}{d\theta^2} + \lambda \cos\theta, \quad 0 \leqslant \theta < 2\pi, \tag{6.2}$$

its eigenfunctions obeying $\Psi_n(\theta + 2\pi) = \Psi_n(\theta)$. In the very high field strength limit ($\lambda \to \infty$), the system becomes a one-dimensional oscillator, and the eigenvalues can be expanded in the form (2.1) with $\beta = 1/2$ and $\alpha = -1/2$. Many other very interesting physical models like a linear (or symmetric-top) polar molecule (in the rigid-rotator approximation) in a uniform electric field[59] behave in this way and can be easily handled by means of the FM. In all these cases, the potential-energy operator is not an homogeneous function of the angular coordinate, and so it does not obey any scaling law. This fact and the periodic boundary conditions make it difficult to apply the convergence techniques described in Sec. III. On the contrary, all this does not matter when using the FM for this purpose, because it only takes into consideration the power series expansions.

*Note added in proof*: Here we briefly prove that the stationary points of $F$ are $\lambda$ independent.

By differentiating Eqs. (2.10) and (2.11), where $\lambda$ is held constant and $B = B_M$, with respect to $b$ we obtain

$$\left(\frac{\partial F}{\partial b}\right)_\lambda = -(1-bu)^{\beta/\alpha}\alpha^{-1}b^{2\alpha+1}\{(\alpha+1)bu - \alpha\}^{-1} \times \{P_0 + P_1 bu + \cdots + P_{M+2}(bu)^{M+2}\},$$

where

$$P_i = iJ_i + \frac{b^2}{\alpha}\frac{\partial J_i}{\partial b} - \left(\frac{\beta}{\alpha} + i - 1\right)J_{i-1} - \frac{b^2}{\alpha}(\alpha+1)\frac{\partial J_{i-1}}{\partial b},$$

$$J_i = d_i/b^i$$

[see Eq. (3.39)], and

$$J_i = 0 \quad \text{if } i<0 \quad \text{or} \quad i > M+1.$$

By virtue of Eq. (2.12) and the properties of the binomial coefficients $\binom{a}{i}$, it follows that $P_i = 0$ if $i < M+2$. We then conclude that the stationary points of $F$ are the roots of $P_{M+2}$ and, therefore, $\lambda$ independent.

A similar reasoning shows that the inflexion points of $F$ remain finite for all $\lambda$ values.


[1] J. J. Loeffel, A. Martin, B. Simon, and A. S. Wightman, Phys. Lett. B **30**, 656–8 (1969).
[2] B. Simon, Ann. Phys. (N.Y.) **58**, 79–136 (1970).
[3] S. Graffi, V. Grecchi, and B. Simon, Phys. Lett. B **32**, 631–4 (1970).
[4] P. Pascual, Ann. Fis. **75**, 77–80 (1979).
[5] I. K. Dmitrieva and G. I. Plindov, Phys. Lett. A **79**, 47–50 (1980).
[6] I. K. Dmitrieva and G. I. Plindov, Phys. Scr. **22**, 386–8 (1980).
[7] K. Bhattacharyya, J. Phys. B **14**, 783–93 (1981).
[8] K. Bhattacharyya, Int. J. Quantum Chem. **20**, 1273–84 (1981).
[9] J. Killingbeck, J. Phys. A **14**, 1005–8 (1981).
[10] E. J. Austin and J. Killingbeck, J. Phys. A **15**, L443–5 (1982).
[11] W. E. Caswell, Ann. Phys. (N.Y.) **123**, 153–84 (1979).
[12] R. Seznec and J. Zinn-Justin, J. Math. Phys. **20**, 1398–1408 (1979).
[13] I. G. Halliday and P. Suranyi, Phys. Lett. B **85**, 421–3 (1979).
[14] I. G. Halliday and P. Suranyi, Phys. Rev. D **21**, 1529–37 (1980).
[15] I. D. Feranchuck and L. I. Komarov, Phys. Lett. A **88**, 211–4 (1982).
[16] J. Katriel, Phys. Lett. A **72**, 94–6 (1979).
[17] G. Rosen, Phys. Rev. A **20**, 1287–8 (1979).
[18] H. Orland, Phys. Rev. Lett. **42**, 285–7 (1979).
[19] F. M. Fernández and E. A. Castro, Phys. Rev. A **27**, 2735–7 (1983).
[20] G. A. Arteca, F. M. Fernández, and E. A. Castro, J. Math. Phys. (to be published).
[21] F. M. Fernández, G. A. Arteca, and E. A. Castro, Physica A **122**, 37–49 (1983).
[22] B. Simon, Int. J. Quantum Chem. **21**, 3–25 (1982).
[23] J. Čížek and E. R. Vrscay, Int. J. Quantum Chem. **21**, 27–68 (1982).
[24] C. M. Bender and T. T. Wu, Phys. Rev. **189**, 1231–60 (1969).
[25] P. O. Löwdin, Int. J. Quantum Chem. **21**, 69–92 (1982).
[26] E. Feenberg, Ann. Phys. (N.Y.) **3**, 292–303 (1958).
[27] S. Wilson, J. Phys. B **12**, 1623–31, L599–L601 (1979).
[28] G. N. Hardy, *Divergent Series* (Oxford U. P., Oxford, 1949).
[29] K. Bhattacharyya, Int. J. Quantum Chem. **22**, 307–330 (1982).
[30] F. M. Fernández and E. A. Castro, Phys. Rev. A **27**, 663–9 (1983).
[31] F. M. Fernández and E. A. Castro, J. Chem. Phys. **79**, 321–4 (1983).
[32] F. M. Fernández and E. A. Castro, Phys. Lett. A **91**, 339–40 (1982).
[33] V. M. Vaĭnberg, Y. L. Eletskiĭ, and V. S. Popov, Zh. Eksp. Teor. Fiz. **81**, 1567 (1981) [Sov. Phys. JETP **54**, 833–40 (1981)].
[34] V. S. Popov and V. M. Weinberg, Phys. Lett. A **90**, 107–9 (1982).
[35] C. Quigg and J. L. Rosner, Phys. Rep. C **56**, 167–235 (1979).
[36] P. J. Ellis and E. Osnes, Phys. Lett. B **45**, 425–8 (1973).
[37] J. M. Leinaas and E. Osnes, Phys. Scr. **22**, 193–6 (1980).
[38] D. McLachlan and W. R. Foster, J. Solid State Chem. **30**, 257–9 (1977).
[39] A. Morita and D. G. Frood, J. Phys. D **11**, 2409–13 (1978).
[40] A. D. Booth, J. Comput. Phys. **46**, 423–8 (1982).
[41] Y. Onodera, Progr. Theoret. Phys. (Kyoto) **44**, 1477–99 (1970).
[42] J. K. Vij and W. G. S. Scaife, J. Chem. Phys. **64**, 226–8 (1976).
[43] J. Zinn-Justin, Phys. Rep. **70**, 109–67 (1981).
[44] J. C. Le Guillou and J. Zinn-Justin, Phys. Rev. B **21**, 3976–8 (1980).
[45] J. Zinn-Justin, J. Math. Phys. **22**, 511–20 (1981).
[46] C. M. Bender and R. Z. Roskies, Phys. Rev. D **25**, 427–33 (1982).
[47] F. T. Hioe, D. MacMillen, and E. Montroll, J. Math. Phys. **17**, 1320–37 (1976).
[48] K. Banerjee, Proc. R. Soc. London Ser. A **364**, 265–75 (1978).
[49] S. I. Chan, D. Stelman, and L. E. Thompson, J. Chem. Phys. **41**, 2828–35 (1964).
[50] R. H. Garstang, Rep. Progr. Phys. **40**, 105–154 (1977).
[51] J. E. Avron, I. W. Herbst, and B. Simon, Commun. Math. Phys. **79**, 529–72 (1981).
[52] C. M. Bender, L. D. Mlodinov, and N. Papanicolaou, Phys. Rev. A **25**, 1305–14 (1982).
[53] J. E. Avron, B. G. Adams, J. Čížek, M. Clay, M. Glasser, P. Otto, J. Paldus, and E. Vrscay, Phys. Rev. Lett. **43**, 691–3 (1979).
[54] J. Simola and J. Virtamo, J. Phys. B **11**, 3309–22 (1978).
[55] M. Robnik, J. Phys. A **14**, 3195–3216. (1981).
[56] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions* (Dover, New York, 1970), p. 17.
[57] G. Wunner and H. Ruder, J. Phys. (Paris) **43**, C2–137 (1982).
[58] G. A. Arteca, F. M. Fernández, and E. A. Castro, Chem. Phys. Lett. **102**, 344 (1983).
[59] S. A. Malvendes, F. M. Fernández, and E. A. Castro (unpublished results).

# Poisson structure of the equations of ideal multispecies fluid electrodynamics

Richard G. Spencer[a)]
*Lawrence Berkeley Laboratory, University of California, Berkeley, California 94720*

(Received 19 July 1983; accepted for publication 15 March 1984)

The equations of the two- (or multi-) fluid model of plasma physics are recast in Hamiltonian form, following general methods of symplectic geometry. The dynamical variables are the fields of physical interest, but are noncanonical, so that the Poisson bracket in the theory is not the standard one. However, it is a skew-symmetric bilinear form which, from the method of derivation, automatically satisfies the Jacobi identity; therefore, this noncanonical structure has all the essential properties of a canonical Poisson bracket.

PACS numbers: 02.40.Vh, 47.65. + a, 03.40.Gc, 52.30. + r

## I. INTRODUCTION

In this paper, we present a derivation of the Poisson structure for ideal multifluid electrodynamics, using methods of symplectic geometry. As a corollary, we shall also obtain the structure for the Coulomb case. These models are thus placed in the larger context of modern classical mechanics as Hamiltonian field theories. For further background and another application of the geometric techniques used in this paper see Ref. 1, in which the Poisson structure for the Maxwell–Vlasov system is derived. Our results have already been presented elsewhere,[2] but without derivation.

Since the completion of this work, other methods have been discussed for deriving Eq. (26). Kaufman[3] has done so from a particle Lagrangian, while Kupershmidt and Holm[4] have used Clebsch variables. In addition, Marsden and Weinstein have proved that Eq. (26) can be derived from the equivalent result for the Maxwell–Vlasov plasma, Eq. (7.1) of Ref. 1, by, essentially, averaging over the particle distribution function (see Ref. 5). They have also pointed out similarities with work of Menikoff and Sharp[6] on current algebras, in which, starting with quantum commutation relations, an equivalent semidirect product group is constructed.

We consider a system composed of one or more fluids. Each species, with species label denoted by $s$, is characterized by the mass $m_s$ and charge $q_s$ of the particles of which it is composed; let $a_s = q_s/m_s$. The most common situation encountered in plasma physics is to have only two species present: electrons and singly charged positive ions. In terms of the electric field $\mathbf{E}$, magnetic field $\mathbf{B}$, fluid velocities $\mathbf{u}_s$, mass densities $\rho_s$, and specific entropies $\sigma_s$, the equations of ideal multifluid dynamics, in rationalized units, are

$$\frac{\partial \mathbf{E}}{\partial t} = \nabla\times\mathbf{B} - \sum_s a_s \rho_s \mathbf{u}_s \,, \tag{1a}$$

$$\frac{\partial B}{\partial t} = -\nabla\times\mathbf{E} \,, \tag{1b}$$

$$\nabla\cdot\mathbf{E} = \sum_s a_s \rho_s + \rho_{\text{ext}} \,, \tag{2a}$$

$$\nabla\cdot\mathbf{B} = 0 \,, \tag{2b}$$

$$\frac{\partial \mathbf{u}_s}{\partial t} = -(\mathbf{u}_s\cdot\nabla)\mathbf{u}_s + a_s(\mathbf{E} + \mathbf{u}_s\times\mathbf{B}) - \rho_s^{-1}\nabla p_s \,, \tag{3a}$$

$$\frac{\partial \rho_s}{\partial t} = -\nabla\cdot(\rho_s \mathbf{u}_s) \,, \tag{3b}$$

$$\frac{\partial \sigma_s}{\partial t} = -\mathbf{u}_s\cdot\nabla\sigma_s \,, \tag{3c}$$

where the specific internal energy $U_s(\rho_s,\sigma_s)$, expressed as an equation of state, yields the (partial) pressure $p_s$ according to

$$p_s = \rho_s^2 \frac{\partial U_s}{\partial \rho_s} \,. \tag{4}$$

Equations (1) and (2) are the Maxwell equations, including an external static charge density $\rho_{\text{ext}}(\mathbf{x})$, and Eqs. (3) and (4) are the laws of compressible ideal fluid dynamics. We neglect dissipation and therefore express entropy convection by the adiabatic condition Eq. (3c).

The set of state variables for this system of $k$ fluids consists of the electric and magnetic fields, along with the $3k$ dynamical variables $\{\mathbf{u}_s,\rho_s,\sigma_s\}$, $s = 1,\ldots,k$. The evolution equations for these quantities are Eqs. (1) and (3), while Eqs. (2) may be regarded as constraints.

The energy of the combined system of fluids and fields is equal to the sum of the kinetic and internal energies of the fluids, and the energy of the Maxwell fields. The Hamiltonian can be written

$$H(\mathbf{u}_s,\rho_s,\sigma_s,\mathbf{B},\mathbf{E}) = \sum_s H_s(\mathbf{u}_s,\rho_s,\sigma_s) + H_M(\mathbf{B},\mathbf{E}) \,. \tag{5}$$

[We shall sometimes, as on the left-hand side of Eq. (5), write, e.g., $\rho_s$, using the general species index $s$ to stand for the set of all $k$ species. Whether this is the case or whether $\rho_s$ refers to the single species $s$ will always be clear from the context.]

Before proceeding, we replace the velocity variables $\mathbf{u}_s$ by their corresponding momentum densities $\mathbf{M}_s \equiv \rho_s\mathbf{u}_s$. Then (5) is, explicitly,

$$H(\mathbf{M}_s,\rho_s,\sigma_s,\mathbf{B},\mathbf{E}) = \sum_s \int \left(\frac{1}{2}\rho_s^{-1}|\mathbf{M}_s|^2 + \rho_s U_s(\rho_s,\sigma_s)\right) d^3x + \int \left(\frac{1}{2}|\mathbf{B}|^2 + \frac{1}{2}|\mathbf{E}|^2\right) d^3x \tag{6}$$

[a)] Current address: Harvard Medical School, 25 Shattuck Street, Boston, MA 02115.

 0022-2488/84/082390-07$02.50  

and Eqs. (1) and (3) can likewise easily be written in terms of $\mathbf{M}_s$. We denote the phase space for each fluid by $\mathscr{P}_s \equiv \{(\mathbf{M}_s, \rho_s, \sigma_s)\}$, the Maxwell phase space by $\mathscr{P}_M \equiv \{(\mathbf{B},\mathbf{E})\}$, and the phase space for the combined system by $\mathscr{P} \equiv \{(\mathbf{M}_s, \rho_s, \sigma_s, \mathbf{B}, \mathbf{E})\}$.

In the following, $\mathscr{F}(S)$ will be used to denote real-valued functions defined on the space $S$.

We wish to derive an expression for a Poisson bracket $\{\ ,\ \}: \mathscr{F}(\mathscr{P}) \times \mathscr{F}(\mathscr{P}) \to \mathscr{F}(\mathscr{P})$ with the following two properties.

(i) $(\mathscr{F}(\mathscr{P}), \{\ ,\ \})$ is a Lie algebra.

(ii) The evolution of a phase functional $\mathscr{F}: \mathscr{P} \to \mathbb{R}$ can be written in the form of a Hamiltonian evolution equation

$$\dot{F} = \{F, H\}, \tag{7}$$

where the Hamiltonian $H$ is given by the energy, Eq. (6). $F$ can be, for example, a component of one of the field variables, yielding Eqs. (1) and (3).

In Sec. II we briefly review some of the necessary mathematics and establish our notation. In Sec. III we discuss a derivation of the Hamiltonian structure for the dynamics of ordinary ideal fluids. This discussion is based on an explicit realization by Marsden and Weinstein of ideas contained implicitly in work of Arnold.[7] In Sec. IV we review briefly Marsden and Weinstein's treatment of the vacuum Maxwell equations. This is an essential ingredient both in their formulation of the Maxwell–Vlasov model of plasma dynamics, and in the formulation of the multispecies fluid model to be discussed in the present paper. In Sec. V, the formalisms of the two preceding sections are combined, following the general plan of reduction of symplectic manifolds with symmetry[8] and the coupling of Hamiltonian systems to gauge fields.[9] This produces the desired form, Eq. (7), for Eqs. (1) and (3), and in addition, yields Eqs. (2) as a direct consequence of gauge invariance.

## II. IDEAL FLUID DYNAMICS

The equations of motion of an ideal compressible fluid are Eqs. (3) and (4), with $k = 1$ and $a_1 = 0$. The Hamiltonian is the first integral in (6). Deleting the subscript $s$ for a single fluid, the dynamical variables are $(\mathbf{M}, \rho, \sigma)$, and the phase space consists of the set of all such triples. Therefore, we may define this system by

$$\frac{\partial \mathbf{u}}{\partial t} = -(\mathbf{u}\cdot\nabla)\mathbf{u} - \rho^{-1}\nabla p, \quad p = \rho^2 \frac{\partial U}{\partial \rho}, \tag{8a}$$

$$\frac{\partial \rho}{\partial t} = -\nabla\cdot(\rho \mathbf{u}), \tag{8b}$$

$$\frac{\partial \sigma}{\partial t} = -\mathbf{u}\cdot\nabla\sigma, \tag{8c}$$

$$H(\mathbf{M}, \rho, \sigma) = \int \left(\frac{1}{2}\rho^{-1}|\mathbf{M}|^2 + \rho U(\rho,\sigma)\right) d^3x. \tag{9}$$

We shall now review the construction used to obtain the Hamiltonian structure of ideal fluid dynamics. Let $G$ be a Lie group, $\mathfrak{g}$ its Lie algebra, and $\mathfrak{g}^*$ the dual space to $\mathfrak{g}$. The duality pairing is denoted by $\langle\ ,\ \rangle$. For a function $F \in \mathscr{F}(\mathfrak{g}^*)$, the functional derivative $\delta F/\delta\mu$ with respect to the variable $\mu$ in $\mathfrak{g}^*$ is defined by

$$DF(\mu)\cdot\nu \equiv \left\langle \nu, \frac{\delta F}{\delta \mu} \right\rangle \quad \text{for all } \nu \in \mathfrak{g}. \tag{10}$$

Here, $\mathfrak{g}^{**}$ has been identified with $\mathfrak{g}$ so that $\delta F/\delta\mu \in \mathfrak{g}$. We will be concerned with the case in which $\mathfrak{g}$ is comprised of vector and scalar fields on $\mathbb{R}^3$. Then $\mathfrak{g}^*$ may be considered to consist of vector fields and scalar fields, with the $L^2$ pairing, or of vector field and scalar field densities. In this case, for $f \in \mathfrak{g}$, $g \in \mathfrak{g}^*$, we may write $\langle f, g\rangle = \int g f\, d^3x$ and Eq. (10) may be rewritten

$$\int \frac{\delta F}{\delta f} \hat{f}\, d^3x \equiv \left.\frac{d}{d\epsilon}\right|_{\epsilon=0} F(f + \epsilon \hat{f}), \tag{11}$$

where $F \in \mathscr{F}(\mathfrak{g}^*)$. This is the most convenient form for explicit calculations.

Now consider the orbits $\mathscr{O}_\mu^*$, $\mu \in \mathfrak{g}^*$, of the co-adjoint representation of $G$ in $\mathfrak{g}^*$. In other words, $\mathscr{O}_\mu^* = \{(\mathrm{Ad}_{g^{-1}})^*(\mu) | g \in G\}$, where $\mathrm{Ad}_g^*: \mathfrak{g}^* \to \mathfrak{g}^*$ is the dual map to $\mathrm{Ad}_g: \mathfrak{g} \to \mathfrak{g}$. The latter is defined by $\mathrm{Ad}_g = T_e(r_{g^{-1}} \circ l_g)$, the tangent map at the identity $e \in G$ of the indicated composition of right and left translation. The Kirillov–Kostant–Souriau theorem states that the $\mathscr{O}_\mu^*$ are symplectic manifolds, with a nondegenerate Poisson bracket of functions $F, G \in \mathscr{F}(\mathscr{O}_\mu^*)$ defined by

$$\{F, G\}(\alpha) = -\left\langle \alpha, \left[ \left.\frac{\delta F}{\delta \beta}\right|_\alpha, \left.\frac{\delta G}{\delta \beta}\right|_\alpha \right] \right\rangle, \tag{12}$$

where $\beta$ is the variable in $\mathscr{O}_\mu^*$, $\alpha \in \mathscr{O}_\mu^*$, and $[\ ,\ ]$ is the Lie algebra bracket on $\mathfrak{g}$. This structure can then be extended to a degenerate Poisson bracket $\{\ ,\ \}$ on all of $\mathfrak{g}^*$, which is a disjoint union of the $\mathscr{O}_\mu^*$, given by the same formula.

For the case in which $\mathfrak{g}$ is comprised of vector and scalar fields on $\mathbb{R}^3$, Eq. (12) may be rewritten

$$\{F, G\}(\alpha) = -\int \alpha \left[ \left.\frac{\delta F}{\delta \beta}\right|_\alpha, \left.\frac{\delta G}{\delta \beta}\right|_\alpha \right] d^3x. \tag{13}$$

We shall refer to Eq. (12) or (13) as the Lie–Poisson bracket. A proof of the theorem, as well as a clear exposition of the background mathematics, may be found in Chapter 4 of Ref. 10.

We claim that a suitable Poisson bracket for the system (8), (9) can be derived as the Lie–Poisson symplectic structure on the co-adjoint orbits of a certain Lie group $G$. The spaces of vector and scalar fields on $\mathbb{R}^3$ will be denoted by $\mathscr{V}$ and $\mathscr{F}$, and the spaces of field densities will be denoted by $\mathscr{V}^*$ and $\mathscr{F}^*$.

In order to guess what $G$ might be, we are motivated by the physical set-up in the following way. The phase space $\{(\mathbf{M}, \rho, \sigma)\}$ must be the dual $\mathfrak{g}^*$ to the Lie algebra $\mathfrak{g}$ of the group $G$. The set of momentum densities may be regarded as the set of vector field densities on $\mathbb{R}^3$, and mass and entropy densities are scalar field densities. Thus the fluid phase space is the direct product

$$\mathscr{P}_{\mathfrak{fl}} \equiv \mathfrak{g}^* = \mathscr{V}^* \times \mathscr{F}^* \times \mathscr{F}^*.$$

An isomorphism between $\mathfrak{g}$ and $\mathfrak{g}^*$ is established by the $L^2$ pairing; it is therefore immediate that $\mathfrak{g} = \mathscr{V} \times \mathscr{F} \times \mathscr{F}$. $\mathscr{V}$, which may be thought of as containing velocity fields on $\mathbb{R}^3$, is the set of generators of displacements of the fluid.

Hence, the part of $G$ which corresponds to the factor $\mathscr{K}$ in $\mathfrak{g}$ is the group of diffeomorphisms on $\mathbb{R}^3$, denoted $\mathscr{D}$ (see Ref. 11 for a discussion of $\mathscr{D}$ as a Lie group). On the other hand, $\mathscr{F}$ is a vector space, so that the parts of $G$ corresponding to these factors in $\mathfrak{g}$ are $\mathscr{F}$ itself.

Throughout this paper, we assume that all fields are $C^\infty$. One can then show[11] that $\mathscr{D}$ and $G$ are smooth manifolds. $G$ may then be regarded as sets of triples of parameters labeling positions, densities, and entropies of fluid elements, and one obvious possibility is that $G = \mathscr{D} \times \mathscr{F} \times \mathscr{F}$. This direct product structure implies complete independence of the groups involved. Physically, however, we note that the dynamical variables are related in the following way. Let $\eta$ denote a diffeomorphism representing the change in position of fluid elements from the set of their initial positions $\{\mathbf{x}_0\}$, to their positions $\{\mathbf{x}\}$ at some later time. Then for the values of the density and entropy of the fluid element at $\mathbf{x}$, we have, respectively, $\rho(\mathbf{x}) = \rho \circ \eta(\mathbf{x}_0)$ and $\sigma(\mathbf{x}) = \sigma \circ \eta(\mathbf{x}_0)$. In other words, we keep track of these quantities by composition with $\eta$. We postulate then that the group $\mathscr{D}$ acts on $\mathscr{F} \times \mathscr{F}$ according to

$$\eta \cdot (f,g) = (f \circ \eta, g \circ \eta), \quad \text{where } (f,g) \in \mathscr{F} \times \mathscr{F}, \tag{14}$$

and therefore conjecture that

$$G = \mathscr{D} \circledS (\mathscr{F} \times \mathscr{F}). \tag{15}$$

The semidirect product structure, indicated by $\circledS$, is specified by the action (14), linking the field variables.

The rest of the calculation is the explicit construction of the Lie–Poisson structure for $\mathfrak{g}^*$. First, we recall two facts from Lie group theory. Let $H = H_1 \circledS H_2$ be a group defined as the semidirect product of the groups $H_1$ and $H_2$, with the action $\Phi: H_1 \times H_2 \to H_2$ of $H_1$ on $H_2$ specified. Then the induced infinitesimal (Lie-algebraic) action $\phi$ of the Lie algebra $\mathfrak{h}_1$ of $H_1$ on the Lie algebra $\mathfrak{h}_2$ of $H_2$ is defined as the tangent map of $\Phi$ at the identity

$$\phi \equiv T_e \Phi: \mathfrak{h}_1 \times \mathfrak{h}_2 \to \mathfrak{h}_2. \tag{16}$$

Furthermore, the Lie bracket on $\mathfrak{h}_1 \times \mathfrak{h}_2$ is given by

$$[(\xi_1,\xi_2),(\zeta_1,\zeta_2)] = ([\xi_1,\zeta_1],[\xi_2,\zeta_2] + \phi(\xi_1,\zeta_2) - \phi(\xi_2,\zeta_1)), \tag{17}$$

where $(\xi_1,\xi_2)$ and $(\zeta_1,\zeta_2)$ are in $\mathfrak{h}_1 \times \mathfrak{h}_2$. The following lemma follows from (16).

*Lemma*: The infinitesimal action $\phi$ of $\mathscr{K}$ on $\mathscr{F} \times \mathscr{F}$ induced by the action (14) is

$$\phi(X,(f,g)) = (X \cdot df, X \cdot dg).$$

Using this and the fact that $\mathscr{F} \times \mathscr{F}$ is an abelian Lie algebra, we find from (17):

*Lemma*: The bracket in the Lie algebra $\mathfrak{g}$ of $G$ is

$$[(X,(f_1,g_1)),(Y,(f_2,g_2))] = (X \cdot dY - Y \cdot dX, X \cdot df_2 - Y \cdot df_1, X \cdot dg_2 - Y \cdot dg_1). \tag{18}$$

With this, we can now write down the Lie–Poisson structure on $\mathfrak{g}^*$.

*Proposition*: For two phase functionals $F, G: \mathfrak{g}^* \to \mathbb{R}$, the Lie–Poisson bracket on $\mathfrak{g}^*$ is

$$\begin{aligned}\{F,G\}(\mathbf{M},\rho,\sigma) = &-\int \mathbf{M} \cdot \left(\frac{\delta F}{\delta \mathbf{M}} \cdot \nabla \frac{\delta G}{\delta \mathbf{M}} - \frac{\delta G}{\delta \mathbf{M}} \cdot \nabla \frac{\delta F}{\delta \mathbf{M}}\right) d^3x \\ &- \int \rho \left(\frac{\delta F}{\delta \mathbf{M}} \cdot \nabla \frac{\delta G}{\delta \rho} - \frac{\delta G}{\delta \mathbf{M}} \cdot \nabla \frac{\delta F}{\delta \rho}\right) d^3x \\ &- \int \sigma \left(\frac{\delta F}{\delta \mathbf{M}} \cdot \nabla \frac{\delta G}{\delta \sigma} - \frac{\delta G}{\delta \mathbf{M}} \cdot \nabla \frac{\delta F}{\delta \sigma}\right) d^3x. \end{aligned} \tag{19}$$

*Proof*: Using the result of the second of the above lemmas in Eq. (13), we compute

$$\begin{aligned}\{F,G\}(\mathbf{M},\rho,\sigma) &= -\int (\mathbf{M},\rho,\sigma) \cdot \left[\left(\frac{\delta F}{\delta \mathbf{M}}, \frac{\delta F}{\delta \rho}, \frac{\delta F}{\delta \sigma}\right), \left(\frac{\delta G}{\delta \mathbf{M}}, \frac{\delta G}{\delta \rho}, \frac{\delta G}{\delta \sigma}\right)\right] d^3x \\ &= \int (\mathbf{M},\rho,\sigma) \cdot \left(\frac{\delta F}{\delta \mathbf{M}} \cdot \nabla \frac{\delta G}{\delta \mathbf{M}} - \frac{\delta G}{\delta \mathbf{M}} \cdot \nabla \frac{\delta F}{\delta \mathbf{M}}, \frac{\delta F}{\delta \mathbf{M}} \cdot \nabla \frac{\delta G}{\delta \rho} - \frac{\delta G}{\delta \mathbf{M}} \cdot \nabla \frac{\delta F}{\delta \rho}, \right. \\ &\qquad \left. \frac{\delta F}{\delta \mathbf{M}} \cdot \nabla \frac{\delta G}{\delta \sigma} - \frac{\delta G}{\delta \mathbf{M}} \cdot \nabla \frac{\delta F}{\delta \sigma}\right) d^3x,\end{aligned}$$

from which the result follows. This bracket is identical to the one obtained by Morris and Greene for ideal fluids [Ref. 12, Eq. (9)], using different methods.

The final step is the verification of the following:

*Proposition*: Equations (8) follow from Eqs. (9) and (19) in the form (7).

*Proof*: In this case, the functional derivatives of $H$ are just the partial derivatives of the integrand in Eq. (9). In other words, writing

$$\mathscr{H} = \tfrac{1}{2}\rho^{-1}|\mathbf{M}|^2 + \rho U(\rho,\sigma)$$

for the Hamiltonian density, so that $H = \int \mathscr{H}\, d^3x$, we have

$$\frac{\delta H}{\delta \mathbf{M}} = \frac{\partial \mathscr{H}}{\partial \mathbf{M}}, \quad \frac{\delta H}{\delta \rho} = \frac{\partial \mathscr{H}}{\partial \rho}, \quad \frac{\delta H}{\delta \sigma} = \frac{\partial \mathscr{H}}{\partial \sigma}.$$

Then Eq. (7) yields equations of motion equivalent to Eqs. (8), which are obtained from them by making the change of variables $\mathbf{M} \equiv \rho \mathbf{u}$.

Note that the validity of this proposition does not follow automatically from what preceded it. Although one tries to motivate a guess of the correct group from physical considerations, the Lie–Poisson structure can of course always be constructed, even if the wrong group is used. Therefore, the equations of motion must always be checked.

## III. THE VACUUM MAXWELL EQUATIONS

The vacuum Maxwell equations have been treated as a Hamiltonian system by Pauli[13] and by Born and Infeld,[14] but

the discussion by Marsden and Weinstein[1] from the viewpoint of the reduction procedure has additional noteworthy aspects. It is appropriate now to recall some relevant definitions and theorems. (See Chap. 4 of Ref. 10 for proofs and further details.)

For an action $\Phi: G \times \mathscr{P} \to \mathscr{P}$ of a Lie group $G$ on a manifold $\mathscr{P}$, the infinitesimal generator corresponding to $\xi \in \mathfrak{g}$, $\xi_{\mathscr{P}}$, is the vector field defined by

$$\xi_{\mathscr{P}}(p) = \frac{d}{dt}\bigg|_{t=0} \Phi(\exp(t\xi), p) .$$

We assume that there is an associated momentum map $\hat{J}: \mathfrak{g} \to \mathscr{F}(\mathscr{P})$ with the property $X_{\hat{J}(\xi)} \equiv \xi_{\mathscr{P}}$, where $X_{\hat{J}(\xi)}$ is the Hamiltonian vector field on $\mathscr{P}$ with energy function $\hat{J}(\xi)$. Then the momentum map $J: \mathscr{P} \to \mathfrak{g}^*$ for the action $\Phi$ is defined in terms of $\hat{J}$ according to $\langle J(p), \xi \rangle \equiv \hat{J}(\xi)(p)$. Now suppose $H: \mathscr{P} \to \mathbb{R}$ is $G$-invariant, that is, $H(\Phi(g,p)) = H(p)$ for all $p \in \mathscr{P}, g \in G$. Then $\hat{J}(\xi)$ is a constant of motion for the dynamics generated by $H$. In other words, to every one-parameter subgroup of $G$, that is, to every $\xi \in \mathfrak{g}$, there is associated a constant of the motion $\hat{J}(\xi)$ under the dynamics of any $G$-invariant Hamiltonian.

It follows from this that the momentum map $J$ is a constant of motion for any $G$-invariant Hamiltonian $H$; in other words, the dynamics take place on level sets of $J$. Now let $G$ be abelian and define an equivalence relation on $\mathscr{P}$ by $p_1 \sim p_2$ if and only if there is a $g \in G$ such that $\Phi_g(p_1) = p_2$. The quotient $\mathscr{P}/G$ is a well-defined Poisson manifold, provided $\Phi$ is proper and free, with the Poisson bracket of functions $F_1, F_2 \in \mathscr{F}(\mathscr{P}/G)$ obtained by lifting $F_1, F_2$ from $\mathscr{P}/G$ to $\mathscr{P}$. The quotient spaces $J^{-1}(\mu)/G \equiv \mathscr{P}_\mu$, $\mu \in \mathfrak{g}^*$, form symplectic leaves of $\mathscr{P}/G$, in the same way that the $\mathscr{O}_\mu^*$ are symplectic leaves of $\mathfrak{g}^*$ (for $\mu$ a regular value of $J$).

The phase space for electromagnetism is the cotangent bundle $T^*\mathfrak{A}$ to the space $\mathfrak{A}$ of vector potential fields $\mathbf{A}$ on $\mathbb{R}^3$. This bundle may be identified with pairs $(\mathbf{A}, \mathbf{Y})$, where $\mathbf{Y}$ is a vector field density on $\mathbb{R}^3$. With the $L^2$ pairing, one has the canonical (cotangent bundle) Poisson bracket on functions $F, G: T^*\mathfrak{A} \to \mathbb{R}$

$$\{F,G\}(\mathbf{A},\mathbf{Y}) = \int \left( \frac{\delta F}{\delta \mathbf{A}} \cdot \frac{\delta G}{\delta \mathbf{Y}} - \frac{\delta F}{\delta \mathbf{Y}} \cdot \frac{\delta G}{\delta \mathbf{A}} \right) d^3x .$$

Defining the electric and magnetic fields by

$$\mathbf{E} = -\mathbf{Y} ,$$
$$\mathbf{B} = \nabla \times \mathbf{A} ,$$

and transforming by this change of variables, one obtains

$$\{F,G\}(\mathbf{B},\mathbf{E}) = \int \left( \frac{\delta F}{\delta \mathbf{E}} \cdot \left( \nabla \times \frac{\delta G}{\delta \mathbf{B}} \right) - \frac{\delta G}{\delta \mathbf{E}} \cdot \left( \nabla \times \frac{\delta F}{\delta \mathbf{B}} \right) \right) d^3x . \tag{20}$$

This, together with the Maxwell field energy

$$H_M(\mathbf{B},\mathbf{E}) = \int \left( \frac{1}{2} |\mathbf{B}|^2 + \frac{1}{2} |\mathbf{E}|^2 \right) d^3x$$

yields the vacuum evolution equations

$$\frac{\partial \mathbf{E}}{\partial t} = \nabla \times \mathbf{B} , \quad \frac{\partial \mathbf{B}}{\partial t} = -\nabla \times \mathbf{E}$$

in the form of Eq. (7). The auxiliary equations

$$\nabla \cdot \mathbf{B} = 0 ,$$
$$\nabla \cdot \mathbf{E} = 0 ,$$

follow from the invariance of electromagnetism under the gauge transformation

$$\mathbf{A} \to \mathbf{A} + \nabla \Lambda , \quad \Lambda \in \mathscr{F} , \tag{21}$$

in a way which will be made explicit in the more general calculation of the next section. The crucial point is that (21), which defines an action of $\mathscr{F}$, considered as a Lie group under addition, on $\mathfrak{A}$, preserves Poission brackets, and leaves $H_M$ invariant.

When $\mathscr{F}$ is thought of as consisting of scalar functions $\Lambda$ defining gauge transformations, we denote it by $\mathscr{G}$, to indicate the gauge group. We state for later use that the momentum map for the action (21) of $\mathscr{G}$ on $T^*\mathfrak{A}$ can be shown to be

$$J_M(\mathbf{A},\mathbf{Y}) = -\nabla \cdot \mathbf{Y} .$$

## IV. MULTIFLUID ELECTRODYNAMICS

We now couple the systems discussed in the two preceding sections, and use the reduction procedure to derive the Poisson structure for the combined system.

The canonical momentum density $\mathbf{N}$ for species $s$ in the presence of the Maxwell fields is

$$\mathbf{N}_s = \mathbf{M}_s + a_s \rho_s \mathbf{A} ,$$

where $\mathbf{A}$ is the vector potential. This is the classical relation between particle momentum and velocity, $\mathbf{p} = m(\mathbf{v} + a\mathbf{A})$, suitably modified for fluids. It may also easily be derived from the Lagrangian for the fluid system.

The prereduction state space is therefore taken to be

$$\tilde{\mathscr{P}} = \{(\mathbf{N}_s, \rho_s, \sigma_s, \mathbf{A}, \mathbf{Y})\} \equiv \mathfrak{g}_s^* \times T^*\mathfrak{A} ,$$

where

$$\mathfrak{g}_s^* \equiv \mathscr{X}_s^* \times \mathscr{F}_s^* \times \mathscr{F}_s^* ,$$

and the Hamiltonian $\tilde{H}: \tilde{\mathscr{P}} \to \mathbb{R}$ is just the energy written in these variables:

$$\tilde{H}(\mathbf{N}_s, \rho_s, \sigma_s, \mathbf{A}, \mathbf{Y}) = \sum_s \int \left( \frac{1}{2} \rho_s^{-1} |\mathbf{N}_s - a_s \rho_s \mathbf{A}|^2 + \rho_s U_s \right) d^3x + \int \frac{1}{2} (|\nabla \times \mathbf{A}|^2 + |\mathbf{Y}|^2) \, d^3x . \tag{22}$$

The Poisson structure on $\tilde{\mathscr{P}}$ is just the sum of that on $\mathfrak{g}_s^*$, Eq. (19) for each fluid, and the canonical bracket on $T^*\mathfrak{A}$. For $\tilde{F}, \tilde{G}: \tilde{\mathscr{P}} \to \mathbb{R}$, then

$$\begin{aligned} \{\tilde{F},\tilde{G}\}(\mathbf{N}_s, \rho_s, \sigma_s, \mathbf{A}, \mathbf{Y}) &= -\sum_s \int \mathbf{N}_s \cdot \left( \frac{\delta \tilde{F}}{\delta \mathbf{N}_s} \cdot \nabla \frac{\delta \tilde{G}}{\delta \mathbf{N}_s} - \frac{\delta \tilde{G}}{\delta \mathbf{N}_s} \cdot \nabla \frac{\delta \tilde{F}}{\delta \mathbf{N}_s} \right) d^3x \\ &\quad - \sum_s \int \rho_s \left( \frac{\delta \tilde{F}}{\delta \mathbf{N}_s} \cdot \nabla \frac{\delta \tilde{G}}{\delta \rho_s} - \frac{\delta \tilde{G}}{\delta \mathbf{N}_s} \cdot \nabla \frac{\delta \tilde{F}}{\delta \rho_s} \right) d^3x \\ &\quad - \sum_s \int \sigma_s \left( \frac{\delta \tilde{F}}{\delta \mathbf{N}_s} \cdot \nabla \frac{\delta \tilde{G}}{\delta \sigma_s} - \frac{\delta \tilde{G}}{\delta \mathbf{N}_s} \cdot \nabla \frac{\delta \tilde{F}}{\delta \sigma_s} \right) d^3x \\ &\quad + \int \left( \frac{\delta \tilde{F}}{\delta \mathbf{A}} \cdot \frac{\delta \tilde{G}}{\delta \mathbf{Y}} - \frac{\delta \tilde{G}}{\delta \mathbf{A}} \cdot \frac{\delta \tilde{F}}{\delta \mathbf{Y}} \right) d^3x . \end{aligned} \tag{23}$$

The coupling between $\mathfrak{g}_s^*$ and $T^*\mathfrak{A}$ appears only in the Hamiltonian. Reduction reverses the situation, so that the Hamiltonian appears as a sum of terms, each of which may be identified with either the fluids or the Maxwell fields, while the Poisson bracket exhibits a mixing of these two types of variables.

The reduction of Sec. III was carried out with the action (21) on $T^*\mathfrak{A}$. The requirements of the reduction procedure are that this action be extended to an action $\Phi$ of the gauge group $\mathscr{G}$ on all of $\tilde{\mathscr{P}}$ in such a way that the Hamiltonian, Eq. (22), is invariant, and that Poisson brackets of functions on $\tilde{\mathscr{P}}$ are preserved. Hence, we require (i) $\tilde{H}\circ\Phi = \tilde{H}$, and (ii) $\{\tilde{F}\circ\Phi, \tilde{G}\circ\Phi\} = \{\tilde{F},\tilde{G}\}\circ\Phi$, where $\tilde{F},\tilde{G}\in\mathscr{F}(\tilde{\mathscr{P}})$. It is obvious from Eq. (22) that the action

$$\Phi(\Lambda,(\mathbf{N}_s,\rho_s,\sigma_s,\mathbf{A},\mathbf{Y})) = (\mathbf{N}_s + a_s\rho_s\nabla\Lambda,\rho_s,\sigma_s,\mathbf{A}+\nabla\Lambda,\mathbf{Y}) \tag{24}$$

satisfies the first requirements, and we must now show that it satisfies the second.

*Lemma*: Let $G$ be a Lie group, with Lie algebra $\mathfrak{g}$. Let $L^*:\mathfrak{g}^*\to\mathfrak{g}^*$ be linear. Then $L^*$ preserves the Lie–Poisson bracket $\{\ ,\ \}$ of Eq. (12) on co-adjoint orbits if $L:\mathfrak{g}\to\mathfrak{g}$, the adjoint of $L^*$, preserves the Lie algebra bracket $[\ ,\ ]$ on $\mathfrak{g}$. In other words, for $\xi,\zeta\in\mathfrak{g}$ and $E,F:\mathfrak{g}^*\to\mathbb{R}$,

$$[L\xi,L\zeta] = L[\xi,\zeta] \text{ implies } \{E\circ L^*,F\circ L^*\} = \{E,F\}\circ L^* .$$

*Proof*: Using the linearity of $L^*$, and with $\mu,\nu\in\mathfrak{g}^*$,

$$\left\langle \nu, \frac{\delta(F\circ L^*)}{\delta\mu}\right\rangle = D(F\circ L^*)(\mu)\cdot\nu = DF(L^*(\mu))\cdot DL^*(\mu)\cdot\nu$$
$$= DF(L^*(\mu))\cdot L^*\nu = \left\langle L^*\nu, \frac{\delta F}{\delta\mu}(L^*\mu)\right\rangle$$
$$= \left\langle \nu, L\frac{\delta F}{\delta\mu}(L^*\mu)\right\rangle ,$$

so that $(\delta(F\circ L^*)/\delta\mu)(\mu) = L(\delta F/\delta\mu)(L^*\mu)$. Hence,

$$\{F\circ L^*,G\circ L^*\}(\mu)$$
$$= -\left\langle \mu, \left[L\frac{\delta F}{\delta\mu}(L^*\mu), L\frac{\delta G}{\delta\mu}(L^*\mu)\right]\right\rangle$$
$$= \left\langle L^*\mu, \left[\frac{\delta F}{\delta\mu}(L^*\mu), \frac{\delta G}{\delta\mu}(L^*\mu)\right]\right\rangle = \{F,G\}\circ L^*(\mu) .$$

It is clear from consideration of the action (21) that the converse of this lemma does not hold.

*Proposition*: The action (24) preserves the Poisson structure Eq. (23) on $\tilde{\mathscr{P}}$.

*Proof*: The action (24) is not linear in the last two variables $(\mathbf{A},\mathbf{Y})$, but we know from Sec. III that it nevertheless preserves the Poisson structure on $T^*\mathfrak{A}$. It is linear in the first three variables $(\mathbf{N}_s,\rho_s,\sigma_s)$, so we check the conditions of the above lemma, with $L^*\equiv\Phi_\Lambda$ restricted to $\mathfrak{g}_s^*$. We first calculate $\Phi_\Lambda^*$, temporarily suppressing the subscript $s$ on the phase space variable. Let $(\mathbf{N},\rho,\sigma)\in\mathfrak{g}_s^*$, $(\mathbf{X},f_1,g_2)\in\mathfrak{g}_s$. Then

$$\langle(\mathbf{N},\rho,\sigma),\Phi_\Lambda^*(\mathbf{X},f_1,g_1)\rangle$$
$$= \langle\Phi_\Lambda(\mathbf{N},\rho,\sigma),(\mathbf{X},f_1,g_1)\rangle$$
$$= \langle(\mathbf{N}+a\rho\nabla\Lambda,\rho,\sigma),(\mathbf{X},f_1,g_1)\rangle$$
$$= \int\mathbf{X}\cdot\mathbf{N}\,d^3x + \int a\rho\mathbf{X}(\Lambda)d^3x + \int(\rho f_1+\sigma g_1)\,d^3x$$
$$= \langle(\mathbf{N},\rho,\sigma),(\mathbf{X},a\mathbf{X}(\Lambda)+f_1,g_1)\rangle .$$

Thus, $\Phi_\Lambda^*(\mathbf{X},f_1,g_1) = (\mathbf{X},a\mathbf{X}(\Lambda)+f_1,g_1)$. Using Eq. (18), and with $(\mathbf{Y},f_2,g_2)\in\mathfrak{g}_s$,

$$[\Phi_\Lambda^*(\mathbf{X},f_1,g_1),\Phi_\Lambda^*(\mathbf{Y},f_2,g_2)]$$
$$= ([\mathbf{X},\mathbf{Y}],\mathbf{X}\cdot(a\mathbf{Y}(\Lambda)+f_2)$$
$$- \mathbf{Y}\cdot(a\mathbf{X}(\Lambda)+f_1),\mathbf{X}\cdot g_2 - \mathbf{Y}\cdot g_1)$$
$$= ([\mathbf{X},\mathbf{Y}],a[\mathbf{X},\mathbf{Y}](\Lambda)+\mathbf{X}(f_2)-\mathbf{Y}(f_1),\mathbf{X}(g_2)-\mathbf{Y}(g_1))$$
$$= \Phi_\Lambda^*[(\mathbf{X},f_1,g_1),(\mathbf{Y},f_2,g_2)] .$$

In order now to calculate the momentum map $J:\tilde{\mathscr{P}}\to\mathscr{G}^*$ for $\Phi$, it suffices to calculate $J_s:\mathfrak{g}_s^*\to\mathscr{G}^*$, the momentum map on $\mathfrak{g}_s^*$. Applying the definition of the momentum map, we require a map $\hat{J}_s:\mathscr{G}\to\mathscr{F}(\mathfrak{g}_s^*)$ such that

$$X_{\hat{J}_s}(\Lambda) = \Lambda_{\mathfrak{g}_s^*} ,$$

where the right-hand side denotes the infinitesimal generator of the action on $\mathfrak{g}_s^*$ corresponding to $\Lambda\in\mathscr{G}$. In the following, species subscripts will be suppressed. We use the above to find $\Lambda_{\mathfrak{g}_s^*}$, by equating

$$\langle\mathbf{X}_{\hat{J}(\Lambda)}(\mathbf{N},\rho,\sigma),(\mathbf{X},f,g)\rangle = \langle\Lambda_{\mathfrak{g}_s^*}(\mathbf{N},\rho,\sigma),(\mathbf{X},f,g)\rangle , \tag{25}$$

where $\mathbf{X}\in\mathscr{X}$, and $f,g\in\mathscr{F}$.

For the left-hand side, we use the form of the Hamiltonian vector field $\mathbf{X}_F(\beta) = \mathrm{ad}_\xi^*(\beta)$, where $\xi = \delta F/\delta\mu$, and $\mathrm{ad}_\xi^*:\mathfrak{g}^*\to\mathfrak{g}^*$ denotes the dual to the map $\mathrm{ad}_\xi:\mathfrak{g}\to\mathfrak{g}$ defined by $\mathrm{ad}_\xi(\zeta) = [\xi,\zeta]$. Then, temporarily writing $\hat{J}_s(\Lambda)\equiv F$ and $\delta\hat{J}_s(\Lambda)/\delta(\mathbf{N},\rho,\sigma)\equiv\xi$ for notational simplicity, and using the $L^2$ isomorphism $\mathfrak{g}^*\simeq\mathfrak{g}$,

$$\langle\mathbf{X}_F(\mathbf{N},\rho,\sigma),(\mathbf{Y},f,g)\rangle = \langle\mathrm{ad}_\xi^*(\mathbf{N},\rho,\sigma),(\mathbf{Y},f,g)\rangle$$
$$= \langle(\mathbf{N},\rho,\sigma),\mathrm{ad}_\xi(\mathbf{Y},f,g)\rangle$$
$$= \left\langle(\mathbf{N},\rho,\sigma),\left[\left(\frac{\delta F}{\delta\mathbf{N}},\frac{\delta F}{\delta\rho},\frac{\delta F}{\delta\sigma}\right),(\mathbf{Y},f,g)\right]\right\rangle$$
$$= \left\langle(\mathbf{N},\rho,\sigma),\left(\left[\frac{\delta F}{\delta\mathbf{N}},\mathbf{Y}\right],\frac{\delta F}{\delta\mathbf{N}}(f,g) - \mathbf{Y}\left(\frac{\delta F}{\delta\rho},\frac{\delta F}{\delta\sigma}\right)\right)\right\rangle$$
$$= \int\mathbf{N}\cdot\left[\frac{\delta F}{\delta\mathbf{N}},\mathbf{Y}\right] + \int\rho\left(\frac{\delta F}{\delta\mathbf{N}}(f) - \mathbf{Y}\left(\frac{\delta F}{\delta\rho}\right)\right)d^3x$$
$$+ \int\sigma\left(\frac{\delta F}{\delta\mathbf{N}}(g) - \mathbf{Y}\left(\frac{\delta F}{\delta\sigma}\right)\right)d^3x .$$

On the other hand, the infinitesimal generator is

$$\left.\frac{d}{dt}\right|_{t=0}\Phi(\exp(t\Lambda),(\mathbf{N},\rho,\sigma)) = (a\rho\nabla\Lambda,0,0) ,$$

so that the right-hand side of Eq. (25) is

$$\langle(a\rho\nabla\Lambda,0,0),(\mathbf{X},f,g)\rangle = \int a\rho\mathbf{X}(\Lambda)d^3x .$$

Clearly then Eq. (25) is satisfied by taking $\delta F/\delta\mathbf{N} = 0$,

$\delta F/\delta\rho = -a\Lambda$, and $\delta F/\delta\sigma = 0$ so that

$$J_s(\mathbf{N},\rho,\sigma)(\Lambda) = \hat{J}_s(\Lambda)(\mathbf{N},\rho,\sigma)$$

$$\equiv F(\mathbf{N},\rho,\sigma) = -\int a\Lambda\rho\, d^3x\,.$$

Thus, $J_s(\mathbf{N}_s,\rho_s,\sigma_s) = -a_s\rho_s$, and the momentum map on $\tilde{\mathscr{P}}$, obtained by summing

$$J(\mathbf{N}_s,\rho_s,\sigma_s,\mathbf{A},\mathbf{Y}) = J_M(\mathbf{A},\mathbf{Y}) + \sum_s J_s(\mathbf{N}_s,\rho_s,\sigma_s)$$

is

$$J(\mathbf{N}_s,\rho_s,\sigma_s,\mathbf{A},\mathbf{Y}) = -\nabla\cdot\mathbf{Y} - \sum_s a_s\rho_s\,.$$

Coordinates on the reduced phase space $\mathscr{P} \equiv J^{-1}(\rho_{\text{ext}})/\mathscr{G}$ are now given by the following.

*Proposition*:

$$\mathscr{P} \equiv J^{-1}(\rho_{\text{ext}})/\mathscr{G}$$

$$= \{\{(\mathbf{M}_s,\rho_s,\sigma_s,\mathbf{B},\mathbf{E})\}$$

$$|\nabla\cdot\mathbf{E} = \rho_{\text{ext}} + \sum_s a_s\rho_s, \nabla\cdot\mathbf{B} = 0\}\,.$$

*Proof*: Associate to elements $(\mathbf{N}_s,\rho_s,\sigma_s,\mathbf{A},\mathbf{Y})$ of $\tilde{\mathscr{P}}$ quintuples $(\mathbf{M}_s,\rho_s,\sigma_s,\mathbf{B},\mathbf{E})$, where $\mathbf{M}_s = \mathbf{N}_s - a_s\rho_s\mathbf{A}, \mathbf{B} = \nabla\times\mathbf{A}$, and $\mathbf{E} = -\mathbf{Y}$. Then the proposition follows from the momentum map constructed above, and a simple verification that two elements of $J^{-1}(\rho_{\text{ext}})$ are associated to the same quintuple if and only if they are related by a gauge transformation (24). Notice that reduction at the external charge density $\rho_{\text{ext}} \in \mathscr{G}^*$ specifies automatically that the dynamics takes place on the level set of $J$, $J^{-1}(\rho_{\text{ext}})$, implying Eq. (2a). Further, $\nabla\cdot\mathbf{B} = 0$ follows automatically from gauge invariance and the definition of these coordinates.

It remains now to compute the Poisson bracket on $\mathscr{P}$.

**Theorem**: For two functionals $F$, $G$ of the field variables $(\mathbf{M}_s,\rho_s,\sigma_s,\mathbf{B},\mathbf{E})$, the Poisson bracket is given by

$$\{F,G\}(\mathbf{M}_s,\rho_s,\sigma_s,\mathbf{B},\mathbf{E})$$

$$= \sum_s \{F,G\}(\mathbf{M}_s,\rho_s,\sigma_s) + \{F,G\}(\mathbf{B},\mathbf{E})$$

$$+ \int \sum_s \left(\frac{\delta F}{\delta\mathbf{M}_s}\cdot\frac{\delta G}{\delta\mathbf{E}} - \frac{\delta G}{\delta\mathbf{M}_s}\cdot\frac{\delta F}{\delta\mathbf{E}}\right.$$

$$\left. + \mathbf{B}\cdot\left(\frac{\delta F}{\delta\mathbf{M}_s}\times\frac{\delta G}{\delta\mathbf{M}_s}\right)\right) a_s\rho_s\, d^3x\,, \tag{26}$$

where the first and second terms are defined in Eqs. (19) and (20).

*Proof*: Given $F$ and $G$, define $\tilde{F}$ on $\tilde{\mathscr{P}}$ according to $\tilde{F}(\mathbf{N}_s,\rho_s,\sigma_s,\mathbf{A},\mathbf{Y}) = F(\mathbf{M}_s,\rho_s,\sigma_s,\mathbf{B},\mathbf{E})$, where the relation between the two sets of variables is as in the proof of the preceding proposition. Define $\tilde{G}$ similarly. Then $\{F,G\}(\mathbf{M}_s,\rho_s,\sigma_s,\mathbf{B},\mathbf{E})$ is found by rewriting Eq. (23) in terms of the variables on $\mathscr{P}$, using the following.

*Lemma*:

(i) $$\frac{\delta\tilde{F}}{\delta\mathbf{A}} = -\sum_s a_s\rho_s\frac{\delta F}{\delta\mathbf{M}_s} + \nabla\times\frac{\delta F}{\delta\mathbf{B}_s}\,,$$

(ii) $$\frac{\delta\tilde{F}}{\delta\mathbf{N}_s} = \frac{\delta F}{\delta\mathbf{M}_s}\,,$$

(iii) $$\frac{\delta\tilde{F}}{\delta\rho_s} = -\sum_s a_s\mathbf{A}\frac{\delta F}{\delta\mathbf{M}_s} + \frac{\delta F}{\delta\rho_s}\,,$$

(iv) $$\frac{\delta\tilde{F}}{\delta\sigma_s} = \frac{\delta F}{\delta\sigma_s}\,,$$

(v) $$\frac{\delta\tilde{F}}{\delta\mathbf{Y}} = -\frac{\delta F}{\delta\mathbf{E}}\,.$$

*Proof*: We have, for example,

$$\int\frac{\delta\tilde{F}}{\delta\mathbf{A}}\cdot\tilde{\mathbf{A}}\,d^3x = \frac{d}{d\epsilon}\bigg|_{\epsilon=0}\tilde{F}(\mathbf{N}_s,\rho_s,\sigma_s,\mathbf{A}+\epsilon\tilde{\mathbf{A}},\mathbf{Y})$$

$$= \frac{d}{d\epsilon}\bigg|_{\epsilon=0}F(\mathbf{N}_s - a_s\rho_s(\mathbf{A}+\epsilon\tilde{\mathbf{A}}),\rho_s,\sigma_s,\nabla\times(\mathbf{A}+\epsilon\tilde{\mathbf{A}}),\mathbf{E})$$

$$= D_{\mathbf{M}_s}F(\mathbf{M}_s,\rho_s,\sigma_s,\mathbf{B},\mathbf{E})\cdot(-a_s\rho_s\tilde{\mathbf{A}}) + D_{\mathbf{B}}F(\mathbf{M}_s,\rho_s,\sigma_s,\mathbf{B},\mathbf{E})\cdot\nabla\times\tilde{\mathbf{A}}$$

$$= \int\frac{\delta F}{\delta\mathbf{M}_s}\cdot(-a_s\rho_s\tilde{\mathbf{A}})\,d^3x + \int\frac{\delta F}{\delta\mathbf{B}}\cdot\nabla\times\mathbf{A}\,d^3x$$

$$= -\int\frac{\delta F}{\delta\mathbf{M}_s}\cdot a_s\rho_s\tilde{\mathbf{A}}\,d^3x + \int\nabla\times\frac{\delta F}{\delta\mathbf{B}}\cdot\tilde{\mathbf{A}}\,d^3x\,,$$

which implies (i).

The other formulas of the lemma follow in a similar fashion. After making the substitutions prescribed in the lemma, and using vector identities to eliminate the vector potential $\mathbf{A}$ in favor of the magnetic field $\mathbf{B} = \nabla\times\mathbf{A}$, one obtains the result of the theorem.

We observe that the first term of Eq. (26) involves only the fluid variables and that the second is purely electromagnetic, while the third provides the coupling of the fluids to the electric and magnetic fields. Bilinearity, skew symmetry, and the Jacobi identity all follow for Eq. (26) by the methods used in its derivation. In addition, it is readily verified that the correct evolution equations for the phase space variables, in the form of Eq. (7), follow from Eqs. (26) and (6). Additional body forces, such as gravity, can easily be incorporated into Eq. (3a) by the inclusion of an appropriate term in the Hamiltonian. Finally, Eqs. (2), rather than being postulated separately as initial conditions, follow from the gauge invariance of electromagnetism.

The restriction of multispecies electrodynamics to the Coulomb case, in which $\mathbf{B} = 0$, can also be treated. The scalar potential $\phi$ is expressed in terms of the mass densities $\rho_s$

by

$$\phi(x) = \frac{1}{4\pi} \int \frac{\Sigma_s a_s \rho_s(\mathbf{x}')}{|\mathbf{x} - \mathbf{x}'|} d^3x'$$

and $\mathbf{E} = \nabla\phi$. Equations (1) and (2) are then replaced by the Poisson equation $\nabla^2\phi = -\Sigma_s a_s \rho_s$, and the Lorentz force term in Eq. (3a) is replaced by $a_s \rho_s \mathbf{E}$. The Hamiltonian structure is obtained by taking the Hamiltonian on the phase space of sets of triples $(\mathbf{M}_s, \rho_s, \sigma_s)$ to be the total energy

$$\begin{aligned} H(\mathbf{M}_s, \rho_s, \sigma_s) &= \sum_s \int \left( \frac{1}{2} \rho_s^{-1} |\mathbf{M}_s|^2 + \rho_s U_s(\rho_s, \sigma_s) \right) d^3x \\ &+ \frac{1}{8\pi} \iint \frac{1}{|\mathbf{x} - \mathbf{x}'|} \left( \sum_s a_s \rho_s(\mathbf{x}) \right) \\ &\cdot \left( \sum_{s'} a_{s'} \rho_{s'}(\mathbf{x}') \right) d^3x \, d^3x' \end{aligned}$$

and letting the Poisson bracket on phase functionals be given by the first term of (26). The correct equations of motion for the dynamical variables $(\mathbf{M}_s, \rho_s, \sigma_s)$ now follow in the form (7).

The application of this formalism to the problem of mode coupling will be the subject of a later publication.

## ACKNOWLEDGMENTS

It is a pleasure to acknowledge valuable discussions with Professor Allan Kaufman, who suggested this problem, and with Professors Jerrold Marsden and Alan Weinstein, who provided guidance at critical stages of this work. We also thank Professor Rudolf Schmid for his interest and comments. This work was supported by the Office of Basic Energy Sciences and the Office of Fusion Energy of the U. S. Department of Energy under Contract Number DE-AC03-76SF00098.

[1] J. E. Marsden and A. Weinstein, Physica D **4**, 394 (1982).
[2] R. G. Spencer and A. N. Kaufman, Phys. Rev. A **25**, 2437 (1982).
[3] A. N. Kaufman, Phys. Fluids **25**, 1993 (1982).
[4] B. Kupershmidt and D. Holm, Physica D **6**, 347 (1983).
[5] J. Marsden, A. Weinstein, T. Ratiu, R. Schmid, and R. Spencer, Proc. IUTAM Symposium on Modern Developments in Analytical Mechanics, Turin, Italy, 7–11 June (1982).
[6] R. Menikoff and D. Sharp, J. Math. Phys. **18**, 471 (1977).
[7] V. I. Arnold, Ann. Inst. Fourier (Grenoble) **16**, 319 (1966).
[8] J. Marsden and A. Weinstein, Rep. Math. Phys. **5**, 121 (1974).
[9] A. Weinstein, Lett. Math. Phys. **2**, 417 (1978).
[10] R. Abraham and J. Marsden, *Foundations of Mechanics* (Benjamin, New York, 1978), 2nd ed.
[11] D. Ebin and J. Marsden, Ann. Math. **92**, 102 (1970).
[12] P. Morrison and J. Greene, Phys. Rev. Lett. **45**, 790 (1980).
[13] W. Pauli, *General Principles of Quantum Mechanics*, English translation (Springer-Verlag, New York, 1981).
[14] M. Born and L. Infeld, Proc. R. Soc. London Ser. A **150**, 141 (1935).

# Probability manifolds

Stanley P. Gudder
*Department of Mathematics and Computer Science, University of Denver, Denver, Colorado 80208*

(Received 30 November 1983; accepted for publication 23 March 1984)

The concept of a probability manifold $M$ is introduced. The global properties of $M$ inherited from its local structure are then considered. It is shown that a deterministic spin model due to Pitowski falls within this general framework. Finally, we construct a phase-space model for nonrelativistic quantum mechanics. These two models give the same global description as conventional quantum mechanics. However, they also give a local description which is not possible in conventional quantum mechanics.

PACS numbers: 02.50. + s, 03.65.Bz

## I. INTRODUCTION

Let $X$ be a nonempty set, and let $(Y,\Sigma_Y,\nu)$ be a probability space. Suppose that for every $(x,y)\in X\times Y$ there exists a probability space $(X(x,y),\ \Sigma(x,y),\ \mu_{x,y})$ such that (a) $X=\cup_{y\in Y}X(x,y)$ for every $x\in X$, and (b) $X(x,y_1)\cap X(x,y_2)=\phi$, $y_1\neq y_2$, for every $x\in X$. We then call the triple

$$M=(X,Y,\{X(x,y):(x,y)\in X\times Y\})$$

a *probability manifold* and we call $\{X(x,y):(x,y)\in X\times Y\}$ the *local structure of* $M$.

Roughly speaking, $X$ is covered by local probability spaces which are joined together in a stochastic fashion. In general, $X$ itself does not form a probability space, but as we shall later see, becomes a generalized probability space.

A set $A\subseteq X$ is *locally measurable* if $A\cap X(x,y)\in\Sigma(x,y)$ for every $(x,y)\in X\times Y$ and in this case we call

$$\mu(A\,|x,y)=\mu_{x,y}[A\cap X(x,y)]$$

the *probability of* $A$ *along* $X(x,y)$. A set of $A\subseteq X$ is *globally measurable* if $A$ is locally measurable and for every $x\in X$, the function $y\to\mu(A\,|x,y)$ is $\Sigma_Y$-measurable. In this case, we call

$$\mu(A\,|x)=\int\mu(A\,|x,y)\nu(dy)$$

the *probability of* $A$ *at* $x$. Finally, a set $A\subseteq X$ is *totally measurable* if $A$ is globally measurable and $\mu(A\,|x)$ is independent of $x$. In this case, we call $\mu(A)=\mu(A\,|x)$ for every $x\in X$ the *total probability of* $A$. We define $\Sigma_L,\Lambda_G,\Lambda_T$, respectively, as the collections of locally, globally, and totally measurable subsets of $X$. If $X$ is the base set of a probability space $(X,\Sigma,\mu_0)$ such that $\Sigma\subseteq\Lambda_T$ and $\mu|\Sigma=\mu_0$, then we say that $M$ is *coherent*. It is easy to show that $\Sigma_L$ is a $\sigma$-algebra and $\mu(\cdot|x,y)$ is a probability measure on $\Sigma_L$ for every $(x,y)\in X\times Y$. Hence, $(X,\Sigma_L,\mu(\cdot|x,y))$ is a probability space for every $(x,y)\in X\times Y$. In general, we have $\Lambda_T\subseteq\Lambda_G\subseteq\Sigma_L$. However, $\Lambda_T$ and $\Lambda_G$ need not be $\sigma$-algebras.

We call a collection of subsets $\Lambda$ of $X$, a *$\sigma$-class* (or *$\lambda$-field*) if (a) $X\in\Lambda$; (b) $A\in\Lambda$ implies the complement $A'\in\Lambda$; and (c) $A_i\in\Lambda$ mutually disjoint, $i=1,2,\ldots$, implies $\cup A_i\in\Lambda$. A *probability measure* on a $\sigma$-class is defined in the same way as a probability measure on a $\sigma$-algebra. If $\Lambda$ is a $\sigma$-class of subsets of $X$, then $(X,\Lambda)$ is called a *generalized measurable space*, and if $\mu$ is a probability measure on $\Lambda$ we call $(X,\Lambda,\mu)$ a *generalized probability space*.[1,2]

**Theorem 1**: If $(X,Y\{X(x,y):(x,y)\in X\times Y\})$ is a probability manifold, then $(X,\Lambda_G,\mu(\cdot|x))$ and $(X,\Lambda_T,\mu)$ are generalized probability spaces for every $x\in X$.

*Proof*: We first show that $(X,\Lambda_G,\mu(\cdot|x))$ is a generalized probability space for $x\in X$. Clearly, $X\in\Lambda_G$ and $\mu(X\,|x)=1$. Suppose $A\in\Lambda_G$. Then

$$A'\cap X(x,y)=X(x,y)-A\cap X(x,y)\in\Sigma(x,y),$$

for every $(x,y)\in X\times Y$ so $A'\in\Sigma_L$. Also,

$$\mu(A'|x,y)=1-\mu(A\,|x,y),$$

so $y\to\mu(A'|x,y)$ is $\Sigma_Y$-measurable for every $x\in X$. Hence, $A'\in\Lambda_G$. Let $A_i\in\Lambda_G$, $i=1,2,\ldots$, be mutually disjoint, and let $A=\cup A_i$. Then

$$A\cap X(x,y)=\cup[A_i\cap X(x,y)]\in\Lambda_G,$$

for every $(x,y)\in X\times Y$ so $A\in\Sigma_L$. Moreover,

$$\mu(A\,|x,y)=\Sigma\,\mu(A_i\,|x,y),$$

so for every $x\in X$, $y\to\mu(A\,|x,y)$ is the limit of $\Sigma_Y$-measurable functions. Hence, $y\to\mu(A\,|x,y)$ is $\Sigma_Y$-measurable for every $x\in X$. By the monotone convergence theorem,

$$\begin{aligned}\mu(A\,|x)&=\int\mu(A\,|x,y)\nu(dy)\\&=\int\mu_{x,y}[A\cap X(x,y)]\nu(dy)\\&=\Sigma\int\mu_{x,y}[A_i\cap X(x,y)]\ \nu(dy)\\&=\Sigma\mu(A_i\,|x).\end{aligned}$$

It follows that $(X,\Lambda_G,\mu(\cdot|x))$ is a generalized probability space for every $x\in X$. It now easily follows that $(X,\Lambda_T,\mu)$ is also a generalized probability space. □

Let $(X,\Lambda)$ be a generalized measurable space and let $f$: $X\to R$. We say that $f$ is *measurable* if $f^{-1}(B)\in\Lambda$ for every Borel set $B\in B(R)$. It is easy to show that $\Lambda_f=\{f^{-1}(B):B\in B(R)\}$ is a sub $\sigma$-algebra of $\Lambda$. If $(X,\Lambda,\mu)$ is a generalized probability space and $f:X\to R$ is measurable, we call $f$ a *random variable*, as usual. In this case, $(X,\Lambda_f,\mu|\Lambda_f)$ is an ordinary probability space and we define the *integral* or *expectation* $E_\mu(f)=\int f\,du$ of $f$ in the usual way. We denote the set of integrable random variables on $(X,\Lambda,\mu)$ by $L^1(X,\Lambda,\mu)$.

 0022-2488/84/082397-05$02.50  

**Theorem 2.** (a) If $f \in L^1(X, \Sigma_L, \mu(\cdot|x,y))$, then $f|X(x,y) \in L^1(X(x,y), \Sigma(x,y), \mu_{x,y})$.

(b) If $f \in L^1(X, \Lambda_G, \mu(\cdot|x,y))$, then $f \in L^1(X, \Sigma_L, \mu(\cdot|x,y))$ for every $y \in Y$, $y \to \int_{X(x,y)} f(w)\,\mu_{x,y}(dw)$ is integrable on $(Y, \Sigma_Y, \nu)$ and

$$\int_Y \left[\int_{X(x,y)} f(w)\,\mu_{x,y}(dw)\right]\nu(dy) = \int_X f(w)\,\mu(dw|x).$$

(c) If $f \in L^1(X, \Lambda_T, \mu)$ then $f \in L^1(X, \Lambda_G, \mu(\cdot|x))$ for every $x \in X$ and

$$\int_X f(w)\,\mu(dw|x) = \int_X f(w)\,\mu(dw),$$

for every $x \in X$.

*Proof*: Part (a) is straightforward and (c) easily follows from (b). We therefore only prove part (b). Let $f \in L^1(X, \Lambda_G, \mu(\cdot|x))$ be a simple function, $f = \Sigma_{i=1}^n c_i \chi_{A_i}$, $A_i \in \Lambda_G$. Then $A_i \cap X(x,y) \in \Sigma(x,y)$ so $f|X(x,y)$ is $\Sigma(x,y)$-measurable. Since $y \to \mu_{x,y}[A_i \cap X(x,y)]$ is $\Sigma_Y$-measurable,

$$y \to \int_{X(x,y)} f(w)\,\mu_{x,y}(dw) = \Sigma c_i\,\mu_{x,y}[A_i \cap X(x,y)]$$

is $\Sigma_Y$-measurable. Now

$$\int_Y \left[\int_{X(x,y)} f(w)\,\mu_{x,y}(dw)\right]\nu(dy)$$

$$= \Sigma c_i \int_Y \mu_{x,y}[A_i \cap X(x,y)]\,\nu(dy)$$

$$= \Sigma c_i\,\mu(A_i|x) = \int_X f(w)\,\mu(dw|x).$$

Hence, the result holds for simple functions. Now let $f \in L^1(X, \Lambda_G, \mu(\cdot|x))$ be non-negative. Then there exists an increasing sequence of simple functions $f_i \in L^1(X, \Lambda_G, \mu(\cdot|x))$ such that $f_i \to f$ almost everywhere. Since $f|X(x,y) = \lim f_i|X(x,y)$, we have that $f|X(x,y)$ is $\Sigma(x,y)$-measurable. By the monotone convergence theorem

$$\int_{X(x,y)} f(w)\,\mu_{x,y}(dw) = \lim \int_{X(x,y)} f_i(w)\,\mu_{x,y}(dw).$$

Since $y \to \int_{X(x,y)} f_i(w)\,\mu_{x,y}(dw)$ is $\Sigma_Y$-measurable, so is its limit $y \to \int_{X(x,y)} f(w)\,\mu_{x,y}(dw)$. Again, by the monotone convergence theorem,

$$\int_X f(w)\,\mu(dw|x) = \lim \int_X f_i(w)\,\mu(dw|x)$$

$$= \lim \int_Y \left[\int_X f_i(w)\,\mu_{x,y}(dw)\right]\nu(dy)$$

$$= \left[\int_Y X_{(x,y)} f(w)\,\mu_{x,y}(dw)\right]\nu(dy).$$

Finally, if $f \in L^1(X, \Lambda_G, \mu(\cdot|x))$ is arbitrary, we can write $f = f_+ - f_-$ where $f_+, f_- \geqslant 0$, $f_+, f_- \in L^1(X, \Lambda_G, \mu(\cdot|x))$ and the ranges of $f_+$ and $f_-$ are contained in a common sub-$\sigma$-algebra of $\Lambda_G$. The result now follows from the additivity of the integral.[2] □

If $f|X(x,y) \in L^1(X(x,y), \Sigma(x,y), \mu(x,y))$, we call

$$E(f|x,y) = \int_{X(x,y)} f(w)\,\mu(dw|x,y)$$

the *local expectation of f along* $X(x,y)$. If $f \in L^1(X, \Lambda_G, \mu(\cdot|x))$ we call

$$E(f|x) = \int_X f(w)\,\mu(dw|x) = \int_Y E(f|x,y)\nu(dy)$$

the *global expectation of f at x*. If $f \in L^1(X, \Lambda_T, \mu)$ we call $E(f) = \int_X f\,d\mu$ the *total expectation of f*.

## II. EXAMPLES

We now give some simple examples which illustrate the above framework. As we shall see, this framework includes product probability spaces and conditional expectations relative to countable measurable partitions.

*Example 1:* Let $(Y, \Sigma_Y, \nu)$ and $(Z, \Sigma_Z, w)$ be probability spaces and let $X = Y \times Z$. For $(x,y) \in X \times Y$, let $X(x,y) = y \times Z$, $\Sigma(x,y) = \{y \times B : B \in \Sigma_Z\}$, $\mu_{x,y}(y \times B) = w(B)$. Then $M = (X, Y\{X(x,y): (x,y) \in X \times Y\})$ is a probability manifold. Denoting the power set of $Y$ by $P(Y)$, it is clear that

$$\Sigma_Y \times \Sigma_Z \subseteq P(Y) \times \Sigma_Z \subseteq \Sigma_L.$$

Let $C \times D \in \Sigma_Y \times \Sigma_Z$ be a measurable rectangle. Then $C \times D \in \Sigma_L$ and

$$\mu_{x,y}[C \times D \cap X(x,y)]$$

$$= \mu_{x,y}[C \times D \cap (y \times Z)] = \chi_C(y) w(D).$$

This is a $\Sigma_Y$-measurable function of $y$. Hence, $C \times D \in \Lambda_G$. Since the intersection of two measurable rectangles is a measurable rectangle, it can be shown[1,2] that

$$\Sigma_Y \times \Sigma_Z \subseteq \Lambda_G \subseteq \Sigma_L.$$

Moreover,

$$\mu(C \times D|x) = \int_Y \mu_{x,y}[C \times D \cap X(x,y)]\,\nu(dy) = \nu(C) w(D)$$

is independent of $x$, so $\Lambda_T = \Lambda_G$. If we place the product measure $\nu \times w$ on $(X, \Sigma_Y \times \Sigma_Z)$, then $M$ is coherent. Hence,

$$(X, \Sigma_Y \times \Sigma_Z, \nu \times w) \subseteq (X, \Lambda_G, \mu)$$

and we may think of $M$ as a generalization of the Cartesian product of two probability spaces. The probability of $C \times D$ along $X(x,y)$ becomes $\mu(C \times D|x,y) = \chi_C(y) w(D)$ and the probability of $C \times D$ at $x$ becomes

$$\mu(C \times D|x) = \mu(C \times D) = \nu(C) w(D) = (\nu \times w)(C \times D).$$

*Example 2:* Let $X = Y = R$, $\Sigma_Y = B(R)$, and $\nu(B) = (1/\sqrt{\pi}) \int_B e^{-y^2}\,dy$, $B \in \Sigma_Y$. For $(x,y) \in X \times Y$, let $X(x,y) = \{x + y\}$, $\Sigma(x,y) = \{\phi, \{x,y\}\}$, and $\mu_{x,y}(\{x+y\}) = 1$. Then $M = (X, Y, \{X(x,y): (x,y) \in X \times Y\})$ is a probability manifold. It is clear that $\Sigma_L = P(X)$. If $A \in \Sigma_L$, then

$$\mu(A|x,y) = \mu_{x,y}[A \cap X(x,y)] = \chi_{A-x}(y).$$

It follows that $\Lambda_G = B(R)$. For $A \in \Lambda_G$, we have

$$\mu(A|x) = \int \mu(A|x,y)\nu(dy) = \nu(A - x)$$

$$= \frac{1}{\sqrt{\pi}} \int_{A-x} e^{-y^2}\,dy.$$

It follows that $\Lambda_T = \{\phi, R\}$.

*Example 3:* Let $(X, \Sigma, \mu)$ be a probability space and let

$B_i, i = 1,2,...$, be a countable measurable partition of $X$, with $\mu(B_i) \neq 0$, $i = 1,2,...$. Let $Y = \{y_1, y_2, ...\}$ be a countable set and define $\Sigma_Y = P(Y)$, $\nu(\{y_i\}) = \mu(B_i)$. For any $x \in X$, define $X(x, y_i) = B_i$, $\Sigma(x, y_i) = \{A \cap B_i : A \in \Sigma\}$ and $\mu_{x,y_i}(B) = \mu(B)/\mu(B_i)$ for $B \in \Sigma(x, y_i)$, $i = 1,2,...$. It is clear that $\Sigma_L = \Sigma$. For $A \in \Sigma$, we have

$$\mu(A \mid x, y_i) = \mu(A \cap B_i)/\mu(B_i),$$

which is a $\Sigma_Y$-measurable function, so $\Lambda_G = \Sigma$. Moreover, $\mu(A \mid x, y_i)$ is the usual conditional probability $\mu(A \mid B_i)$, $i = 1,2,...$. Also for $A \in \Sigma$, we have

$$\mu(A \mid x) = \int_Y \mu(A \mid x, y)\nu(dy) = \Sigma\, \mu(A \cap B_i) = \mu(A).$$

It follows that $\Lambda_T = \Sigma$.

## III. LIMIT THEOREMS

In this section we consider some simple global limit theorems which follow from local properties. We restrict our attention to the simplest version of the strong law of large numbers. One can prove more general versions of the law of large numbers and also the central limit theorem. Moreover, we only consider what we call "$g$-harmonic" functions. Relaxing this condition would be an interesting direction for further research. The results of this section generalize some theorems due to Pitowski[3,4] which are considered in the next section.

Let $(X, Y, \{X(x, y): (x, y) \in X \times Y\})$ be a probability manifold and let $f_i: X \to R$, $i = 1, 2, ...$, be a sequence of locally measurable functions. If $f_i \mid X(x, y)$ are stochastically independent for every $(x, y) \in X \times Y$, we say that $f_i$, $i = 1,2,...$, are *locally independent*. If $f_i \mid X(x, y)$ are identically distributed, we say that $f_i$, $i = 1,2,...$, are *locally identically distributed*. Let $g \in L^1(Y, \Sigma_Y, \nu)$ be a function which is not identically zero. A locally integrable function $f: X \to R$ is *$g$-harmonic* if $E(f \mid x, y) = f(x)\, g(y)$ for every $(x, y) \in X \times Y$.

Notice that if $f$ is $g$-harmonic, then $f$ is globally integrable and

$$E(f \mid x) = \int_Y E(f \mid x, y)\nu(dy) = \int_Y f(x)\, g(y)\nu(dy)$$

$$= f(x) \int_Y g(y)\nu(dy).$$

If follows that $f$ is totally measurable if and only if $\int g(y)\nu(d\lambda) = 0$ or $f$ is a constant function. For a fixed $x \in X$, each $w \in X$ is contained in a unique set $X(x, y_w)$. Define $\Pi_x: X \to Y$ by $\Pi_x(w) = y_w$.

**Theorem 3.** Let $f_i$, $i = 1,2,...$, be a sequence of locally independent, locally identically distributed, $g$-harmonic functions.

(a) If $f_i(x_0) = \alpha$, $i = 1, 2, ...$, then

$$\mu\left(\left\{w \in X: \frac{1}{n}\sum_{i=1}^{n} f_i(w) \to g[\Pi_{x_0}(w)]\right\} \middle| x_0\right) = 1.$$

(b) If $g(y) = 0$, then for every $x \in X$

$$\mu\left(\left\{w \in X: \frac{1}{n}\sum_{i=1}^{n} f_i(w) \to 0\right\} \middle| x_0, y\right) = 1.$$

(c) If $1/n\, \Sigma_{i=1}^{n} f_i(x_0) \to 0$, then

$$\mu\left(\left\{w \in X: \frac{1}{n}\sum_{i=1}^{n} f_i(w) \to 0\right\} \middle| x_0\right) = 1.$$

*Proof*: (a) Since $f_i(x_0) = \alpha$, we have $E(f_i \mid x_0, y) = \alpha g(y)$ for $i = 1,2...$. Applying the law of large numbers to $X(x_0, y)$ we obtain

$$\frac{1}{n}\sum_{i=1}^{n} f_i(w) \to \alpha g(y) = \alpha g\left[\prod_{x_0}(w)\right],$$

for almost every $w \in X(x_0, y)$. Hence,

$$\mu\left(\left\{w \in X: \frac{1}{n}\sum_{i=1}^{n} f_i(w) \to \alpha g\left[\prod_{x_0}(w)\right]\right\} \middle| x_0, y\right) = 1,$$

for every $y \in Y$. Integrating this equation over $Y$ gives the result.

(b) This follows from the law of large numbers on $X(x_0, y)$ using the fact that

$$E(f_i \mid x, y) = f_i(x_0)\, g(y) = 0.$$

(c) The proof is similar to part (a) and uses the fact that

$$\frac{1}{n}\sum_{i=1}^{n} E(f_i \mid x_0, y) = \frac{g(y)}{n}\sum_{i=1}^{n} f_i(x_0) \to 0. \qquad \square$$

The next result shows that under fairly general conditions, $g$-harmonic functions exist. A probability space is called a *Lebesgue space* if it is isomorphic to $[0,1]$ with Lebesgue measure.

**Theorem 4**: Let $(X, Y, \{X(x, y): (x, y) \in X \times Y\})$ be a probability manifold, with card $(X)$ = card $(R)$, and let $g$: $Y \to R$ be $\Sigma_Y$-measurable with $|g(y)| \leq 1$ for every $y \in Y$. If $X(x, y)$ are Lebesque spaces such that $X(x_1, y_1) \cap X(x_2, y_2)$ is at most countable whenever $x_1 \neq x_2$, then there exists a locally measurable $g$-harmonic function.

*Proof*: The proof is similar to the methods used by Pitowski in his proof of Theorem 1 (see Ref. 4) so we only give an outline. We assume the axiom of choice and the continuum hypothesis. Well-order the family of sets

$$F = \{\{x\} \cup X(x, y): (x, y) \in X \times Y\}.$$

Since card $(F)$ = card $(R)$, it follows from the continuum hypothesis that there is such a well-ordering in which every element has only countably many predecessors. We define $f$ by induction on this ordering.

On the first element $\{x_1\} \cup X(x_1, y_1)$ of the ordering define $f$ as follows: $f(x_1) = 1$; partition $X(x_1, y_1)$ into two measurable sets, one of measure $[1 - g(y_1)]/2$ and one of measure $[1 + g(y_1)]/2$; define $f$ to be $-1$ on the first set and $+1$ on the second set. Then

$$\int_{X(x_1, y_1)} f(w)\, \mu_{x_1, y_1}(dw) = \frac{1 + g(y_1)}{2} - \frac{1 - g(y_1)}{2}$$

$$= g(y_1) = f(x_1)\, g(y_1).$$

Suppose we have defined $f$ on all the elements of $F$ up to but not including $\{x_\alpha\} \cup X(x_\alpha, y_\alpha)$ in the order. If $x_\alpha$ does not belong to any preceding element of $F$, define $f(x_\alpha) = 1$. Since $\{x_\alpha\} \cup X(x_\alpha, y_\alpha)$ is preceded by at most countably many elements and since the intersection of two different $X(x, y)$'s is at most countable, $f$ has already been defined on at most countably many points of $X(x_\alpha, y_\alpha)$, which is a set of measure zero. Hence, we can proceed as in the original step. Otherwise, $x_\alpha$ belongs to a preceding element. If $f(x_\alpha) = 1$, pro-

ceed as before. If $f(x_\alpha) = -1$, reverse the measures of the two sets. □

## IV. DETERMINISTIC SPIN MODEL

In this section we give an application to a deterministic spin model due to Pitowski.[3,4] However, unlike Pitowski, we pose this model in the framework of a probability manifold. This provides a unification of Pitowski's work and points the way toward generalizations. We shall outline some of Pitowski's main results and we refer the reader to his works for more details.

Let $X = S^{(2)} \subseteq R^3$ be the unit sphere. For $x \in S^{(2)}$, $0 \leqslant \theta < \Pi$, let

$$X(x,\theta) = \{u \in S^{(2)} : u \cdot x = \cos\theta\}.$$

Let $m_{x,\theta}$ be Lebesgue measure on $X(x,\theta)$ and let $\mu_{x,\theta} = (2\Pi \sin\theta)^{-1} m_{x,\theta}$ be the uniform distribution on $X(x,\theta)$. Define $Y = [0,\Pi]$, $\Sigma_Y = B([0,\Pi])$ and let $\nu$ be the uniform distribution on $Y$. Then $M = (X, Y, \{X(x,\theta) : (x,\theta) \in X \times Y\})$ is a probability manifold. The next theorem shows that a totally measurable set need not be Lebesgue measurable. Moreover, if $\mu_0$ denotes the normalized Lebesgue measure on $B(S^{(2)})$ then $M$ becomes a coherent probability manifold.

**Theorem 5:** (Pitowski) $B(S^{(2)}) \subsetneq \Lambda_T$ and $\mu | B(S^{(2)}) = \mu_0$.

The next theorem gives the basic result for a deterministic spin model.

**Theorem 6:** (Pitowski) There exists a function $f \in L^1(X, \Lambda_T, \mu)$ satisfying the following.

(a) $f: X \to \{-\frac{1}{2}, \frac{1}{2}\}$.

(b) $f(-w) = -f(w)$, for every $w \in X$.

(c) $E(f|x,\theta) = f(x)\cos\theta$, for every $(x,\theta) \in X \times Y$.

We call $f$ a *spin-$\frac{1}{2}$ function*. Notice that $f$ is $g$-harmonic, where $g(\theta) = \cos\theta$. Spin-$\frac{1}{2}$ functions may be interpreted as giving a deterministic spin model. In fact, if a system is described by a spin-$\frac{1}{2}$ function $f$, then $f(x)$ specifies the spin in the direction $x$. It is not hard to show that

$$\mu(f^{-1}\{\tfrac{1}{2}\}|x,\theta) = \begin{cases} \cos^2\theta/2, & \text{if } f(x) = \frac{1}{2}, \\ \sin^2\theta/2, & \text{if } f(x) = -\frac{1}{2}. \end{cases}$$

This result shows that spin-$\frac{1}{2}$ functions produce the usual quantum mechanical probabilities.

Pitowski has shown that there exist many spin-$\frac{1}{2}$ functions, in fact, there exist sequences of locally independent spin-$\frac{1}{2}$ functions. The next result shows that spin-$\frac{1}{2}$ functions give the usual quantum expectations. The proof of this result follows from Theorem 3.

**Theorem 7:** (Pitowski) Let $f_1, f_2, \ldots$ be a locally independent sequence of spin-$\frac{1}{2}$ functions satisfying $f_i(x_0) = \frac{1}{2}$. Then

$$\mu\left(\left\{w \in X : \frac{1}{n}\sum_{i=1}^{n} f_i(w) \to \frac{w \cdot x_0}{2}\right\} \middle| x_0\right) = 1.$$

## V. PHASE-SPACE MODEL

In this section we present a phase-space model for nonrelativistic quantum mechanics. Although we restrict our attention to two-dimensional phase space, our results are easily generalized to higher dimensions.

Let $R^2 = \{(x,k) : x, k \in R\}$ be a two-dimensional phase space. A function $f: R^2 \to C$ is a *probability amplitude function* if (a) for every $k$, $f(\cdot,k)$ is Lebesgue integrable, $k \to \int_R f(x,k)\,dx$ is Lebesgue measurable, and

$$\int_R \left| \int_R f(x,k)\,dx \right|^2 dk = 1;$$

and (b) for every $x$, $f(x,\cdot)$ is Lebesgue integrable, $x \to \int_R f(x,k)\,dk$ is Lebesgue measurable, and

$$\int_R \left| \int_R f(x,k)\,dk \right|^2 dx = 1.$$

Let $f$ be a probability amplitude function. For $A \in B(R)$ we call $F(A,k) = |\int_A f(x,k)\,dx|^2$ the *P-density following an A-measurement of Q*, and $F(x,A) = |\int_A f(x,k)\,dk|^2$ the *Q-density following an A-measurement of P*. For $A, B \in B(R)$ we call $\mu_P(B,A) = \int_B F(A,k)\,dk$ the *P-probability of B following an A-measurement of Q* and $\mu_Q(B,A) = \int_B F(x,A)\,dx$ the *Q-probability of B following an A-measurement of P*. The *P-probability of B given an A-measurement of Q* is $\mu_P(B|A) = \mu_P(B,A)/\mu_P(R,A)$ and the *Q-probability of B given an A-measurement of P* is $\mu_Q(B|A) = \mu_Q(B,A)/\mu_Q(R,A)$.

Notice, that for all $B \in B(R)$

$$\mu_Q(B,R) = \int_B F(x,R)\,dx = \int_B \left| \int_R f(x,k)\,dk \right|^2 dx,$$

$$\mu_P(B,R) = \int_B F(R,k)\,dk = \int_B \left| \int_R f(x,k)\,dx \right|^2 dk$$

are measures on $B(R)$. If we let $Y = Z = R$, $\Sigma_Y = \Sigma_Z = B(R)$, $\nu(B) = \mu_Q(B,R)$, $w(B) = \mu_P(B,R)$, then $(Y, \Sigma_Y, \nu)$ and $(Z, \Sigma_Z, w)$ become probability spaces. Let $X = Y \times Z = R^2$. As in Example 1, for $(\mathbf{x}, y) \in X \times Y$, let $X(\mathbf{x}, y) = y \times Z$, $\Sigma(\mathbf{x}, y) = \{y \times B : B \in \Sigma_Z\}$, $\mu_{\mathbf{x},y}(y \times B) = w(B)$. Then

$$M_1 = (X, Y, \{X(\mathbf{x}, y) : (\mathbf{x}, y) : (\mathbf{x}, y) \in X \times Y\})$$

is a probability manifold. In a similar way, for $(\mathbf{x}, z) \in X \times Z$, let $X(\mathbf{x}, z) = Y \times z$, $\Sigma(\mathbf{x}, z) = \{B \times z : B \in \Sigma_Z\}$, $\mu_{\mathbf{x},z}(B \times z) = \nu(B)$. Then

$$M_2 = (X, Z, \{X(\mathbf{x}, z) : (\mathbf{x}, z) \in X \times Z\})$$

is a probability manifold. In this way, the phase space $X = R^2$ can be thought of as a probability manifold with two different local structures, given by $M_1$ and $M_2$. The probability amplitude function $f$ allows us to consider the two local structures simultaneously. Thus position and momentum are considered simultaneously.

If $\psi \in L^2(R,dx)$, then of course $\psi$ corresponds to a pure one-dimensional quantum state. The next theorem shows that there is a corresponding probability amplitude function.

**Theorem 8:** If $\psi \in L^2(R,dx)$ with $\|\psi\| = 1$, then there exists an $f: R^2 \to C$ satisfying

(a) for every $k \in R$,

$$f(x,k) = (2\pi)^{-1}\psi(x)e^{-ikx}, \text{ a.e.}[x],$$

(b) for every $x \in R$,

$$f(x,k) = (2\pi)^{-1/2}\hat{\psi}(k)e^{ikx}, \text{ a.e.}[k],$$

where $\hat{\psi}$ denotes the Fourier transform of $\psi$.

*Proof*: The proof is similar to that of Theorem 4. Define the family of sets

$$F = \{x \times R, R \times k : x, k \in R\}.$$

Now $F$ can be well-ordered as in Theorem 4. We now define $f$ by induction. If the first set in this ordering has the form $x_1 \times R$, define $f$ on this set by $f(x_1,k) = (2\pi)^{-1/2}\hat{\psi}(k)e^{ikx_1}$ and if the first set has the form $R \times k_1$, define $f$ on this set by $f(x,k_1) = (2\pi)^{-1/2}\psi(x)e^{-ik_1x}$. The proof now proceeds as in Theorem 4. □

It is easy to show that $f$ in Theorem 8 is a probability amplitude function. We call $f$ a probability amplitude function *corresponding* to $\psi$. The next theorem shows that for $f$ to correspond to a quantum state $\psi$, $f$ must satisfy some strict constraints. This indicates that there may be quantum systems that are not described by conventional states.

**Theorem 9:** A probability amplitude function $f$ corresponds to a $\psi \in L^2(R,dx)$ if and only if for every $A \in B(R)$ we have

$$\text{(a)}\quad \int_A f(x,k)dx = \left[\chi_A(v)\int_R f(v,u)du\right]^{\wedge}(k),$$

$$\text{(b)}\quad \int_A f(x,k)dk = \left[\chi_A(v)\int_R f(u,v)du\right]^{\vee}(x).$$

*Proof*: Suppose $f$ corresponds to the quantum state $\psi$ Then for any $A \in B(R)$ we have

$$\left[\chi_A(v)\int_R f(v,u)du\right]^{\wedge}(k) = (\chi_A\,\psi)^{\wedge}(k) = \int_A f(x,k)dx,$$

and

$$\left[\chi_A(v)\int_R f(u,v)du\right]^{\vee}(x) = (\chi_A\,\hat{\psi})^{\vee}(k) = \int_A f(x,k)dk.$$

Conversely, suppose that (a) and (b) hold. Define $\psi(x) = \int_R f(x,u)du$. Then for $A \in B(R)$ we have

$$\int_A f(x,k)dx = (\chi_A\,\psi)^{\wedge}(k) = (2\pi)^{-1/2}\int_A \psi(x)e^{-ikx}\,dx\,.$$

Hence, for every $k \in R$, $f(x,k) = (2\pi)^{-1/2}\psi(x)e^{-ikx}$, a.e.[x]. Also, for $A \in B(R)$ we have

$$\int_A f(x,k)dk = (\chi_A\,\hat{\psi})^{\vee}(x) = (2\pi)^{-1/2}\int_A \hat{\psi}(k)e^{-ikx}\,dk.$$

Hence, for every $x \in R$, $f(x,k) = (2\pi)^{-1/2}\hat{\psi}(k)e^{-ikx}$, a.e.[k]. □

The next two results show that a probability amplitude function corresponding to a quantum state produces the usual quantum marginal distributions and conditional expectations. We denote the spectral measures for $P$ and $Q$ by $E^P$ and $E^Q$, respectively.

*Lemma 10:* Let $f$ be a probability amplitude function corresponding to $\psi \in L^2(R,dx)$. Then

$$\text{(a)}\quad F(A,k) = |(\psi\chi_A)^{\wedge}(k)|^2, \quad F(R,k) = |\hat{\psi}(k)|^2;$$

$$\text{(b)}\quad F(x,A) = |(\hat{\psi}\chi_A)^{\vee}(x)|^2, \quad F(x,k) = |\psi(x)|^2;$$

$$\text{(c)}\quad \mu_P(B,A) = \int_B |(\psi\chi_A)^{\wedge}(k)|^2 dk,$$

$$\mu_P(R,A) = \int_A |\psi(x)|^2\,dx,$$

$$\mu_P(A,R) = \int_A |\hat{\psi}(k)|^2\,dk;$$

$$\text{(d)}\quad \mu_Q(B,A) = \int |(\hat{\psi}\chi_A)^{\vee}(x)|^2\,dx,$$

$$\mu_Q(R,A) = \int |\hat{\psi}(k)|^2\,dk,$$

$$\mu_Q(A,R) = \int |\psi(x)|^2\,dx\,.$$

*Proof*: Straightforward application of definitions. □

**Theorem 11:** If $f$ is a probability amplitude function corresponding to $\psi \in L^2(R,dx)$, and $P_\psi$ denotes the one-dimensional projection onto $\psi$, then

$$\text{(a)}\quad \mu_P(B\,|A) = \frac{\operatorname{tr}\left[E^P(B)E^Q(A)P_\psi E^Q(A)\right]}{\operatorname{tr}\left[E^Q(A)P_\psi\right]},$$

$$\text{(b)}\quad \mu_Q(B\,|A) = \frac{\operatorname{tr}\left[E^Q(B)E^P(A)P_\psi E^P(A)\right]}{\operatorname{tr}\left[E^P(A)P_\psi\right]}.$$

*Proof*: (a) Denote the Fourier transform by $F$. Using the fact that $E^Q(B)$ is multiplication by $\chi_B$ and $E^P(B) = F^*E^Q(B)F$ we obtain by Lemma 10,

$$\begin{aligned}
\mu_P(B,A) &= \int_B |(\psi\chi_A)^{\wedge}(k)|^2\,dk = \|\chi_B(\psi\chi_A)^{\wedge}\|^2 \\
&= \|F^*\chi_B\,F(\psi\chi_A)\|^2 = \|E^P(B)(\psi\chi_A)\|^2 \\
&= \langle E^P(B)\psi\chi_A, \psi\chi_A\rangle \\
&= \langle E^P(B)E^Q(A)\psi, E^Q(A)\psi\rangle \\
&= \langle E^Q(A)E^P(B)E^Q(A)P_\psi\,\psi,\psi\rangle \\
&= \operatorname{tr}\left[E^Q(A)E^P(B)E^Q(A)P_\psi\right] \\
&= \operatorname{tr}\left[E^P(B)E^Q(A)P_\psi E^Q(A)\right].
\end{aligned}$$

Hence,

$$\begin{aligned}
\mu_P(B\,|A) &= \frac{\mu_P(B,A)}{\mu_P(R,A)} \\
&= \frac{\operatorname{tr}\left[E^P(B)E^Q(A)P_\psi E^Q(A)\right]}{\operatorname{tr}\left[E^Q(A)P_\psi\right]}.
\end{aligned}$$

(b) As in part (a) we have

$$\begin{aligned}
\mu_Q(B,A) &= \int_B |(\hat{\psi}\chi_A)^{\vee}(x)|^2\,dx = \|\chi_B(\chi_A\,\hat{\psi})^{\vee}\|^2 \\
&= \|\chi_B F^*E^Q(A)F\psi\|^2 = \|E^Q(B)E^P(A)\psi\|^2 \\
&= \langle E^Q(B)E^P(A)\psi,\quad E^P(A)\psi\rangle \\
&= \langle E^P(A)E^Q(B)E^P(A)P_\psi\,\psi,\psi\rangle \\
&= \operatorname{tr}\left[E^P(A)E^Q(B)E^P(A)P_\psi\right] \\
&= \operatorname{tr}\left[E^Q(B)E^P(A)P_\psi E^P(A)\right].
\end{aligned}$$

Hence,

$$\begin{aligned}
\mu_Q(B\,|A) &= \frac{\mu_Q(B,A)}{\mu_Q(R,A)} \\
&= \frac{\operatorname{tr}\left[E^Q(B)E^P(A)P_\psi E^P(A)\right]}{\operatorname{tr}\left[E^P(A)P_\psi\right]}.
\end{aligned}$$

□


[1] S. Gudder, *Stochastic Methods in Quantum Mechanics* (North Holland, New York, 1979).
[2] S. Gudder, SIAM Rev. **26**, 71 (1984).
[3] I. Pitowski, Phys. Rev. Lett. **48**, 1299 (1982).
[4] I. Pitowski, Phys. Rev. D **27**, 2316 (1983).

# Probability distributions with given multivariate marginals

Leon Cohen
*Hunter College of the City University, New York, New York 10021*

(Received 25 October 1983; accepted for publication 23 March 1984)

A method is presented for obtaining joint probability density functions which satisfy given multivariate marginal densities.

PACS numbers: 02.50.Cw, 02.30. + g

## I. INTRODUCTION

Suppose we have the marginal probability densities $P_1$ and $P_2$ of the two sets of random variables $X_1 = (x_1,x_2,...,x_N)$ and $X_2 = (x_{N_1+1},x_{N_1+2},...,x_{N_1+N_2})$, respectively. The purpose of this paper is to present a method for obtaining a joint distribution $P(x_1,x_2,...,x_{N_1+N_2})$ consistent with the marginals. The case where the marginals are a function of only one variable has been extensively studied[1] and a general method has been given to generate an infinite number of joint distributions for that situation.[2–4] We shall here consider the case where the marginals are in general multivariate, that is, functions of more than one random variable. We assume that the given marginals do not have any variables in common.

What we seek are positive functions such that

$$\int_{-\infty}^{\infty} P(x_1,x_2,...,x_{N_1+N_2})dx_{N_1+1}\cdots dx_{N_1+N_2} = P_1(x_1,x_2,...,x_{N_1}), \tag{1.1}$$

$$\int P(x_1,x_2,...,x_{N_1+N_2})dx_1\,dx_2\cdots dx_{N_1} = P_2(x_{N_1+1},...,x_{N_1+N_2}). \tag{1.2}$$

## II. JOINT DISTRIBUTION

Choose *any* positive function, $h(u_1,u_2,...,u_N)$ of $N(=N_1+N_2)$ variables defined in the $N$-dimensional unit cube and normalize it to 1,

$$\int_0^1 h(u_1,u_2,...,u_N)du_1\,du_2\cdots du_N = 1, \tag{2.1}$$

and let

$$h_1(u_1,...,u_{N_1}) = \int_0^1 h(u_1,...,u_{N_1+N_2})du_{N_1+1}\cdots du_{N_1+N_2}, \tag{2.2}$$

$$h_2(u_{N_1+1},...,u_{N_1+N_2}) = \int_0^1 (u_1,...,u_{N_1+N_2})du_1\cdots du_{N_1}, \tag{2.3}$$

$$\rho(u_1,...,u_N) = h - h_1 - h_2 + 1. \tag{2.4}$$

An example of such a function is

$$h = 2^N \prod_{i=1}^{N} u_i,$$

$$h_1 = 2^{N_1} \prod_{i=1}^{N_1} u_i,$$

$$h_2 = 2^{N_2} \prod_{i=N_1+1}^{N} u_i,$$

$$\rho = \left\{2^{N_1}\prod_{i=1}^{N_1} u_i - 1\right\}\left\{2^{N_2}\prod_{i=N_1+1}^{N} u_i - 1\right\}. \tag{2.5}$$

An infinite number of joint distributions satisfying the marginals may be generated by

$$P(x_1,x_2,...,x_N) = P_1P_2[1 + c\rho(u_1,u_2,...,u_N)], \tag{2.6}$$

where now the $u$'s are defined in terms of the $x$'s

$$u_i = u_i(x_1,x_2,...,x_{N_1}), \quad i = 1,2,...,N_1 \tag{2.7}$$

$$u_i = u_i(x_{N_1+1},...,x_{N_1+N_2}), \quad i = N_1+1,...,N_1+N_2, \tag{2.8}$$

in such a way thta the Jacobian of the transformations satisfy

$$J_1 = P_1, \tag{2.9}$$

$$J_2 = P_2. \tag{2.10}$$

Also, the transformation must map the infinite $N_1$-dimensional space into the $N_1$-dimensional unit cube and similarly for $N_2$. InSec. III we will give the explicit transfomation.

The constant $c$ is chosen so that $P$ will be positive. Identical arguments to that previously given[2] implies that $c$ may be chosen to be any number in the range

$$-1/l_2 \leqslant c \leqslant 1/l_1, \tag{2.11}$$

where $-l_1$ and $l_2$ are the absolute minimum and maximum of $\rho$.

To show that the marginals are satisfied consider

$$\int_{-\infty}^{\infty} P_2(x_{N_1+1},...,x_{N_1+N_2})h(u_1,...,u_N)dx_{N_1+1}\cdots dx_{N_1+N_2} \tag{2.12}$$

$$= \int_0^1 h(u_1,...,u_N)du_{N_1+1}\cdots du_N = h_1.$$

Furthermore

$$\int_{-\infty}^{\infty} P_2\,h_2\,dx_{N_1+1}\cdots dx_N = \int_0^1 h\,du_{N_1+1}\cdots du_N = 1, \tag{2.13}$$

and hence

$$\int P_2\rho\,dx_{N_1+1}\cdots dx_{N_1+N_2} = 0, \tag{2.14}$$

which proves that the first marginal condition is satisfied. Similarly for the second.

## III. TRANSFORMATION AND JACOBIAN

What remains to be shown is that the transformations can be found such that the Jacobians are given by Eqs. (2.9) and (2.10). Define

 0022-2488/84/082402-02$02.50  

$$P_1^{(N_1-1)}(x_{i+1},\ldots,x_{N_1}) = \int_{-\infty}^{\infty} P_1(x_1,\ldots,x_{N_1})dx_1\cdots dx_i, \tag{3.1}$$

then a transformation satisfying Eq. (2.9) is

$$u_i = \frac{\int_{-\infty}^{x_i} P_1^{(N_1-i+1)}(x_i',\ldots,x_{N_1})dx_i'}{P_1^{(N_1-i)}(x_{i+1},\ldots,x_{N_1})}, \quad i \leqslant N_1. \tag{3.2}$$

Since

$$\frac{\partial u_i}{\partial x_j} = 0, \quad j<i, \quad i \leqslant N_1, \tag{3.3}$$

The Jacobian is the product of the diagonal elements which are given by

$$\frac{\partial u_i}{\partial x_i} = \frac{P_1^{(N_1-i+1)}}{P_1^{(N_1-i)}}, \quad i \leqslant N, \tag{3.4}$$

and therefore

$$J_1 = \prod_{i=1}^{N_1} \frac{\varphi_1^{(N_i-i+1)}}{P_1^{(N_1-i)}} = P_1. \tag{3.5}$$

For the other set of variables the identical procedure is followed. Also, it s clear that as the $x$'s range over all space the range of the $u$'s is from zero to one.

## IV. CONCLUSION AND GENERALIZATION

An infinite number of $P$'s can be generated by choosing different $h$'s and $c$'s. We point out that sometimes different $H$'s will lead to the same $\rho$ and that also for some choices of $h$ one could get a $\rho$ identical to zero.

Generalization to more than two marginals is as follows. Suppose we have $M$ marginals each having $N_i$ random variables then the joint density will be a function of $MN$ random variables where $N$ is the sum of the $N_i$'s. For each marginal the $u$ transformations are chosen as above. The joint density is given by

$$P(x_1,x_2,\ldots,x_{MN}) = \prod_{i=1}^{M} P_i(X_i)[1 + c\rho(U_1,U_2,\ldots,U_M)], \tag{4.1}$$

where now h is a function of $MN$ variables and

$$\begin{aligned} U_i &= (u_{K_i+1},\ldots,u_{K_i+N_i}), \\ X_i &= (x_{K_i+1},\ldots,x_{K_i+N_i}), \\ K_i &= \prod_{j=1}^{i} N_j, \end{aligned} \tag{4.2}$$

with

$$\begin{aligned} \rho(U_1,\ldots,U_M) &= h(U_1,\ldots,U_M) \\ &\quad - \sum_{i=1}^{M} h_i(V_i) + (M-1), \\ h_i(U_i) &= \int_0^1 h\, dU_1\cdots dU_{i-1}\, dU_{i+1}\cdots dU_M. \end{aligned} \tag{4.3}$$

Finch and Groblicki[5] have shown that for the case where the marginals are a function of only one variable, the procedure given above encompasses all solutions. A similar proof holds for the multivariate case. We note that in general $h$ can be a functional of the marginals.


## ACKNOWLEDGMENT

This research was supported in part by a grant from the City University of New York Research Award Program.



[1]See, for example, the following papers and the references therein: D. Morgenstern, Mitt. Math. Stat. **8**, 234 (1956); E. J. Gumbel, Rev. Fac. Ci. Univ. Lisboa Ser. 2A Ci. Mat. 7, 179 (1959); D. J. G. Farlie, Biometrica **50**, 499 (1963).

[2]L. Cohen and Y. I. Zaparovanny, J. Math. Phys. **21**, 794 (1980).

[3]L. Cohen and T. Posch, "Positive Time-Frequency Distribution Functions," submitted to IEEE Trans. Acoust., Speech Signal Process.

[4]L. Cohen, Proc. ICASSP **84**, 41B1.1 (1984).

[5]P. D. Finch and R. Groblicki, "Bivariate Probability Densities with Given Margins," Found. Phys. (to appear).

# Maximum entropy in the problem of moments

Lawrence R. Mead and N. Papanicolaou
*Department of Physics, Washington University, St. Louis, Missouri 63130*

(Received 8 November 1983; accepted for publication 13 January 1984)

The maximum-entropy approach to the solution of underdetermined inverse problems is studied in detail in the context of the classical moment problem. In important special cases, such as the Hausdorff moment problem, we establish necessary and sufficient conditions for the existence of a maximum-entropy solution and examine the convergence of the resulting sequence of approximations. A number of explicit illustrations are presented. In addition to some elementary examples, we analyze the maximum-entropy reconstruction of the density of states in harmonic solids and of dynamic correlation functions in quantum spin systems. We also briefly indicate possible applications to the Lee–Yang theory of Ising models, to the summation of divergent series, and so on. The general conclusion is that maximum entropy provides a valuable approximation scheme, a serious competitor of traditional Padé-like procedures.

PACS numbers: 02.60. + y, 75.10.Jm

## I. INTRODUCTION

The maximum-entropy approach to the solution of underdetermined inverse problems was introduced some time ago.[1,2] Following the original contributions, there has been a long debate concerning the conceptual foundations of maximum entropy for problems outside the traditional domain of thermodynamics. The debate is currently more meaningful than ever in view of the augmenting list of successful practical applications[3] which have become possible because of the steadily increasing computing power available today. While our aim is not to engage in further conceptual ramifications of the rationale of maximum entropy,[4] we shall attempt to sharpen its mathematical foundations and to extend its applicability to various concrete problems encountered in quantum physics.

We consider the classical moment problem where a positive density $P(x)$ is sought from knowledge of its power moments

$$\int_a^b x^n P(x)dx = \mu_n, \quad n = 0,1,2,\ldots . \tag{1.1}$$

The extent to which the density $P(x)$ may be determined from its moments has been extensively discussed in the mathematical literature. In practice, only a finite number of moments, say $N + 1$, is usually available. Clearly then there exists an infinite variety of functions whose first $N + 1$ moments coincide and a unique reconstruction of $P(x)$ is impossible. Nevertheless, various approximation procedures exist which aim at constructing specific sequences of functions $P_N(x)$, such that

$$\int_a^b x^n P_N(x)dx = \mu_n, \quad n = 0,1,\ldots, N, \tag{1.2}$$

which eventually converge to the true distribution $P(x)$ as $N$ approaches infinity. It is often mathematically expedient, and physically useful, to abandon the requirement of pointwise convergence and, instead, require weaker convergence for averages of the form

$$\langle F \rangle = \int_a^b F(x)P(x)dx = \lim_{N\to\infty} \int_a^b F(x)P_N(x)dx, \tag{1.3}$$

where $F(x)$ is some known function of physical interest.

The maximum-entropy approach offers a definite procedure for the construction of a sequence of approximations. The positive density $P(x)$ is interpreted as a probability density and the corresponding entropy is maximized under the condition that the first $N + 1$ moments be equal to the true moments $\mu_n, n = 0,1,\ldots, N$. Introducing appropriate Lagrange multipliers, one seeks maximization of the entropy functional $S = S(P)$ defined from

$$S = -\int_a^b [P(x)\ln P(x) - P(x)]dx + \sum_{n=0}^{N} \lambda_n \left( \int_a^b x^n P(x)dx - \mu_n \right). \tag{1.4}$$

Notice that we have incorporated in the definition of the entropy a term linear in $P(x)$, mostly for notational convenience. The linear term may be absorbed by a trivial redefinition of the Lagrange multiplier $\lambda_0$ in Eq. (1.4). Returning to the main point, the maxima $P = P_N(x)$ of (1.4) for $N = 1,2,\ldots$ will be taken as a sequence of approximations for the true density $P(x)$. It is customary to say that the maximum-entropy sequence $P_N(x)$ is the least-biased sequence of approximations.

The mathematical problem posed in the preceding paragraph was already considered in standard works on maximum entropy and concrete applications were also worked out in certain cases.[3,5,6] Nonetheless, recent reviews of a wealth of moment problems in quantum physics[7] do not even acknowledge possible use of the maximum-entropy approach. This situation is understandable because the more popular methods, such as polynomial expansions, Padé approximants, and the like, have had the advantage of extensive mathematical scrutiny over a period of a century or so. It is clear that a similar status for maximum entropy could be achieved only by an equally thorough study of its mathematial basis, by widening the scope of concrete applications, and by explicit comparison with the best approximation procedures currently in use.

For comparison purposes, it seems appropriate to brief-

 0022-2488/84/082404-14$02.50  

ly outline here some of the better known methods for approximate solutions of the moment problem. A simple possibility is to expand $P(x)$ in some set of orthogonal polynomials. The resulting series is truncated after $N+1$ terms and the expansion coefficients are determined by requiring that the first $N+1$ moments be correct. This entails the solution of a $(N+1)\times(N+1)$ system of linear equations. Judicious choices of weighted orthogonal polynomials could lead to rapidly convergent sequences. In practice, the choice of a suitable weight is usually difficult, so the resulting sequences often produce notoriously oscillating approximations to $P(x)$ which are further impaired by lack of positivity at each finite stage of iteration.

Alternative, usually more powerful, procedures have been developed over the years, most of which are classified under the generic name of Padé approximations.[8] For instance, one may attempt to approximate the positive function $P(x)$ by finite sums of $\delta$-functions of the form

$$P_N(x)=\sum_{i=1}^{(N+1)/2} m_i\delta(x-x_i), \tag{1.5a}$$

when $N$ is odd, and

$$P_N(x)=\sum_{i=0}^{N/2} m_i\delta(x-x_i), \quad x_0\equiv a, \tag{1.5b}$$

when $N$ is even. In a language preferred by mathematicians, one writes $P(x)dx=d\mu(x)$, where the nondecreasing measure $\mu(x)$ is approximated by multistep functions. The parameters $m_i$ and $x_i$ in (1.5) are again determined by the requirement that the first $N+1$ moments be correct:

$$\sum_i m_i x_i^n=\mu_n, \quad n=0,1,\ldots,N. \tag{1.6}$$

These are nonlinear equations but their solution may be reduced to the diagonalization of a tridiagonal Jacobi matrix.[9,10] The corresponding numerical procedure is apparently very stable and is often quoted in the literature as the Lanczos algorithm.[11] While the preceding method does not directly address a pointwise construction of $P(x)$, it is well suited for the computation of averages of the form (1.3) for which approximations may be obtained from

$$\langle F\rangle_N=\sum_i m_i F(x_i). \tag{1.7}$$

For the special case where $F(x)=1/(1+zx)$, Eq. (1.7) is but the standard Padé approximant associated with Stieljes integrals of the form

$$\langle F\rangle=\int_a^b \frac{P(x)dx}{1+zx}, \quad \langle F\rangle_N=\sum_i \frac{m_i}{1+zx_i}. \tag{1.8}$$

The distinction between even and odd $N$ implicit in Eq. (1.5) results in two independent sequences of approximation which are the familiar diagonal and off-diagonal Padé sequences.

A number of questions raised in the preceding general introduction will be addressed in the following to varying degrees of completeness. In Sec. II, we briefly review well-known facts about maximum entropy and present some new mathematical results. In important special cases, we are able to derive necessary and sufficient conditions for the existence of a maximum-entropy solution and to some extent study the convergence of the resulting sequence. The numerical procedure and some elementary examples are also discussed in Sec. II. More interesting applications are described in the remainder of the paper. A detailed calculation of the density of states for a harmonic face-centered-cubic (fcc) crystal is presented in Sec. III and the results are compared with the earlier work of Gordon and Wheeler[10] using the Padé-like procedure outlined above; Sec. IV presents a similar calculation for dynamic correlation functions in some typical quantum spin systems. Further applications are contemplated in Sec. V and are illustrated by some simple exercises in the context of the Lee–Yang theory for Ising models. The same section contains a number of concluding remarks and some suggestions for possible generalizations.

## II. FORMULATION AND ELEMENTARY EXAMPLES

The starting point for our discussion is the entropy defined by Eq. (1.4) for which we seek a maximum. Functional variation with respect to the unknown density $P(x)$ yields

$$\frac{\delta S}{\delta P(x)}=0\Rightarrow P=P_N(x)=\exp\Big(-\lambda_0-\sum_{n=1}^{N}\lambda_n x^n\Big), \tag{2.1}$$

to be supplemented by the condition that the first $N+1$ moments be given by $\mu_n$:

$$\int_a^b x^n P_N(x)dx=\mu_n, \quad n=0,1,\ldots,N. \tag{2.2}$$

Equations (2.2) should be viewed as a (nonlinear) system of $N+1$ equations for the $N+1$ unknown Lagrange multipliers $\lambda_0,\lambda_1,\ldots,\lambda_N$. Without loss of generality, we may assume in the following that the density $P(x)$ is normalized such that $\mu_0=1$. The first equation ($n=0$) in (2.2) then reads

$$\int_a^b P_N(x)dx=\int_a^b dx\exp\Big(-\lambda_0-\sum_{n=1}^{N}\lambda_n x^n\Big)=1, \tag{2.3}$$

and may be used to express $\lambda_0$ in terms of the remaining Lagrange multipliers:

$$e^{\lambda_0}=\int_a^b dx\exp\Big(-\sum_{n=1}^{N}\lambda_n x^n\Big)\equiv Z. \tag{2.4}$$

The system of equations (2.2) reduces to

$$\langle x^n\rangle=\mu_n, \quad n=1,2,\ldots,N,$$
$$\langle x^n\rangle\equiv\frac{\int_a^b dx\, x^n\exp(-\Sigma_{n=1}^{N}\lambda_n x^n)}{\int_a^b dx\exp(-\Sigma_{n=1}^{N}\lambda_n x^n)}. \tag{2.5}$$

An analytical solution of (2.5) is, of course, impossible except for the simple case $N=1$. For numerical as well as theoretical purposes, one introduces a potential $\Gamma=\Gamma(\lambda_1,\lambda_2,\ldots,\lambda_N)$ through the Legendre transformation[5]

$$\Gamma=\ln Z+\sum_{n=1}^{N}\mu_n\lambda_n, \tag{2.6}$$

there the $\mu_n$'s are the actual numerical values of the known moments. Stationary points of the potential $\Gamma$ are solutions to the equations

$$\frac{\partial\Gamma}{\partial\lambda_n}=0\Rightarrow\langle x^n\rangle=\mu_n, \quad n=1,2,\ldots,N, \tag{2.7}$$

which are precisely Eqs. (2.5). The first important property of $\Gamma$ is summarized in the following lemma.

*Lemma 1*: The potential $\Gamma = \Gamma(\lambda_1,\lambda_2,...,\lambda_N)$ is everywhere convex. The proof of this result is already given in the literature[5] and proceeds by explicit construction of the Hessian

$$H_{nm} = \frac{\partial^2 \Gamma}{\partial\lambda_n\,\partial\lambda_m} = \langle x^{n+m}\rangle - \langle x^n\rangle\langle x^m\rangle, \tag{2.8}$$

which may be proven to be positive definite for any generic set of $\lambda$'s, not necessarily satisfying Eqs. (2.5). A more direct demonstration obtains by treating all Lagrange multipliers, including $\lambda_0$, on a common basis. Thus we introduce the potential $\Delta = \Delta(\lambda_0,\lambda_1,...,\lambda_N)$ from

$$\Delta = \int_a^b \left[\exp\left(-\lambda_0 - \sum_{n=1}^{N}\lambda_n x^n\right) - 1\right]dx + \sum_{n=0}^{N}\mu_n\lambda_n, \tag{2.9}$$

whose stationary points are given by

$$\frac{\partial\Delta}{\partial\lambda_n} = 0 \Rightarrow \langle x^n\rangle = \mu_n, \quad n = 0,1,...,N, \tag{2.10}$$

which recombine Eqs. (2.5) with (2.4). Had we eliminated $\lambda_0$ using (2.4), the first term in (2.9) would vanish and the remaining terms would give (with $\mu_0 = 1$)

$$\Delta = \sum_{n=0}^{N}\mu_n\lambda_n = \mu_0\lambda_0 + \sum_{n=1}^{N}\mu_n\lambda_n = \ln Z + \sum_{n=1}^{N}\mu_n\lambda_n, \tag{2.11}$$

which is the potential $\Gamma$ introduced earlier. However, one may directly work with $\Delta = \Delta(\lambda_0,\lambda_1,...,\lambda_N)$ which also satisfies Lemma 1. The corresponding Hessian reads

$$\Theta_{nm} = \frac{\partial^2\Delta}{\partial\lambda_n\,\partial\lambda_m} = \int_a^b dx\, x^{n+m}\exp\left(-\sum_{n=0}^{N}\lambda_n x^n\right) \equiv \langle x^{n+m}\rangle, \quad n,m = 0,1,.., \tag{2.12}$$

and its positive definiteness is trivially established noting that

$$\int_a^b dx\,(u_0 + u_1x + \cdots + u_kx^k)^2\exp\left(-\sum_{n=0}^{N}\lambda_n x^n\right) \geqslant 0, \tag{2.13}$$

for any nonnegative integer $k$ and for any real $u_0,u_1,...,u_k$. Equation (2.13) may be rewritten as

$$\sum_{n,m=0}^{k}\langle x^{n+m}\rangle u_n u_m = \sum_{n,m=0}^{k}\Theta_{nm}u_n u_m \geqslant 0, \tag{2.14}$$

which establishes that the Hessian $\Theta_{nm}$ is positive definite.

In practice, the potentials $\Gamma$ or $\Delta$ may be used with comparable efficiency. We shall therefore proceed using the potential $\Gamma$. However, the potential $\Delta$ will prove more flexible for some generalizations discussed in Sec. V.

The convexity of $\Gamma$ guarantees that if a stationary point is found for some finite values of $\lambda_1,\lambda_2,...,\lambda_N$, it must be a unique absolute minimum. Notice, however, that convexity alone does not imply that such a minimum should exist. This is not surprising because the convexity of $\Gamma$ was established without any reference to the specific properties of the actual moments $\mu_n$. A simple illustration may be given taking $N = 1$ and $[a,b] = [0,1]$, so that

$$Z = \int_0^1 dx\, e^{-\lambda_1 x} = \frac{1 - e^{-\lambda_1}}{\lambda_1},$$

$$\Gamma = \ln[(1 - e^{-\lambda_1})/\lambda_1] + \mu_1\lambda_1. \tag{2.15}$$

It is not difficult to see that the convex function $\Gamma = \Gamma(\lambda_1)$ possesses a minimum at some finite $\lambda_1$ only if $\mu_1 < 1\ (= \mu_0)$. It is clear that this is the first of a set of conditions that the actual moments must satisfy in order to guarantee a minimum for $\Gamma = \Gamma(\lambda_1,...,\lambda_N)$. What are those conditions?

In order to answer the above question, it is pertinent at this point to review the general conditions under which the full moment problem shall have a solution, independently of the method of approximation. We restrict our discussion to the moment problem defined over a finite interval, say [0,1], which is the so-called Hausdorff moment problem.[12] Let $P(x)$ be a nonnegative density integrable in [0,1] and let $\{\mu_n, n = 0,1,2,...\}$ be the associated moment sequence. Noting that

$$\int_0^1 x^n(1-x)^k P(x)dx > 0, \tag{2.16}$$

and working out the integrand using the binomial expansion, the left side of (2.16) may be expressed in terms of the moments $\mu_n$:

$$\Delta^k\mu_n \equiv \sum_{m=0}^{k}(-1)^m\binom{k}{m}\mu_{n+m} > 0, \quad n,k = 0,1,2,... . \tag{2.17}$$

It is evident that the set of inequalities (2.17) is a set of necessary conditions for the moment sequence. Such a sequence is called completely monotonic. More importantly, Hausdorff established the sufficiency of the above conditions. Namely, given a completely monotonic moment sequence, there exists a nonnegative density $P(x)$ integrable in [0,1] whose moments coincide with $\mu_n$.

Applying (2.17) for $k = 1$ and $n = 0$ one finds that $\mu_1 < \mu_0$, which is the condition we found earlier so that the potential $\Gamma = \Gamma(\lambda_1)$ of Eq. (2.15) will have a minimum. It is tempting to presume that the general potential $\Gamma = \Gamma(\lambda_1,\lambda_2,...,\lambda_N)$ will have a minimum if and only if the full set of conditions (2.17) is satisfied. That this is indeed so is guaranteed by the following theorem.

**Theorem 1:** A necessary and sufficient condition that the potential $\Gamma$ should have a unique absolute minimum at some finite $\lambda_1,\lambda_2,...,\lambda_N$ for any $N$ is that the moment sequence $\{\mu_n, n = 0,1,2,...\}$ should be completely monotonic.

We first establish sufficiency which is obviously the most relevant aspect of Theorem 1 for practical applications. In view of the convexity, it is clear that the essence of the proof should evolve around the asymptotic behavior of $\Gamma$ at large $\lambda$. Hence the Lagrange multipliers are written as

$$\lambda_n = \lambda\alpha_n, \quad \sum_{n=1}^{N}\alpha_n^2 = 1, \tag{2.18}$$

where $\lambda$ is positive and the $\alpha_n$'s are the familiar direction cosines. One then obtains

$$\Gamma = \ln Z + \sum_{n=1}^{N}\mu_n\lambda_n = \ln\left[\int_0^1 dx\exp\left(-\lambda\sum_{n=1}^{N}\alpha_n x^n\right)\right] + \lambda\sum_{n=1}^{N}\mu_n\alpha_n. \tag{2.19}$$

Our aim is to study the behavior of $\Gamma$ as $\lambda \to \infty$ for an arbitrary choice of the direction cosines. It is convenient to combine both terms in (2.19) and write

$$\Gamma = \ln J, \quad J = \int_0^1 dx\, e^{\lambda \Pi_N(x)},$$

$$\Pi_N(x) = \sum_{n=1}^{N} \alpha_n(\mu_n - x^n), \tag{2.20}$$

so our task reduces to the study of the asymptotic behavior of the Laplace integral $J = J(\lambda)$ at $\lambda \to \infty$. The general procedure is explained in standard textbooks[13] and the result depends on the behavior of the $N$ th degree polynomial $\Pi_N(x)$ in [0,1]. The relevant property of $\Pi_N(x)$ for our current purposes is summarized in the following lemma.

*Lemma 2*: If $\{\mu_n, n = 0,1,2,...\}$ is a completely monotonic moment sequence, the $N$ th-degree polynomial $\Pi_N(x) = \Sigma_{n=1}^{N} \alpha_n(\mu_n - x^n)$ is strictly positive in a nontrivial subinterval of [0,1], for arbitrary real $\alpha_1, \alpha_2, ..., \alpha_N$ not all of which vanish.

The proof of the lemma proceeds by contradiction. Let us assume that $\Pi_N(x)$ is not positive anywhere in [0,1], i.e.,

$$\Pi_N(x) \leqslant 0, \quad x \in [0,1]. \tag{2.21}$$

The polynomial $\Pi_N(x)$ may not be identically equal to zero because not all of the coefficients $\alpha_1, \alpha_2, ..., \alpha_N$ vanish. It is therefore evident that the polynomial $\Pi_N(x)$ may not vanish but at a finite number of points not exceeding its degree $N$. Furthermore, the general theory of Hausdorff guarantees that given a completely monotonic moment sequence there exists a nonnegative density $P(x)$ whose moments are $\mu_0 = 1$, $\mu_1, \mu_2, ...$. While $P(x)$ may vanish over nontrivial subintervals of [0,1], it must also be strictly positive over some nontrivial regions in [0,1]. Hence, in view of (2.21) and the ensuing remarks, the product $\Pi_N(x)P(x) \leqslant 0$ may vanish over nontrivial regions but its integral over [0,1] is strictly negative:

$$\int_0^1 \Pi_N(x)P(x)dx < 0. \tag{2.22}$$

Some implicit smoothness assumptions about $P(x)$ are inherent in the above argument; $P(x)$ cannot be a $\delta$-function, for instance. On introducing the explicit expression for the polynomial $\Pi_N(x)$ in (2.22), one finds that

$$\sum_{n=1}^{N} \alpha_n \int_0^1 (\mu_n - x^n)P(x)dx < 0. \tag{2.23}$$

Recalling that $\int_0^1 P(x)dx = \mu_0 = 1$ and $\int_0^1 x^n P(x)dx = \mu_n$ for $n = 1,2,...$, the left side of (2.23) vanishes. We have thus reached a contradiction implying that (2.21) cannot be true over the entire interval [0,1]. Hence there exist nontrivial regions in [0,1] where $\Pi_N(x)$ is strictly positive, establishing the validity of Lemma 2.

The conditions of Lemma 2 are valid for the polynomial $\Pi_N(x)$ defined in Eq. (2.20) because not all of the direction cosines $\alpha_1, \alpha_2, ..., \alpha_N$ may vanish simultaneously in view of the normalization constraint (2.18). Let $x_0$ be the point where $\Pi_N(x)$ achieves its maximum which is positive:

$$\max_{x \in [0,1]} \Pi_N(x) = \Pi_N(x_0) > 0. \tag{2.24}$$

The behavior of $J(\lambda)$ at $\lambda \to \infty$ is governed by the behavior of $\Pi_N(x)$ in the neighborhood of $x_0$. There are several cases to consider depending on whether $x_0$ lies at one of the endpoints or not, and whether the corresponding derivatives $\Pi'_N(x_0), \Pi''_N(x_0), ...$ vanish or not. A complete analysis of the various cases may be found in Ref. 13, which we will not repeat here. The general result is that $J(\lambda)$ grows exponentially as $\lambda \to \infty$ for an arbitrary choice of the direction cosines. Hence the potential $\Gamma = \ln J$ grows linearly with $\lambda$ in all directions. A convex function with the above asymptotic behavior must possess a unique absolute minimum at some finite $\lambda_1, \lambda_2, ..., \lambda_N$, which establishes sufficiency in Theorem 1.

The necessity of the conditions stated in Theorem 1 does not have a direct bearing on the practical aspects of this problem, but we briefly sketch the proof of its validity for the sake of completeness. Let us assume that the sequence of real numbers $\{\mu_n, n = 0,1,2,...\}$ is such that the potential $\Gamma$ possess a minimum at some finite $\lambda_1, \lambda_2, ..., \lambda_N$. We are to prove that the sequence $\{\mu_n\}$ must be completely monotonic. Recall that $\Gamma$ is convex everywhere for arbitrary values of $\mu_n$. Since $\Gamma$ possesses a minimum, by our hypothesis, the minimum is unique and absolute. Therefore, moving away from the minimum in any direction should result in monotonically increasing values for $\Gamma$. This behavior is compatible only with a polynomial $\Pi_N(x)$ in Eq. (2.20) that achieves a positive maximum at some point $x_0$ in [0,1] for any value of the direction cosines. We write

$$\phi_N(x) \equiv \lambda \Pi_N(x) = \sum_{s=1}^{N} \lambda\alpha_s(\mu_s - x^s), \quad \phi_N(x_0) > 0, \tag{2.25}$$

for any real $\lambda\alpha_1, \lambda\alpha_2, ..., \lambda\alpha_N$ not all of which vanish. In particular, choose

$$\lambda\alpha_s = \begin{cases} 0, & 1 \leqslant s < n, \\ (-1)^{s-n}\binom{k}{s-n}, & n \leqslant s \leqslant n+k, \\ 0, & n+k < s \leqslant N, \end{cases} \tag{2.26}$$

so that, using the notation of Eq. (2.17),

$$\phi_N(x) = \Delta^k \mu_n - x^n(1-x)^k. \tag{2.27}$$

The only stationary point of $\phi_N(x)$ in [0,1] is a local minimum at the interior point $x = n/(n+k)$. Therefore the maximum of $\phi_N(x)$ occurs at one of the endpoints ($x_0 = 0$ or 1) where the second term in (2.27) vanishes, and $\phi_N(x_0) > 0$ implies that $\Delta^k \mu_n > 0$. The sequence $\{\mu_n, n = 0,1,2,...\}$ is thus completely monotonic.

To summarize, it is gratifying that the conditions for the existence of a maximum-entropy solution are identical to Hausdorff's conditions addressing the full moment problem. Given a completely monotonic moment sequence, Theorem 1 guarantees the existence of a maximum-entropy solution $P_N(x)$ for any $N$, however large. The solution $P_N(x)$ is nonnegative and integrable (in fact, absolutely continuous) in [0,1] and satisfies the moment conditions

$$\int_0^1 x^n P_N(x)dx = \mu_n, \quad n = 0,1,...,N. \tag{2.28}$$

A sequence of functions $P_N(x)$, $N = 1,2,...$ with the above general properties converges in the sense of the following theorem.

**Theorem 2:** Let $P(x)$ be a nonnegative function integrable in [0,1] whose moments are $\mu_0,\mu_1,...$, and let $P_N(x)$, $N = 1,2,...$ be the maximum-entropy sequence associated with the same moments. If $F(x)$ is some continuous function in [0,1] then

$$\lim_{N\to\infty}\int_0^1 F(x)P_N(x)dx = \int_0^1 F(x)P(x)dx. \tag{2.29}$$

The proof of the above theorem can be obtained by putting together some standard results of analysis which may be found in the book of Widder[12] and are freely used in the following. Consider the sequence of functions

$$\psi_N(x) = \int_0^x [P(t) - P_N(t)]dt, \quad N = 1,2,..., \tag{2.30}$$

each member of which is a function of bounded variation because both $P(x)$ and $P_N(x)$ are nonnegative. The variation of $\psi_N(x)$ is given by

$$V[\psi_N(x)]_0^x = \int_0^x [P(t) + P_N(t)]dt \leqslant \int_0^1 [P(t) + P_N(t)]dt = 2. \tag{2.31}$$

Had we maintained arbitrary normalization for $P(x)$ and $P_N(x)$, the right side of (2.31) would read $2\mu_0$. In all cases, the right side of (2.31) is $N$-independent. Therefore the sequence (2.30) is of uniform bounded variation. Hence there exists a subsequence $\{\psi_{N_j}(x)\}$, and a function of bounded variation $\psi(x)$, such that

$$\lim_{j\to\infty}\psi_{N_j}(x) = \psi(x). \tag{2.32}$$

It follows from (2.28) and (2.30) that

$$\int_0^1 x^n\, d\psi_N(x) = 0, \quad n = 0,1,...,N, \tag{2.33}$$

which combines with (2.32) to yield

$$\int_0^1 x^n\, d\psi(x) = 0, \quad n = 0,1,...,\infty. \tag{2.34}$$

This and the uniqueness theorem for moment sequences associated with functions of bounded variation gives

$$\psi(x) = 0, \tag{2.35}$$

for every $x$ in [0,1]. Furthermore, every subsequence of (2.30) has in it a subsequence that converges to a function of bounded variation, and all convergent subsequences may be shown to converge to the same limit $\psi(x) = 0$ by iterating the uniqueness theorem. Therefore the original sequence (2.30) also converges and

$$\lim_{N\to\infty}\psi_N(x) = \lim_{N\to\infty}\int_0^x [P(t) - P_N(t)]dt = 0, \tag{2.36}$$

for every $x$ in [0,1]. This result implies that Eq. (2.29) holds for every continuous function $F(x)$.

The weak convergence established by Theorem 2 was obtained using only general properties of the maximum-entropy sequence, notably, the positivity of each approximant $P_N(x)$. In principle, it may prove possible to establish stronger forms of convergence by incorporating the finer details of the actual construction and by imposing further constraints on the original density $P(x)$. Rigorous results on the latter subject are not available at this point, so our subsequent discussion will be based on extensive empirical evidence. Similarly, the maximum-entropy solution for moment problems on an infinite interval has not been understood to any degree of mathematical rigor. To indicate some of the complications, we briefly consider here the two-moment solution for a semi-infinite interval, which is obtained by minimizing the potential

$$\Gamma = \ln Z + \mu_1\lambda_1 + \mu_2\lambda_2,$$
$$Z = \int_0^\infty dx\, e^{-\lambda_1 x - \lambda_2 x^2}. \tag{2.37}$$

The integral in (2.37) is expressed in terms of the usual error function and its asymptotic behavior may be studied explicitly. It turns out that a minimum exists only if

$$\mu_1^2 < \mu_2 < 2\mu_1^2, \quad \mu_0 = 1. \tag{2.38}$$

In practice, violation of the second inequality in (2.38) is the rule rather than the exception. Nevertheless numerical studies indicate that a solution always exists if the moments are incorporated in odd numbers. Explicit examples will be discussed later in the text. We would like to add here that conditions analogous to (2.38) were also discussed in a recent preprint[14] which was brought to our attention during the preparation of this work.

### A. Numerical procedure

The method for numerical computations is clearly suggested by the convexity of the potential $\Gamma$. One may use the familiar Newton minimization procedure. We briefly review the procedure in order to indicate possible differences between our numerical work and earlier calculations.[5] Starting with some initial choices for $\lambda_1,\lambda_2,...,\lambda_N$, updated $\lambda$'s are defined from

$$\lambda_n' = \lambda_n - \sum_{m=1}^N (H^{-1})_{nm}\frac{\partial\Gamma}{\partial\lambda_m} = \lambda_n - \sum_{m=1}^N (H^{-1})_{nm}[\mu_m - \langle x^m\rangle], \tag{2.39}$$

where $H = (H_{nm})$ is the Hessian matrix defined in Eq. (2.8) and $H^{-1}$ is its inverse. All quantities on the right side of (2.39) are calculated at the input values of $\lambda$. This entails the evaluation of $2N$ expected values of the form $\langle x^k\rangle, k = 1,2,..., 2N$, and the solution of a $N\times N$ system of linear equations. Write $\lambda_n' = \lambda_n - a_n$, so that the Newton shift $a_n$ satisfies the linear system

$$\sum_{m=1}^N H_{nm}a_m = \mu_n - \langle x^n\rangle, \quad n = 1,2,...,N. \tag{2.40}$$

Because $H$ is positive definite, the linear system (2.40) was solved with the standard IMSL routine LEQT2P (used here with double-precision arithmetic).

The demanded accuracy for the evaluation of the integrals involved in $\langle x^k\rangle$, see Eq. (2.5), was better than 12 significant figures. Two independent routines were used for the evaluation of the integrals thus obtaining an important

check of consistency. We first describe the calculation for the finite interval [0,1]. We found that four double-precision 24-point Gaussian quadratures evenly distributed over [0,1] very efficiently produce the demanded accuracy for $\langle x^k \rangle$. An independent check was performed with an adaptive Newton–Cotes integration algorithm. While the results of the Gaussian quadrature were always confirmed, the Newton–Cotes algorithm proved slower. However, there were two instances in which the Newton–Cotes algorithm proved indispensable. First, while the Gaussian quadrature is very accurate for the evaluation of integrals involving polynomials as well as exponentials of polynomials, its accuracy is impaired somewhat in the calculation of averages of the form $\langle F \rangle$, where $F = F(x)$ is some function containing square roots, integrable singularities, etc. Second, Gaussian quadratures proved inadequate for calculations on an infinite interval in view of the high degree of precision requested in our calculations. On the contrary, the adaptive Newton–Cotes algorithm could be adjusted to handle such complications, as is described in more detail in Sec. IV.

For finite-interval calculations, the updating procedure (2.39) may be initialized setting all $\lambda$ 's equal to zero, and using the previously computed $\lambda$ 's as input for all higher iterations. Two criteria were used to stop the procedure. First, the updated moments $\langle x^n \rangle, n = 1,2,...,N$ were requested to agree with the actual moments to an accuracy of one part in $10^{12}$ or better. Second, the updated values for $\lambda_1,\lambda_2,...,\lambda_N$ were stabilized to at least one part in $10^8$. It is an empirical fact that averages such as $\langle x^n \rangle$ were consistently reproduced with higher relative accuracy than the corresponding accuracy for individual $\lambda$ 's. All calculations were performed interactively on an IBM 360/70 system. Five to six Newton iterations were typically sufficient. Under the precision standards set above, we were able to handle $10 \sim 12$ moments without difficulty.

Incorporating higher moments introduces instabilities in the algorithm reflected in failures of the LEQT2P routine, apparently due to accumulation of roundoff error. For most practical calculations described in the following we did not find it necessary to go beyond 12 moments. We include, however, some comments concerning possible procedures to remedy the numerical difficulties with higher moments. An obvious possibility is to increase the precision requirements at the cost of considerably slowing down the algorithm. Other remedies are (i) initiate the algorithm with a simple gradient method, or with a gradient and/or Newton algorithm with step adjustment, and (ii) introduce a line-search routine of the type discussed in Ref. 5. We have in fact built in our procedure all of the above options. As far as calculations involving up to $10 \sim 12$ moments are concerned, those options are a hindrance rather than an improvement. Nevertheless, they may prove valuable for higher-moment calculations and for calculations over an infinite interval.

To conclude this discussion, we mention that the preceding algorithm may be trivially adjusted to calculate the minimum of the potential $\Delta$ introduced in Eq. (2.9). The Hessian $H_{nm}$ in Eq. (2.39) is replaced by $\Theta_{nm} = \langle x^{n+m} \rangle$ and the summation is extended to include $n,m = 0$. Our numerical results were thus reproduced with equal efficiency.

### B. Elementary examples

Consider the elementary density

$$P(x) = \alpha + 2(1 - \alpha)x, \tag{2.41}$$

where the parameter $\alpha$ varies in $0 \leqslant \alpha \leqslant 1$, so that $P(x)$ is positive in [0,1]. The corresponding moments are given by

$$\mu_n = \alpha/(n + 1) + 2(1 - \alpha)/(n + 2). \tag{2.42}$$

It is not clear whether or not the maximum-entropy sequence will produce a good approximation for (2.41), because one attempts to approximate a simple monomial by an exponential of a polynomial. For $\alpha = 1$, $P(x)$ reduces to a constant and the correct result is, of course, obtained after a single iteration. The results of a pointwise comparison with the true density for $\alpha = \frac{1}{2}$ and $\alpha = 0$ are summarized in Table I using eight moments as input. (The value of $N$, instead of $N + 1$, will hereafter be referred to as the number of input moments.)

It is evident that the pointwise fit is excellent for $\alpha = \frac{1}{2}$, while the accuracy diminishes for $\alpha = 0$ apparently because $P = 2x$ contains a zero. Much more stable is, however, the convergence of averages of the form (2.29) even for $P = 2x$. Choosing

$$F(x) = \tfrac{1}{2}\sqrt{x}, \tag{2.43}$$

the corresponding average $\langle F \rangle = U_0$ will be called the zero-temperature internal energy for reasons explained in Sec. III. The sequence of approximations to $U_0$ is given in Table II for $P = 2x$ and is shown to accurately reproduce the exact answer. Such behavior is in fact typical for densities $P(x)$ with sophisticated structure.

We have further studied the behavior of individual $\lambda$ 's. Typically, they alternate in sign while their absolute value increases substantially with increasing $N$. We found it instructive to rearrange the polynomial

$$-\ln P_N(x) = \sum_{n=0}^{N} \lambda_n(N)x^n \tag{2.44}$$

as a linear superposition of the Legendre polynomials defined in [0,1],

$$L_n(x) = \frac{(-1)^n}{n!}\frac{d^n}{dx^n}[x^n(1 - x)^n], \tag{2.45}$$

and write

$$\sum_{n=0}^{N} \lambda_n(N)x^n = \sum_{n=0}^{N} \nu_n(N)L_n(x). \tag{2.46}$$

It is not difficult to see that

$$\nu_n(N) = (2n + 1)\sum_{m=n}^{N} \frac{\Gamma^2(m + 1)}{\Gamma(m - n + 1)\Gamma(m + n + 2)}\lambda_m(N). \tag{2.47}$$

We found through numerical experimentation that the $\nu_n(N)$'s approach the ordinary Legendre coefficients of $-\ln P(x)$ when $N$ increases. It is, of course, incorrect to presume that Eq. (2.46) for finite $N$ is directly related to the Legendre expansion of $-\ln P(x)$.

The preceding comparison relies on the assumption

that both $P(x)$ and $-\ln P(x)$ are integrable in $[0,1]$. We were thus prompted to consider the density

$$P(x) = \mu_0^{-1} e^{-1/x},$$
$$\mu_0 = \int_0^1 dt\, e^{-1/t} = 0.148\,495\,51, \tag{2.48}$$

whose logarithm is not integrable. Nevertheless, the pointwise fit of the corresponding maximum-entropy solution presented in Table I, as well as the average $\langle \sqrt{x}/2 \rangle$ given in Table II, is in good agreement with the exact answer. This may not be surprising because functions with nonintegrable singularities, such as $-\ln P(x) \sim 1/x$, may also be approximated by polynomials closely related to the Bernstein polynomials.[15] The preceding remarks should, however, make one cautious in comparing the maximum-entropy sequence with standard polynomial expansions.

We conclude this section by mentioning an amusing calculation. Setting $\alpha = -1$ in Eqs. (2.41) and (2.42), the resulting distribution is not positive everywhere in $[0,1]$. The moments

$$\mu_n = -1/(n+1) + 4/(n+2) \tag{2.49}$$

are still positive and monotonically decreasing. However, the moment sequence (2.49) is not completely monotonic. For instance, the Hausdorff inequality (2.17) is violated for $n = 0$ and $k = 2$. According to Theorem 1, the potential $\Gamma$ cannot have a finite minimum in this case, a fact that was easily demonstrated numerically feeding (2.49) into our algorithm.

## III. DENSITY OF STATES IN HARMONIC SOLIDS

We now turn our attention to some more sophisticated moment problems that arise in physical applications. The first example will be the discussion of the density of states in a harmonic solid. Upon the assumption that anharmonic forces are relatively insignificant, the problem of calculating various quantities of physical interest reduces to the diagonalization of a matrix. In most cases, however, an analytical diagonalization is impossible and, even today, direct numerical procedures for 3-D lattices are difficult to implement with sufficient accuracy. On the contrary, indirect moment methods have long been known to provide very valuable substitutes for exact solutions.

TABLE I. Maximum-entropy solution using eight moments and comparison with the exact density for $P = x + \frac{1}{2}$, $P = 2x$, and $P = \mu_0^{-1} e^{-1/x}$.

| | $P = x + \frac{1}{2}$ | | $P = 2x$ | | $P = \mu_0^{-1} e^{-1/x}$ | |
|---|---|---|---|---|---|---|
| $x$ | Maxent | Exact | Maxent | Exact | Maxent | Exact |
| 0.0 | 0.500 003 5 | 0.5 | 0.0172 | 0.0 | $10^{-9}$ | 0.0000 |
| 0.1 | 0.600 000 6 | 0.6 | 0.2032 | 0.2 | 0.0004 | 0.0003 |
| 0.2 | 0.699 999 9 | 0.7 | 0.3998 | 0.4 | 0.0450 | 0.0454 |
| 0.3 | 0.799 999 3 | 0.8 | 0.5946 | 0.6 | 0.2409 | 0.2402 |
| 0.4 | 0.900 000 9 | 0.9 | 0.8062 | 0.8 | 0.5513 | 0.5528 |
| 0.5 | 1.000 000 0 | 1.0 | 1.0005 | 1.0 | 0.9131 | 0.9114 |
| 0.6 | 1.099 999 1 | 1.1 | 1.1934 | 1.2 | 1.2711 | 1.2719 |
| 0.7 | 1.200 000 7 | 1.2 | 1.4047 | 1.4 | 1.6130 | 1.6139 |
| 0.8 | 1.300 000 2 | 1.3 | 1.6015 | 1.6 | 1.9314 | 1.9294 |
| 0.9 | 1.399 999 3 | 1.4 | 1.7949 | 1.8 | 2.2143 | 2.2168 |
| 1.0 | 1.499 996 3 | 1.5 | 1.9736 | 2.0 | 2.4689 | 2.4774 |

TABLE II. Maximum-entropy sequence of approximations of the zero-temperature internal energy, for the model densities $P = 2x$ and $P = \mu_0^{-1} e^{-1/x}$, and for a harmonic fcc crystal.

| Number of moments | Internal energy $U_0$ $P = 2x$ | $P = \mu_0^{-1} e^{-1/x}$ | fcc crystal |
|---|---|---|---|
| 1 | 0.396 913 3 | 0.422 396 9 | 0.333 333 3 |
| 2 | 0.399 500 9 | 0.425 637 5 | 0.339 133 9 |
| 3 | 0.399 864 8 | 0.425 870 2 | 0.340 615 7 |
| 4 | 0.399 951 8 | 0.425 894 1 | 0.341 081 2 |
| 5 | 0.399 979 4 | 0.425 897 2 | 0.340 775 9 |
| 6 | 0.399 990 0 | 0.425 897 7 | 0.340 867 6 |
| 7 | 0.399 994 7 | 0.425 897 8 | 0.340 879 6 |
| 8 | 0.399 997 0 | 0.425 897 8 | 0.340 865 3 |
| 9 | 0.399 998 2 | | 0.340 884 5 |
| 10 | 0.399 998 8 | | 0.340 886 4 |
| Accurate | 0.400 000 0 | 0.425 897 8 | 0.340 887 2 |

The quantity of prime interest is the (positive) density of states $g = g(\omega)$, where $\omega$ is the frequency varying in a finite interval $[0,\omega_0]$. The maximum frequency $\omega_0$ is usually known and may be scaled out so that the interval becomes $[0,1]$. It is generally possible to explicitly determine a number of power moments of the form

$$\mu_n = \int_0^1 \omega^{2n} g(\omega) d\omega, \tag{3.1}$$

which involve computations of traces rather than eigenvalues of a given matrix. Only even moments appear, but the moment problem may be reduced to its standard form by introducing a new density $P = P(x)$ from

$$x = \omega^2, \quad g(\omega)d\omega = P(x)dx, \tag{3.2}$$

so that $g(\omega) = 2\omega P(\omega^2)$ and Eqs. (3.1) read

$$\mu_n = \int_0^1 x^n P(x) dx, \quad n = 0,1,2,\ldots . \tag{3.3}$$

If the density of states were known, various thermodynamic quantities could be computed as suitable averages with respect to $P(x)$. For instance, the internal energy and specific heat are given by

$$U = \tau \int_0^1 \left[ \frac{\sqrt{x}}{2\tau} \coth\left( \frac{\sqrt{x}}{2\tau} \right) \right] P(x) dx \tag{3.4}$$

and

$$C = \int_0^1 \left( \frac{\sqrt{x}/2\tau}{\sinh(\sqrt{x}/2\tau)} \right)^2 P(x) dx. \tag{3.5}$$

All units have been scaled out of (3.4) and (3.5) and $\tau$ stands for a suitably normalized temperature. It will prove useful to consider, in particular, a special case of (3.4), the zero-temperature limit of the internal energy:

$$U_0 = \frac{1}{2} \int_0^1 \sqrt{x} P(x) dx. \tag{3.6}$$

Notice that (3.4)–(3.6) are concrete examples of averages of the form considered in Eq. (2.29).

The specific problem we study here is that of a harmonic fcc crystal with nearest-neighbor interactions. This problem has attracted interest over a long period of time,[16] espe-

TABLE III. The $Z$ and $\lambda$ coefficients for the density of states in a harmonic fcc crystal (using ten moments), and for the Fourier transform of a two-point correlation function in the 1-D isotropic Heisenberg and $XY$ models (using five moments).

| | fcc crystal | Heisenberg | $XY$ model |
|---|---|---|---|
| $Z$ | 1.471 329 4 + 1 | 1.915 731 738 + 0 | 2.121 488 856 + 0 |
| $\lambda_1$ | − 6.470 649 9 + 1 | 5.610 630 136 − 1 | 7.624 279 265 − 1 |
| $\lambda_2$ | 9.945 635 3 + 2 | − 8.991 552 268 − 2 | − 2.871 904 750 − 1 |
| $\lambda_3$ | − 8.642 201 9 + 3 | 6.070 910 692 − 3 | 4.830 147 309 − 2 |
| $\lambda_4$ | 4.235 568 8 + 4 | − 1.619 301 504 − 4 | − 3.552 353 395 − 3 |
| $\lambda_5$ | − 1.247 859 5 + 5 | 1.507 717 407 − 6 | 9.382 829 849 − 5 |
| $\lambda_6$ | 2.287 607 1 + 5 | | |
| $\lambda_7$ | − 2.606 395 1 + 5 | | |
| $\lambda_8$ | 1.772 624 8 + 5 | | |
| $\lambda_9$ | − 6.474 692 4 + 4 | | |
| $\lambda_{10}$ | 9.504 657 1 + 3 | | |

cially after Isenberg[17] was able to explicitly compute a large number of moments (35 moments). A thorough analysis of the associated moment problem was later given by Wheeler and Gordon[10] using the Padé-like procedure outlined in the Introduction. We thought it appropriate to calculate the density of states and some thermodynamic averages by the maximum-entropy approach, so stringent comparisons can be made with independent powerful methods. The explicit values of the moments used here are known to high precision and may be found summarized in Ref. 10. Our usual convention $\mu_0 = 1$ is implicit in the following. Hence we consider the $N$-moment maximum-entropy approximation of the form

$$P_N(x) = \frac{1}{Z}\exp\Big(-\sum_{n=1}^{N}\lambda_n x^n\Big), \tag{3.7}$$

for various values of $N$. The calculated values for $Z,\lambda_1,\lambda_2,\ldots$ with $N = 10$ are given in Table III. It is then straightforward to plot the corresponding approximation to the density of states, as is done in Fig. 1. The general features of the emerging curve are consistent with the known van-Hove critical points [where $P(x)$ has infinite derivative] which are located at

$$x = \tfrac{1}{4},\tfrac{1}{2} \quad \text{and} \quad \tfrac{1}{4}(2+\sqrt{2}) = 0.853\ldots. \tag{3.8}$$

The obtained curve is also consistent with the general form of $P(x)$ anticipated by root-sampling techniques.[10]

How good is the approximation furnished by Fig. 1? Since the exact solution is unknown, an indirect assessment can be made by examining the sequence of successive approximations for $N = 1$–$10$. We found that the gross features of the curve given in Fig. 1 have stabilized, but changes are still underway in the neighborhood of the critical points (especially at $x = 0.853$) and in the neighborhood of the endpoints. It should be mentioned here that direct comparison with the calculation of Wheeler and Gordon cannot be made at this level because they approximate the density by a sum of $\delta$-functions. However, such comparison is possible for thermodynamic averages. Table II contains the maximum entropy sequence of approximations for the internal energy at zero temperature obtained by substituting the approximation (3.7) in Eq. (3.6). The result obtained here using ten moments is consistent with the upper and lower bounds derived by Wheeler and Gordon[10] using 30 moments:

$$U_0 = 0.340\ 88^{83}_{07}. \tag{3.9}$$

This is an encouraging fact especially because the authors of that reference note that using 12 moments in their procedure yields a result accurate only to one part in $10^3$. Some caution should be exercised, however, because it is not presently known whether incorporating higher moments in the maximum-entropy calculation will rapidly lead to significant improvements. We noted earlier that our numerical procedure is not yet capable of handling a large number of moments. From the practical point of view, it is interesting that we obtain good results for averages using a relatively small number of moments.

Further important tests are obtained by calculating the specific heat from Eq. (3.5) for various temperatures. It is evident that the coefficients of the high-temperature expansion of the specific heat may be directly expressed in terms of the moments of $P(x)$. On the other hand, the high-temperature expansion is known to diverge for $\tau < \tau_c = 1/2\pi = 0.159$. Hence the maximum-entropy sequence may be thought of as a method of summation of the high-temperature series outside its radius of convergence. It is to be expected that the same procedure will also yield a faster sequence of approximations in the region $\tau > \tau_c$. Our results are summarized in Table IV. A few iterations suffice to yield very accurate results for $\tau > \tau_c$. The sequence is also convergent in the region $\tau < \tau_c$, albeit at a slower rate. For very low temperatures the rate of convergence becomes poor because the integrand in Eq. (3.6) is concentrated over a small region around $x \simeq 0$. This situation could be remedied by incorporating the information from a short-frequency expansion of $P(x)$.[10]

To summarize, the maximum-entropy approach gives good results for thermodynamic averages even when a relatively small number of moments is included. An equally satisfactory pointwise fit of the density of states generally requires a larger number of moments. However, wild oscillations typical of polynomial expansions are less likely to occur in a maximum-entropy calculation. Some more demanding examples where gaps appear in the spectrum will be discussed in later sections.

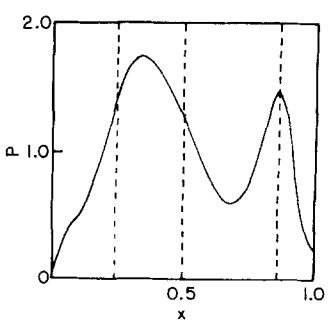



FIG. 1. Maximum-entropy calculation of the density of states for a harmonic fcc crystal using ten moments ($N = 10$). The vertical dashed lines indicate the location of the van-Hove critical points.

## IV. DYNAMIC CORRELATION FUNCTIONS IN QUANTUM SPIN SYSTEMS

Dynamic correlation functions provide the most direct tool for comparisons between theory and experiment in the study of quantum spin systems.[18] Nevertheless, a variety of

TABLE IV. Results for the specific heat in a harmonic fcc crystal at various temperatures.

| Number of moments | Specific Heat | | | |
|---|---|---|---|---|
| | $\tau = 0.05$ | 0.1 | 0.5 | 1.0 |
| 3 | 0.011 899 47 | 0.100 879 49 | 0.852 041 352 2 | 0.959 599 573 6 |
| 4 | 0.007 687 81 | 0.097 249 15 | 0.852 040 548 4 | 0.959 599 569 6 |
| 5 | 0.011 754 13 | 0.098 997 75 | 0.852 040 575 8 | 0.959 599 569 6 |
| 6 | 0.010 289 76 | 0.098 685 65 | 0.852 040 575 5 | 0.959 599 569 6 |
| 7 | 0.010 086 38 | 0.098 661 53 | 0.852 040 575 5 | 0.959 599 569 6 |
| 8 | 0.010 317 67 | 0.098 676 25 | 0.852 040 575 5 | 0.959 599 569 6 |
| 9 | 0.010 036 05 | 0.098 666 40 | 0.852 040 575 5 | 0.959 599 569 6 |
| 10 | 0.010 011 97 | 0.098 665 93 | 0.852 040 575 5 | 0.959 599 569 6 |

theoretical methods that are suitable for the calculation of a wide range of static properties (ground state, spectrum of elementary and collective excitations, thermodynamic averages) often prove inadequate for detailed predictions of dynamical properties. Even for 1-D Heisenberg models where powerful Bethe-ansatz techniques apply and yield exact results for static properties, the computation of time-dependent correlation functions has proved difficult. The best known exception is the $XY$ model for which the two-point longitudinal function is known for all temperatures.[19,20] Some extensions to more complicated cases have also become possible through the continuing effort of a number of authors.[19–25]

Needless to say, the preceding remarks apply also to various semiclassical methods whose limitations for calculations of dynamical properties were often emphasized in the literature.[26] On the other hand, it is generally agreed that indirect moment methods can provide a very valuable tool for the calculation of dynamic correlations. To illustrate the ideas, let us consider a simple 1-D anisotropic Heisenberg model described by the Hamiltonian

$$H = -\sum_{n,\delta} [J_\perp S_n^- S_{n+\delta}^+ + J_\parallel S_n^z S_{n+\delta}^z], \tag{4.1}$$

where standard notation has been used. Setting $J_\parallel = 0$ in (4.1) yields the $XY$ model, while the isotropic Heisenberg model corresponds to $J_\perp = J_\parallel \equiv J$. A typical correlation function is given by

$$\tfrac{1}{4} G_{n-m}(t) = \langle S_n^z(t) S_m^z(0) \rangle, \tag{4.2}$$

where $\langle \cdots \rangle$ denotes the usual thermodynamical average, which reduces to the vacuum-expectation value at zero temperature. Translation invariance is reflected in the notation of Eq. (4.2) and the factor of $\frac{1}{4}$ was introduced for future convenience.

The difficulties in the explicit computation of the average of Eq. (4.2) become apparent on writing

$$S_n^z(t) = e^{iHt} S_n^z e^{-iHt}, \tag{4.3}$$

where $S_n^z = S_n^z(0)$. However, the computation of the coefficients of a short-time expansion is feasible. An exact calculation of a modestly large number of coefficients is possible at infinite temperature, where the averages $\langle \cdots \rangle$ reduce to simple traces, and at zero temperature when the exact ground state is known.[27] For our illustrations, we will consider the infinite-temperature limit for which $6 (= 5 + 1)$ moments have been computed[28] in some important cases [for the spin-$\frac{1}{2}$ model described by the Hamiltonian (4.1), for instance]. The correlation function (4.2) is then symmetric under time reversal ($t \to -t$), so the short-time expansion contains only even terms:

$$G_{n-m}(t) = \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k)!} \mu_{n-m}^{(2k)} t^{2k}, \tag{4.4}$$

where the moments $\mu_{n-m}^{(2k)}$ are static averages (traces) of the form

$$\begin{aligned} \tfrac{1}{4}\mu_{n-m}^{(0)} &= \langle S_n^z S_m^z \rangle, \\ \tfrac{1}{4}\mu_{n-m}^{(2)} &= \langle [H,[H,S_n^z]] S_m^z \rangle, \end{aligned} \tag{4.5}$$

and so on.

A further restriction in our calculation will be to consider the correlation function for zero-space separation ($n - m = 0$). The Fourier transform defined from

$$G_0(t) = \int_0^\infty \cos(\omega t) P(\omega) d\omega \tag{4.6}$$

yields a nonnegative density $P = P(\omega)$ whose power moments

$$\int_0^\infty \omega^{2n} P(\omega) d\omega = \mu_0^{(2n)} \equiv \mu_n, \quad n = 0,1,2,\ldots \tag{4.7}$$

are equal to the perturbation coefficients appearing in (4.4) with $n - m = 0$. Explicit values for the first few moments in the spin-$\frac{1}{2}$ $XY$ model ($J_\perp = 0, J_\parallel \equiv 1$), and in the isotropic Heisenberg model ($J_\parallel = J_\perp \equiv 1$), may be extracted from the work of Morita[28] and are summarized in Table V of our paper. While $G_0(t)$ for the Heisenberg model is not known, an exact solution for the $XY$ model is available[20]:

$$G_0(t) = [J_0(2t)]^2, \tag{4.8}$$

where $J_0$ is the familiar Bessel function. The Fourier transform reads

$$P(\omega) = \begin{cases} (1/\pi^2) K'(\omega/4), & \text{for } \omega < 4, \\ 0, & \text{for } \omega > 4, \end{cases} \tag{4.9}$$

where $K'(k) = K[(1 - k^2)^{1/2}]$ and $K(k)$ is the complete elliptic integral of the first kind.

The preceding exact result for the $XY$ model will provide important tests for our approximate calculations. Our task is to reconstruct the density $P(\omega)$ by a maximum-en-

TABLE V. Moments of a two-point function for the 1-D spin-$\frac{1}{2}$ $XY$ and isotropic Heisenberg models at infinite temperature.

| | $XY$ model | Heisenberg |
|---|---|---|
| $\mu_0$ | 1 | 1 |
| $\mu_1$ | 4 | 4 |
| $\mu_2$ | 36 | 44 |
| $\mu_3$ | 400 | 652 |
| $\mu_4$ | 4 900 | 11 636 |
| $\mu_5$ | 63 504 | 242 816 |

tropy calculation using as input the moments of Table V. The calculation will be carried out for both the $XY$ and the isotropic Heisenberg model. We will further compute $G_0(t)$ from Eq. (4.6), which is an average of the general form considered earlier in Eq. (2.29). Hence the maximum-entropy calculation may be thought of as a resummation procedure for the perturbative series (4.4).

We would like to stress from the outset that the current examples furnish demanding tests of the efficiency of maximum-entropy techniques. As is suggested by the exact solution in the $XY$ model, Eq. (4.9), the Fourier transform $P(\omega)$ typically vanishes outside a certain region whose extent is not known. We are thus forced to perform a semi-infinite-interval calculation and eventually predict the support along with other details of $P(\omega)$. Furthermore, $P(\omega)$ often develops (integrable) singularities in the infrared region ($\omega \simeq 0$) whose precise nature is unknown.

By analogy with the calculation of Sec. III, one could reduce the moment problem (4.7) to a more standard form by defining a new density $Q = Q(x)$ from

$$P(\omega) = 2\omega Q(\omega^2), \quad \omega^2 = x. \tag{4.10}$$

Equations (4.7) then read

$$\int_0^\infty x^n Q(x)dx = \mu_n, \quad n = 0,1,2,..., \tag{4.11}$$

and the maximum-entropy solution is given by

$$Q_N(x) = \frac{1}{Z}\exp\left(-\sum_{n=1}^{N}\lambda_n x^n\right),$$
$$P_N(\omega) = \frac{2\omega}{Z}\exp\left(-\sum_{n=1}^{N}\lambda_n \omega^{2n}\right). \tag{4.12}$$

However, we take this opportunity to point out that the procedure is not unique. For instance, the maximum-entropy sequence obtained directly from (4.7) would read

$$P_N(\omega) = \frac{1}{Z}\exp\left(-\sum_{n=1}^{N}\lambda_n \omega^{2n}\right), \tag{4.13}$$

where $Z,\lambda_1,\lambda_2,...$ are, of course, different from those appearing in (4.12). This fact is obviated by the appearance of the factor $\omega$ in Eq. (4.12). Had we taken into account the general expectation that $P(\omega)$ diverge rather than vanish at $\omega = 0$, the option (4.13) would be the least-biased candidate. Nevertheless, additional information about the detailed behavior of $P(\omega)$ is often unavailable in practice. How serious then is the nonuniqueness of the maximum entropy approximation?

To obtain some empirical evidence, we studied the current problem using both (4.12) and (4.13). While the details of the corresponding pointwise approximations to $P(\omega)$ differ, the computation of Fourier averages of the form (4.6) proved much less sensitive. Such behavior is in fact typical with many examples we considered, and with more general types of nonuniqueness hinted at in Sec. V. Hence judicious choices of weights and other *a priori* information could lead to significant improvements in the pointwise approximation of the density, but are less likely to affect the behavior of averages.

We present here the explicit results obtained through Eq. (4.13) using five moments ($N = 5$) as input. The values for $Z$ and $\lambda_1,...,\lambda_5$ are given in Table II for both the $XY$ and the isotropic Heisenberg model. The corresponding approximations to the density $P(\omega)$ are plotted in Fig. 2 together with the exact result for the $XY$ model given by Eq. (4.9). It is observed that significant oscillations are still present which could be attributed to the low number of input moments and to the combined effect of the singularity at $\omega = 0$ and the appearance of a gap in the spectrum for $\omega > \omega_0$. However, the appearance of a gap at $\omega_0 \sim 4$ in the $XY$ model is indicated clearly in Fig. 2. Similarly, a maximum frequency in the region $5 \leqslant \omega_0 \leqslant 6$ is also indicated for the Heisenberg model in agreement with the earlier semiclassical estimates and finite-lattice calculations.[29]

Despite the limited success of the pointwise approximation of $P(\omega)$, a decisive improvement is obtained for the Fourier transform (4.9) over the earlier Gaussian fit of Morita.[28] Thus the maximum-entropy approximation (4.13) with $N = 5$ is inserted in Eq. (4.9) to produce the results for $G_0(t)$ depicted in Fig. 3. For the $XY$ model, we plot the exact solution (4.8) by a solid line and superimpose a number of dots calculated by maximum entropy. The agreement with the exact answer, even within the third period of oscillation of

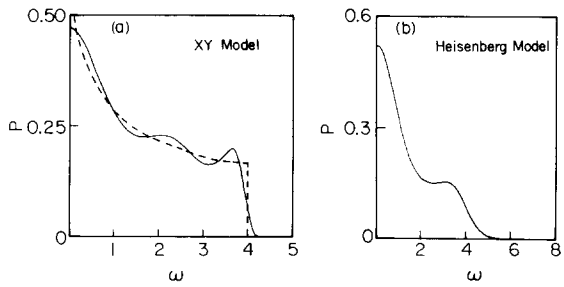



FIG. 2. A five-moment calculation ($N = 5$) of the Fourier transform of a two-point correlation function. (a) Results for the $XY$ model and comparison with the exact answer (dashed line). (b) Corresponding results for the isotropic Heisenberg model.

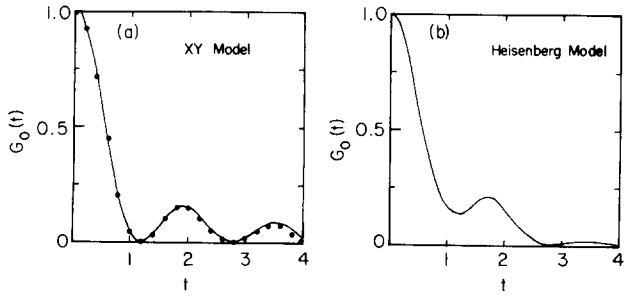



FIG. 3. A five-moment calculation ($N = 5$) of a two-point correlation function. (a) The exact result for the $XY$ model is depicted by a solid line while the dots correspond to the results from maximum entropy. (b) Maximum-entropy prediction for the isotropic Heisenberg model.

the Bessel function, is impressive. The actual numbers are summarized in Table VI together with the calculated numbers for the isotropic Heisenberg model for which an exact solution is not known. If such a solution is found in the future, we have no doubt that the result will be graphically indistinguishable from the plot given in Fig. 3(b). These results are in concert with our earlier remarks that maximum entropy consistently gives good results for averages, even when the pointwise approximation of the actual density is relatively poor.

We conclude this section with a few remarks. Generalizations of the calculation to the completely anisotropic Heisenberg model, to nonvanishing space separation ($n - m \neq 0$), to zero-temperature calculations, and to higher-dimensional lattices are relatively straightforward. A modest number of moments necessary for those generalizations have already appeared in the literature.[26–28] Independent knowledge of the maximum frequency $\omega_0$ would reduce the calculation to a moment problem over a finite interval and would improve the pointwise approximation for $P(\omega)$. Experimentation with the $XY$ model incorporating the fact that $\omega \leqslant \omega_0 = 4$ has indeed shown that the oscillations of $P(\omega)$ are reduced significantly but the results for $G_0(t)$ are essentially unchanged. In general $\omega_0$ is not known. However, assuming that the density does vanish for $\omega > \omega_0$, the limit

$$\lim_{n\to\infty} \mu_n^{1/n} = \Lambda \tag{4.14}$$

may be shown to be finite. The frequency $\omega_0$ is then given by $\omega_0 = \sqrt{\Lambda}$. Equation (4.14) may be used as the starting point for a numerical determination of $\omega_0$.

Our actual calculation was performed on a semi-infinite interval. It was thus necessary to use the adaptive Newton–Cotes integration algorithm mentioned in Sec. II setting an upper limit $\Omega$ which we could vary to as large a value as 50 000. For the problems discussed in this section, a modest value in the region $\Omega = 50 \sim 100$ was sufficient. Additional complications of the type indicated by Eq. (2.38) are absent if the moments are used in odd numbers.

## V. FURTHER EXAMPLES AND DISCUSSION

The explicit examples studied so far are but a few representatives of a host of moment problems that arise in practical applications. In this last section, we sketch some potential applications of the maximum-entropy approach, discuss possible generalizations, and summarize our conclusions.

(1) The Lee–Yang theory[30] of ferromagnetic Ising models provides a very elegant setting for the description of the mechanism by which phase transitions may occur. The same authors indicated that their procedure could also prove useful for practical calculations. In the interim, however, most of the calculational effort has gone into the renormalization-group approach addressing directly the critical region. Nevertheless, several authors have considered the moment problem naturally associated with the Lee–Yang theory and

TABLE VI. Detailed numerical results for the correlation function $G_0 = G_0(t)$ in the 1-D isotropic Heisenberg and $XY$ models (using five moments), and comparison with exact result $G_0 = [J_0(2t)]^2$ for the $XY$ model.

| | Correlation function $G_0(t)$ | | |
|---|---|---|---|
| $t$ | Heisenberg | $XY$ model | $[J_0(2t)]^2$ |
| 0.0 | 1.000 000 00 | 1.000 000 00 | 1.000 000 00 |
| 0.2 | 0.922 876 11 | 0.922 364 75 | 0.922 364 75 |
| 0.4 | 0.723 406 49 | 0.716 202 28 | 0.716 202 28 |
| 0.6 | 0.479 818 14 | 0.450 419 16 | 0.450 419 16 |
| 0.8 | 0.275 542 35 | 0.207 391 15 | 0.207 391 13 |
| 1.0 | 0.160 012 87 | 0.050 127 34 | 0.050 127 08 |
| 1.2 | 0.133 030 | 0.000 008 | 0.000 006 |
| 1.4 | 0.157 033 | 0.034 248 | 0.034 238 |
| 1.6 | 0.184 765 | 0.102 559 | 0.102 520 |
| 1.8 | 0.184 296 | 0.153 600 | 0.153 483 |
| 2.0 | 0.149 565 | 0.158 018 | 0.157 728 |
| 2.2 | 0.095 432 | 0.117 734 | 0.117 345 |
| 2.4 | 0.044 075 | 0.058 805 | 0.057 804 |
| 2.6 | 0.011 684 | 0.013 488 | 0.012 164 |
| 2.8 | 0.001 825 | 0.001 868 | 0.000 727 |
| 3.0 | 0.007 064 | 0.022 462 | 0.022 694 |
| 3.2 | 0.015 935 | 0.055 607 | 0.059 200 |
| 3.4 | 0.020 049 | 0.076 353 | 0.085 905 |
| 3.6 | 0.017 295 | 0.068 951 | 0.087 067 |
| 3.8 | 0.010 458 | 0.034 964 | 0.063 303 |
| 4.0 | 0.003 645 | − 0.008 826 | 0.029 464 |

there is little doubt that considerable progress in that direction is to be expected supplementing the renormalization-group approach.

For a ferromagnetic system described by the familiar Ising Hamiltonian

$$H = -J\sum \sigma_i\sigma_j - B\sum \sigma_i, \quad J\geqslant 0, \tag{5.1}$$

Lee and Yang showed that the free energy may be expressed as a dispersion integral:

$$F = -B - \frac{c}{2}J - \frac{1}{\beta}\int_0^\pi \ln(1 - 2z\cos\theta + z^2)g(\theta,x)d\theta, \tag{5.2}$$

where $c$ is the number of neighbors, $\beta$ is the inverse temperature, $x = e^{-2\beta J}$, and $z = e^{-2\beta B}$ is the activity. For temperatures below a critical value $T_c$, the support of the positive density $g(\theta,x)$ extends over the entire interval $0\leqslant\theta\leqslant\pi$. For $T > T_c$, $g(\theta,x)$ vanishes in a finite region $0\leqslant\theta\leqslant\theta_0(x)$ whose precise extent is not known. Calculational progress can be made noting that several moments of the form

$$\int_0^\pi [\cos(\theta/2)]^{2n}g(\theta,x)d\theta = \mu_n(x), \quad n = 0,1,2,\ldots \tag{5.3}$$

may be computed explicitly for all temperatures. The results of the work of several authors are summarized in Ref. 31 where the moments $\mu_n = \mu_n(x)$ were shown to be polynomials in $x$ with remarkable properties.

Our task is to approximately construct $g(\theta,x)$ from a finite number of moments. The maximum-entropy approximation associated with (5.3) reads

$$g_N(\theta,x) = \frac{1}{2Z}\exp\left\{-\sum_{n=1}^{N}\lambda_n\left[\cos\left(\frac{\theta}{2}\right)\right]^{2n}\right\}. \tag{5.4}$$

For purposes of illustration, we give in Fig. 4 the solution with $N = 7$ and $x = 0.5$ for the 1-D Ising model ($c = 2$) for which the moments (5.3) read

$$\mu_n(x) = \frac{1}{2^{2n+1}}\frac{\Gamma(2n+1)}{\Gamma^2(n+1)}(1 - x^2)^n. \tag{5.5}$$

The result is compared with the known exact answer

$$2\pi g(\theta,x) = \frac{\sin(\theta/2)}{[\sin^2(\theta/2) - x^2]^{1/2}}, \quad \sin\left(\frac{\theta}{2}\right)\geqslant x, \tag{5.6}$$

which exhibits a gap for $0\leqslant\theta < 2\arcsin(x)$ and a singularity at $\theta_0 = 2\arcsin(x)$. While a gap is clearly indicated by the maximum-entropy solution, significant oscillations are still present. However, thermodynamical averages calculated from (5.2) are obtained with much greater accuracy. As an example, consider the intensity of magnetization

$$I = -\frac{\partial F}{\partial B} = 2(1 - z^2)\int_0^\pi \frac{g(\theta,x)d\theta}{1 - 2z\cos\theta + z^2}, \tag{5.7}$$

whose exact form for the 1-D model reads

$$I = (1 - z)/[1 - 2z(1 - 2x^2) + z^2]^{1/2}. \tag{5.8}$$

Notice incidentally that the spontaneous magnetization $M = I(z\to 1^-)$ is given by $M = 2\pi g(0,x)$ in any dimension.

The exact result (5.8) is plotted in Fig. 5(a) for various temperatures together with results from the $N = 7$ maximum-entropy approximation. The agreement is excellent for essentially all values of the activity $z$ and the temperature variable $x$. To push the calculation further, we examine the slope of the magnetization $I(z)$ at vanishing magnetic field ($z = 1^-$) for various temperatures:

$$\chi = \left.\frac{\partial I}{\partial B}\right|_{B=0^+} = \beta\int_0^\pi \frac{g(\theta,x) - g(0,x)}{\sin^2(\theta/2)}d\theta. \tag{5.9}$$

This is but the magnetic susceptibility whose exact form in the 1-D model reads $\chi = \beta/x$. It is evident from Fig. 5(b) that departures from the exact answer occur for small temperatures, that is, temperatures near the critical point of the 1-D model.

One may conclude from the preceding elementary calculation that application of the maximum-entropy approach to higher-dimensional Ising models will provide useful information for finite magnetic fields, and for vanishing magnetic fields but temperatures away from the critical point. The extent to which the critical region may be reached depends on the number of input moments and other technical details which we will not discuss further in this paper.

(2) Implicit in the discussion of Secs. III and IV was the suggestion that the maximum-entropy approach may be used to sum divergent series. A typical example is a Stieljes integral of the form

$$F(z) = \int_0^\infty \frac{P(x)dx}{1 + zx}, \tag{5.10}$$

whose formal expansion in $z$ may be expressed in terms of the power moments of the density $P(x)$:

$$F(z) = \sum_{n=0}^{\infty}(-1)^n\mu_n z^n. \tag{5.11}$$

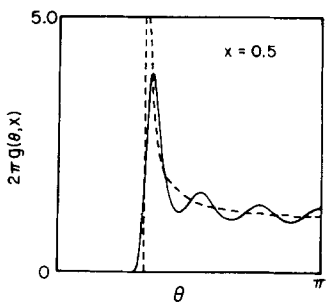



FIG. 4. Maximum-entropy prediction of the Lee–Yang density in the 1-D Ising model using seven moments ($N = 7$) and comparison with the exact answer (dashed line).

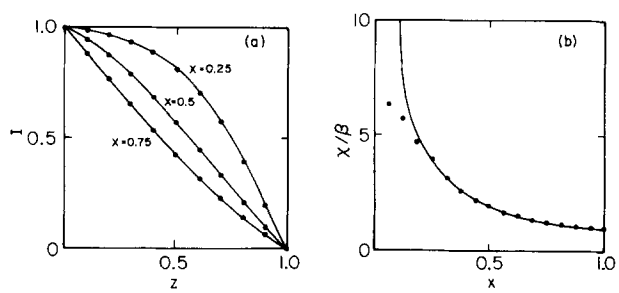



FIG. 5. Results from a $N = 7$ maximum-entropy calculation in the 1-D Ising model. (a) The solid lines correspond to the exact answer for the intensity of magnetization ($I$) as a function of the activity ($z$) at various temperatures ($x$). The dots are results from maximum entropy. (b) Similar results for the magnetic susceptibility.

In most practical applications the formal series (5.11) is divergent for all $z$. Nevertheless various methods exist for summing divergent asymptotic series, notably, Padé approximants.[8] Bender urged us to study the maximum-entropy sequence for series of the form (5.11) for which the asymptotic behavior of the moments for large $n$ reads[32]

$$\mu_n \sim (\nu n)! \tag{5.12}$$

Series with $\nu > 2$ are said to violate the Carleman criterion[8,13] which states that an essentially unique reconstruction of the measure $P(x)dx = d\mu(x)$ is possible from its moments if $\nu < 2$. While for $\nu > 2$ the moments do not uniquely determine $\mu(x)$, violation of the Carleman condition does not necessarily imply nonuniqueness for the average (5.10). It is also known that the diagonal and off-diagonal Padé sequences converge but their limits may not coincide with each other and with the true average $F(z)$. Bender further provided us with 20 moments for a nontrivial example with $\nu = 3$ (the ground-state energy of an octic oscillator), and with the associated Padé sequences which, indeed, stabilize away from the true answer, the latter being known from independent numerical calculation.[33]

In contrast, the first few iterations of the maximum-entropy approach showed substantial improvement over the Padé sequence, which indicates that the maximum-entropy sequence of approximations for averages of the form (5.10) might converge to the true answer even though $\nu > 2$. Because of the tremendous growth of successive moments in the current problem, we have not yet been able to incorporate a reasonably large number of moments in our numerical procedure. We thus postpone further discussion for a future occasion.

(3) We now return to the question of the inherent nonuniqueness of the maximum-entropy procedure mentioned in Sec. IV. A possible generalization may be obtained by defining a new density $Q(x)$ from

$$P(x) = w(x)Q(x), \tag{5.13}$$

where $w(x)$ is a known positive weight whose specific form is dictated by some *a priori* knowledge of detailed properties of $P(x)$. The usual power moments of $P(x)$ are then interpreted as weighted moments of the density $Q(x)$:

$$\int_a^b x^n Q(x) w(x) dx = \mu_n. \tag{5.14}$$

The entropy functional (1.4) is replaced by

$$S = -\int_a^b [Q(x)\ln Q(x) - Q(x)]dx + \sum_{n=0}^{N} \lambda_n \left( \int_a^b x^n Q(x) w(x) dx - \mu_n \right), \tag{5.15}$$

whose extrema are of the form

$$Q(x) = \exp\left( -w(x) \sum_{n=0}^{N} \lambda_n x^n \right), \tag{5.16}$$

and the original density reads

$$P(x) = w(x)\exp\left( -w(x) \sum_{n=0}^{N} \lambda_n x^n \right). \tag{5.17}$$

In order to construct a potential whose minimum determines $\lambda_0, \lambda_1, \ldots, \lambda_N$, one should treat all Lagrange multipliers, including $\lambda_0$, on a common basis. We thus seek generalization of the potential $\Delta$ defined in Eq. (2.9) rather than the potential $\Gamma$ of Eq. (2.6). The correct generalization reads

$$\Delta = \int_a^b \left[ \exp\left( -w(x) \sum_{n=0}^{N} \lambda_n x^n \right) - 1 \right] dx + \sum_{n=0}^{N} \mu_n \lambda_n. \tag{5.18}$$

Notice that the (positive-definite) Hessian of the above potential

$$\Theta_{nm} = \int_a^b dx\, x^{n+m} w^2(x) \exp\left( -w(x) \sum_{n=0}^{N} \lambda_n x^n \right) \tag{5.19}$$

is expressed in terms of moments that are weighted by $w^2(x)$ rather than $w(x)$.

Judicious choices of the weight $w(x)$ may lead to improvements in the pointwise approximation for $P(x)$. For instance, knowledge of the asymptotic behavior of the moments as in Eq. (5.12) may be used to determine $w(x)$ so that $P(x) \sim w(x)$ at large distances. However, our experience shows that averages of the form (2.29) are substantially stable against variations of $w(x)$.

(4) We believe to have presented ample evidence for the potential as well as the limitations of the maximum-entropy approach. It is important to keep in mind that the maximum-entropy approach, just as any other approximation procedure, should not be looked upon as a panacea for the solution of all moment problems that may arise in practice. After all, a polynomial expansion would be ideal if the true density happened to be a finite polynomial, the Padé-like procedure outlined in the Introduction would be ideal if the density were a finite sum of $\delta$-functions, maximum entropy would be ideal if the density were the exponential of a finite polynomial, and so on.

The merits of the current approach should be searched for in the larger context of the frequency of successful applications to a wide disparity of actual problems. While the limited ensemble of problems treated in this paper may not qualify for a genuine random sample, it nevertheless suggests the following appealing features for the maximum-entropy approach.

(i) Accurate averages are obtained, even when a low number of moments are available, and are substantially stable against variations in the specific mode of calculation. For instance, incorporation of suitable weights and/or other detailed information about the density (the appearance of singularities, the actual support, and so on) typically does not improve or diminish the accuracy of averages.

(ii) However, the method is flexible enough to incorporate such additional information which may result in significant improvements of the pointwise approximation of the density.

(iii) Maximum-entropy results compare favorably with corresponding results from independent powerful methods. It is fair to mention, however, that the Padé procedure often has the advantage of furnishing rigorous upper and lower bounds for the approximated averages. To the extent that the maximum-entropy technique has been developed, an accurate estimate of the error is not possible at this point.

## ACKNOWLEDGMENTS

We are grateful to E. T. Jaynes for his direct as well as indirect influence, to C. M. Bender for some suggestions mentioned in the text, and to S. H. Margolis for providing us with the Newton–Cotes routine. We also thank M. Baernstein, L. Benofy, M. Hughes, E. Shpiz, and G. Tiktopoulos for very useful discussions.

This work was supported in part by the U. S. Department of Energy.

[1]C. Shannon, Bell. Syst. Tech. J. **27**, 379, 623 (1948).
[2]E. T. Jaynes, Phys. Rev. **106**, 620 (1957); **108**, 171 (1957).
[3]*The Maximum Entropy Formalism*, edited by R. D. Levine and M. Tribus (MIT, Cambridge, MA, 1979).
[4]E. T. Jaynes, Proc. IEEE **70**, 939 (1982).
[5]N. Agmon, Y. Alhassid, and R. D. Levine, J. Comp. Phys. **30**, 250 (1979).
[6]R. D. Levine, J. Phys. A: Math. Gen. **13**, 91 (1980).
[7]*Theory and Applications of Moment Problems in Many-Fermion Systems*, edited by B. J. Dalton, S. M. Grimes, J. P. Vary, and S. A. Williams (Plenum, New York, 1979).
[8]G. A. Baker and P. Graves-Morris, *Padé Approximants* (Addison–Wesley, Reading, MA, 1980).
[9]R. G. Gordon, J. Math. Phys. **9**, 655 (1968).
[10]J. G. Wheeler and R. G. Gordon, J. Chem. Phys. **51**, 5566 (1969).
[11]R. R. Whitehead, in Ref. 7, p. 235.
[12]D. V. Widder, *The Laplace Transform* (Princeton U. P., Princeton, 1946).
[13]C. M. Bender and S. A. Orszag, *Advanced Mathematical Methods for Scientists and Engineers* (McGraw–Hill, New York, 1978).
[14]Z. Gburski, C. G. Gray, and D. E. Sullivan, "Information Theory of Line-Shape in Collision-Induced Absorption," University of Guelph, Ontario, preprint, 1983.
[15]G. G. Lorentz, *Bernstein Polynomials* (Toronto U. P., Toronto, 1953).
[16]R. B. Leighton, Rev. Mod. Phys. **20**, 165 (1948).
[17]C. Isenberg, Phys. Rev. **132**, 2427 (1963); **150**, 712 (1966).
[18]M. Steiner, J. Villain, and C. G. Windsor, Adv. Phys. **25**, 87 (1976).
[19]Th. Niemeijer, Physica (Utrecht) **36**, 377 (1967); **39**, 313 (1968).
[20]S. Katsura, T. Horiguchi, and M. Suzuki, Physica (Utrecht) **46**, 67 (1970).
[21]E. Barouch and B. M. McCoy, Phys. Rev. A **3**, 2137 (1971).
[22]H. W. Capel, E. J. Van Dongen, and Th. J. Sishens, Physica (Utrecht) **76**, 445 (1974).
[23]A. Sur, D. Jasnow, and I. J. Love, Phys. Rev. B **12**, 3845 (1975).
[24]B. M. McCoy, J. H. H. Perk, and R. E. Shrock, Nucl. Phys. B **220**, 269 (1983).
[25]G. Müller and R. E. Shrock, "Dynamic Correlation Functions in Quantum Spin Chains," Stony Brook preprint, 1983.
[26]J. H. Taylor and G. Müller, "On the Limitations of Spin-Wave Theory in $T = 0$ Spin Dynamics," Stony Brook preprint, 1983.
[27]G. Müller, Phys. Rev. B **26**, 1311 (1982).
[28]T. Morita, J. Math. Phys. **12**, 2062 (1971); **13**, 714 (1972).
[29]F. Carboni and P. M. Richards, Phys. Rev. **177**, 889 (1969).
[30]T. D. Lee and C. N. Yang, Phys. Rev. **87**, 410 (1952).
[31]J. D. Bessis, J. M. Drouffe, and P. Moussa, J. Phys. A: Math. Gen. **9**, 2105 (1976).
[32]C. M. Bender and T. T. Wu, Phys. Rev. Lett. **27**, 461 (1971).
[33]F. T. Hioe, D. MacMillan, and E. W. Montroll, J. Math. Phys. **17**, 1320 (1976).

# The connection between variational principles in Eulerian and Lagrangian descriptions

F. Bampi
*Institute of Engineering Mathematics, Pl. Kennedy, 16129 Genoa, Italy*

A. Morro
*Department of Biophysical and Electronic Engineering, Viale Causa 13, 16145 Genoa, Italy*

(Received 7 June 1983; accepted for publication 2 December 1983)

The question of whether there exists a connection between variational principles in Eulerian and Lagrangian descriptions is investigated. By having recourse to a proper view of the Eulerian description it is shown that a variational principle in one description holds whenever a corresponding variational principle in the other description is given. This theoretical conclusion is operative in that a precise rule for writing the new Lagrangian is exhibited. As an application, a new Lagrangian for fluid dynamics in the Eulerian description is determined.

PACS numbers: 02.70. + d, 03.40.Gc

## 1. INTRODUCTION

Often a physical system consists of a set of fields some of which are naturally accounted for in the spatial or Eulerian ($E$) description while the other ones are more easily expressed in the material or Lagrangian ($L$) description. In fact we are accustomed to field theories involving the $E$-description, the fields under consideration being viewed as functions of space coordinates $\mathbf{r}$ and time $t$. Usually, however, continuum mechanics deals with the $L$-description; typically, the motion is represented by the function $\mathbf{r} = \mathbf{x}(\mathbf{R},t)$ providing the position $\mathbf{r}$ of any particle at time $t$ in terms of the position $\mathbf{R}$ it occupies at a reference time $t_0$. Furthermore, there are systems whose Lagrangian density is known when they are represented through the $L$-description while the Lagrangian density of other systems is known when the $E$-description is involved.

The choice of a description is often a matter of convenience. In many cases, however, the two descriptions occur simultaneously as it happens, for example,in the variational formulation of the electromagnetic field interacting with matter.[1] Moreover, as Lax and Nelson[1] do, one could think that, apart from genuine interaction terms, the total Lagrangian density of a system is a proper sum of Lagrangian densities pertaining to subsystems which are expressed in different descriptions (Eulerian or Lagrangian). This procedure raises a conceptual problem: does the sum of free Lagrangian densities result in the appearance of (unphysical) interaction terms just because of the reduction to a common description? In fact, the essence of the problem consists in finding the connection between Eulerian and Lagrangian variational principles. Of course, the connection must render the Euler–Lagrange equations with respect to the $E$-variables equivalent to the Euler–Lagrange equations with respect to the $L$-variables.

It is the aim of this paper to set up such a connection. In this regard we observe that, as well-known, noticeable difficulties hinder finding variational principles in the $E$-description. Accordingly, we first revisit the $E$-description and we conclude that the essential features are preserved if the velocity is replaced with the motion as an unknown function (Sec. 2). Then, upon looking at the motion as an unknown function, we are able to arrive at the main result of this paper: there is a simple relation between the Lagrangian (density) in the $E$-description and that in the $L$-description and, moreover, such a relation allows a variational principle in one description to hold whenever a variational principle in the other description is given (Sec. 3). The operative character of this result is illustrated (Sec. 4) in conjunction with some variational principles for fluid dynamics which is a remarkable field of application of the connection between variational principles in Eulerian and Lagrangian descriptions.

## 2. THE EULERIAN DESCRIPTION REVISITED

When dealing with continuum physics, the $E$-description consists in representing the behavior of the body at hand through the velocity field $\mathbf{v}(\mathbf{r},t)$, besides other suitable fields, as a function of the position $\mathbf{r}$ and time $t$, without consideration of which particle currently occupies the position $\mathbf{r}$. It is well known that there are conceptual difficulties in setting up a variational principle in the $E$-description, this problem having attracted considerable attention especially in conjunction with fluids; we refer the interested reader to Refs. 2 and 3 for some details on this topic. Here we content ourselves with observing that the main difficulty originates from the velocity $\mathbf{v}$ being an unknown function.

Within the approach pertaining to the inverse problem of the calculus of variations,[3–6] the fact that the velocity, instead of the motion, is an unknown function changes drastically the mathematical structure of the balance equations, with respect to the counterpart in the $L$-description, and ascribes a central role to odd-order derivatives. On the one hand, this gives reason for the difficulty to find variational principles in the $E$-description. On the other, it suggests that decisive improvements could be attained by replacing the field $\mathbf{v}$ with another unknown field.

To our mind, the essential feature of the $E$-description is the dependence of the fields under consideration on the spatial position $\mathbf{r}$ and time $t$, not the fact that just the velocity is an unknown function. Accordingly, we attempt to select the most suitable unknown function by relying on a duality criterion with respect to the $L$-description.

Assume that, for each time $t$, the motion $\mathbf{x}$ is a smooth

 0022-2488/84/082418-04$02.50  

homeomorphism and let $\mathbf{X} = \mathbf{x}^{-1}$, i.e., $\mathbf{X}(\mathbf{r},t) = \mathbf{R}$. In the $L$-description the unknown is the motion $\mathbf{x}(\mathbf{R},t)$. Motivated by the duality criterion we take the view that, in the $E$-description, $\mathbf{X}(\mathbf{r},t)$ instead of $\mathbf{v}(\mathbf{r},t)$ is the unknown. To make this view operative we proceed as follows.

We use throughout Cartesian tensor notation; on adhering to the current literature[7,8] we denote the components of $\mathbf{r},\mathbf{x},\mathbf{v}$ through lower case indices and the components of $\mathbf{R},\mathbf{X}$ through capital indices. A comma followed by an index denotes partial differentiation with respect to a coordinate, a superposed dot the material time derivative. So $\dot{x}_a, a = 1,2,3$, are the components of the velocity and $x_{a,Q}, a,Q = 1,2,3$, are the components of the deformation gradient. To obtain the expression of $\mathbf{v}$ in terms of $\mathbf{X}(\mathbf{r},t)$ we start from the identity

$$\mathbf{r} = \mathbf{x}(\mathbf{X}(\mathbf{r},t),\mathbf{t});$$

differentiation with respect to $t$ while $\mathbf{r}$ is held constant gives

$$\frac{\partial x_a}{\partial R_Q}\frac{\partial X_Q}{\partial t} + \frac{\partial x_a}{\partial t} = 0,$$

whence

$$v_a = -x_{a,Q}X_{Q,t}. \tag{2.1}$$

Of course $x_{a,Q}$ must be viewed as a function of $X_{Q,b}$; in fact $x_{a,Q}(X_{P,b})$ follows from $x_{a,Q}$ being the inverse matrix of $X_{P,b}$, namely

$$x_{a,Q}X_{Q,b} = \delta_{ab}. \tag{2.2}$$

As a consequence, letting $\alpha$ stand for $a = 1,2,3$ or $t$, we have

$$(x_{a,P})_{,\alpha} = -x_{a,Q}X_{Q,b\alpha}x_{b,P}. \tag{2.3}$$

Before investigating the conceptual consequences of the new viewpoint we observe that the standard $E$-description is usually simpler because viewing $\mathbf{v}$ as unknown lowers the order of the differential equations; physically this means that we content ourselves with determining the velocity field rather than the motion. Nevertheless, within the variational framework the standard description results in formidable difficulties. As we shall see shortly, instead, our viewpoint, based on a strict duality between the Lagrangian and the Eulerian descriptions, make the variational formulation simpler and then the sought connection immediate.

## 3. THE CONNECTION

With each material system we associate a spatial Lagrangian (density) $L^{\mathrm{s}}$ or a material Lagrangian $L^{\mathrm{m}}$ according as we describe the system through the spatial or material description. Specifically, consider a material system interacting with a set of fields $\phi_\sigma$, $\sigma = 1,\ldots,n$; the Langrangians $L^{\mathrm{s}}$ and $L^{\mathrm{m}}$ are given as

$$L^{\mathrm{s}} = L^{\mathrm{s}}(\phi_\sigma,\mathbf{X}), \tag{3.1}$$

$$L^{\mathrm{m}} = L^{\mathrm{m}}(\phi_\sigma(\mathbf{x}),\mathbf{x}); \tag{3.2}$$

the relations (3.1) and (3.2) mean that the unknown functions are $\phi_\sigma(\mathbf{r},t),\mathbf{X}(\mathbf{r},t)$ and $\phi_\sigma(\mathbf{x}(\mathbf{R},t),t)$, $\mathbf{x}(\mathbf{R},t)$, respectively, while, to save writing, the dependence on the derivatives of the unknown functions is unspecified and understood.

In connection with $L^{\mathrm{s}}$ and $L^{\mathrm{m}}$ we define the actions $A^{\mathrm{s}}$ and $A^{\mathrm{m}}$ as follows. Let $\mathscr{V}_0$ be a three-dimensional domain occupied by the system at a reference time $t_0$; for any time $t > t_0$ the domain $\mathscr{V}_0$ is mapped into the spatial domain $\mathscr{V}(t)$ by the smooth homeomorphism $\mathbf{r} = \mathbf{x}(\mathbf{R},t)$. On letting $J = \det(x_{a,Q})$, the volume elements $d\nu$ and $dV$, of $\mathscr{V}$ and $\mathscr{V}_0$, are related by

$$d\nu = J\,dV. \tag{3.3}$$

Accordingly, for any two given times $t_2,t_1$ $(t_2 > t_1)$ we set

$$A^{\mathrm{s}} = \int_{t_1}^{t_2}\int_{\mathscr{V}} L^{\mathrm{s}}(\phi_\sigma,\mathbf{X})\,d\nu\,dt, \tag{3.4}$$

$$A^{\mathrm{m}} = \int_{t_1}^{t_2}\int_{\mathscr{V}_0} L^{\mathrm{m}}(\phi_\sigma(\mathbf{x}),\mathbf{x})\,dV\,dt. \tag{3.5}$$

We are now in a position to establish our main result.

**Theorem**: Look at the actions (3.4) and (3.5) and let $\delta L^{\mathrm{s}}/\delta X_Q, \delta L^{\mathrm{s}}/\delta\phi_\sigma, \delta L^{\mathrm{m}}/\delta x_a$, and $\delta L^{\mathrm{m}}/\delta\phi_\sigma$ denote variational derivatives. If

$$A^{\mathrm{s}} = A^{\mathrm{m}} \tag{3.6}$$

holds for any pair of functions $\mathbf{x},\mathbf{X}$ such that

$$\mathbf{X}(\mathbf{x}(\mathbf{R},t),\mathbf{t}) = \mathbf{R}, \quad \mathbf{R}\in\mathscr{V}_0, \tag{3.7}$$

then

$$\frac{\delta L^{\mathrm{m}}}{\delta\phi_\sigma} = J\frac{\delta L^{\mathrm{s}}}{\delta\phi_\sigma}, \quad \sigma = 1,\ldots,n, \tag{3.8}$$

$$\frac{\delta L^{\mathrm{m}}}{\delta x_a} + \frac{\delta L^{\mathrm{m}}}{\delta\phi_\sigma}\phi_{\sigma,a} = -JX_{Q,a}\frac{\delta L^{\mathrm{s}}}{\delta X_Q}. \tag{3.9}$$

*Proof*: For the sake of simplicity we subdivide the proof in two steps. Consider first the change of functions

$$\phi_\sigma \to \phi_\sigma + \eta_\sigma \tag{3.10}$$

while the functions $\mathbf{x}$ and $\mathbf{X}$ are held fixed. The (Gâteaux) differentials of $A^{\mathrm{s}}$ and $A^{\mathrm{m}}$ under (3.10) are given by

$$\begin{aligned} dA^{\mathrm{s}}(\phi_\sigma|\eta_\sigma) &= \int_{t_1}^{t_2}\int_{\mathscr{V}}\frac{\delta L^{\mathrm{s}}}{\delta\phi_\sigma}\eta_\sigma\,d\nu\,dt + \text{b.t.},\\ dA^{\mathrm{m}}(\phi_\sigma|\eta_\sigma) &= \int_{t_1}^{t_2}\int_{\mathscr{V}_0}\frac{\delta L^{\mathrm{m}}}{\delta\phi_\sigma}\eta_\sigma\,dV\,dt + \text{b.t.}, \end{aligned} \tag{3.11}$$

where b.t. stands for inessential boundary terms. Expressing (3.11) in terms of the variables $\mathbf{R}$ (so that $\mathscr{V}$ is mapped onto $\mathscr{V}_0$) and use of (3.3) yield

$$dA^{\mathrm{s}}(\phi_\sigma|\eta_\sigma) = \int_{t_1}^{t_2}\int_{\mathscr{V}_0}\frac{\delta L^{\mathrm{s}}}{\delta\phi_\sigma}\eta_\sigma J\,dV\,dt + \text{b.t.}$$

Of course (3.6) implies that $dA^{\mathrm{s}}(\phi_\sigma|\eta_\sigma) = dA^{\mathrm{m}}(\phi_\sigma|\eta_\sigma)$; hence, owing to the arbitrariness of $\eta_\sigma$ we arrive at (3.8).

As the second step consider the change of functions $\mathbf{X}\to\mathbf{X}+\mathbf{H}$ and $\mathbf{x}\to\mathbf{x}+\mathbf{h}$ while the fields $\phi_\sigma(\mathbf{r},t)$ are unchanged. Owing to (3.7), $\mathbf{H}$ and $\mathbf{h}$ must be related by

$$(\mathbf{X}+\mathbf{H})((\mathbf{x}+\mathbf{h})(\mathbf{R},t),t) = \mathbf{R}, \quad \mathbf{R}\in\mathscr{V}_0; \tag{3.12}$$

on keeping only first-order terms with respect to $\mathbf{h}$ and $\mathbf{H}$, (3.12) provides

$$H_Q(\mathbf{x}(\mathbf{R},t),t) = -(X_{Q,a}h_a)(\mathbf{R},t). \tag{3.13}$$

With this in mind, observe first that, in view of (3.1),

$$dA^{\mathrm{s}}(\mathbf{X}|\mathbf{H}) = \int_{t_1}^{t_2}\int_{\mathscr{V}}\frac{\delta L^{\mathrm{s}}}{\delta X_Q}H_Q\,d\nu\,dt + \text{b.t.}$$

According to (3.2), under the change $\mathbf{x}\to\mathbf{x}+\mathbf{h}$, the action functional $A^{\mathrm{m}}$ is influenced by the explicit dependence on $\mathbf{x}$

and, moreover, by the dependence on $\mathbf{x}$ through the fields $\phi_\sigma$. So (3.10) still applies with

$$\eta_\sigma = \phi_{\sigma,a} h_a + O(|\mathbf{h}|).$$

In conclusion we obtain

$$dA^{\mathrm{m}}(\mathbf{x}|\mathbf{h}) = \int_{t_1}^{t_2} \int_{\mathscr{V}_0} \left( \frac{\delta L^{\mathrm{m}}}{\delta x_a} + \frac{\delta L^{\mathrm{m}}}{\delta \phi_\sigma} \phi_{\sigma,a} \right) h_a \, dV \, dt + \text{b.t.} \tag{3.14}$$

Now, because of (3.13), it follows that

$$dA^{\mathrm{s}}(\mathbf{x}|\mathbf{h}) = -\int_{t_1}^{t_2} \int_{\mathscr{V}_0} \frac{\delta L^{\mathrm{s}}}{\delta X_Q} J X_{Q,a} h_a \, dV \, dt + \text{b.t.}$$

Then, following along the previous procedure, the condition $dA^{\mathrm{m}}(\mathbf{x}|\mathbf{h}) = dA^{\mathrm{s}}(\mathbf{x}|\mathbf{h})$ leads immediately to (3.9). □

On the basis of this theorem we can assert that whenever a variational principle is given, both in the Eulerian and in the Lagrangian description, and (3.6) is verified, then the corresponding Euler–Lagrange equations are mapped into each other according to the homeomorphism $\mathbf{r} = \mathbf{x}(\mathbf{R},t)$ or $\mathbf{R} = \mathbf{X}(\mathbf{r},t)$. So we have the noticeable result that a variational principle in one description delivers, via (3.6), a variational principle in the other description. These observations lend operative meaning to the following connection between variational principles in Eulerian and Lagrangian descriptions.

**Action axiom**: The spatial action $A^{\mathrm{s}}$ and the material action $A^{\mathrm{m}}$ of a physical system are equal.

As an immediate consequence, this axiom implies that, owing to (3.4) and (3.5), the Lagrangians $L^{\mathrm{s}}$ and $L^{\mathrm{m}}$ are related by

$$L^{\mathrm{s}} = j L^{\mathrm{m}}, \tag{3.15}$$

where $j = J^{-1}$. Of course the equality (3.15), as well as (3.8) and (3.9), holds provided that a common coordinate system is employed. For example, if $L^{\mathrm{m}} = L^{\mathrm{m}}(\dot{x}_a)$ then

$$L^{\mathrm{s}}(X_{Q,a}, X_{Q,t}) = j(X_{Q,a}) L^{\mathrm{m}}(-x_{a,Q}(X_{R,b})).$$

*Remark*: According to the previous theorem the relations (3.8) and (3.9) are necessary for the validity of (3.6) under (3.7). The addition of suitable conditions, relating the boundary terms b.t. in the two descriptions, make (3.8) and (3.9) also sufficient. However, we are not interested at all in the sufficiency (and in the boundary terms) for two reasons. First, it is the necessary condition of the theorem that is important for arriving at the action axiom and for rendering it operative. Second, as to the applications we have in mind, concerning the inverse problem of the calculus of variations,[3] the boundary terms are inessential for determining the Lagrangian of a given physical (or mathematical) system.

## 4. NEW VARIATIONAL PRINCIPLES FROM KNOWN ONES

The physical relevance of the action axiom consists in the feature that a variational principle in one description is obtained as soon as a variational principle in the other description is given. This feature, true for any material system, becomes particularly outstanding within the framework of fluid dynamics. Accordingly, so as to emphasize the operative meaning of the action axiom, we investigate some examples of variational principles in fluid dynamics.

In essence, in the $L$-description we are given the following variational principles: the behavior of a fluid is governed by any of the three Lagrangians[3]

$$L^{\mathrm{m}}(\mathbf{x}) = \tfrac{1}{2} \rho_0 \dot{x}_a \dot{x}_a - \rho_0 E(\rho_0 j, S_0), \tag{4.1}$$

$$L^{\mathrm{m}}(\mathbf{x},\rho) = \tfrac{1}{2} \rho_0 \dot{x}_a \dot{x}_a - \rho_0 E(\rho, S_0) + (J - \rho_0/\rho) p(\rho, S_0), \tag{4.2}$$

$$L^{\mathrm{m}}(\mathbf{x},\rho,\mu) = \tfrac{1}{2} \rho_0 \dot{x}_a \dot{x}_a - \rho_0 E(\rho, S_0) + \mu (J - \rho_0/\rho), \tag{4.3}$$

where $\rho, \rho_0$ are the mass densities in the spatial and the reference configuration, rspectively, while $S_0 = S_0(\mathbf{R})$ is the entropy, $E$ is the internal energy, $p$ is the pressure ($p/\rho^2 = \partial E / \partial \rho$), and $\mu = \mu(\mathbf{R},t)$ is a Lagrange multiplier. As is immediately seen, the Euler–Lagrange equations relative to (4.1) are the equations of motion, those relative to (4.2) are the equations of motion and the continuity equation, namely

$$\rho_0 \ddot{x}_a + J X_{Q,a} p_{,Q} = 0, \tag{4.4}$$

$$J - \rho_0/\rho = 0. \tag{4.5}$$

Finally, the Euler–Lagrange equations relative to (4.3) are the equations of motion with $\mu$ as the pressure, the continuity equation, and the condition specifying that $\mu$ coincides with the pressure $\rho^2 \, \partial E / \partial \rho$.

In a sense the Lagrangian (4.1) has already been given by Seliger and Whitham[2] while (4.3) has been given by Lanczos[9] and Eckart[10] in conjunction with incompressible fluids. To our knowledge the Lagrangian (4.2) has been determined very recently by ourselves for the first time.[3] It is then worth confining our attention to (4.2) also because it embraces all typical aspects of the subject.

In view of (3.15) and (2.1), with the material Lagrangian (4.2) there is associated the spatial Lagrangian

$$L^{\mathrm{s}}(\mathbf{X},\rho) = \tfrac{1}{2} j \rho_0 x_{a,Q} x_{a,B} X_{Q,t} X_{B,t} - j\rho_0 E(\rho, S_0) + (1 - j\rho_0/\rho) p(\rho, S_0). \tag{4.6}$$

Upon observing that[7]

$$(j x_{a,B})_{,a} = 0$$

and that

$$j_{,t} + j_{,b} v_b + j v_{b,b} = 0,$$

$$v_{a,b} x_{b,B} - x_{a,Bt} - x_{a,Bb} v_b = 0,$$

we obtain

$$\frac{\delta L^{\mathrm{s}}}{\delta X_B} = j\rho_0 x_{a,B} \left( \dot{v}_a + \frac{1}{\rho} p_{,a} \right) + \left( 1 - \frac{j\rho_0}{\rho} \right) \frac{\partial p}{\partial S} S_{0,B},$$

$$\frac{\delta L^{\mathrm{s}}}{\delta \rho} = \left( 1 - \frac{j\rho_0}{\rho} \right) \frac{\partial p}{\partial \rho}.$$

On the other hand, it is a simple matter to see that

$$\frac{\delta L^{\mathrm{m}}}{\delta x_a} = -(\rho_0 \ddot{x}_a + J X_{Q,a} p_{,Q}),$$

$$\frac{\delta L^{\mathrm{m}}}{\delta \rho} = \left( J - \frac{\rho_0}{\rho} \right) \frac{\partial p}{\partial \rho}.$$

Accordingly, (3.8) and (3.9) are obviously satisfied. Furthermore the Euler–Lagrange equations $\delta L^{\mathrm{s}}/\delta \rho = 0$, $\delta L^{\mathrm{s}}/\delta X_{\mathrm{B}} = 0$ yield

$$1 - j\rho_0/\rho = 0,$$

$$\dot{v}_a + (1/\rho) p_{,a} = 0,$$

namely the balance equations in the $E$-description. Thus we have ascertained that the expression (4.6) is just the Lagrangian for fluid dynamics in the $E$-description.

There is an aspect of the theorem that deserves a final comment. In view of (3.15), $L^{\mathrm{m}} = JL^{\mathrm{s}}$ depends on $\mathbf{x}$ even if $L^{\mathrm{s}}$ does not depend on $\mathbf{X}$. At first sight the presence of $J(x_{a,Q})$ suggests that an additional equation for the motion $\mathbf{x}(\mathbf{R},t)$ occurs, which, however, would reflect merely a change of coordinates but not the behavior of a physical system. In fact, this additional equation results in an identity because of (3.9). By way of example, look at the electromagnetic field whose spatial Lagrangian is

$$L^{\mathrm{s}}(\mathbf{A},\Phi) = \tfrac{1}{2}(1/c\mathbf{A}_{,t} + \nabla\Phi)^2 - \tfrac{1}{2}(\nabla\times\mathbf{A})^2,$$

where $\mathbf{A}$ is the vector potential, $\Phi$ is the scalar potential, and $c$ is the light speed. The material Lagrangian is given by

$$L^{\mathrm{m}}(\mathbf{A},\Phi,\mathbf{x}) = \tfrac{1}{2}J((1/c)\dot{A}_a - (1/c)\dot{x}_b A_{a,b} + \Phi_{,Q}X_{Q,a})$$

$$\times((1/c)\dot{A}_a - (1/c)x_b A_{a,b} + \Phi_{,B}X_{B,a})$$

$$- \tfrac{1}{2}J\epsilon_{abc}\epsilon_{ade}A_{C,Q}A_{e,B}X_{Q,b}X_{B,d},$$

where $\epsilon_{abc}$ is the Levi-Cività tensor. The dependence of $L^{\mathrm{m}}$ on $\mathbf{x}$ through $J$, $\ddot{x}_b$, and $X_{Q,a}$ is evident. Nevertheless, coherently with the general theory, a direct calculation shows that

$$\frac{\delta L^{\mathrm{m}}}{\delta x_a} + \frac{\delta L^{\mathrm{m}}}{\delta A_b}A_{b,a} + \frac{\delta L^{\mathrm{m}}}{\delta\Phi}\Phi_{,a} = 0$$

is an identity. Accordingly, the vanishing of $\delta L^{\mathrm{m}}/\delta x_a$ is due to the vanishing of $\delta L^{\mathrm{m}}/\delta A_b$ and $\delta L^{\mathrm{m}}/\delta\Phi$, namely, to the field equations for $\mathbf{A}$ and $\Phi$.

## 5. FINAL REMARKS

As shown in Sec. 3, the relation

$$L^{\mathrm{s}} = jL^{\mathrm{m}}$$

allows us to obtain a variational principle in the $E$-description whenever a variational principle in the $L$-description is available and vice versa. This result is based upon an unusual view of the $E$-description: the motion is described through the function $\mathbf{X}(\mathbf{r},t)$ instead of $\mathbf{v}(\mathbf{r},t)$. In this respect, however, we observe that, in spite of the generality of our theorem, customary variational principles in the $E$-description need not have their counterpart in the $L$-description. This is so because usually variational principles in the $E$-description involve $\mathbf{v}$ as an unknown function; upon substitution of

$$v_a = -x_{a,Q}X_{Q,t},$$

the equations under consideration need no longer be the Euler–Lagrange equations of the Lagrangian

$$L^{\mathrm{s}}(-x_{a,Q}X_{Q,t}).$$

As an example, we mention that the Lagrangian $L^{\mathrm{s}}_{\mathrm{sw}}(v_a)$ for fluid dynamics in the $E$-description set up by Seliger and Whitham[2] does not allow the pertinent equations to arise from the Lagrangian $L^{\mathrm{s}}_{\mathrm{sw}}(-x_{a,Q}X_{Q,t})$.

The connection between variational principles in Eulerian and Lagrangian descriptions is crucially based on the transformation $\mathbf{R}\to\mathbf{r} = \mathbf{x}(\mathbf{R},t)$ being a homeomorphism, which is usually true in continuum mechanics. In this regard, however, it is worth pointing out that sometimes nonlinear boundary-value problems are solved in closed form via an analogous (von Mises) transformation

$$x = x(\psi,t), \tag{5.1}$$

the stream function $\psi$ playing the role of reference coordinate.[11,12] As remarked by Wilhelm in his nonlinear theory of electron neutralization waves in ion beams,[12] there are physical conditions rendering (5.1) noninvertible (neutralization shock waves and multivalued flows). In such conditions our connection does not apply.


[1] M. Lax and D. F. Nelson, Phys. Rev. B **13**, 1759 (1976).
[2] R. L. Seliger and G. B. Whitham, Proc. R. Soc. London, Ser. A **305**, 1 (1968).
[3] F. Bampi and A. Morro, J. Math. Phys. **23**, 2312 (1982).
[4] E. Tonti, Acad. R. Belg. (Classe des Sci.) **55**, 137 (1969).
[5] R. W. Atherton and G. M. Homsy, Studies Appl. Math. **54**, 31 (1975).
[6] R. M. Santilli, *Foundations of Theoretical Mechanics I* (Springer, New York, 1978).
[7] C. Truesdell and R. A. Toupin, in *Encyclopedia of Physics*, edited by S. Flügge (Springer, Berlin, 1960), Vol. III/1.
[8] A. C. Eringen, in *Continuum Physics*, edited by A. C. Eringen (Academic, New York, 1971), Vol. II.
[9] C. Lanczos, *The Variational Principles of Mechanics* (University of Toronto, Toronto, 1970), p. 360.
[10] C. Eckart, Phys. Fluids **3**, 421 (1960).
[11] G. Kalman, Ann. Phys. (N.Y.) **10**, 1 (1960); **10**, 29 (1960).
[12] H. E. Wilhelm, Phys. Fluids **17**, 1841 (1974).

# A generalization of the Siewert–Burniston method for the determination of zeros of analytic functions

E. G. Anastasselou
*Division of Applied Mechanics, The National Technical University of Athens, P.O. Box 61028, GR-151.10 Amaroussion, Greece*

N. I. Ioakimidis
*Division of Applied Mathematics and Mechanics, School of Engineering, University of Patras, P.O. Box 1120, GR-261.10 Patras, Greece*

(Received 26 January 1984; accepted for publication 6 April 1984)

The Siewert–Burniston method for the derivation of closed-form formulas for the zeros of sectionally analytic functions with a discontinuity interval along the real axis (based on the Riemann–Hilbert boundary value problem in complex analysis) is generalized to apply to the determination of the zeros of analytic functions (without discontinuity intervals) inside or outside simple smooth contours. An application of this method to the closed-form solution of the transcendental equation $ze^z = be^b$, appearing in the theory of neutron moderation in nuclear reactors, is also made.

PACS numbers: 02.90. + p, 02.60.Lj, 28.20.Lh

## I. INTRODUCTION

An efficient method for the derivation of closed-form formulas for the zeros of sectionally analytic functions (with a discontinuity interval along the real axis) was proposed in 1971 by Siewert and Burniston. This method was applied to the determination of the zeros of a series of sectionally analytic functions and was used for the solution of many physical problems of practical interest.[1–15] The principle of the method was the solution of an appropriate Riemann–Hilbert boundary value problem along the discontinuity interval.[16] Unfortunately, in cases where the function whose zeros were sought did not possess a discontinuity interval, it should be transformed into another function with such an interval (and, clearly, with the same zeros or, at least, the same sought zeros).[2–5,7,8,10–12,14] This approach is not very convenient and, quite frequently, impossible. Here we generalize the method of Siewert and Burniston to apply to the determination of zeros of analytic functions inside or outside a simple smooth contour $C$. This generalization is quite simple, but also important for the aforementioned reason. After the illustration of this generalization in a trivial case, we will apply it to the closed-form solution of the transcendental equation $ze^z = be^b$, appearing in a problem of the theory of neutron moderation.[12] Numerical results, verifying the validity of the closed-form formulas for the nontrivial root $z_0$ of this equation, will be also presented and will be seen to be quite satisfactory.

Moreover, we can mention that competitive methods for the analytical determination of the zeros of analytic functions are reported by Abd-Elall, Delves, and Reid,[17] Anastasselou,[18] and Anastasselou and Ioakimidis.[19] Approximate numerical and graphical methods are also available,[20,21] but they have the disadvantage that they do not lead to closed-form formulas for the zeros of general validity and independent of the numerical values of the parameters (coefficients) of the analytic function whose zeros are sought. Finally, recently, Ioakimidis[22] proceeded to another generalization of the method of Siewert and Burniston and applied it to the determination of curves in the complex plane satisfying an equation of the form $|\Phi(z)| = D$, where $\Phi(z)$ is a sectionally analytic function with a discontinuity interval along the real axis and $D$ is a constant.

## II. ANALYSIS

We consider an analytic function $\Phi(z)$ inside a simple smooth contour $C$ in the complex plane $z = x + iy$. Essentially without loss of generality, we assume that the origin of the Cartesian coordinate system, $z = 0$, lies inside $C$. We take also into account the definition of the index $\varkappa$ of the boundary values of $\Phi(z)$, $\Phi(t)$, as $t$ varies along $C$ (see Ref. 16)

$$\varkappa = \mathrm{Ind}\,\Phi(t) = (1/2\pi)[\arg \Phi(t)]_C. \tag{2.1}$$

It is well known from the argument principle that $\varkappa$ is equal to the number of zeros of $\Phi(z)$ inside $C$ [under the assumptions that $\Phi(z)$ does not possess poles inside $C$ and on $C$ and zeros on $C$].[16] Methods for the computation of the index $\varkappa$, which is always an integer, were also reported by Gakhov.[16]

We define now a sectionally analytic function $F(z)$, discontinuous along $C$,

$$F(z) \equiv F^+(z) \equiv \Phi(z),\ z \in S^+,\ F(z) \equiv F^-(z) \equiv 1,\ z \equiv S^-, \tag{2.2}$$

where $S^+$ and $S^-$ denote the open regions inside and outside the contour $C$, respectively. Then, clearly,

$$F^+(t)/F^-(t) = \Phi(t), \quad t \in C, \tag{2.3}$$

where $F^{\pm}(t)$ denote the boundary values of $F(z)$ as $z \to t$ from $S^{\pm}$. Thus, we have reached a very simple, homogeneous Riemann–Hilbert boundary value problem[16] along $C$ for the "unknown" function $F(z)$. Its solution is very simple[16]

$$\begin{aligned} &F^+(z) \equiv \Phi(z) = X(z)P_\varkappa(z), \quad z \in S^+, \\ &F^-(z) \equiv 1 = X(z)P_\varkappa(z), \quad z \in S^- \end{aligned} \tag{2.4}$$

[where Eqs. (2.2) have been taken into consideration]. In (2.4) $X(z)$ is the canonical function of our problem, given by[16]

$$\begin{aligned} &X(z) = \exp[\Gamma(z)], \quad z \in S^+, \\ &X(z) = z^{-\varkappa} \exp[\Gamma(z)], \quad z \in S^-, \end{aligned} \tag{2.5}$$

with

$$\Gamma(z) = \frac{1}{2\pi i} \int_C \frac{\log[t^{-\varkappa}\Phi(t)]}{t - z}\, dt. \tag{2.6}$$

 0022-2488/84/082422-04$02.50  

Clearly, both $\Gamma(z)$ and $X(z)$ are discontinuous along $C$. Moreover, $X(z)$ does not possess zeros or poles in the finite complex plane.

Now, knowing the given function $\Phi(z)$ and evaluating $X(z)$ from (2.5) [together with (2.6)], we obtain directly from (2.4)

$$P_\varkappa(z) = \Phi(z)/X(z), \quad z\in S^+, \quad P_\varkappa(z) = 1/X(z), \quad z\in S^-. \tag{2.7}$$

These are identities in $S^\pm$ and permit us to determine the polynomial $P_\varkappa(z)$, the $\varkappa$ zeros of which coincide with the zeros of $\Phi(z)$ as is quite clear. We have thus reduced the problem of determination of the zeros of $\Phi(z)$ to that of determination of the zeros of the polynomial $P_\varkappa(z)$. If $\varkappa \leqslant 4$, these zeros can be obtained in closed form. Otherwise, we have to choose another contour $C^*$ with a corresponding index $\varkappa^*$ in our problem not exceeding 4. Frequently, it happens that $\Phi(z)$ has just one zero $z_0$ in $S^+$. Then $\varkappa = 1$, $P_\varkappa(z) = z - z_0$, and (2.7) give for $z_0$:

$$z_0 = z - \Phi(z)/X(z), \quad z\in S^+, \tag{2.8a}$$

or

$$z_0 = z - 1/X(z), \quad z\in S^-. \tag{2.8b}$$

These identities determine directly $z_0$. Any point $z$ in $S^\pm$ can be used, but the choices $z = 0$ in the first of (2.8) and $z\to\infty$ in the second of (2.8) (we have already assumed that $0\in S^+$) are the most common ones.

Clearly, if $\Phi(z)$ possesses poles instead of zeros in $S^+$, being a meromorphic function, we have simply to consider the analytic function $\Psi(z) = 1/\Phi(z)$ and use the previous method. Alternatively, we can appropriately modify this method to apply to meromorphic functions with poles in $S^+$.

More important is the case when we seek the zeros of an analytic function $\Phi(z)$ (with a possible pole of finite order at infinity) outside a simple smooth contour $C$, that is, in our notation, in $S^-$. In this case, $F(z)$ is defined by

$$F(z)\equiv F^+(z)\equiv 1, \quad z\in S^+, \quad F(z)\equiv F^-(z)\equiv \Phi(z), \quad z\in S^-. \tag{2.9}$$

The index $\varkappa$ is determined again from (2.1), but with a negative sign in the right-hand side. On the other hand, the number $\lambda$ of the sought zeros of $\Phi(z)$ in $S^-$ is equal to $\varkappa$ plus the order of the possible pole of $\Phi(z)$ at infinity or minus the order of the possible zero of the same function at infinity.

Moreover, in the present case (2.3) takes the form

$$F^+(t)/F^-(t) = 1/\Phi(t), \quad t\in C, \tag{2.10}$$

whereas (2.4) are modified as

$$F^+(z)\equiv 1 = X(z)P_\lambda(z), \quad z\in S^+,$$
$$F^-(z)\equiv \Phi(z) = X(z)P_\lambda(z), \quad z\in S^-, \tag{2.11}$$

with $X(z)$ determined again from (2.5), but with[16]

$$\Gamma(z) = \frac{1}{2\pi i}\int_C \frac{\log[t^{-\varkappa}/\Phi(t)]}{t - z}\,dt. \tag{2.12}$$

The zeros of $\Phi(z)$ in $S^-$ will be determined again as the zeros of $P_\lambda(z)$. Because of (2.11), we have

$$P_\lambda(z) = 1/X(z), \quad z\in S^+, \quad P_\lambda(z) = \Phi(z)/X(z), \quad z\in S^-. \tag{2.13}$$

For example, for $\lambda = 1$, whence $P_\lambda(z) = z - z_0$, we find from (2.13)

$$z_0 = z - 1/X(z), \quad z\in S^+, \text{ or } z_0 = z - \Phi(z)/X(z), \quad z\in S^-. \tag{2.14}$$

[These formulas are analogous to (2.8).]

We proceed now to an illustration of the present method, by applying it to a trivial analytic function.

## III. AN ILLUSTRATION OF THE METHOD

To illustrate the validity of the method described in the previous section, we consider the trivial analytic function

$$\Phi(z) = z - c, \tag{3.1}$$

with a zero $z_0 = c$. For convenience, we assume that the contour $C$ is the circumference of the unit circle with center the point $z = 0$ and that $c$ is a real constant. We will verify the validity of (2.8) and (2.14).

At first, we assume that $|c| < 1$. Then (2.8) hold true for the zero $z_0$ of $\Phi(z)$. Since $\varkappa = 1$ in our case [as can be verified from (2.1)],

$$t^{-\varkappa}\Phi(t) = 1 - c/t = (1 - c\cos\vartheta) + ic\sin\vartheta, \tag{3.2}$$

where $\vartheta$ is the polar angle [$t = \exp(i\vartheta)$]. Moreover, $dt/t = i\,d\vartheta$. Then we obtain easily from (2.6)

$$\Gamma(0) = \frac{1}{2\pi}\int_0^\pi \log(1 + c^2 - 2c\cos\vartheta)\,d\vartheta. \tag{3.3}$$

Now (2.8a) for $z = 0$ takes the form

$$z_0 = -\Phi(0)/X(0) = -\Phi(0)\exp[-\Gamma(0)]. \tag{3.4}$$

In our case, because of (3.1) and (3.3), we find from (3.4)

$$z_0 = c\exp\left[-\frac{1}{2\pi}\int_0^\pi \log(1 + c^2 - 2c\cos\vartheta)\,d\vartheta\right]. \tag{3.5}$$

But it is well known that[23]

$$\int_0^\pi \log(1 + c^2 - 2c\cos\vartheta)\,d\vartheta = \begin{cases} 0, & \text{if } |c| < 1, \\ 2\pi\log|c|, & \text{if } |c| > 1. \end{cases} \tag{3.6}$$

Therefore, in our case, $|c| < 1$, (3.5) reduces to

$$z_0 = c, \tag{3.7}$$

as was expected.

Now we will reach the same result by using (2.8b) for $z\to\infty$. Then this equation reduces to

$$z_0 = -\frac{1}{2\pi i}\int_C \log\left[\frac{\Phi(t)}{t}\right]dt. \tag{3.8}$$

Because of (3.2) and the fact that $dt = (-\sin\vartheta + i\cos\vartheta)\,d\vartheta$, we find from (3.8) in our case

$$z_0 = \frac{1}{2\pi}\int_0^\pi \left[2\tan^{-1}\left(\frac{c\sin\vartheta}{1 - c\cos\vartheta}\right)\sin\vartheta - \log(1 + c^2 - 2c\cos\vartheta)\cos\vartheta\right]d\vartheta. \tag{3.9}$$

We further take into account that[23]

$$\int_0^\pi \tan^{-1}\left(\frac{c\sin\vartheta}{1 - c\cos\vartheta}\right)\sin\vartheta\,d\vartheta = \frac{\pi c}{2}, \quad |c| < 1, \tag{3.10}$$

$$\int_0^\pi \log(1 + c^2 - 2c\cos\vartheta)\cos\vartheta\,d\vartheta = \begin{cases} -\pi c, & \text{if } |c| < 1, \\ -\pi/c, & \text{if } |c| > 1, \end{cases} \tag{3.11}$$

and we find directly that (3.9) reduces again to (3.7) for $|c| < 1$.

Let us proceed now to the case when $|c| > 1$. In this case, we will use (2.14). Since $\Phi(z)$ presents now a first-order pole at infinity and, furthermore, $\varkappa = 0$, we observe that $\lambda = 1$ as was expected. In the present case, since $\varkappa = 0$ and

$$\Phi(t) = (\cos\vartheta - c) + i\sin\vartheta, \tag{3.12}$$

we find easily from the first of (2.14), applied here for $z = 0$, that is

$$z_0 = -1/X(0) = -\exp[-\Gamma(0)], \tag{3.13}$$

that

$$z_0 = \operatorname{sgn} c \times \exp\left[\frac{1}{2\pi}\int_0^{\pi} \log(1 + c^2 - 2c\cos\vartheta)d\vartheta\right]. \tag{3.14}$$

Because of the second of (3.6), (3.14) reduces to (3.7) again.

Similarly, by applying the second of (2.14) for $z \to \infty$ (and $|c| > 1$), we find

$$z_0 = c + \frac{1}{2\pi i}\int_C \log[\Phi(t)]\,dt, \tag{3.15}$$

or, further, because of (3.12),

$$z_0 = c - \frac{1}{2\pi}\int_0^{\pi}\left[2\tan^{-1}\left(\frac{\sin\vartheta}{\cos\vartheta - c}\right)\sin\vartheta - \log(1 + c^2 - 2c\cos\vartheta)\cos\vartheta\right]d\vartheta. \tag{3.16}$$

Because of (3.10), applied here with $1/c$ instead of $c$, and of the second of (3.11), we observe directly that (3.16) reduces to (3.7) once more. In this way, we have verified all four formulas (2.8) and (2.14). We proceed now to a nontrivial application of the proposed method.

## IV. AN APPLICATION TO A PHYSICAL PROBLEM

In this section we consider the transcendental equation

$$ze^z = be^b, \tag{4.1}$$

appearing in a problem of neutron moderation in nuclear reactors. This equation was already considered by Siewert and Burkart,[12] who applied the method of Siewert and Burniston and the corresponding previous results[10] to the solution of this equation, that is, to the determination of the zeros of the function

$$\Phi(z) = ze^z - c, \tag{4.2}$$

where

$$c = be^b. \tag{4.3}$$

The determination of the zeros of (4.2), but not accompanied by (4.3), appears also during the solution of a critical condition, described by age-diffusion theory, for a bare nuclear reactor.[5]

Here we will determine the nontrivial zero $z = z_0$ of (4.2), (4.3) taken also into consideration, for $b$ restricted to real negative values.[12] The trivial zero $z = b$ of the same function is not of interest. We assume again that $C$ is the circumference of the unit circle with center the point $z = 0$. For $b \in (-\infty, -1)$, $z_0 \in (-1, 0)$, that is, $z_0 \in S^+$. We restrict our attention to this particular case and we use (2.8) or, better, (3.4) and (3.8) (for $z = 0$ and $z \to \infty$, respectively). Since $\varkappa = 1$ and along $C$

$$t^{-\varkappa}\Phi(t) = e^t - c/t = R(\vartheta) + iI(\vartheta), \tag{4.4}$$

where

$$R(\vartheta) = \exp(\cos\vartheta)\cos(\sin\vartheta) - c\cos\vartheta, \tag{4.5a}$$

$$I(\vartheta) = \exp(\cos\vartheta)\sin(\sin\vartheta) - c\sin\vartheta, \tag{4.5b}$$

(3.4) takes the form

$$z_0 = c\exp\left\{-\frac{1}{2\pi}\int_0^{\pi}\log[R^2(\vartheta) + I^2(\vartheta)]d\vartheta\right\}. \tag{4.6}$$

Similarly, (3.8) takes the form

$$z_0 = \frac{1}{2\pi}\int_0^{\pi}\left\{2\tan^{-1}\left[\frac{I(\vartheta)}{R(\vartheta)}\right]\sin\vartheta - \log[R^2(\vartheta) + I^2(\vartheta)]\cos\vartheta\right\}d\vartheta. \tag{4.7}$$

For $R(\vartheta)$ and $I(\vartheta)$ defined by (4.5), we do not expect to simplify (4.6) and (4.7) analytically; these are our final formulas although equivalent since both determine the same zero $z_0$ of (4.2) [or root of (4.1)]. Just to check these equations, we present in Table I numerical results obtained by the classical

TABLE I. Numerical results for the nontrivial zero $z_0$ of $ze^z - be^b$ for $b = -1.1, -1.5, -3, -5$, and $-10$, obtained both from (4.6), $z_{01}$, and from (4.7), $z_{02}$. The integrals were computed by the trapezoidal rule with $n = 5(5)25$ nodes; the "exact" values of $z_0$ were determined by the Newton–Raphson method.

| $n$ | $z_{01}$ | $z_{02}$ |
|---|---|---|
| | $b = -1.1$ | |
| 5 | −0.849 765 00 | −0.840 087 42 |
| 10 | −0.894 132 13 | −0.892 583 35 |
| 15 | −0.903 038 31 | −0.902 651 43 |
| 20 | −0.905 320 30 | −0.905 209 19 |
| 25 | −0.905 967 17 | −0.905 933 02 |
| "Exact" value: | −0.906 252 44 | |
| | $b = -1.5$ | |
| 5 | −0.624 135 13 | −0.622 394 35 |
| 10 | −0.625 770 47 | −0.625 756 27 |
| 15 | −0.625 782 41 | −0.625 782 25 |
| 20 | −0.625 782 53 | −0.625 782 53 |
| 25 | −0.625 782 53 | −0.625 782 53 |
| "Exact" value: | −0.625 782 53 | |
| | $b = -3$ | |
| 5 | −0.178 560 32 | −0.178 554 98 |
| 10 | −0.178 560 63 | −0.178 560 63 |
| 15 | −0.178 560 63 | −0.178 560 63 |
| 20 | −0.178 560 63 | −0.178 560 63 |
| 25 | −0.178 560 63 | −0.178 560 63 |
| "Exact" value: | −0.178 560 63 | |
| | $b = -5$ | |
| 5 | −0.034 885 768 | −0.034 885 712 |
| 10 | −0.034 885 768 | −0.034 885 768 |
| 15 | −0.034 885 768 | −0.034 885 768 |
| 20 | −0.034 885 768 | −0.034 885 768 |
| 25 | −0.034 885 768 | −0.034 885 768 |
| "Exact" value: | −0.034 885 768 | |
| | $b = -10$ | |
| 5 | −0.000 454 206 | −0.000 454 206 |
| 10 | −0.000 454 206 | −0.000 454 206 |
| 15 | −0.000 454 206 | −0.000 454 206 |
| 20 | −0.000 454 206 | −0.000 454 206 |
| 25 | −0.000 454 206 | −0.000 454 206 |
| "Exact" value: | −0.000 454 206 | |

trapezoidal rule [with $n = 5(5)25$ nodes] for periodic functions.[24] In our case, this rule may take the form

$$\frac{1}{\pi}\int_0^{\pi} g(\vartheta)d\vartheta \simeq \frac{1}{n}\sum_{i=1}^{n} g(\vartheta_i),$$

$$\vartheta_i = (i-0.5)\pi/n, \quad i = 1(1)n. \tag{4.8}$$

The numerical results obtained from (4.6), $z_{01}$, are presented in the second column of Table I and the corresponding results obtained from (4.7), $z_{02}$, in the third column of the same table. All five cases $b = -1.1, -1.5, -3, -5$, and $-10$ were considered. The "exact" values of the zero $z_0$, determined by the Newton–Raphson method, are also displayed in Table I. The numerical results of this table are quite satisfactory, especially for large values of $-b$. As $-b$ tends to 1, the accuracy of these results decreases. The remedy is simply either to use another *appropriate* contour $C$ or to use a more sophisticated quadrature rule instead of the trapezoidal rule. We will not enter into such computational details, since our aim was to derive the analytical formulas (4.6) and (4.7) for the nontrivial zero $z_0$ of (4.2) and, moreover, we are satisfied from the numerical results of Table I. For even more accurate results in the same table, we can simply use a larger number of nodes $n$ in (4.8).

## ACKNOWLEDGMENT

The present results were obtained in the course of a research project supported by the National Hellenic Research Foundation. The authors gratefully acknowledge the financial support of this Foundation.

[1]C. E. Siewert and E. E. Burniston, Astrophys. J. **173**, 405 (1972).
[2]C. E. Siewert and E. E. Burniston, Celest. Mech. **6**, 294 (1972).
[3]E. E. Burniston and C. E. Siewert, Proc. Cambridge Philos. Soc. **73**, 111 (1973).
[4]E. E. Burniston and C. E. Siewert, SIAM J. Appl. Math. **24**, 460 (1973).
[5]C. E. Siewert, Nucl. Sci. Eng. **51**, 78 (1973).
[6]C. E. Siewert, E. E. Burniston, and J. R. Thomas, Jr., Phys. Fluids **16**, 1532 (1973).
[7]C. E. Siewert and C. J. Essig, Z. Angew. Math. Phys. **24**, 281 (1973).
[8]C. E. Siewert and A. R. Burkart, Z. Angew. Math. Phys. **24**, 435 (1973).
[9]C. E. Siewert and J. T. Kriese, J. Nucl. Energy **27**, 831 (1973).
[10]C. E. Siewert and E. E. Burniston, J. Math. Anal. Appl. **43**, 626 (1973).
[11]C. E. Siewert and E. E. Burniston, J. Math. Anal. Appl. **46**, 329 (1974).
[12]C. E. Siewert and A. R. Burkart, Nucl. Sci. Eng. **54**, 455 (1974).
[13]J. R. Thomas, Jr., Plasma Phys. **18**, 715 (1976)
[14]C. E. Siewert, J. Math. Phys. **19**, 434 (1978).
[15]C. E. Siewert, J. Math. Phys. **21**, 2468 (1980).
[16]F. D. Gakhov, *Boundary Value Problems* (Pergamon and Addison–Wesley, Oxford, 1966), pp. 85–96.
[17]L. F. Abd-Elall, L. M. Delves, and J. K. Reid, in *Numerical Methods for Nonlinear Algebraic Equations*, edited by P. Rabinowitz (Gordon and Breach, London, 1970), pp. 47–59.
[18]E. G. Anastasselou, Ph.D. thesis, National Technical University of Athens, 1984 (in preparation).
[19]E. G. Anastasselou and N. I. Ioakimidis, Lett. Math. Phys. **8**, 135 (1984).
[20]A. S. Householder, *The Numerical Treatment of a Single Nonlinear Equation* (McGraw–Hill, New York, 1970).
[21]W. Pfeiffer, J. Comput. Phys. **33**, 397 (1979).
[22]N. I. Ioakimidis, J. Franklin Inst. **317**, 27 (1984).
[23]I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series, and Products* (Academic, New York, 1965), pp. 527, 593, 613.
[24]P. J. Davis and P. Rabinowitz, *Numerical Integration* (Blaisdell, Waltham, MA, 1967), pp. 53–59.

# Closing and abelizing operatorial gauge algebra generated by first class constraints

I. A. Batalin and E. S. Fradkin
*Physical Lebedev Institute, Moscow, USSR*

(Received 4 October 1983; accepted for publication 6 January 1984)

A canonical transformation of operators of relativistic phase space is constructed, which abelizes the basis of the operator-valued gauge algebra of first class constraints. The new operators of contraints commute among themselves. The new Hamiltonian commutes with the constraints. Dynamics of the new operators is physically equivalent to the initial one.

PACS numbers: 03.20. + i, 03.30. + p, 02.30.Tb, 02.20.Km

## I. INTRODUCTION

In our paper,[1] a method was developed for operatorial quantization of dynamical systems subject to first class constraints. A basic role in this method is played by generating equations of an operator-valued gauge algebra. When expanded into power series in ghost operators, these equations produce, in the lowest order, operator involution relations for the first class constraints and the initial Hamiltonian and, in higher orders, the complete set of structural Jacobi relations, each providing the formal compatability of the previous one. The gauge algebra generated by the operators of the first class constraints is, generally, open (unclosed). In the present paper we build the ghost-operator-dependent canonical transformation which makes the gauge algebra of the first class constraints closed and even abelian, i.e., the new constraint operators commute with each other. The same as in Ref. 1, we are, in fact, considering the case of irreducible, i.e., linearly independent first class constraints. The generalization of our result for constraints of any stage of reducibility reduces evidently to the extension of the ghost sector, indicated in Ref. 2.

## II. NOTATIONS

$\epsilon(A) = \epsilon(\hat{A})$ is the designation for the Grassmann parity of a classical quantity $A$ and of the corresponding operator $\hat{A}$. The equation

$$[\hat{A},\hat{B}] = \hat{A}\hat{B} - (-1)^{\epsilon(\hat{A})\epsilon(\hat{B})}\hat{B}\hat{A} \tag{1}$$

denotes the supercommutator of operators $\hat{A}$ and $\hat{B}$. The ghost number of an operator $\hat{A}$ is denoted as $\mathrm{gh}(\hat{A})$. As usual, $\partial_r(\partial_l)$ stands for the right (left) derivatives.

## III. GENERATING EQUATIONS OF THE GAUGE ALGEBRA

Let $(\hat{P}_i,\hat{Q}^i)$, where $i = 1,...,n$,

$$\epsilon(\hat{P}_i) = \epsilon(\hat{Q}^i), \quad \mathrm{gh}(\hat{P}_i) = -\mathrm{gh}(\hat{Q}^i) = 0,$$

be operators of the initial dynamical variables which obey the equal-time canonical commutation relations, so that the only nonvanishing commutator is

$$[\hat{Q}^j,\hat{P}_k] = i\hbar\delta^j_k\hat{1}.$$

Denote as

$$\hat{H}_0(\hat{P},\hat{Q}), \quad \hat{T}_\alpha(\hat{P},\hat{Q}), \quad \alpha = 1,...,m < n,$$

where $\epsilon(\hat{T}_\alpha) \equiv \epsilon_\alpha$, the initial operator-valued Hamiltonian and first class constraints, respectively. They are in involution

$$[\hat{T}_\alpha,\hat{T}_\beta] = i\hbar\hat{T}_\alpha\hat{U}^\gamma_{\alpha\beta}, \quad [\hat{H}_0,\hat{T}_\beta] = i\hbar\hat{T}_\gamma\hat{V}^\gamma_\beta \tag{2}$$

and are irreducible

$$\hat{T}_\alpha\hat{Z}^\alpha = 0 \Rightarrow \hat{Z}^\alpha = \hat{T}_\beta\hat{F}^{\alpha\beta} - \tfrac{1}{2}i\hbar\hat{U}^\alpha_{\beta\gamma}\hat{F}^{\beta\gamma}. \tag{3}$$

Here

$$\hat{U}^\gamma_{\alpha\beta} = -(-1)^{\epsilon_\alpha\epsilon_\beta}\hat{U}^\gamma_{\beta\alpha}, \hat{V}^\gamma_\beta$$

are operator-valued structural coefficients of the involution (2);

$$\hat{F}^{\alpha\beta} = -(-1)^{\epsilon_\alpha\epsilon_\beta}\hat{F}^{\beta\alpha}$$

are any operators, "skew-symmetrical" in the upper indices. All of them are functions of the initial operators $(\hat{P}_i,\hat{Q}^i)$.

Let, further $(\hat{\bar{\mathscr{P}}}_\alpha,\hat{C}^\alpha)$, where $\alpha = 1,...,m$,

$$\epsilon(\hat{\bar{\mathscr{P}}}_\alpha) = \epsilon(\hat{C}^\alpha) = \epsilon_\alpha + 1, \quad \mathrm{gh}(\hat{\bar{\mathscr{P}}}_\alpha) = -\mathrm{gh}(\hat{C}^\alpha) = -1,$$

be ghost operators that, too, satisfy canonical equal-time commutation relations, so that the only nonzero commutator is

$$[\hat{C}^\alpha,\hat{\bar{\mathscr{P}}}_\beta] = i\hbar\delta^\alpha_\beta\hat{1}.$$

The generating equations for the operatorial gauge algebra for the fermion $\hat{\Omega}_{\min}$, $\mathrm{gh}(\hat{\Omega}_{\min}) = 1$, and the boson $\hat{H}_{\min}$, $\mathrm{gh}(\hat{H}_{\min}) = 0$, are formulated in Ref. 1 to be

$$[\hat{\Omega}_{\min},\hat{\Omega}_{\min}] = 0, \quad [\hat{H}_{\min},\hat{\Omega}_{\min}] = 0. \tag{4}$$

Solution for these equations is looked for in the form of $\hat{\bar{\mathscr{P}}}\hat{C}$-normally ordered ghost-operator power series:

$$\hat{\Omega}_{\min} = \hat{T}_\alpha\hat{C}^\alpha + \sum_{n=1}^{\infty}\hat{\bar{\mathscr{P}}}_{\alpha_n}\cdots\hat{\bar{\mathscr{P}}}_{\alpha_1}\hat{U}^{\alpha_1\cdots\alpha_n}, \tag{5}$$

$$\hat{H}_{\min} = \hat{H}_0 + \sum_{n=1}^{\infty}\hat{\bar{\mathscr{P}}}_{\alpha_n}\cdots\hat{\bar{\mathscr{P}}}_{\alpha_1}\hat{V}^{\alpha_1\cdots\alpha_n}, \tag{6}$$

where

$$\hat{U}^\alpha = \tfrac{1}{2}\hat{U}^\alpha_{\beta\gamma}\hat{C}^\gamma\hat{C}^\beta(-1)^{\epsilon_\beta}, \quad \hat{V}^\alpha = \hat{V}^\alpha_\beta\hat{C}^\beta,$$

etc. In the zeroth order in the ghost canonical momentum $\hat{\bar{\mathscr{P}}}$ Eqs. (4) are identically fulfilled owing to the involution (2). Equations for the other operator-valued coefficients in (5) and (6) are written explicitly in Ref. 1. Following Ref. 1, we assume that the ghost operators behave under the Hermitian conjugation as

 0022-2488/84/082426-04$02.50  

$$(\hat{\bar{\mathscr{P}}}_\alpha)^+ = -\hat{\bar{\mathscr{P}}}_\alpha(-1)^{\epsilon_\alpha}, \quad (\hat{C}^\alpha)^+ = \hat{C}^\alpha,$$

and require the hermiticity of the operators (5) and (6).

## IV. ABELIZATION OF THE GAUGE GROUP

Define new operator-valued constraints $\hat{\mathscr{T}}_\alpha(\hat{P},\hat{Q})$ by the equation

$$\hat{\Omega}_{\min}\hat{\Gamma} = \hat{\Gamma}(\hat{\mathscr{T}}_\alpha \hat{C}^\alpha), \tag{7}$$

where $\hat{\Gamma}$ is a unitary operator, depending on $\hat{P}_i$, $\hat{Q}^i$ and the ghosts $\hat{\bar{\mathscr{P}}}_\alpha$, $\hat{C}^\alpha$. $\hat{\Gamma}$ is a boson with zero ghost number.

Equation (7) performs a canonical transformation which converts $\hat{\Omega}_{\min}$ into $(\hat{\mathscr{T}}_\alpha \hat{C}^\alpha)$. By virtue of (4) and (7) one has

$$[\hat{\mathscr{T}}_\alpha, \hat{\mathscr{T}}_\beta] = 0. \tag{8}$$

Therefore, if Eq. (7) is solvable as that for a unitary operator $\hat{\Gamma}$, the new operatorial constraints $\hat{\mathscr{T}}_\alpha$ are abelian.

Let us be looking for the solution of (7) as a series of normal $\hat{\bar{\mathscr{P}}}\hat{C}$-products of the powers of the ghost operators $\hat{\bar{\mathscr{P}}}_\alpha$, $\hat{C}^\alpha$:

$$\hat{\Gamma} = \hat{F}_0 + \sum_{n=1}^{\infty} \hat{\bar{\mathscr{P}}}_{\alpha_n}\cdots\hat{\bar{\mathscr{P}}}_{\alpha_1}\hat{F}^{\alpha_1\cdots\alpha_n}. \tag{9}$$

Here $\hat{F}_0(\hat{P},\hat{Q})$ is a reversible operator,

$$\hat{F}^\alpha = \hat{F}^\alpha_\mu(\hat{P},\hat{Q})\hat{C}^\mu, \quad \hat{F}^{\alpha\beta} = \tfrac{1}{2}\hat{F}^{\alpha\beta}_{\mu\nu}(\hat{P},\hat{Q})\hat{C}^\nu\hat{C}^\mu(-1)^{(\epsilon_\alpha+\epsilon_\mu)},$$

etc. We assume, as usual, that the generalized symmetry properties, with respect to the upper indices, of every coefficient operator in (5), (6), and (9) are the same as those of the monomial:

$$\hat{C}^{\alpha_1}\cdots\hat{C}^{\alpha_n}.$$

Substituting expansions (5) and (9) into (7), we obtain the following relations for the coefficient operators $\hat{F}_0$, $\hat{F}^{\alpha_1\cdots\alpha_n}$: at $n = 0$:

$$(\hat{T}_\alpha\hat{C}^\alpha)\hat{F}_0 + i\hbar\hat{T}_\alpha\hat{F}^\alpha = \hat{F}_0(\hat{\mathscr{T}}_\alpha\hat{C}^\alpha), \tag{10}$$

at $n \geqslant 1$:

$$\begin{aligned}
&\hat{U}^{\alpha_1\cdots\alpha_n}\hat{F}_0 + (\hat{T}_\alpha\hat{C}^\alpha)\hat{F}^{\alpha_1\cdots\alpha_n}(-1)^{\epsilon_0^n} \\
&\quad + (n+1)i\hbar\hat{T}_\alpha\hat{F}^{\alpha\alpha_1\cdots\alpha_n}(-1)^{\epsilon_0^n} \\
&\quad + \sum_{k=1}^{n+1}(i\hbar)^k\frac{\partial_r^k\hat{U}^{\alpha_1\cdots\alpha_n}}{\partial\hat{C}^{\beta_1}\cdots\partial\hat{C}^{\beta_k}}\hat{F}^{\beta_1\cdots\beta_k} \\
&\quad + (\hat{M}_{\mathrm{sym}})^{\alpha_1\cdots\alpha_n} = \hat{F}^{\alpha_1\cdots\alpha_n}(\hat{\mathscr{T}}_\alpha\hat{C}^\alpha),
\end{aligned} \tag{11}$$

$$\begin{aligned}
\hat{M}^{\alpha_1} \equiv 0, \quad \hat{M}^{\alpha_1\cdots\alpha_n} \equiv \sum_{m=1}^{n-1}(-1)^{\epsilon_m^n}\Bigg(&\hat{U}^{\alpha_1\cdots\alpha_m}\hat{F}^{\alpha_{m+1}\cdots\alpha_n} \\
&+ \sum_{k=1}^{m+1}\frac{(n-m+k)!}{k!(n-m)!}(i\hbar)^k\frac{\partial_r^k\hat{U}^{\alpha_1\cdots\alpha_m}}{\partial\hat{C}^{\beta_1}\cdots\partial\hat{C}^{\beta_k}} \\
&\times\hat{F}^{\beta_1\cdots\beta_k\alpha_{m+1}\cdots\alpha_n}\Bigg),
\end{aligned} \tag{12}$$

$$\epsilon_m^n \equiv n - m + \sum_{j=1+m}^{n}\epsilon_{\alpha_j}, \tag{13}$$

$$(\hat{M}_{\mathrm{sym}})^{\alpha_1\cdots\alpha_n} \equiv S^{\alpha_1\cdots\alpha_n}_{\beta_1\cdots\beta_n}\hat{M}^{\beta_1\cdots\beta_n}, \tag{14}$$

$$n!S^{\alpha_1\cdots\alpha_n}_{\beta_1\cdots\beta_n} \equiv \left(\frac{\partial_l}{\partial\bar{\mathscr{P}}_{\alpha_1}}\cdots\frac{\partial_l}{\partial\bar{\mathscr{P}}_{\alpha_n}}\bar{\mathscr{P}}_{\beta_n}\cdots\bar{\mathscr{P}}_{\beta_1}\right). \tag{15}$$

The lowest equation (10) leads to the following expression for the new constraint operators:

$$\hat{\mathscr{T}}_\alpha = \hat{F}_0^{-1}\hat{T}_\alpha\hat{F}_0 + i\hbar(\hat{F}_0^{-1}\hat{T}_\beta\hat{F}_0)(\hat{F}_0^{-1}\hat{F}^\beta_\alpha). \tag{16}$$

The first term on the rhs of (16) is induced by the "bare" transformation, which is performed by the operator $\hat{F}_0(\hat{P},\hat{Q})$ and acts only on the initial operators $\hat{P}_i$ and $\hat{Q}^i$. This transformation does not affect the ghost operators and does not hence give rise, effectively, to rotation of the basis of the constraint operators. It is the second term on the rhs of (16) that includes the effect of the rotation, caused by the operator $(\hat{F}_0^{-1}\hat{F}^\beta_\alpha)$. On the other hand, it is more or less evident that, for closing and abelizing the gauge algebra, only the effect of the basis rotation is of value. So we may study the special case of $\hat{F}_0 = \hat{1}$. Equation (16) then gives

$$\hat{\mathscr{T}}_\alpha = \hat{T}_\alpha + i\hbar\hat{T}_\beta\hat{F}^\beta_\alpha. \tag{17}$$

Consider now Eq. (11) at $n = 1$. Setting $\hat{F}_0 = 1$ in it and using (17), we have

$$\begin{aligned}
&\hat{U}^\alpha_{\mu\nu} + [\hat{F}^\alpha_\mu,\hat{T}_\nu] - [\hat{F}^\alpha_\nu,\hat{T}_\mu](-1)^{\epsilon_\nu\epsilon_\mu} \\
&\quad + i\hbar\hat{U}^\alpha_{\mu\beta}\hat{F}^\beta_\nu - i\hbar\hat{U}^\alpha_{\nu\beta}\hat{F}^\beta_\mu(-1)^{\epsilon_\nu\epsilon_\mu} \\
&\quad + \hbar\hat{F}^\alpha_\mu\hat{T}_\beta\hat{F}^\beta_\nu - i\hbar\hat{F}^\alpha_\nu\hat{T}_\beta\hat{F}^\beta_\mu(-1)^{\epsilon_\nu\epsilon_\mu} \\
&= i\hbar(\delta^\alpha_\beta\hat{T}_\gamma - \delta^\alpha_\gamma\hat{T}_\beta(-1)^{\epsilon_\beta\epsilon_\gamma} \\
&\quad - i\hbar\hat{U}^\alpha_{\beta\gamma})\hat{F}^{\gamma\beta}_{\mu\nu}(-1)^{(\epsilon_\beta+\epsilon_\gamma)}.
\end{aligned} \tag{18}$$

Multiplying (18) by $\hat{T}_\alpha$ to the left and using (2), we obtain (8) immediately. The rest of Eqs. (11) are, generally, an infinite sequence of equations, each of which, after being multiplied by the constraint to the left, gives the condition of formal compatability of the previous equation.

If we take operator (5) for granted, i.e., consider all the structural operators of the gauge algebra generated by the initial constraints $\hat{T}_\alpha$ as known, Eqs. (10)–(12) determine, one after another, all the coefficient operators $\hat{F}^{\alpha_1\cdots\alpha_n}$ in the expansion (9) for $\hat{\Gamma}$. In particular, for $\hat{F}_0 = \hat{1}$ Eq. (18) determines [after multiplying (18) by $\hat{T}_\alpha$ to the left] the operators $\hat{F}^\alpha_\mu$; the next equation (11) at $n = 2$ determines the operators $\hat{F}^{\alpha\beta}_{\mu\nu}$, etc.

Assume now that the operator $\hat{\Gamma}$ is known, and define a new operator $\hat{\mathscr{H}}_{\min}$ by the equation

$$\hat{H}_{\min}\hat{\Gamma} = \hat{\Gamma}\hat{\mathscr{H}}_{\min}. \tag{19}$$

From (4), (7), and (19) it follows that

$$[\hat{\mathscr{H}}_{\min}, (\hat{\mathscr{T}}_\alpha\hat{C}^\alpha)] = 0. \tag{20}$$

Unfortunately, this does not imply yet that $\hat{\mathscr{H}}_{\min}$ commutes with $\hat{\mathscr{T}}_\alpha$, since $\hat{\mathscr{H}}_{\min}$ depends on ghost operators, generally. Expand, therefore, $\hat{\mathscr{H}}_{\min}$ into the $\hat{\bar{\mathscr{P}}}\hat{C}$-normal power series in ghosts:

$$\hat{\mathscr{H}}_{\min} = \hat{\mathscr{H}}_0 + \sum_{n=1}^{\infty}\hat{\bar{\mathscr{P}}}_{\alpha_n}\cdots\hat{\bar{\mathscr{P}}}_{\alpha_1}\hat{\mathscr{V}}^{\alpha_1\cdots\alpha_n}, \tag{21}$$

where $\hat{\mathscr{H}}_0(\hat{P},\hat{Q})$ does not depend on ghost operators,

$$\hat{\mathscr{V}}^\alpha = \hat{\mathscr{V}}^\alpha_\mu(\hat{P},\hat{Q})\hat{C}^\mu,$$

$$\hat{\mathscr{V}}^{\alpha\beta} = \tfrac{1}{2}\hat{\mathscr{V}}^{\alpha\beta}_{\mu\nu}(\hat{P},\hat{Q})\hat{C}^\nu\hat{C}^\mu(-1)^{(\epsilon_\alpha+\epsilon_\mu)},$$

etc. Then we obtain from (20) that

$$[\hat{\mathscr{H}}_0, (\hat{\mathscr{T}}_\alpha\hat{C}^\alpha)] = i\hbar\hat{\mathscr{T}}_\alpha\hat{\mathscr{V}}^\alpha, \tag{22}$$

$$[\hat{\mathscr{V}}^{\alpha_1\cdots\alpha_n}, (\hat{\mathscr{T}}_\alpha \hat{C}^\alpha)] = i\hbar(n+1)\hat{\mathscr{T}}_\alpha \hat{\mathscr{V}}^{\alpha\alpha_1\cdots\alpha_n}(-1)^{\epsilon_0^n}. \quad (23)$$

Define, further, the operator

$$\hat{\mathscr{H}}'_{\min} = \hat{\mathscr{H}}_{\min} + (i\hbar)^{-1}[(\hat{\mathscr{T}}_\alpha \hat{C}^\alpha), \hat{\Lambda}], \quad (24)$$

where $\hat{\Lambda}$ is a Fermionic operator which depends on $\hat{P}_i$, $\hat{Q}^i$ and $\hat{\bar{\mathscr{P}}}_\alpha$, $\hat{C}^\alpha$, and $\mathrm{gh}(\hat{\Lambda}) = -1$. By virtue of (8) and (20), we have

$$[\hat{\mathscr{H}}'_{\min}, (\hat{\mathscr{T}}_\alpha \hat{C}^\alpha)] = 0. \quad (25)$$

Expand $\hat{\Lambda}$ into $\hat{\bar{\mathscr{P}}}$ $\hat{C}$-normal ghost-power series:

$$\hat{\Lambda} = \sum_{n=1}^{\infty} \hat{\bar{\mathscr{P}}}_{\alpha_n} \cdots \hat{\bar{\mathscr{P}}}_{\alpha_1} \hat{\mathscr{L}}^{\alpha_1\cdots\alpha_n}. \quad (26)$$

Here $\hat{\mathscr{L}}^\alpha(\hat{P}, \hat{Q})$ does not depend on ghosts,

$$\hat{\mathscr{L}}^{\alpha\beta} = \hat{\mathscr{L}}^{\alpha\beta}_\mu(\hat{P}, \hat{Q})\hat{C}^\mu(-1)^{\epsilon_\alpha},$$
$$\hat{\mathscr{L}}^{\alpha\beta\gamma} = \tfrac{1}{2}\hat{\mathscr{L}}^{\alpha\beta\gamma}_{\mu\nu}(\hat{P}, \hat{Q})\hat{C}^\nu \hat{C}^\mu(-1)^{(\epsilon_\beta+\epsilon_\mu)},$$

etc. Substituting (21) and (26) into (24), we obtain

$$\hat{\mathscr{H}}'_0 = \hat{\mathscr{H}}_0 + \hat{\mathscr{T}}_\alpha \hat{\mathscr{L}}^\alpha, \quad (27)$$

$$\hat{\mathscr{V}}'^{\alpha_1\cdots\alpha_n} = \hat{\mathscr{V}}^{\alpha_1\cdots\alpha_n} + (n+1)\hat{\mathscr{T}}_\alpha \hat{\mathscr{L}}^{\alpha\alpha_1\cdots\alpha_n}(-1)^{\epsilon_0^n} + (i\hbar)^{-1}[\hat{\mathscr{L}}^{\alpha_1\cdots\alpha_n}, (\hat{\mathscr{T}}_\alpha \hat{C}^\alpha)], \quad (28)$$

where $\hat{\mathscr{H}}'_0$, $\hat{\mathscr{V}}'^{\alpha_1\cdots\alpha_n}$ are the coefficient operators in the expansion of $\hat{\mathscr{H}}'_{\min}$ into $\hat{\bar{\mathscr{P}}}$ $\hat{C}$-ghost-power series, like (21).

Now make the operators $\hat{\mathscr{L}}^{\alpha_1\cdots\alpha_n}$ obey the condition that every operator (28) should vanish,

$$\hat{\mathscr{V}}'^{\alpha_1\cdots\alpha_n} = 0. \quad (29)$$

The formal compatability of Eqs. (29) is provided by (8) and (23). By virtue of (29) we have

$$\hat{\mathscr{H}}'_{\min} = \hat{\mathscr{H}}'_0, \quad (30)$$

whence it follows due to (25) that

$$[\hat{\mathscr{H}}'_0, \hat{\mathscr{T}}_\alpha] = 0, \quad (31)$$

since the operator (27) is ghost-independent.

This completes the total abeliation of the gauge algebra, generated by irreducible constraints. It has resulted in the new constraints $\hat{\mathscr{T}}_\alpha$ which commute with each other and the new Hamiltonian $\hat{\mathscr{H}}'_0$ which commutes with the new constraints.

## V. TRANSFORMATION OF THE UNITARIZING HAMILTONIAN

Following Ref. 1, let us write the expression for the total unitarizing Hamiltonian:

$$\hat{H}_{\hat{\Psi}} = \hat{H}_{\min} + (i\hbar)^{-1}[\hat{\Psi}, \hat{\Omega}], \quad (32)$$

where

$$\hat{\Psi} = \hat{\bar{C}}_\alpha \chi^\alpha + \hat{\bar{\mathscr{P}}}_\alpha \hat{\lambda}^\alpha \quad (33)$$

is the gauge fermion which contains the operator of gauge fixing. The operator $\hat{\Omega}$ is defined as

$$\hat{\Omega} = \hat{\Omega}_{\min} + \hat{\pi}_\alpha \hat{\bar{\mathscr{P}}}^\alpha. \quad (34)$$

The operators $(\hat{\pi}_\alpha, \hat{\lambda}^\alpha)$ and $(\hat{\bar{C}}_\alpha, \hat{\bar{\mathscr{P}}}^\alpha)$, where $\alpha = 1,\ldots,m$,

$$\epsilon(\hat{\pi}_\alpha) = \epsilon(\hat{\lambda}^\alpha) = \epsilon_\alpha, \quad \mathrm{gh}(\hat{\pi}_\alpha) = -\mathrm{gh}(\hat{\lambda}^\alpha) = 0,$$
$$\epsilon(\hat{\bar{C}}_\alpha) = \epsilon(\hat{\bar{\mathscr{P}}}^\alpha) = \epsilon_\alpha + 1, \quad \mathrm{gh}(\hat{\bar{C}}_\alpha) = -\mathrm{gh}(\hat{\bar{\mathscr{P}}}^\alpha) = -1$$

satisfy canonical commutation relations, with the only commutators different from zero being

$$[\hat{\lambda}^\alpha, \hat{\pi}_\beta] = i\hbar\delta^\alpha_\beta \hat{1}, \quad [\hat{\bar{\mathscr{P}}}^\alpha, \hat{\bar{C}}_\beta] = i\hbar\delta^\alpha_\beta \hat{1}.$$

Introduce, further, the condensed notation for the complete set of operators of the relativistic phase space,

$$\hat{X}^A \equiv (\hat{P}_i, \hat{\bar{\mathscr{P}}}_\alpha, \hat{\pi}_\alpha, \hat{\bar{C}}_\alpha; \hat{Q}^i, \hat{C}^\alpha, \hat{\lambda}^\alpha, \hat{\bar{\mathscr{P}}}^\alpha). \quad (35)$$

The time evolution of these operators is described by the Heisenberg equations induced by the Hamiltonian (32):

$$i\hbar\partial_t \hat{X}^A = [\hat{X}^A, \hat{H}_{\hat{\Psi}}(\hat{X})]. \quad (36)$$

Define the transformed operators $\hat{\mathscr{X}}^A$ by the equation

$$\hat{\Gamma}\hat{X}^A = \hat{\mathscr{X}}^A\hat{\Gamma}. \quad (37)$$

Due to (36), they obey the equation

$$i\hbar\partial_t \hat{\mathscr{X}}^A = [\hat{\mathscr{X}}^A, \hat{H}_{\hat{\Psi}}(\hat{X})]. \quad (38)$$

Define, finally, the transformed operator $\hat{\mathscr{H}}_{\hat{\Psi}}$ by the equation

$$\hat{H}_{\hat{\Psi}}(\hat{X})\hat{\Gamma} = \hat{\Gamma}\hat{\mathscr{H}}_{\hat{\Psi}}(\hat{X}), \quad (39)$$

so that

$$\hat{H}_{\hat{\Psi}}(\hat{X}) = \hat{\mathscr{H}}_{\hat{\Psi}}(\hat{\mathscr{X}}). \quad (40)$$

Substituting (40) into the rh of (38), we get

$$i\hbar\partial_t \hat{\mathscr{X}}^A = [\hat{\mathscr{X}}^A, \hat{\mathscr{H}}_{\hat{\Psi}}(\hat{\mathscr{X}})]. \quad (41)$$

We may express now the operator $\hat{\mathscr{H}}_{\hat{\Psi}}(\hat{X})$, with the use of (7), (19), (24), (30), (32), and (34), as

$$\hat{\mathscr{H}}_{\hat{\Psi}}(\hat{X}) = \hat{\mathscr{H}}'_0 + (i\hbar)^{-1}[\hat{\Phi}, (\hat{\mathscr{T}}_\alpha \hat{C}^\alpha + \hat{\pi}_\alpha \hat{\bar{\mathscr{P}}}^\alpha)], \quad (42)$$

where the new gauge fermion $\hat{\Phi}$ is defined by the equation

$$\hat{\Gamma}(\hat{\Phi} + \hat{\Lambda}) = \hat{\Psi}\hat{\Gamma}. \quad (43)$$

The Hamiltonian $\hat{\mathscr{H}}_{\hat{\Psi}}(\hat{\mathscr{X}})$ involved on the rhs of Eq. (41) is to be obtained from (42) merely by the formal replacement $\hat{X}^A \to \hat{\mathscr{X}}^A$ of the operators (35) by the transformed operators from (37).

Thus, we may see from (41) and (42) that the evolution of the new operators $\hat{\mathscr{X}}^A$ is governed by the abelized unitarizing Hamiltonian $\hat{\mathscr{H}}_{\hat{\Psi}}(\hat{\mathscr{X}})$ built in the same way as (32)–(34) was, but using the abelian constraints $\hat{\mathscr{T}}_\alpha$ and the "initial" Hamiltonian $\hat{\mathscr{H}}'_0$, which commutes with them. Simultaneously, the initial gauge fermion $\hat{\Psi}$ turns into the new operator $\hat{\Phi}$ from (43), which totally absorbs the dependence on the fermion $\hat{\Lambda}$, involved in the redefinition (24). Since the abelization of the operatorial gauge algebra results, in fact, only in the replacement $\hat{\Psi} \to \hat{\Phi}$ of the gauge fermion, while physical quantities are independent of any choice of the latter, we may state that the dynamics, governed by the abelized Hamiltonian $\hat{\mathscr{H}}'_0$ and constraints $\hat{\mathscr{T}}_\alpha$, is physically equivalent to that of the initial Hamiltonian $\hat{H}_0$ and constraints $\hat{T}_\alpha$.

## VI. CONCLUSION

The general procedure of abelization of irreducible gauge algebra presented above shows that, in principle, there always exists the possibility of passing to a set of constraint operators which commute among themselves and to a Hamiltonian operator which commutes with them all. In this sense, the nonabelianness and unclosedness are consequences of an "unlucky" choice of the basis of the operatorial gauge algebra, rather than the intrinsic property of the

theory. However, it is this "unlucky" choice of the generally nonabelian and unclosed basis, which is imperatively dictated, in the relativistic context, by requirement of four-dimensional locality and covariance of the dynamical description. We encounter here principally the same situation as the one faced when abelizing the Lagrangian gauge algebra in the covariant functional approach.[3,4] We conclude that the method of operatorial quantization of dynamical systems subject to the first class constraints, developed in Ref. 1, retains its significance, since it remains applicable within the general basis of the gauge algebra.

[1]I. A. Batalin and E. S. Fradkin, Preprint N51, P. N. Lebedev Institute 1983;. Phys. Lett. B **128**, 303 (1983).
[2]I. A. Batalin and E. S. Fradkin, Phys. Lett. B **122**, 157 (1983).
[3]I. A. Batalin and G. A. Vilkovisky, Preprint N 144, P. N. Lebedev Institute 1983, Nucl. Phys. B **234**, 106 (1984).
[4]B. L. Voronov and I. V. Tyutin, Teor. Mat. Fiz. (Theor. Math. Phys. USSR) **50**, 333 (1982).

# Blow-up regularization of singular Lagrangians

José F. Cariñena and Manuel F. Rañada
*Departamento de Física Teórica, Facultad de Ciencias, Universidad de Zaragoza, Zaragoza, Spain*

(Received 20 July 1983; accepted for publication 2 March 1984)

We propose a method for regularizing singular Lagrangians by adding new degrees of freedom. We illustrate this regularization method with some particular examples. This procedure is shown to be very useful when the Lagrangian is homogeneous of degree one in the velocities giving rise to an identically vanishing Hamiltonian.

PACS numbers: 03.20. + i, 03.50. − z

## I. INTRODUCTION

One of the most important developments in theoretical mechanics in recent times has certainly been the study of singular Lagrangians initiated by Dirac[1,2] in his systematic treatment of constrained Hamiltonian systems. We recall that the dynamical evolution of a physical system can be described using two different approaches. The first one is the Lagrangian formulation which is constructed over a function $L$ defined on the tangent bundle $T\mathcal{Q}$ on the configuration space $\mathcal{Q}$, $L \in \mathcal{F}(T\mathcal{Q})$. The second approach is the Hamiltonian formulation constructed over a function $H$ defined on the cotangent bundle $T^*\mathcal{Q}$, $H \in \mathcal{F}(T^*\mathcal{Q})$. Both approaches are related by the Legendre transformation $D_L : T\mathcal{Q} \to T^*\mathcal{Q}$ and they are equivalent in the hyper-regular case in which $D_L$ is a diffeomorphism, or even in the regular case, at least locally, in which $D_L$ is a local diffeomorphism. On the contrary if the Lagrangian $L$ is singular, the prescription for going from the Lagrangian form of the dynamics to the Hamiltonian one fails.

The attractive theory developed by Dirac provides a method for dealing with such singular systems by adding to the canonical Hamiltonian $H_c$ new terms of the form $v^a\phi_a(q, p)$ corresponding to the constraints $\phi_a(q, p) = 0$ weighted with some coefficients $v^a$. This treatment has been successful to develop methods for the study of gauge field theories,[3] where the very important role played by the coefficients $v^a$ is known. However a rigorous geometrical counterpart has not been developed so adequately.

Usually the geometrical approach[4] is based in the search of submanifolds and regular foliations to which the restriction of the Legendre transformation recovers the regularity properties. Actually this geometrical treatment has been found to be more difficult to develop and not so successful as the proper Dirac theory. Consequently it is worthwhile to study different approaches. In this paper we present a study of singular Lagrangians similar in the philosophy of its principles to the blowing up of an algebraic variety[5] at a point for the regularization of singularities. In other words, instead of reducing the dimension of the manifold $T\mathcal{Q}$ we will try to introduce new auxiliary degrees of freedom. Obviously we will also modify appropriately the Lagrangian in such a way that it will be regular and will lead to similar Euler–Lagrange equations. The meaning of the word similar will be made more precise later.

The regularization we propose is by no means unique and it only gives general prescriptions to be satisfied by the new configuration space and Lagrangian. We shall also propose a concrete way of regularization which will be shown to be useful in some later examples. The new Lagrangian being regular, it defines a correct Legendre transformation and we can obtain a well-defined Hamiltonian. This regularization procedure turns out to be very adequate for the study of those cases in which the Lagrangian is a homogeneous function of degree one in the velocities and therefore the canonical Hamiltonian vanishes identically.

The paper is organized as follows. In Sec. II, we present the notations and we explain the main characteristics of the singular Lagrangians. In Sec. III, we present the method of regularization and in Sec. IV, we apply this method to one example. The construction of the Hamiltonian is discussed in Sec. V, and in Sec. VI and VII the relativistic point particle and the homogeneous formalism are analyzed. Finally in Sec. VIII, we discuss briefly what has been done and point out some problems that remain to be solved.

## II. SINGULAR LAGRANGIANS

We consider a dynamical system with a $N$-dimensional configuration manifold $\mathcal{Q}$. The tangent and cotangent bundles over $\mathcal{Q}$ will be denoted as usual by $T\mathcal{Q}$ and $T^*\mathcal{Q}$, respectively. Let $q^r$, $r = 1,..., N$, be a coordinate system for $\mathcal{Q}$ and $\dot{q}^r$ the corresponding velocities. On $T\mathcal{Q}$ is defined the Lagrangian $L$ that we suppose to be time independent

$$L = L(q^r, \dot{q}^r). \tag{2.1}$$

The equations of motion are

$$E_r L = W_{rs}\ddot{q}^s + V_{rs}\dot{q}^s + U_r = 0, \quad r, s = 1,..., N, \tag{2.2}$$

where $E_r$ is the Euler–Lagrange operator and

$$W_{rs} = \frac{\partial^2 L}{\partial \dot{q}^r \partial \dot{q}^s}, \quad V_{rs} = \frac{\partial^2 L}{\partial \dot{q}^r \partial q^s}, \quad U_r = -\frac{\partial L}{\partial q^r}. \tag{2.3}$$

These equations can be solved for the accelerations $\ddot{q}^r$ if and only if the Hessian matrix $W = [W_{rs}]$ is invertible and in this case the Lagrangian is said to be regular. The Lagrangian $L$ is singular if $W$ has less than maximal rank and therefore the inverse matrix $W^{-1}$ is not defined.

Equations (2.2) can always be rewritten in two separated sets, up to a reordering of coordinates

$$E_i L = W_{ij}\ddot{q}^j + W_{ia}\ddot{q}^a + V_{is}\dot{q}^s + U_i = 0, \quad i, j = 1,..., R, \tag{2.4a}$$

$$E_b L = W_{bj}\ddot{q}^j + W_{ba}\ddot{q}^a + V_{bs}\dot{q}^s + U_b = 0, \quad a, b = R + 1,..., N, \tag{2.4b}$$

where $[W_{ij}]$ is a submatrix of maximal rank, rank

 0022-2488/84/082430-06$02.50  

$[W_{ij}] = R$. As $[W_{ij}]$ is not singular, there is an inverse matrix $[W^{ki}]$ such that

$$W^{ki}W_{ij} = \delta^k_j \tag{2.5}$$

and we can solve in Eq. (2.4a) for $q^1,...,q^R$. We then have equations of the form

$$\ddot{q}^k = -W^{ki}(W_{ia}\ddot{q}^a + V_{is}\dot{q}^s + U_i) \quad k = 1,..., R. \tag{2.6}$$

Because the matrix $W$ is singular, its rank $R$ computed by treating all the variables $q^1,...,q^N,\dot{q}^1,...,\dot{q}^N$ as independent is less than $N$. Consequently there exist $N-R$ linearly independent null eigenvectors $\lambda_b$

$$\lambda_b{}^r(q,\dot{q})W_{rs}(q,\dot{q}) = 0, \quad b = 1,..., N-R. \tag{2.7}$$

If Eqs. (2.2) are rewritten in the form

$$W_{rs}\ddot{q}^s = f_r(q,\dot{q}) \tag{2.8}$$

we get a set of algebraic relations involving the $q$'s and $\dot{q}$'s

$$f'_b(q,\dot{q}) \equiv \lambda_b{}^r(q,\dot{q})f_r(q,\dot{q}) = 0, \quad b = 1,..., N-R. \tag{2.9}$$

These relations, and all other relations among the $q$'s and $\dot{q}$'s that may turn up in the analysis, are called constraint equations in the Lagrangian sense.

As summary we can state that what in the standard case are $N$ different equations for the accelerations, in the singular situation presents three possibilities.

(1) There are identities between Eq. (2.2). There are only $R$ independent equations of motion and therefore the remaining $N-R$ equations are simple algebraic consequences. In this case all the constraint functions $f'_b(q,\dot{q})$ are identically null and there are no constraint equations.

(2) $f'_b(q,\dot{q}) \not\equiv 0$, $\forall\, b$ and the system presents the maximum number of constraint equations in addition to the equations for the accelerations.

(3) The intermediate situations with identities and constraint equations simultaneously.

## III. THE REGULARIZATION OF THE LAGRANGIAN

Let us now consider a new configuration space $\mathcal{Q}'$ in the form $\mathcal{Q}' = \mathcal{Q} \times \mathbb{R}^{N'-N}$ in such a way that $\mathcal{Q}'$ is a $N'$-dimensional manifold characterized by the coordinates $q^\rho$, $\rho = 1,..., N'$, where $q^\rho = q^r$ if $\rho = 1,..., N$ and $q^\rho = q^\alpha$ if $\rho = N+1,..., N'$. We denote by $\pi_{\mathcal{Q}}$ the projection $\pi_{\mathcal{Q}} : \mathcal{Q}' \to \mathcal{Q}$ and by $i_{\mathcal{Q}} : \mathcal{Q} \to \mathcal{Q}'$ the natural inclusion map $i_{\mathcal{Q}}(q^r) = (q^r,0)$.

A regular extension of $L$ on the tangent bundle $T\mathcal{Q}'$ is defined to be a Lagrangian $\overline{L} \in \mathcal{F}(T\mathcal{Q}')$ such that

(i) $\overline{L}$ is regular in almost any point of $T\mathcal{Q}'$,

(ii) the image under $i_{\mathcal{Q}}$ of any solution of the equations of motion associated to $L$ is a particular solution of the Euler–Lagrange equations from $\overline{L}$.

Note that the two Lagrangians $L$ and $\overline{L}$ are defined on different manifolds and consequently they are neither equivalent nor even $s$-equivalent[6] but $\overline{L} \circ Ti_{\mathcal{Q}} = L$. The rank of the new Hessian matrix $\overline{W} = [\overline{W}_{\rho\sigma}]$ is maximum (rank $\overline{W} = N'$) because it is regular and we can solve for $\ddot{q}^\tau$ in the corresponding Euler–Lagrange equations. If $[\overline{W}^{\tau\rho}]$ denotes the inverse matrix $\overline{W}^{-1}$, then

$$\ddot{q}^\tau = -\overline{W}^{\tau\rho}(\overline{V}_{\rho\sigma}\dot{q}^\sigma + \overline{U}_\rho). \tag{3.1}$$

It is to be remarked that the two above conditions do not guarantee the uniqueness of $\overline{L}$ and in consequence, given a singular Lagrangian $L$ it would be possible to associate to it different regular extensions. Two such extensions $\overline{L}_1$ and $\overline{L}_2$ must verify $\overline{L}_1 \circ Ti_{\mathcal{Q}} = \overline{L}_2 \circ Ti_{\mathcal{Q}}$.

As main points for the construction of an appropriate $\overline{L}$ in the case that all the constraint equations are given by (2.9), we will assume the following simplicity conditions:

(a) The dimension of the manifold $\mathcal{Q}'$ will be $2N-R$, that is to say, the number $N'-N$ of new coordinates $q^\alpha$ is given by the corank of the matrix $W$.

(b) The new Lagrangian $\overline{L}: T\mathcal{Q}' \to \mathbb{R}$ is obtained from $L$ by adding a new term $L'$ depending only on $q^\alpha, q^\alpha$ and their velocities

$$\overline{L}(q^r,q^\alpha,\dot{q}^r,\dot{q}^\alpha) = L(q^r,\dot{q}^r) + L'(q^a,q^\alpha,\dot{q}^a,\dot{q}^\alpha) \tag{3.2}$$

in such a way that $\overline{L} = L \circ T\pi_{\mathcal{Q}} + L'$.

In the simple case of being rank $W = N-1$ it is enough to introduce only a new variable $q^{N+1}$ and by adding the restriction of being $L'$ quadratic on the velocities, we will study

$$\overline{L}(q^r,q^{N+1},\dot{q}^r,\dot{q}^{N+1}) = L(q^r,\dot{q}^r) + F(q^N,q^{N+1})\dot{q}^N\dot{q}^{N+1} \tag{3.3}$$

which represents a coupling in $L'$ between $\dot{q}^N$ and $\dot{q}^{N+1}$ with $F$ as coupling function. The matrix $\overline{W}$ is given by

$$\begin{array}{lll} \overline{W}_{ij} = W_{ij} & \overline{W}_{iN} = W_{iN} & \overline{W}_{iN+1} = 0 \\ \overline{W}_{Nj} = W_{Nj} & \overline{W}_{NN} = W_{NN} & \overline{W}_{NN+1} = F \\ \overline{W}_{N+1j} = 0 & \overline{W}_{N+1N} = F & \overline{W}_{N+1N+1} = 0 \end{array} \tag{3.4}$$

and it is easy to check that

$$\det \overline{W} = -F^2 \det[W_{ij}] \quad i,j = 1,..., N-1 \tag{3.5}$$

and consequently $\overline{L}$ is regular whenever $F \neq 0$. The Euler–Lagrange equations for $\overline{L}$ are

$$E_i\overline{L} = W_{ir}\ddot{q}^r + V_{ir}\dot{q}^r + U_i = 0, \tag{3.6a}$$

$$E_N\overline{L} = W_{Nr}\ddot{q}^r + F\ddot{q}^{N+1} + V_{Nr}\dot{q}^r + F'_{N+1}(\dot{q}^{N+1})^2 + U_N = 0, \tag{3.6b}$$

$$E_{N+1}\overline{L} = F\ddot{q}^N + F'_N(\dot{q}^N)^2 = 0, \tag{3.6c}$$

where $F'_p = \partial F/\partial q^p$; $p = N, N+1$. The equations $E_i\overline{L} = 0$ are the same ones that $E_iL = 0$, and the equation $E_N\overline{L} = 0$ can be arranged in the form

$$Fq^{N+1} + F'_{N+1}(\dot{q}^{N+1})^2 + f'(q^r,\dot{q}^r) = 0. \tag{3.7}$$

Consequently, in order to recover the equations associated to $L$ in the restriction to $T\mathcal{Q}$, and therefore $\overline{L}$ be a regular extension, the function $F$ must verify the following conditions:

$$F(q^N,0) = 0, \quad F'_N(q^N,0) = 0. \tag{3.8}$$

Simple particular solutions are those on the form $F = (q^N)^n(q^{N+1})^m$ with $m > 0$.

In the more general situation of being $N-R > 1$, and generalizing (3.3), we will study the replacement of the singular Lagrangian $L$ by the following regularization $\overline{L}$

$$L = L(q^r,\dot{q}^r) + \sum_{a=R+1}^{N} Q^\alpha\dot{q}^a\dot{Q}^\alpha\delta^\alpha_a, \tag{3.9}$$

where we have introduced capital letters $Q^\alpha$ for the new coordinates taking the index $\alpha$ the same values as $a$. $L'$ repre-

sents a one-to-one coupling between $\dot{q}^\alpha$'s and the $\dot{Q}^\alpha$'s taking the $Q^\alpha$'s as coupling function $F$. We have for $\overline{W}$

$$\det \overline{W} = \left( \prod_\alpha Q^\alpha \right)^2 \det [W_{ij}] \tag{3.10}$$

and therefore $L$ is regular whenever that $Q^\alpha \neq 0$, $\forall \alpha$. The meaning of this fact is the following: We have introduced a new coordinate $Q$ for each old degree of freedom associated to the singularity, and thus, over those regions in which some of the $Q^\alpha$ is null, it appears over $L$ and over its equations the corresponding singularity associated to $q^a$. The equations $E_i \overline{L} = 0$ are the same ones as in Eq. (2.6) and, as far as $E_b \overline{L} = 0$ is concerned, they can be arranged in the form

$$\frac{d}{dt}(Q^\beta \dot{Q}^\beta) + f'_b(q^r, \dot{q}^r) = 0, \quad b = \beta, \tag{3.11}$$

showing that the equations of motion of the $Q$'s are coupled to the constraint functions $f'$, and that each time any of the new coordinates vanishes we recover the corresponding constraint equation. The normal form of $E_\beta \overline{L} = 0$, $\beta = R+1,\ldots, N$ is $\ddot{q}^b = 0$, $b = \beta$, but for the restriction to $T\mathcal{Q}$ these equations do not contribute and we obtain

$$\ddot{q}^k = -W^{ki}(W_{ia}\ddot{q}^a + V_{ij}\dot{q}^j + U_i), \tag{3.12a}$$

$$f'_b(q^r, \dot{q}^r) = 0 \tag{3.12b}$$

recovering the situation associated to $L$.

Although the general prescription of regularization does not provide a unique way of constructing the new regular Lagrangian $\overline{L}$ (the ambiguity being for instance the function $F$) for every such regularization the Legendre transformation $D_{\overline{L}}$ is well behaved: the pull-back $\omega_{\overline{L}}$ of the canonical 2-form $\omega_0$ on $T^*\mathcal{Q}'$ is a symplectic form on the submanifold $\mathcal{M}$ of $T\mathcal{Q}$ where the points in which someone of the new coordinates vanishes have been removed. The dynamical vector field $X_{\overline{L}}$ is defined on $\mathcal{M}$ by $i(X_{\overline{L}})\omega_{\overline{L}} = dE_{\overline{L}}$ with $E_{\overline{L}}$ being the energy function defined by $E_{\overline{L}} = \Delta(\overline{L}) - \overline{L}$, where the Euler vector field $\Delta$ is that one generating dilatations along the fibers in $T\mathcal{Q}'$. A Hamiltonian function $\overline{H}$ on $D_{\overline{L}}(\mathcal{M})$ is defined $\overline{H} = E_{\overline{L}} \circ D_{\overline{L}}^{-1}$ in such a way that the Hamiltonian dynamical systems $\{\mathcal{M}, \omega_{\overline{L}}, E_{\overline{L}}\}$ and $\{D_{\overline{L}}(\mathcal{M}), \omega_0, \overline{H}\}$ are isomorphic. The advantage of the regularization procedure is therefore that it permits us to define a well-behaved Legendre transformation and consequently to develop a Hamiltonian formalism.

## IV. EXAMPLE

As an example of the results obtained in the previous section we will construct a Lagrangian $\overline{L}$ regularizing the following singular Lagrangian

$$L = \tfrac{1}{2}(\dot{q}^1 - \dot{q}^3)^2 + \tfrac{1}{2}(\dot{q}^2 - \dot{q}^4)^2 \tag{4.1}$$

defined on the tangent bundle $T\mathcal{Q}$ of a four-dimensional configuration space $\mathcal{Q}$. It is easy to check that the Hessian matrix $W$ has rank $W = 2$, either through direct calculation or by taking into account that the singular character is an intrinsic property remaining invariant under point transformations. Particularly under the change

$$\begin{aligned} &\overline{Q}^1 = q^1 - q^3, \quad \overline{Q}^3 = q^1 + q^3, \\ &\overline{Q}^2 = q^2 - q^4, \quad \overline{Q}^4 = q^2 + q^4, \end{aligned} \tag{4.2}$$

the Lagrangian becomes

$$L = \tfrac{1}{2}(\overline{Q}^1)^2 + \tfrac{1}{2}(\overline{Q}^2)^2 \tag{4.3}$$

which manifests explicitly its singular character. In spite of its apparent simplicity this particular system reflects the general situation: All the problems related with singular Lagrangians are caused by the existence of apparently superfluous degrees of freedom.

The four equations of Euler–Lagrange are reduced to only two

$$\frac{d}{dt}(\dot{q}^1 - \dot{q}^3) = 0, \tag{4.4a}$$

$$\frac{d}{dt}(\dot{q}^2 - \dot{q}^4) = 0, \tag{4.4b}$$

and we are in the situation (1) in which there are no constraint equations since $f_r = 0$, $r = 1,\ldots,4$. It is therefore a dynamical system with four degrees of freedom governed by only two equations for the accelerations.

In the regularization we are proposing, the new configuration space $\mathcal{Q}'$ will be a six-dimensional manifold and the new Lagrangian $\overline{L}$ is defined by

$$\overline{L} = \tfrac{1}{2}(\dot{q}^1 - \dot{q}^3)^2 + \tfrac{1}{2}(\dot{q}^2 - \dot{q}^4)^2 + Q^5\dot{q}^3Q^5 + Q^6\dot{q}^4Q^6. \tag{4.5}$$

The Hessian matrix $\overline{W}$ has a determinant given by

$$\det \overline{W} = (Q^5 Q^6)^2 \tag{4.6}$$

which shows that $\overline{L}$ is regular when $Q^5 \neq 0$ and $Q^6 \neq 0$. The new equations for the accelerations are

$$\ddot{q}^1 - \ddot{q}^3 = 0, \tag{4.7a}$$

$$\ddot{q}^2 - \ddot{q}^4 = 0, \tag{4.7b}$$

$$Q^5\ddot{Q}^5 + (\dot{Q}^5)^2 = 0, \tag{4.7c}$$

$$Q^6\ddot{Q}^6 + (\dot{Q}^6)^2 = 0, \tag{4.7d}$$

$$Q^5\ddot{q}^3 = 0, \tag{4.7e}$$

$$Q^6\ddot{q}^4 = 0. \tag{4.7f}$$

These equations are perfectly defined and their number corresponds to the dimension of the manifold $\mathcal{Q}'$. In the particular situation of motions restricted to the submanifold $\mathcal{Q}$, we recover the equations associated to $L$.

## V. HAMILTONIAN TREATMENT

The Dirac theory points out that if the Lagrangian is singular there are functions $\phi_a \in \mathcal{F}(T^*\mathcal{Q})$ introducing constraint relations among the $q$'s and $p$'s

$$\phi_a(q^r, p_r) = 0 \tag{5.1}$$

in such a way that the Hamiltonian defined by

$$H_c(q^r, p_r) = \dot{q}^r \frac{\partial L}{\partial \dot{q}^r} - L(q^r, \dot{q}^r) \tag{5.2}$$

has a functional form not unique. Thus, one can always add to $H_c$ any linear combination of $\phi_a$ with coefficients $v^a$

$$H = H_c(q^r, p_r) + v^a\phi_a(q^r, p_r). \tag{5.3}$$

We suppose, without loss of generality, that (5.1) allows us to express the momenta $p_{R+1},\ldots, p_N$ in terms of $p_1,\ldots, p_R$ and the $q$'s

$$p_a = \psi_a(q^r, p_j) \tag{5.4}$$

and $H$ takes the form

$$H = H_c(q^r, p_r) + v^a\{ p_a - \psi_a(q^r, p_j)\}. \tag{5.5}$$

In the aforementioned regularization we can obtain, at least locally, the Hamiltonian $\overline{H}$ corresponding to $\overline{L}$ by means of the Legendre transformation $D_{\overline{L}}: T\mathcal{Q}' \to T^*\mathcal{Q}'$

$$\overline{H}(q^\rho, p_\rho) = \dot{q}^\rho \frac{\partial \overline{L}}{\partial \dot{q}^\rho} - \overline{L}(q^\rho, \dot{q}^\rho). \tag{5.6}$$

For obtaining the expression of $\overline{H}$ we will use the following lemma.

*Lemma:* Let $F$ be a differentiable map $F: \mathcal{Q}_1 \to \mathcal{Q}_2$ and $D_{L_i}$ $i = 1,2$, the Legendre transformations $D_{L_i}: T\mathcal{Q}_i \to T^*\mathcal{Q}_i$ defined by the Lagrangians $L_i \in \mathcal{F}(T\mathcal{Q}_i)$. If $L_1 = L_2 \circ TF$, both Legendre transformations are related by $D_{L_1} = T^*F \circ D_{L_2} \circ TF$.

*Proof:* We recall that the Legendre transformations $D_{L_i}$ are defined[7] by $D_{L_i}(q,\dot{q}) = (q, dL_{iq}(\dot{q}))$, where $L_{iq}: T_q\mathcal{Q}_i \to \mathbb{R}$ are given by $L_{iq}(\dot{q}) = L_i(q,\dot{q})$. Since $L_1 = L_2 \circ TF$, we have $L_{1q} = L_{2F(q)} \circ F_{*q}$ and therefore $dL_{1q}(\dot{q}) = dL_{2F(q)}(F_{*q}(\dot{q})) \circ dF_{*q}(\dot{q}) = dL_{2F(q)}(F_{*q}(\dot{q})) \circ F_{*q}$, where we have taken into account that $F_{*q}$ is linear and its differential does not depend on the point $\dot{q}$. The explicit expression for $D_{L_1}(q,\dot{q})$ will be

$$D_{L_1}(q,\dot{q}) = (q, dL_{2F(q)}(F_{*q}(\dot{q})) \circ F_{*q}).$$

On the other hand, the map $T^*F \circ D_{L_2} \circ TF$ is given by

$$\begin{aligned}(T^*F \circ D_{L_2} \circ TF)(q,\dot{q}) &= (T^*F \circ D_{L_2})(F(q), F_{*q}(\dot{q})) \\ &= T^*F(F(q), dL_{2F(q)}(F_{*q}(\dot{q}))).\end{aligned}$$

Finally, both expressions coincide if we recall that $T^*F$ is defined by $T^*F(F(q),\alpha) = (q, \alpha \circ F_{*q}), \forall \alpha \in T^*_{F(q)}\mathcal{Q}_2$.

The application of this lemma to $L_2 = \overline{L}$, $F = i_{\mathcal{Q}}$, and to $L_2 = L$, $F = \pi_{\mathcal{Q}}$, using in the second case the additivity of the Legendre transformations, leads to

$$D_L = T^*i_{\mathcal{Q}} \circ D_{\overline{L}} \circ Ti_{\mathcal{Q}}, \tag{5.7a}$$

$$D_{\overline{L}} = T^*\pi_{\mathcal{Q}} \circ D_L \circ T\pi_{\mathcal{Q}} + D_{L'}. \tag{5.7b}$$

The Hamiltonian $\overline{H} \in \mathcal{F}(T^*\mathcal{Q}')$ and the energy function $E_{\overline{L}} \in \mathcal{F}(T\mathcal{Q}')$ are related by $\overline{H} = E_{\overline{L}} \circ D_{\overline{L}}^{-1}$, where $E_{\overline{L}}$ is given by

$$E_{\overline{L}} = \Delta(\overline{L}) - \overline{L} = E_L + E_{L'}, \tag{5.8a}$$

$$E_L = \Delta(L \circ T\pi_{\mathcal{Q}}) - L \circ T\pi_{\mathcal{Q}}, \tag{5.8b}$$

$$E_{L'} = \Delta(L') - L', \tag{5.8c}$$

and consequently $\overline{H}$ can be written as the sum of two terms

$$\begin{aligned}\overline{H} &= E_L \circ D_{\overline{L}}^{-1} + E_{L'} \circ D_{\overline{L}}^{-1} \\ &= H_c + H'_c \circ \{1 - (T^*\pi_{\mathcal{Q}} \circ D_L \circ T\pi_{\mathcal{Q}}) \circ D_{\overline{L}}^{-1}\},\end{aligned} \tag{5.9}$$

where $H_c$ corresponds to the canonical Hamiltonian associated to $L \circ T\pi_{\mathcal{Q}}$ and $H'_c$ is related with $E_{L'}$ as usual by $H'_c \circ D_{L'} = E_{L'}$.

For the particular choice (3.9) of $L'$ we find that $H'_c(q, Q; p, P) = \Sigma(P_\alpha/Q^\alpha)p_a\delta^\alpha_a$ and therefore the decomposition (5.9) for $\overline{H}$ becomes

$$\begin{aligned}\overline{H}(q, Q; p, P) &= H_c(q^r; p_j, \psi_a) \\ &\quad + \sum_{a=r+1}^{N} \frac{P_\alpha}{Q^\alpha}(p_a - \psi_a)\delta^\alpha_a,\end{aligned} \tag{5.10}$$

where the functions $\psi_a$ are defined by $D_{L \circ T\pi_{\mathcal{Q}}} \circ D_{\overline{L}}^{-1}(q, Q; p, P) = (q^r, 0; p_j, \psi_a, 0)$, that is to say $\psi_a = (\partial L/\partial \dot{q}^a)$ in the point $T\pi_{\mathcal{Q}} \circ D_{\overline{L}}^{-1}(q, Q; p, P)$.

The separation in two parts (5.10) obtained for $\overline{H}$, shows its relation with the Hamiltonian (5.5) of the Dirac theory, since (i) the first term $H_c$ represents in $T^*\mathcal{Q}'$ the canonical Hamiltonian associated to $L$; and (ii) the remaining terms on the form $(P_\alpha/Q^\alpha)(p_a - \psi_a)\delta^\alpha_a$, where $p_a - \psi_a \in \mathcal{F}(T^*\mathcal{Q}')$ are similar to the additional terms in (5.5) introduced by the constraints, where $\psi_a$ play the role of those and $P_\alpha/Q^\alpha$ are replacing the multipliers $v^a$ causing gauge indeterminations.

In the previous example the constraints are

$$p_3 + p_1 = 0, \quad p_4 + p_2 = 0 \tag{5.11}$$

so that the Dirac Hamiltonian $H$ may now be taken to be

$$H = \tfrac{1}{2}p_1^2 + \tfrac{1}{2}p_2^2 + v^1(p_3 + p_1) + v^2(p_4 + p_2). \tag{5.12}$$

In order to have consistent equations of motion, the coefficients $v^1$ and $v^2$ must take the values

$$v^1 = \dot{q}^3, \quad v^2 = \dot{q}^4. \tag{5.13}$$

In this way, we have obtained an expression for $H$ that although it generates correct Hamilton equations is ill defined since $H$ must be a function on $T^*\mathcal{Q}$ and we have obtained a velocity dependent function.

For the regularized Lagrangian (4.5) we will obtain a Hamiltonian $\overline{H} \in (T^*\mathcal{Q}')$ correctly defined. In this case the functions $\psi_a$ take the values $\psi_3 = -p_1$, $\psi_4 = -p_2$ and thus $\overline{H}$ is given by

$$\overline{H} = \frac{1}{2}p_1^2 + \frac{1}{2}p_2^2 + \frac{P_5}{Q^5}(p_3 + p_1) + \frac{P_6}{Q^6}(p_4 + p_2). \tag{5.14}$$

It is to be remarked that we will obtain, when we make use of the Hamilton equations associated to $\overline{H}$, values for $P_5/Q^5$ and $P_6/Q^6$ in agreement with those previously found in (5.13) for the coefficients $v^1$ and $v^2$.

## VI. RELATIVISTIC POINT PARTICLE

The example studied above is fundamentally of rather academic interest. We will now analyze two cases of more practical importance. As our first application of the regularization procedure and of its relation to the Dirac approach to constrained Hamiltonian system, we examine the relativistic spinless point particle whose Lagrangian is

$$L = -m(-\dot{q}^\mu \dot{q}_\mu)^{1/2}, \tag{6.1}$$

where our metric convention is

$$-g^{00} = g^{11} = g^{22} = g^{33} = 1.$$

It is easy to check that $\det W = \det(\partial^2 L/\partial \dot{q}_\mu \partial \dot{q}_\nu) = 0$ and $\operatorname{rank} W = 3$, so that $L$ is a singular Lagrangian requiring the use of Dirac's methods to define the Hamiltonian dynamics.

The canonical momenta are

$$p^\mu = \frac{\partial L}{\partial \dot{q}_\mu} = \frac{m\dot{q}^\mu}{(-\dot{q}^2)^{1/2}} \tag{6.2}$$

and the Euler–Lagrange equations and their solutions are

$$\frac{d}{d\tau}p^\mu = 0, \tag{6.3}$$

$$p^{\mu}(\tau) = p^{\mu}(0). \tag{6.4}$$

Remark that the momentum $p^{\mu}$ is a homogeneous function of degree zero in the velocities and consequently the formulas (6.2) are not invertible. Moreover, as the Lagrangian (6.1) is a homogeneous function of degree one in the velocities, the canonical Hamiltonian vanishes identically

$$H_c = p_{\mu}\dot{q}^{\mu} - L = 0. \tag{6.5}$$

The Dirac theory assumes that the total Hamiltonian is given by

$$H = H_c + v(p^2 + m^2), \tag{6.6}$$

where $v$ is to be obtained by imposing that $H$ generates correctly the Hamilton equations of motion. In this way, expressing $v$ in terms of the velocities $\dot{q}^{\mu}$, $H$ finally becomes

$$H = (1/2m)(-\dot{q}^2)^{1/2}(p^2 + m^2). \tag{6.7}$$

This Hamiltonian being proportional to the constraint is identically null, but it results useful for working formally with it and, as it was said above, it generates correctly the equations. However it is convenient to emphasize that it depends on the velocities and consequently is ill defined.

Let us now consider the regularizing Lagrangian $\overline{L}$ defined on $T\mathcal{Q}'$, where $\mathcal{Q}'$ is the new five-dimensional configuration space

$$\overline{L} = -m(-\dot{q}_{\mu}\dot{q}^{\mu})^{1/2}(1 + Q\dot{Q}). \tag{6.8}$$

The form of $\overline{L}$, which is not exactly that proposed in (3.2), has been motivated because the homogeneity of (6.1) permit us to construct the quadratic term $L'$ by multiplying $L$ by $Q\dot{Q}$. This takes advantage of dealing on the same footing with the four coordinates $q^{\mu}$ according to the relativistic character of the theory. It is easy to check that $\det \overline{W} \neq 0$ and consequently $\overline{L}$ is regular whenever the new coordinate $Q \neq 0$.

The canonical momenta are

$$p^{\mu} = [m\dot{q}^{\mu}/(-\dot{q}^2)^{1/2}](1 + Q\dot{Q}), \tag{6.9a}$$

$$P = -m(-\dot{q}^2)Q, \tag{6.9b}$$

and the Euler–Lagrange equations

$$\frac{d}{d\tau}p^{\mu} = 0, \tag{6.10a}$$

$$Q\frac{d}{d\tau}(-\dot{q}^2)^{1/2} = 0. \tag{6.10b}$$

With regard to the Hamiltonian $\overline{H}$, it turns out that $E_{\overline{L}}$ is given by the term of $\overline{L}$ quadratic in the velocities

$$E_{\overline{L}} = E_{L'} = -m(-\dot{q}^2)^{1/2}Q\dot{Q} \tag{6.11}$$

and as $H'_c = (P/mQ)(-p^2)^{1/2}$ we obtain that, according to (5.9), and taking into account that

$$\psi^{\mu} = mp^{\mu}/(-p^2)^{1/2} \tag{6.12}$$

the Hamiltonian $\overline{H}$ is given by

$$\begin{aligned}\overline{H} &= (P/mQ)\{-(p-\psi)^2\}^{1/2} \\ &= (P/mQ)\{(-p^2)^{1/2} - m\}.\end{aligned} \tag{6.13}$$

The most important qualitative feature is that this Hamiltonian, contrary to (6.7), is perfectly well defined and generates correctly the corresponding Hamiltonian equations

$$\dot{q}^{\mu} = \{q^{\mu}, \overline{H}\} = \frac{\partial \overline{H}}{\partial p_{\mu}} = \left(\frac{P}{mQ}\right)\frac{p^{\mu}}{(-p^2)^{1/2}}, \tag{6.14a}$$

$$\dot{Q} = \{Q, \overline{H}\} = \frac{\partial \overline{H}}{\partial P} = \frac{1}{mQ}\{(-p^2)^{1/2} - m\}, \tag{6.14b}$$

$$\dot{p}^{\mu} = \{p, \overline{H}\} = -\frac{\partial \overline{H}}{\partial q_{\mu}} = 0, \tag{6.14c}$$

$$\dot{P} = \{P, \overline{H}\} = -\frac{\partial \overline{H}}{\partial Q} = \frac{P}{mQ^2}\{(-p^2)^{1/2} - m\}. \tag{6.14d}$$

The Hamiltonian $\overline{H}$ is correctly defined but it does not satisfy explicitly the property of

$$H = \overline{H} \circ T^{*}\pi_{\mathcal{Q}}. \tag{6.15}$$

In order to look for a new expression of $\overline{H}$ overcoming this problem, we will consider $\overline{H}$ not as a function defined on $T^{*}\mathcal{Q}'$, but as a function defined on the graph $G(D_{\overline{L}})$ of $D_{\overline{L}}$

$$G(D_{\overline{L}}) = \{(q,\dot{q},p) | (q,p) = D_{\overline{L}}(q,\dot{q})\}. \tag{6.16}$$

$G(D_{\overline{L}})$ is a submanifold of $T\mathcal{Q}' \circ T^{*}\mathcal{Q}'$ of dimension $2N'$ since although it depends on $3N'$ coordinates, only $2N'$ among them are independent. In this way, we will express the Hamiltonian $\overline{H}$ in an incorrect form by introducing velocity-dependent terms. Equations (6.9) show that

$$P/mQ = -(-\dot{q}^2)^{1/2}, \tag{6.17a}$$

$$(-p^2)^{1/2} - m = -(1/2m) \times \{p^{\mu}p_{\mu} + m^2 + m^2(Q\dot{Q})^2\}, \tag{6.17b}$$

so that finally $\overline{H}$ may be rewritten in the form

$$\overline{H} = (1/2m)(-\dot{q}^2)^{1/2}\{p^{\mu}p_{\mu} + m^2 + m^2(Q\dot{Q})^2\} \tag{6.18}$$

showing explicitly its relation with (6.7) and that it verifies property (6.15).

## VII. THE HOMOGENEOUS FORMALISM

The appropriate geometrical setting for the description of time-dependent systems has shown to be the so-called "homogeneous formalism" in which the time does not play any distinguished role.[8,9] The new "extended configuration space" is the space of events $\mathcal{Q} \times \mathbb{R}$. Instead of the traditional Lagrangian function $L(q,t,\dot{q})$ we will make use of the homogeneous Lagrangian function $L^{h}(q,t;\dot{q}',w)$ defined on $T(\mathcal{Q} \times \mathbb{R})$ by means of the relation $L^{h}(q,t;q',w) = wL(q,t,q'/w)$ which is homogeneous of degree one in the velocities, that property guaranteeing that the corresponding variational problem for the accelerations is well defined. This homogeneity property however implies that the corresponding homogeneous Hamiltonian vanishes identically. If $L$ is assumed to be regular, the regularization prescription amounts to taking as the regularizing Lagrangian $\overline{L}(q,t,Q;q',w,\dot{Q}) = L^{h}(q,t;q',w) + wQ\dot{Q}$. The Legendre transformation $D_{\overline{L}}: T\mathcal{Q}' \to T^{*}\mathcal{Q}'$ is well behaved and the Hamiltonian $\overline{H}$ is given by

$$\overline{H}(q,t,Q;p,\pi,P) = \frac{P}{Q}(H + \pi), \tag{7.1}$$

where $H$ is the usual Hamiltonian corresponding to $L$. Notice that the Legendre transformation is

$$p = \left(\frac{\partial L^{h}}{\partial q'}\right)_{(q,t;q',w)} = \left(\frac{\partial L}{\partial \dot{q}}\right)_{(q,t,q'/w)}, \tag{7.2a}$$

$$\pi = \left(\frac{\partial L^{h}}{\partial w}\right)_{(q,t;q',w)} + Q\dot{Q}$$

$$= \left\{ L(q,t,q'/w) - (q'/w)\left(\frac{\partial L}{\partial \dot{q}}\right)_{(q,t,q'/w)} \right\} + Q\dot{Q}\,, \tag{7.2b}$$

$$P = wQ\,, \tag{7.2c}$$

and therefore, when the time is used as parameter $w = 1$ the Hamiltonian $\overline{H}$ reduces to the Hamiltonian $H + \pi$ which has recently been used[10] for studying a symplectic formulation of time-dependent systems allowing a generalization of the concept of canonical transformation for such systems.

## VIII. CONCLUSIONS AND OUTLOOK

We have developed a method for regularizing singular Lagrangians, not for seeking improvement of Dirac's treatment, but for analyzing alternative approaches of dealing with them. This method has permitted us to obtain well-defined Hamiltonians and it has been shown to be useful in some practical examples, particularly in the case of homogeneous Lagrangians of degree one where the canonical Hamiltonian vanishes identically. Moreover the Hamiltonian obtained in the case of the homogeneous formalism is the same one used by Asorey *et al.*[10] in their geometrical description of time-dependent systems.

This paper must be considered as a first attempt at looking in a different way at the theory of singular Lagrangians and some features are deserving a deeper analysis. The study of the physical meaning of the new coordinates introduced in this geometrical method remains, therefore, as an interesting task, especially in connection with the usual gauge functions. This should be a convenient step for the generalization to systems with infinite degrees of freedom.

As a final remark, we must comment on the nonuniqueness of the regularizing Lagrangian $\overline{L}$, a situation that seems to be similar, in a certain sense, to the nonuniqueness of the blowing-up of an algebraic variety.[5] It is well known that the Lagrangian describing a mechanical system is not unique. Not only point transformations and gauge equivalence of Lagrangians are allowed, but there may be two non-gauge-equivalent regular Lagrangians leading to the same equations of motion, i.e., defining the same dynamical vector field, but different symplectic 2-forms. Therefore, in these formulations, two Lagrangians $\overline{L}_1$ and $\overline{L}_2$ defining dynamical vector fields that although different on $T\mathcal{Q}'$ are related on $T\mathcal{Q}$ in the sense that both give the same equations in this submanifold, are also to be considered as equivalent for the singular system. The relation between this equivalence and those[4] introduced by foliations in the usual approach remains to be studied.


## ACKNOWLEDGMENTS

We would like to thank L. A. Ibort for fruitful discussions and the Instituto de Estudios Nucleares of Madrid for partial financial support.



[1] P. A. M. Dirac, Canad. J. Math. **2**, 129 (1950).
[2] P. A. M. Dirac, *Lectures on Quantum Mechanics* (Belfer Graduate School of Science, Yeshiva University, New York, 1964).
[3] A. Hanson, T. Regge, and C. Teitelboim, *Constrained Hamiltonian Systems* (Academia Nazionale dei Lincei, Rome, 1976).
[4] G. M. Marmo, N. Mukunda, and J. Samuel, Riv. Nuovo Cimento **6** (2), 1 (1983).
[5] R. Hartshorne, *Algebraic Geometry* (Springer-Verlag, New York, 1977).
[6] S. Hojman and H. Harleston, J. Math. Phys. **22**, 1414 (1981).
[7] R. Abraham and J. E. Marsden, *Foundations of Mechanics* (Benjamin, Reading, MA, 1978), 2nd ed.
[8] J. Hajdu, G. Gyorgyi, and Th. Kahan, *Symétries et groupes en Mécanique Classique Prérelativiste*, in *Théorie des groupes en Physique Classique et Quantique*, edited by Th. Kahan (Dunod, Paris, 1971), Vol. 2.
[9] N. Woodhouse, *Geometric Quantization* (Clarendon, Oxford, 1980).
[10] M. Asorey, J. F. Cariñena, and L. A. Ibort, J. Math. Phys. **24**, 2745 (1983).

# Anharmonic oscillators and generalized hypergeometric functions

J. P. Codaccioni and R. Caboz
*Laboratoire de Physique Appliquée, Faculté des Sciences, Avenue Louis Sallenave, 64000 PAU, France*

(Received 15 April 1983; accepted for publication 6 January 1984)

A large class of anharmonic oscillators represented by the Hamiltonian $H(q,p)=\frac{1}{2}p^2+\frac{1}{2}\omega_0^2q^2+\lambda q^\alpha$ ($\alpha$ integer $>2$) is considered. Owing to an integration technique using the Lagrange–Bürmann theorem, one can give for the period and the action integral of bounded motions closed expressions in terms of energy in the form of very simple generalized hypergeometric functions. Finally an application of the method is given for doubly anharmonic oscillators.

PACS numbers: 03.20. + i, 02.30. + g

The Hamiltonian $H(q,p)=\frac{1}{2}p^2+\frac{1}{2}\omega_0^2q^2+\lambda q^\alpha$ ($\omega_0,\lambda\in\mathbb{R}$ and $\alpha$ integer $>2$) concerns an important class of nonlinear conservative dynamical systems with one degree of freedom, the Duffing oscillator ($\alpha=4$, $\lambda>0$) being the most famous example. The main characteristic of the motion is the first integral of energy $H(p,q)=E$.

The potential function $U(q)=\frac{1}{2}\omega_0^2q^2+\lambda q^\alpha$ enables us to define a potential well by the conditions $|q|<|q_c|$ and $0<E<E_c$, $q_c$ and $E_c$ being the coordinates of the critical point ($s$) of the function U.[1]

In this paper we shall be concerned with bounded motion only, i.e., motion "inside" the potential well represented in phase space $\{p,q\}$ by a closed curve.

## I. THE PERIOD OF THE ANHARMONIC OSCILLATOR

It is well known[2] that the motion is periodic with period

$$T(E)=\sqrt{2}\int_0^E du(E-u)^{-1/2}\left[\frac{dq_2}{du}-\frac{dq_1}{du}\right], \tag{1}$$

where $q_1(u)$ and $q_2(u)$ are the two real roots of the algebraic equation $U(q)-u=0$ lying in the potential well and ordered such that $q_1<0<q_2$. If we introduce the equation

$$z_i=\frac{2u}{\omega_0^2}-(-1)^{\alpha i}\frac{2\lambda}{\omega_0^2}z_i^{\alpha/2},\quad z_i\in\mathbb{R},\quad i=1,2, \tag{2}$$

then $q_1$ and $q_2$ are given by

$$q_i=(-1)^i z_i^{1/2},\quad i=1,2. \tag{3}$$

Equation (2) can be solved in terms of a power series using Lagrange's theorem.[3] Let us recall that for $\phi$ analytic on and inside a contour $C$ surrounding a point $a$, and $t$ satisfying

$$|t\phi(z)|<|z-a| \tag{4}$$

for all points $z$ on $C$, then the equation

$$z=a+t\phi(z) \tag{5}$$

has one root $z$ in the interior of $C$, and further any function $f$ analytic on and inside $C$ has a convergent power series in $t$ given by the Lagrange formula

$$f(z)=f(a)+\sum_{n=1}^{\infty}\frac{t^n}{n!}\frac{d^{n-1}}{da^{n-1}}[f'(a)\{\phi(a)\}^n]. \tag{6}$$

In the case of Eq. (2)

$$a=2u/\omega_0^2,\quad t\phi(z)=-(-1)^{\alpha i}(2\lambda z^{\alpha/2}/\omega_0^2),\quad z\in\mathbb{R}, \tag{7}$$

and the condition (4), written for all $u\in[0,E]$ as required by the integral of Eq. (1), leads to

$$E<E_0,\quad E_0=[(\alpha-2)/2](\omega_0^{2\alpha}/\lambda^2\alpha^\alpha)^{1/(\alpha-2)}. \tag{8}$$

What is the mechanical meaning of condition (8)? Except for the case $\alpha$ even and $\lambda>0$ for which the potential well is infinite (we shall return to this case later), $E_0=E_c$, and the condition (8) is merely the condition for the motion to be bounded.

Using (6) with $f(z)=z^{1/2}$ one can derive for $q_i$ the convergent power series

$$q_i(u)=(-1)^i\sum_{n=0}^{\infty}Q_{in}u^{(n\alpha-2n+1)/2},\quad u\in[0,E[, \tag{9}$$

where

$$Q_{in}=\frac{(-1)^{n(1+\alpha i)}2^{(n\alpha+1)/2}\lambda^n\Gamma(n\alpha/2+\frac{1}{2})}{\omega_0^{n\alpha+1}n!(n\alpha-2n+1)\Gamma([\alpha/2-1]n+\frac{1}{2})}. \tag{10}$$

Derivation and integration term by term of Eqs. (9) and (1) successively (the condition $E<E_0$ ensuring the uniform convergence of the series under consideration) leads, after some algebraic manipulations, to the following expression for the period:

$$T(E)=\frac{2\pi}{\omega_0}\,{}_{\alpha/2}F_{\alpha/2-1}\left[\begin{matrix}\frac{1}{\alpha},\frac{3}{\alpha},\cdots,\frac{\alpha-1}{\alpha};\\ \\ \frac{2}{\alpha-2},\frac{4}{\alpha-2},\cdots,\frac{\alpha-2}{\alpha-2};\end{matrix}\ x\right],\quad \alpha\ \text{even} \tag{11}$$

 0022-2488/84/082436-03$02.50  

$$T(E)=\frac{2\pi}{\omega_0}\ {}_{\alpha}F_{\alpha-1}\left[\begin{matrix}\frac{1}{2\alpha},\frac{3}{2\alpha},\cdots,\frac{2\alpha-1}{2\alpha};\\ \\ \frac{1}{\alpha-2},\frac{2}{\alpha-2},\cdots,\frac{\alpha-2}{\alpha-2},\frac{1}{2};\end{matrix}\ x^2\right],\quad \alpha \text{ odd},$$

where

$$x=-\frac{\lambda}{|\lambda|}\left(\frac{E}{E_0}\right)^{(\alpha-2)/2}, \tag{12}$$

and

$${}_{a}F_{b}\left[\begin{matrix}(A)\\(B)\end{matrix};x\right]$$

denotes the generalized hypergeometric series.[4,5] These series are convergent for $|x|<1$, i.e., for $E<E_0$. As previously stated, this radius of convergence has no mechanical meaning in the case of an infinite potential well ($\lambda>0$, $\alpha$ even), but in that case the hypergeometric series may be analytically continued[4,5] for $E>E_0$.

Then if we denote by the same symbol the hypergeometric series and its possible analytic continuation, Eq. (11) gives in any case of bounded motion the right formulas for the period as function of energy.[6]

Up to now, with a few exceptions,[7] the period $T$ and pulsation $\omega=2\pi/T$ of the motion were known only in the form of asymptotic expansions limited to a few terms and obtained via perturbative methods.[8–10]

## II. ACTION AND CANONICAL FORM OF THE HAMILTONIAN

We now try to express the action integral $I=(1/2\pi)\oint p\,dq$. Here we have

$$I(E)=\frac{1}{2\pi}\int_0^E T(E')dE'. \tag{13}$$

Owing to the identity

$$\int_0^x x'^{\nu}{}_{a}F_{b}\left[\begin{matrix}(A);\\(B);\end{matrix}cx'^{\mu}\right]dx'=\frac{x^{\nu+1}}{\nu+1}\ {}_{a+1}F_{b+1}\left[\begin{matrix}(A); & \frac{\nu+1}{\mu};\\ \\ (B); & \frac{\nu+1}{\mu}+1;\end{matrix}\ cx^{\mu}\right],$$

we get

$$I(E)=\frac{E}{\omega_0}\ {}_{\alpha/2}F_{\alpha/2-1}\left[\begin{matrix}\frac{1}{\alpha},\frac{3}{\alpha},\cdots,\frac{\alpha-1}{\alpha};\\ \\ \frac{4}{\alpha-2},\frac{6}{\alpha-2},\cdots,\frac{\alpha-2}{\alpha-2},\frac{\alpha}{\alpha-2};\end{matrix}\ x\right],\quad \alpha \text{ even},$$

$$I(E)=\frac{E}{\omega_0}\ {}_{\alpha}F_{\alpha-1}\left[\begin{matrix}\frac{1}{2\alpha},\frac{3}{2\alpha},\cdots,\frac{2\alpha-1}{2\alpha};\\ \\ \frac{2}{\alpha-2},\frac{3}{\alpha-2},\cdots,\frac{\alpha-2}{\alpha-2},\frac{\alpha-1}{\alpha-2},\frac{1}{2};\end{matrix}\ x^2\right],\quad \alpha \text{ odd}. \tag{14}$$

By reversion of these series, one can obtain the Hamilton function in its canonical form $H=H(I)$. Unfortunately one cannot express this function in closed form. Nevertheless the above-mentioned method to reach the canonical form of the Hamiltonian is often easier to implement than perturbation techniques based on Lie transforms[11] or the Kolmogorov super-convergent method.[12]

## III. APPLICATION: A DOUBLY ANHARMONIC OSCILLATOR

The method used in this paper may be easily generalized to systems with polynomial nonlinearity. For instance, in the case of a doubly anharmonic oscillator described by the Hamilton function

$$H = \tfrac{1}{2}p^2 + \tfrac{1}{2}\omega_0^2 q^2 + \lambda q^4 + \mu q^6, \tag{15}$$

one can express the period and the action integral in the form of a double power series of the Horn type[4,5]:

$$T(E) = \frac{2\pi}{\omega_0}\sum_{m,n} A_{mn}\frac{x^m}{m!}\frac{y^n}{n!},$$
$$I(E) = \frac{E}{\omega_0}\sum_{m,n} B_{mn}\frac{x^m}{m!}\frac{y^n}{n!}, \tag{16}$$

where

$$x = -4\lambda E^2/\omega_0^4, \qquad y = -8\mu E^4/\omega_0^6 \tag{17}$$

and

$$A_{mn} = 2\Gamma(2m + 3n + \tfrac{1}{2})/\Gamma(\tfrac{1}{2})(m+n)!,$$
$$B_{mn} = A_{mn}/(2m + 4n + 1). \tag{18}$$

In the absolute plane ($\xi = |x|, \eta = |y|$) the domain of convergence of these series is delimited by the $\xi$ and $\eta$ axis and the curve whose parametric equations are

$$\xi = (1+2t)/(2+3t)^2, \quad \eta = t(1+2t)^2/(2+3t)^3. \tag{19}$$

## IV. CONCLUSION

In this paper we have been able to express, in terms of energy, the period and the action integral for anharmonic oscillators with polynomial nonlinearity. We note that the compact relationship between the action and energy enables, in a very simple way, a semiclassical quantization of the system


## ACKNOWLEDGMENTS

The authors would like to thank H. Exton and P. Claverie for helpful discussions on the subject.


[1]For $\alpha$ even and $\lambda > 0$ the potential well is infinite: $q_c, E_c \to \infty$.


[2]L. D. Landau and E. M. Lifshitz, *Mechanics* (Pergamon, New York, 1960), Sec. 12.

[3]E. T. Whittaker and G. N. Watson, *A Course of Modern Analysis* (Cambridge U.P., London, 1980), Sec. 7.32.

[4]A. Erdelyi, *Higher Transcendental Functions* (McGraw-Hill, New York, 1953), Vol. I.

[5]H. Exton, *Multiple Hypergeometric Functions and Applications* (Ellis Horwood, Chichester, 1976).


[6]Starting from the expression $T(E) = \int \delta(H - E)dp\,dq$, Eqs. (11) were first derived in an heuristic way by a formal Taylor expansion around the unperturbed Hamiltonian $H_0 = \frac{1}{2}p^2 + \frac{1}{2}\omega_0^2 q^2$ of the $\delta$ distribution concentrated on the phase-space orbit $H(p,q) = E$. [R. Caboz and P. Poletti, C. R. Acad. Sci. Paris **296**, 1633 (1983); R. Caboz and J. F. Loiseau, *ibid.* **296**, 1753 (1983); R. Caboz in *Journées d'Analyse Non Linéaire* (Publications de l'U.E.R. Math. pures et appliquées, Lille, 1983), Vol. I, Sec. 2.]The authors are grateful to the referee of this paper for having suggested a more rigorous path.


[7]See L. Pesquera and P. Claverie, J. Math. Phys. **23**, 1315 (1982).

[8]A. H. Nayfeh, *Perturbation Methods* (Wiley, New York, 1973).

[9]A. H. Nayfeh and D. T. Mook, *Non-Linear Oscillations* (Wiley, New York, 1979).

[10]P. Hagedorn, *Non-Linear Oscillations* (Clarendon, Oxford, 1981).

[11]G. E. O. Giagaglia, *Perturbation Methods in Non-Linear Systems* (Springer, New York, 1972).

[12]A. N. Kolmogorov, Dokl. Akad. Nauk SSSR **98**, 527 (1954).

# Hamilton–Jacobi theory for constrained systems

D. Dominici and G. Longhi
*Dipartimento di Fisica, Università di Firenze, Istituto Nazionale di Fisica Nucleare, Sezione di Firenze, Firenze, Italy*

J. Gomis and J. M. Pons
*Departament de Física Teòrica, Universitat de Barcelona, Barcelona, Spain*

(Received 13 October 1982; accepted for publication 6 January 1984)

We extend the Hamilton–Jacobi formulation to constrained dynamical systems. The discussion covers both the case of first-class constraints alone and that of first- and second-class constraints combined. The Hamilton–Dirac equations are recovered as characteristic of the system of partial differential equations satisfied by the Hamilton–Jacobi function.

PACS numbers: 03.20. + i, 03.30. + p, 12.90. + b

## I. INTRODUCTION

In recent years the old problem of a consistent formulation of relativistic Hamiltonian dynamics, for a system of particles in direct interaction, has received new interest.[1,2] It is now widely recognized that, in order to establish manifest covariance, it is convenient to look for a formulation in terms of constrained dynamics, where the constraints on the phase-space variables guarantee the correct number of physical degrees of freedom, and essentially contain the dynamics of the system under consideration.

From this point of view, relativistic Hamiltonian dynamics can be seen as a theory of systems constrained in phase space. Such a theory was developed by Dirac[3] and reformulated in various respects by others,[4] so by now it seems to be a well-developed subject.

The corresponding Hamilton–Jacobi theory has been the subject of various papers, mainly with regard to theoretical field applications.[5] What seems to be lacking in the literature on this subject, in the opinion of the present authors, is a unified systematic approach to the Hamilton–Jacobi theory for a system of particles, especially with regard to the possible applications in the presence of second-class constraints.

For this reason, the present work reviews the Hamilton–Jacobi method in a systematic and almost didactic way, presenting in some detail even those topics which are well known from the mathematical literature on systems of partial differential equations.

In order to give a unified approach for both first- and second-class constraints, the concept of (Cauchy) characteristic vectors will be used, so that in both cases the equations of motion will appear as characteristic equations. In this way Dirac's bracket structure will emerge naturally, and the integrability conditions of the characteristic system will be explicitly verified by using the properly generalized Jacobi identity.

With regard to this last point, the Mayer identity (that is, the generalization of the Jacobi identity to nonhomogeneous Poisson brackets) appears to hold for Dirac's brackets also, as has been verified by explicit calculation.

The analysis of the characteristic equations of the given set of constraints, though interesting by itself, is only preliminary to a Hamilton–Jacobi approach. The latter will be given, in the first-class case, by a review of Jacobi's method of integration.

In the second-class case the integrability conditions for the existence of the Hamilton–Jacobi function are not satisfied, so this function, as a function on all phase-space space, does not exist. Nevertheless, as we shall show in Sec. V, in a number of interesting cases, to which, in principle, the problem may always be reduced, the Hamilton–Jacobi method can be used fruitfully to get the solution of the equations of motion.

Since the authors were especially interested in this case, Sec. V is to be considered the central section of this paper. In that section two possible methods are discussed.

The concept of weak equality, first introduced by Dirac,[3] is used throughout this paper. A careful discussion of its extension to vector fields and differential forms is presented in Appendices A and B.

No applications are discussed in the present work, but the method is best suited for the formulation given by Todorov[6] and Komar[7] of the dynamics of a system of particles, and an application can be found in Ref. 8, where the Hamilton–Jacobi function is calculated for a two-body system. An interesting topic which has not been discussed here, but which deserves further investigation, is the study of symmetries from this point of view.

The paper is organized as follows: In Sec. II we make a general discussion of the extension of the Hamilton–Jacobi theory to constrained systems. Section III is devoted to the study of the first-class constraints systems and Secs. IV and V to the case where second-class constraints are also present. In Sec. VI we show how we can recover the Hamilton–Dirac equations.[3] In Appendix A the problem of classification of constraints into first- and second-class constraints is studied. Finally in Appendices B and C we prove some useful geometric results for our discussion. In Appendix D the transformation properties of the Hamilton–Jacobi function under canonical transformations are briefly reviewed.

## II. A GENERAL DISCUSSION OF THE HAMILTON–JACOBI THEORY FOR CONSTRAINED SYSTEMS

Let us assume that a dynamical system is described in terms of a canonical Hamiltonian $H_c(x^i, p_i)$, where the set

 0022-2488/84/082439-14$02.50  

$(x^i, p_i)$, $i = 1,...,n$, denotes the phase-space variables, with the following Poisson brackets:

$$\{x^i, p_j\} = \delta^i_j \tag{2.1}$$

and a set of constraints $\phi_\rho(x^i, p_i) = 0, \rho = 1,...,m \leqslant n$, which in general will be of both first and second class, in the sense used by Dirac.[3]

If a Lagrangian exists, the canonical Hamiltonian $H_c$ is known and the functions $\phi_\rho$ will be a consequence of the analysis of the Lagrangian equations of motion,[4] but more generally we may assume that the dynamical system is given in terms of the set of functions $H_c$ and $\phi_\rho$.[9]

The Hamilton–Jacobi equations for the given system are expected to be in the form

$$\begin{aligned} &\frac{\partial S}{\partial x^0} + H_c(x^i, p_j) = 0 \\ &\qquad\qquad (q = 1,...,m), \\ &\phi_\rho(x^i, p_j) = 0 \end{aligned} \tag{2.2}$$

where

$$p_i = \frac{\partial S}{\partial x^i}, \tag{2.3}$$

and $S = S(x^i)$ is the Hamilton–Jacobi function.

In the next part of this section we will analyze the set of Eqs. (2.2) from the point of view of the general theory of partial differential equations (PDE) (of first order).

Let us write Eqs. (2.2) in a more compact notation. To this end, it is more convenient to work in an enlarged phase space $(x^\alpha, p_\beta)$, $\alpha, \beta = 0,1,...,n$, where

$$p_0 = \frac{\partial S}{\partial x^0} \tag{2.4}$$

and

$$\{x^0, p_0\} = 1, \tag{2.5}$$

and we will write the set of Eqs. (2.2) as

$$\begin{aligned} &\phi_\rho(x^\alpha, p_\alpha) = 0 \\ &(\alpha = 0,1,...,n, \quad \rho = 0,1,...,m), \end{aligned} \tag{2.6}$$

where

$$\phi_0(x^\alpha, p_\alpha) = p_0 + H_c(x^i, p_i), \tag{2.7}$$

and

$$p_\alpha = \frac{\partial S}{\partial x^\alpha}. \tag{2.8}$$

So the set of equations we want to study is the set (2.6). The choice (2.7) for $\phi_0$ will be reserved for the cases where the canonical Hamiltonian $H_c$ is different from zero. It is necessary to put some restriction on the functions (2.6), in order to develop the subsequent theory. We make the following assumption about the rank[10]:

$$\operatorname{rank}\left|\left| \frac{\partial \phi_\rho}{\partial x^\alpha}, \frac{\partial \phi_\rho}{\partial p_\alpha} \right|\right| = m + 1. \tag{2.9}$$

It is well known from the theory of PDE[12] that if $S(x)$ is a solution of Eqs. (2.6), it must also be a solution of the equations

$$\{\phi_\rho(x, p), \phi_\sigma(x, p)\} = 0 \quad (\rho, \sigma = 0,1,...,m). \tag{2.10}$$

When $\phi_\rho$ are a set of first-class constraints, Eqs. (2.10) are satisfied by virtue of Eq. (2.6), but when $\phi_\rho$ include second-class constraints, this is no longer true. In this last case we have to add the lhs of Eq. (2.10) as new equations, and continue the procedure until we get a complete system of PDE[12] or a set of equations which are inconsistent. (A typical case of this last situation is when the Poisson bracket of two $\phi_\rho$ is a constant.)

If we get a complete system, the Hamilton–Jacobi function for this new set will describe a dynamical system different from the original one which was required to satisfy Eq. (2.6) only. The new system will have more constraints, and will describe a completely different physical situation.

As a consequence, when second-class constraints are present, we cannot consider the set (2.6) as a set of Hamilton–Jacobi equations. Nevertheless, it is known that the characteristic system exists, is completely integrable, and gives the usual Hamilton–Dirac equations of motion[3] for the system.

In order to construct the characteristic system, we can substitute Eqs. (2.6) with the following exterior differential system $A$:

$$A = \begin{cases} \phi_\rho \equiv \phi_\rho(x, p) = 0 \\ \theta \equiv dS - p_\alpha \, dx^\alpha = 0 \end{cases} \tag{2.11}$$

defined in the space $\mathbb{R}^{2(n+1)+1}$ of $(2(n+1)+1)$-tuples

$$(\xi^a;\ a = 1,...,2n+3) \equiv (x^0, x^1, ..., x^n, S, p_0, p_1, ..., p_n).$$

According to the usual procedure,[13] we consider the closure of $A$, $\overline{A}$:

$$\overline{A} = \begin{cases} \phi_\rho = 0, \quad d\phi_\rho = \left(\dfrac{\partial \phi_\rho}{\partial x^\alpha} + p_\alpha \dfrac{\partial \phi_\rho}{\partial S}\right) dx^\alpha \\ \qquad\qquad + \dfrac{\partial \phi_\rho}{\partial p_\alpha} dp_\alpha = 0, \\ \theta = 0, \quad d\theta = -dp_\alpha \wedge dx^\alpha = 0. \end{cases} \tag{2.12}$$

The characteristic system $C$ of $A$ is the associated Pfaff system of the set $\{\theta = 0,\ d\theta = 0,\ d\phi_\rho = 0\}$ and the equations $\phi_\rho = 0$.

As is shown in Appendix C, for the study of this characteristic system it is convenient to consider the space $Q$ of the vector fields satisfying

$$i_v I_A \subsetneq I_A, \tag{2.13}$$

where $I_A$ is the ideal generated by the differential forms $\{\theta, d\theta, d\phi_\rho\}$,[13] and where the notation $\subsetneq$ means weak inclusion as defined in Appendix C.

By considering Eqs. (2.12) for each of the forms $d\theta$, $\theta$, and $d\phi_\rho$, we get the following results:

$$i_v \, d\theta \approx \lambda^\rho d\phi_\rho + \mu\theta, \tag{2.14a}$$

$$i_v \, \theta \approx 0, \tag{2.14b}$$

$$i_v \, d\phi_\rho \approx 0, \tag{2.14c}$$

where the notation $\approx$ means weak equality as defined in Appendix B.

Using the representation of the vector field in the local coordinates $\xi^a$,

$$v \equiv v^a \frac{\partial}{\partial \xi^a} = v^\alpha \frac{\partial}{\partial x^\alpha} + v^{n+1} \frac{\partial}{\partial S} + u_\alpha \frac{\partial}{\partial p_\alpha} \tag{2.15}$$

in Eq. (2.14a), we get

$$v^\alpha\, dp_\alpha - u_\alpha\, dx^\alpha \approx \sum_{\rho=0}^{m} \lambda^\rho(\xi)\, d\phi_\rho(\xi) \tag{2.16}$$

and

$$\mu \approx 0,$$

due to the fact that the lhs of Eq. (2.14a) does not contain $dS$.

Taking into account the explicit expression of $d\phi_\rho$ [Eq. (2.12)], we have[14]

$$v^\alpha = \lambda^\rho \frac{\partial \phi_\rho}{\partial p_\alpha}, \quad u_\alpha = -\lambda^\rho \frac{\partial \phi_\rho}{\partial x^\alpha}. \tag{2.17}$$

Equations (2.14b) and (2.14c) give the conditions

$$v^{n+1} = p_\alpha v^\alpha, \tag{2.18}$$

$$v^\alpha \frac{\partial \phi_\rho}{\partial x^\alpha} + u_\alpha \frac{\partial \phi_\rho}{\partial p_\alpha} \approx 0, \tag{2.19}$$

which, together with Eq. (2.17), determine the components of the characteristic vector fields. Using Eq. (2.17) in Eq. (2.19), we get

$$\sum_{\sigma=0}^{m} \lambda^\sigma \{\phi_\rho, \phi_\sigma\} \approx 0, \tag{2.20}$$

which is a necessary and sufficient condition for the existence of characteristic vector fields.

Let us then discuss Eq. (2.20). We have to consider three different cases corresponding to the value of the rank

$$r = \operatorname{rank} \| \{\phi_\rho, \phi_\sigma\} \| \quad (\rho,\sigma = 0,\ldots,m) \tag{2.21}$$

calculated on the manifold defined by Eqs. (2.6):

(I) $r = 0$: $\lambda^0,\ldots,\lambda^m$ arbitrary;

(II) $0 < r < m+1$: $(m+1-r)$ $\lambda^\rho$ arbitrary $(\rho = 0,1,\ldots,m-r)$;

(III) $r = m+1$: $\lambda^0 = \cdots = \lambda^m = 0$ $(m+1$ even).

Let us recall that $r$ must be even, being the rank of an antisymmetric matrix.

According to the Dirac terminology, case (I) corresponds to a set of first-class constraints, case (II) to a set of first- and second-class constraints, and case (III) to second-class constraints.

With regard to case (II), the problem arises as to the classification of $\phi_\rho$ in first- and second-class constraints. This problem is solved in Appendix A.

In order to construct the characteristic system, it is necessary to consider the space $Q^*$ of 1-forms weakly incident to $\tilde{Q}$, i.e.,

$$i_v \bar{\theta} \approx 0, \quad \bar{\theta} \in Q^*, \quad v \in \tilde{Q}. \tag{2.22}$$

It is obvious that $\dim Q^* = 2n + 3 - (m + 1 - r)$.

As suggested in Ref. 13, if the characteristic vectors are defined by a system of equations of the form

$$v^a b_a^{\cdot b} = 0 \quad (a,b = 1,\ldots,2n+3) \tag{2.23}$$

for some $b_a^{\cdot b}$, then the 1-forms

$$\bar{\theta}^a = dx^b\, b_b^{\cdot a} \tag{2.24}$$

satisfy Eqs. (2.22). Let us observe that the forms (2.24) will not all be independent, as we will verify in the following. Nevertheless, using Eqs. (2.24) we will get, with a proper choice, a basis for $Q^*$.

## III. THE CHARACTERISTIC SYSTEM AND THE HAMILTON–JACOBI FUNCTION FOR A SET OF FIRST-CLASS CONSTRAINTS

In this case, corresponding to $r = 0$, all the $\lambda$'s being arbitrary, a basis for the space of the characteristic vectors $\tilde{Q}$ is given by

$$v_\rho = \frac{\partial \phi_\rho}{\partial p_\alpha}\left(\frac{\partial}{\partial x^\alpha} + p_\alpha \frac{\partial}{\partial S}\right) - \frac{\partial \phi_\rho}{\partial x^\alpha}\frac{\partial}{\partial p_\alpha} \quad (\rho = 0,\ldots,m), \tag{3.1}$$

which are linearly independent due to the hypothesis (2.9) on the rank of the matrix of the components of $v_\rho$.

By using the nonhomogeneous Poisson bracket[15]

$$\{f,g\}_{nh} = \left(\frac{\partial f}{\partial x^\alpha} + p_\alpha \frac{\partial f}{\partial S}\right)\frac{\partial g}{\partial p_\alpha} - \frac{\partial f}{\partial p_\alpha}\left(\frac{\partial g}{\partial x^\alpha} + p_\alpha \frac{\partial g}{\partial S}\right), \tag{3.2}$$

we can write the characteristic vectors [Eq. (3.1)] as

$$v_\rho = -\{\phi_\rho, \cdot\}_{nh}. \tag{3.3}$$

It is now possible to reduce the search for a solution of the system of nonlinear PDE (2.6) to a system of linear homogeneous PDE. In fact, if we consider the system

$$v_\rho(g) = -\{\phi_\rho, g\}_{nh} = 0, \tag{3.4}$$

where $g: R^{2n+3} \to R$, this is a system of linear homogeneous equations which is completely integrable. In fact, using the Mayer identity (Ref. 16, p. 172):

$$\{f_1, \{f_2, f_3\}_{nh}\}_{nh} + \text{cyclic} = -\frac{\partial f_1}{\partial S}\{f_2 f_3\}_{nh} + \text{cyclic}, \tag{3.5}$$

we have

$$[v_\rho, v_\sigma](g) \approx -\{g, \{\phi_\rho, \phi_\sigma\}_{nh}\}_{nh} - \frac{\partial \phi_\rho}{\partial S}\{\phi_\sigma, g\}_{nh} - \frac{\partial \phi_\sigma}{\partial S}\{g, \phi_\rho\}_{nh} - \frac{\partial g}{\partial S}\{\phi_\rho, \phi_\sigma\}_{nh}, \tag{3.6}$$

and using the fact that $\phi_\rho$ are independent of $S$ and that they are a set of first-class constraints:

$$\{\phi_\rho, \phi_\sigma\} = c_{\rho\sigma\tau}\phi_\tau, \tag{3.7}$$

we finally get

$$[v_\rho, v_\sigma] \approx c_{\rho\sigma\tau} v_\tau \quad (\rho,\sigma,\tau = 0,\ldots,m), \tag{3.8}$$

and, using the results of Appendix C, the system (3.4) is completely integrable over the surface of the constraints.

Thus the system (3.4) has $(2n+3) - (m+1)$ independent solutions depending on $x$, $p$, and $S$; $m+1$ of these solutions are nothing but the $\phi_\rho$'s, which are functions only of $x$ and $p$.

Among the solutions of the system (3.4), we can choose one, which we will call $G_1(x,p,S)$, independent of $\phi_\rho$. Then we can add to the set $\phi_\rho = 0$ the new equations $G_1 = C_1$, where $C_1$ is an arbitrary constant, and consider the new set

$$v_\rho(g)=0, \quad v_{G_1}(g)=0. \tag{3.9}$$

We can proceed in this way until we have extended the original set of $m+1$ equations to the new set

$$v_\rho(g)=0, \quad v_{G_1}(g)=0, \quad \ldots, \quad v_{G_{n-m}}(g)=0, \tag{3.10}$$

which is completely integrable and has $2n+3-(m+1+n-m)=n+2$ independent solutions, of which $n+1$ are already known: $\phi_0,\ldots,\phi_m, G_1,\ldots,G_{n-m}$. So the system (3.10) still has one new solution, $G_{n-m+1}$. In this way we get an involutory system $\phi_0,\ldots,\phi_m$, $G_1-C_1,\ldots,G_{n-m+1}-C_{n-m+1}$ of $n+2$ functions $(\phi_\rho=\phi_\rho(x^\alpha,p_\alpha), G_{\tilde{k}}=G_{\tilde{k}}(x^\alpha,P_\alpha,S), \rho=0,\ldots,m$, $\tilde{k}=1,\ldots,n-m+1)$, where $C_{\tilde{k}}$ are $n-m+1$ arbitrary constants.

At this point we can go no further, since if we consider the new set of equations

$$\{\phi_\rho,g\}_{nh}=0, \quad \{G_{\tilde{k}},g\}_{nh}=0, \tag{3.11}$$

we should conclude that it has just $n+1$ independent solutions $[2n+3-(m+1+n-m+1)]$, whereas we already know $n+2$ solutions in involution among themselves! This conclusion is wrong, since one of Eqs. (3.11) is not independent, as an equation, from the others, as we will verify at the end of the discussion.

Jacobi's method of integration now involves considering in place of the original set of constraint equations the new set

$$\begin{aligned}&\phi_\rho=G_\rho(x,p)=0 \quad (\rho=0,\ldots,m),\\ &G_{\tilde{k}}(x,p,S)-C_{\tilde{k}}=0 \quad (\tilde{k}=1,\ldots,n-m+1),\end{aligned} \tag{3.12}$$

from which a solution can be obtained algebraically.

In order to show this, we will follow the usual procedure,[17] here adapted to the use of nonhomogeneous Poisson brackets.

Taking into account Eq. (2.9) and the procedure we have followed, we have the functions $G_i$ $(i=0,1,\ldots,n+1)$ independent by construction, so we may assume

$$\left|\frac{\partial(G_i)}{\partial(p_\alpha,S)}\right|\neq 0, \tag{3.13}$$

apart from a possible reordering of the canonical variables $(x,p)$; in any case it is essential for at least one of the $G_i$ to have the derivative with respect to $S$ different from zero.

It follows that we may solve the equations

$$G_i(x,p,S)-C_i=0 \tag{3.14}$$

(where $C_\rho=0, \rho=0,1,\ldots,m$) in $p_\alpha$ and $S$:

$$P_i \equiv p_i - f_i(x,c)=0, \tag{3.15}$$

where $P_{n+1}=S$.

The functions $P_i$ are in involution with respect to the nonhomogeneous Poisson brackets,

$$\{P_i,P_j\}_{nh}\approx 0. \tag{3.16}$$

In fact, using (3.12) and (3.15), we have identically with respect to $x^\alpha$

$$G_i(x^\alpha,p_\alpha=f_\alpha(x,c),S=f_{n+1}(x,c))-C_i=0, \tag{3.17}$$

and, differentiating in $x^\alpha$,

$$\frac{\partial G_i}{\partial x^\alpha}+\frac{\partial G_i}{\partial p_\beta}\frac{\partial f_\beta}{\partial x^\alpha}+\frac{\partial G_i}{\partial S}\frac{\partial f_{n+1}}{\partial x^\alpha}\approx 0 \tag{3.18}$$

or

$$\frac{\partial G_i}{\partial x^\alpha}\approx\frac{\partial G_i}{\partial p_\beta}\frac{\partial P_\beta}{\partial x^\alpha}+\frac{\partial G_i}{\partial S}\frac{\partial P_{n+1}}{\partial x^\alpha}, \tag{3.19}$$

where the weak equality means that the derivatives of $G_i$ with respect to $p_\beta$ and $S$ must be evaluated on the surface (3.15).

Besides (3.19) we have identically

$$\frac{\partial G_i}{\partial p_j}=\frac{\partial G_i}{\partial p_k}\frac{\partial P_k}{\partial p_j}, \tag{3.20}$$

where $j,k=0,1,\ldots,n+1$, $P_{n+1}=S$.

With (3.19) and (3.20) we have

$$\{G_i,G_j\}_{nh}\approx\frac{\partial G_i}{\partial p_k}\frac{\partial G_j}{\partial p_h}\{p_k-f_k(x,c),p_h-f_h(x,c)\}_{nh}, \tag{3.21}$$

as can easily be verified. Using (3.13), we conclude that

$$\{P_k,P_h\}_{nh}\approx 0, \tag{3.16}$$

$$\{p_\alpha-f_\alpha(x,c),p_\beta-f_\beta(x,c)\}_{nh}\approx 0, \tag{3.22}$$

$$\{p_\beta-f_\beta(x,c),S-f_{n+1}(x,c)\}_{nh}\approx 0, \tag{3.23}$$

from which we get

$$\frac{\partial f_\alpha(x,c)}{\partial x^\beta}\approx\frac{\partial f_\beta(x,c)}{\partial x^\alpha}, \tag{3.24}$$

since in (3.22) the bracket reduces to an ordinary Poisson bracket, and

$$\frac{\partial}{\partial x^\alpha}f_{n+1}(x,c)-p_\alpha\approx 0. \tag{3.25}$$

This last result shows that $f_{n+1}(x,c)$ is determined by $f_\alpha(x,c)$, since the weak equality holds when $p_\alpha=f_\alpha(x,c)$ and $S=f_{n+1}(x,c)$, that is, we can take

$$\frac{\partial f_{n+1}(x,c)}{\partial x^\alpha}=f_\alpha(x,c)$$

or

$$dS=f_\alpha(x,c)\,dx^\alpha. \tag{3.26}$$

On the other hand, in (3.24) $p_\alpha$ does not appear, so it holds identically and not only when $p_\alpha=f_\alpha(x,c)$; thus

$$\frac{\partial f_\alpha}{\partial x^\beta}=\frac{\partial f_\beta}{\partial x^\alpha}, \tag{3.27}$$

which is consistent with $dS$ being an exact form.

For practical purposes it is easier to work with homogeneous (usual) Poisson brackets. In this case we would have to integrate Eq. (3.26) in order to get $S$; this is the way in which this point is usually presented.

Now we may understand the observation made after (3.11). The set of functions $\phi_\rho, G_k$ $(k=1,2,\ldots,n-m+1)$ are in involution, but the set of equations (3.11) are not independent since, in the new form,

$$\{P_i,g\}_{nh}\approx 0 \tag{3.28}$$

[where $\approx$ still means an equality on the surface (3.15)]; when $i=n+1$, it reads

$$\left(p_\alpha - \frac{\partial f_{n+1}(x,c)}{\partial x^\alpha}\right)\frac{\partial g}{\partial p_\alpha} \approx 0$$

or

$$\left(f_\alpha(x,c) - \frac{\partial f_{n+1}(x,c)}{\partial x^\alpha}\right)\frac{\partial g}{\partial p_\alpha} \approx 0, \tag{3.29}$$

which shows that the last of Eqs. (3.28) is already satisfied. So the real number of independent equations is again $n+1$, with $2n+3-(n+1)=n+2$ solutions given by $\phi_\rho$, $G_k$, or $P_i$, $i=0,1,...,n+1$.

Summing up, in the hypothesis that the set of constraints $\phi_\rho$ is first class, a function

$$S = S(x,c)$$

exists, which is the solution of Eqs. (2.6). This function is defined apart from an additive constant. Neglecting this constant, $S$ will in general contain $(n+1)-(m+1)$ constants of integration, and so it is a complete integral. It is known from the theory of systems of PDE that from a complete integral it is possible to get all other integrals by means of differentiations and eliminations only. So the Jacobi method of integration gives a general kind of solution.

The constraints (2.6) are contained implicitly in the set of equations (3.15). Indeed, by eliminating $(n+1)-(m+1)$ constants $C_k$, we again obtain the constraints (2.6). Finally, let us observe that the transformation

$$\begin{aligned} Q^\alpha &= x^\alpha, \\ P_\alpha &= p_\alpha - f_\alpha(x,c) \end{aligned} \tag{3.30}$$

is a canonical transformation due to Eqs. (3.27). It is a phase transformation generated by the function $S(x,c)$:

$$\begin{aligned} Q^\alpha &= e^{-S} * x^\alpha = x^\alpha, \\ P_\alpha &= e^{-S} * p_\alpha = p_\alpha - f_\alpha(x,c), \end{aligned} \tag{3.31}$$

where the operation $*$ is defined by

$$e^A * B = B + \{A,B\} + \tfrac{1}{2}\{A,\{A,B\}\} + \cdots. \tag{3.32}$$

Let us now continue the discussion on the construction of the characteristic system. As anticipated in Sec. II, the 1-forms belonging to the space $Q^*$ can be written using Eqs. (2.24), if Eqs. (2.17), (2.18), and (2.19) can be rewritten in the form (2.23). In the hypothesis

$$\left|\frac{\partial \phi_\rho}{\partial p_\sigma}\right| \neq 0 \tag{3.33}$$

by eliminating the functions $\lambda^\rho$ in Eqs. (2.17), (2.18), and (2.19), we get

$$\begin{aligned} \lambda^\rho &= v^\sigma (A^{-1})^\rho_\sigma, \\ A^{\cdot\sigma}_\rho &= \frac{\partial \phi_\rho}{\partial p_\sigma} \end{aligned} \tag{3.34}$$

and

$$\begin{aligned} v^\alpha - v^\sigma (A^{-1})^\rho_\sigma \frac{\partial \phi_\rho}{\partial p_\alpha} &= 0, \\ u_\alpha + v^\sigma (A^{-1})^\rho_\sigma \frac{\partial \phi_\rho}{\partial x^\alpha} &= 0, \\ v^{n+1} - p_\alpha v^\alpha &= 0, \\ v^\alpha \frac{\partial \phi_\rho}{\partial x^\alpha} + u_\alpha \frac{\partial \phi_\rho}{\partial p_\alpha} &\approx 0, \end{aligned} \tag{3.35}$$

which are in the form (2.23). By the formal substitution $v^\alpha \to dx^\alpha$, $u_\alpha \to d p_\alpha$, $v^{n+1} \to dS$, we get the set of 1-forms:

$$\{\bar\theta^a\} \equiv \begin{cases} \eta^\alpha = dx^\alpha - dx^\sigma (A^{-1})^\rho_\sigma \dfrac{\partial \phi_\rho}{\partial p_\alpha}, \\ \zeta_\alpha = d p_\alpha + dx^\sigma (A^{-1})^\rho_\sigma \dfrac{\partial \phi_\rho}{\partial x^\alpha}, \\ \theta = dS - p_\alpha\, dx^\alpha, \\ d\phi_\rho = dx^\alpha \dfrac{\partial \phi_\rho}{\partial x^\alpha} + d p_\alpha \dfrac{\partial \phi_\rho}{\partial p_\alpha}. \end{cases} \tag{3.36}$$

It is easily verified that only $(2n+3)-(m+1)$ of these forms are independent and hence can be chosen as a basis for $Q^*$. The associated Pfaff system will be given by the exterior differential equations

$$\begin{aligned} dx^\alpha &= dx^\sigma (A^{-1})^\rho_\sigma \{x^\alpha, \phi_\rho\}, \\ d p_\alpha &= dx^\sigma (A^{-1})^\rho_\sigma \{p_\alpha, \phi_\rho\}, \\ dS &= p_\alpha\, dx^\sigma (A^{-1})^\rho_\sigma \{x^\alpha, \phi_\rho\} \end{aligned} \tag{3.37}$$

[where for $\alpha = \sigma$ the first set of equations are identities; see Eqs. (3.34)], to which it is necessary to add the equations

$$\phi_\rho(x,p) = 0, \tag{3.38}$$

in order to get the characteristic system $C$ of our original exterior differential system $A$ [Eqs. (2.11)].

Therefore, since the number of integration constants in (3.30) is $(2n+3)-(m+1)$ and they must satisfy (3.38), we conclude that the characteristic manifold is parametrized by $2(n-m)+1$ constants.

Observe that the equations $d\phi_\rho = 0$ are satisfied owing to (3.37) and (3.7):

$$d\phi_\rho = dx^\sigma (A^{-1})^\tau_\sigma \{\phi_\tau, \phi_\rho\} \quad (\rho,\sigma,\tau = 0,...,m).$$

We may put Eqs. (3.37) in a more explicit form. Under the hypothesis (3.33), Eqs. (3.38) can be solved in terms of $p_\rho$ ($\rho = 0,1,...,m$):

$$p_\rho - \psi_\rho(p_{\rho'}, x^\alpha) = 0 \quad (\rho' = m+1,...,n), \tag{3.39}$$

so that the equations

$$\phi_\rho(x^\alpha, p_{\rho'}, p_\rho = \psi_\rho(p_{\rho'}, x^\alpha)) = 0$$

are identities in $x^\alpha$ and $p_\rho$. On the basis of the same argument applied to Eq. (3.17), we get the following weak equations:

$$\begin{cases} \dfrac{\partial \phi_\rho}{\partial x^\alpha} \approx -A^{\cdot\sigma}_\rho \{p_\alpha, p_\sigma - \psi_\sigma\}, \\ \dfrac{\partial \phi_\rho}{\partial p_\alpha} \approx A^{\cdot\sigma}_\rho \{x^\alpha, p_\sigma - \psi_\sigma\}, \end{cases}$$

so that the characteristic system (3.37) can be rewritten as

$$\begin{aligned} dx^\alpha &= \{x^\alpha, p_\rho - \psi_\rho\}\, dx^\rho, \\ d p_\alpha &= \{p_\alpha, p_\rho - \psi_\rho\}\, dx^\rho, \\ dS &= p_\alpha \{x^\alpha, p_\rho - \psi_\rho\}\, dx^\rho, \\ \phi_\rho &= 0. \end{aligned} \tag{3.40}$$

We already know from Eqs. (3.8) that this system is integrable. This can easily be checked with the use of the Jacobi identity (here we are speaking of local integrability) and remembering that the constraints in the form $p_\rho - \psi_\rho = 0$ are in involution among themselves.

The equation for $S$, after integration of the characteristic equations, will give the Hamilton–Jacobi function evaluated along the characteristic surfaces (which are $m+1$ dimensional). From this it is possible (even when second-class constraints are absent) to recover the function $S = S(x)$ by eliminating half of the integration constants in favor of an equal number of coordinates. We will not discuss this point but rather the inverse procedure which consists of finding from $S$ the solutions of the equations of motion (characteristic equations).

Indeed the knowledge of a complete integral of the Hamilton–Jacobi equations $S = \phi(x^\alpha, c_k) + c$ gives a solution of the equations of motion (3.40) (characteristic manifold). As is well known from analytical mechanics, a complete integral can be used in the following way: from

$$S - \phi(x^\alpha, c_k) - c = 0 \quad (k = m+1,\ldots,n), \tag{3.41}$$

$$\frac{\partial \phi(x^\alpha, c_k)}{\partial c_h} - b^h = 0 \quad (h = m+1,\ldots,n), \tag{3.41'}$$

where $b^h$ are new $n-m$ constants, and adding Eqs. (3.15) to these,

$$p_\alpha - f_\alpha(x^\beta, c_k) = 0, \tag{3.41''}$$

where $f_\alpha(x,c) = \partial\phi(x,c)/\partial x^\alpha$, we get a submanifold of $R^{2n+3}$ with dimension equal to $m+1$ for any given value of the $2(n-m)+1$ constants $c_k$, $b^k$, and $c$. If we can show that the characteristic vector fields (3.3) are tangent to each of these manifolds, that is, for any choice of the constants $c_k$, $b^k$, and $c$, we will have demonstrated that these are the characteristic manifolds.

This is easily checked. In fact we have

$$v_\rho(S - \phi(x,c) - c) = -\{\phi_\rho, S - \phi(x,c) - c\}_{nh}$$
$$= \frac{\partial \phi_\rho}{\partial p_\alpha}\left(p_\alpha - \frac{\partial \phi}{\partial x^\alpha}\right) = 0,$$
$$v_\rho\left(\frac{\partial \phi(x,c)}{\partial c_h} - b^h\right) = -\frac{\partial \phi_\rho}{\partial p_\alpha}\frac{\partial^2 \phi}{\partial c_h \partial x^\alpha}$$
$$= -\frac{\partial}{\partial c_h}\phi_\rho(x^\alpha, p_\alpha = f_\alpha(x,c)) = 0,$$

since the equations $\phi_\rho = 0$ are satisfied identically by $p_\alpha = f_\alpha(x,c)$, and finally

$$v_\rho\left(p_\alpha - \frac{\partial \phi}{\partial x^\alpha}\right) = -\left(\frac{\partial \phi_\rho}{\partial x^\alpha} + \frac{\partial \phi_\rho}{\partial p_\beta}\frac{\partial^2 \phi}{\partial x^\beta \partial x^\alpha}\right)$$
$$= -\frac{\partial}{\partial x^\alpha}\phi_\rho(x^\alpha, p_\alpha = f_\alpha(x,c)) = 0$$

with the same argument.

Let us observe that the argument is the same as in standard Hamilton–Jacobi theory. This can be understood by observing that a constrained system is nothing more than a particular case of a classical system where a certain number of constants of motion, the constants $c_\rho$ of Eq. (3.14), are required to be zero, instead of being arbitrary constants. For this reason, the dimension of the characteristic manifold is not $2(n+1)-1$, but $2(n+1)-(m+1)$.

## IV. THE CHARACTERISTIC SYSTEM FOR THE CASE R≠0

When $r \neq 0$, due to condition (2.10), the set of constraints cannot be interpreted as a set of PDE in the unknown $S$. Strictly speaking, as we pointed out in Sec. II, a Hamilton–Jacobi function does not exist (it can, however, exist in some reduced space; see below).

Nevertheless, the characteristic system exists and is integrable, as we will see in this section. Moreover, some use can again be made of the Hamilton–Jacobi approach developed in the last section.

Let us first discuss the characteristic system, which should be constructed by starting from the same equations (2.17), (2.18), and (2.19), and the condition (2.20).

Since the rank of the matrix $\|\{\phi_\rho, \phi_\sigma\}\|$ is now $r \neq 0$, there will exist a minor of rank $r$ different from zero, which we will assume to be formed with the last $r$ rows and columns of $\|\{\phi_\rho, \phi_\sigma\}\|$:

$$\{\phi_{\mu'}, \phi_{\nu'}\} = C_{\mu'\nu'} \quad (\mu', \nu' = m-r+1,\ldots,m), \tag{4.1}$$

with

$$|C_{\mu'\nu'}| \neq 0, \tag{4.2}$$

where $r$ must be even, due to the antisymmetry of the matrix $\|\{\phi_\rho, \phi_\sigma\}\|$; we will put $r = 2s$.

In the case $r < m+1$, Eqs. (2.20) will give $r$ of $\lambda^\sigma$ in terms of the remaining $m-r+1$:

$$\lambda^{\mu'} \approx -(C^{-1})^{\mu'\nu'}\{\phi_{\nu'}, \phi_\mu\}\lambda^\mu$$
$$(\mu'\nu' = m-r+1,\ldots,m,\ \mu,\nu = 0,\ldots,m-r=f)$$

or

$$\lambda^\rho \approx (\delta^\rho_\mu - \delta^\rho_{\mu'}(C^{-1})^{\mu'\nu'}\{\phi_{\nu'}, \phi_\mu\})\lambda^\mu \quad (\rho = 0,\ldots,m), \tag{4.3}$$

which for $\rho = \nu$ is an identity.

The case $r = m+1$, which is possible only when $m+1$ is even, will be discussed at the end of this section.

If we substitute in Eqs. (2.17), we get

$$v^\alpha = \lambda^\mu(\delta^\rho_\mu - \delta^\rho_{\mu'}(C^{-1})^{\mu'\nu'}\{\phi_{\nu'}, \phi_\mu\})\{x^\alpha, \phi_\rho\},$$
$$u_\alpha = \lambda^\mu(\delta^\rho_\mu - \delta^\rho_{\mu'}(C^{-1})^{\mu'\nu'}\{\phi_{\nu'}, \phi_\mu\})\{p_\alpha, \phi_\rho\}, \tag{4.4}$$

which besides Eqs. (2.18) give the new set of characteristic vector fields:

$$v_\mu = -\{\phi_\mu, \cdot\}^*_{nh}, \tag{4.5}$$

where we have defined the nonhomogeneous Dirac bracket:

$$\{f,g\}^*_{nh} = \{f,g\}_{nh} - \{f,\phi_{\mu'}\}_{nh}(C^{-1})^{\mu'\nu'}\{\phi_{\nu'}, g\}_{nh}. \tag{4.6}$$

The dimension of the space $\tilde{Q}$ is now $m-r+1$.

The system of equations

$$v_\mu(g) = -\{\phi_\mu, g\}^*_{nh} = 0, \tag{4.7}$$

associated with the set of characteristic vector fields (4.5), is completely integrable on the surface (2.6). This can be verified, as in Sec. III, by using a generalization of the Mayer

identity to Dirac brackets. By a very long computation it can be verified that

$$\{f_1,\{f_2,f_3\}^*_{nh}\}^*_{nh} + \frac{\partial f_1}{\partial S}\{f_2,f_3\}^*_{nh} + \text{cyclic} = 0. \quad (4.8)$$

Let us observe that the characteristic vectors (4.5) can be rewritten in an equivalent way as

$$v_\mu = -\{\phi^*_\mu,\cdot\}_{nh}, \quad (4.9)$$

where $\phi^*_\mu$ are the "starred" variables defined in (A1). They are first class and satisfy

$$\{\phi^*_\mu,\phi^*_\nu\} \equiv O(\phi^*_\mu) + O(\phi^2_{\mu'})$$
$$(\lambda,\mu,\nu = 0,\ldots,m-r,\ \mu',\nu' = m-r+1,\ldots,m). \quad (4.10)$$

At this point we could try to look for an involutory system starting from the set $\phi^*_\mu$, in analogy to the method followed in Sec. III. But it is easily realized that, in the presence of second-class constraints, one cannot get an analogous set of $n+1$ functions (or better $n+2$, taking into account the function $S$ as we did in Sec. III). Indeed, the system (4.7) with $v_\mu$ in the form (4.9) has $2n+3-(f+1)$ solutions, of which $f+1+2s$ are already known [the $2s$ second-class functions $\phi_{\mu'}$ satisfy by construction the system (4.7)]. Therefore, by selecting from among these solutions a set of functions to be added to $\phi^*_\mu$: $G_1(x,p,S)$, $G_2(x,p,S)$,...,we can form from these the corresponding starred quantities, $G^*_1, G^*_2, \cdots$, in order to preserve the solutions we already know.

In this way we get the maximum set

$$G^*_1, G^*_2, \ldots, G^*_{n+1-(f+s)}$$

of functions in weak involution with respect to the nonhomogeneous Poisson bracket.

The characteristic system can now be constructed starting from this definition of the characteristic vector fields, by eliminating $\lambda^\mu$ from the following expression of the components of a general characteristic vector field:

$$x^\alpha = \lambda^\mu\{x^\alpha,\phi^*_\mu\},$$
$$u_d = \lambda^\mu\{p_\alpha,\phi^*_\mu\}, \quad (4.11)$$
$$v^{n+1} = p_\alpha v^\alpha.$$

If we put

$$A^{\cdot\nu}_\mu = \{x^\nu,\phi^*_\mu\} \quad (4.12)$$

and if we make the usual assumption that

$$\left|\frac{\partial\phi^*_\mu}{\partial p_\nu}\right| \neq 0,$$

we get

$$v^\alpha - v^\nu(A^{-1})^\mu_\nu\{x^\alpha,\phi^*_\mu\} = 0,$$
$$u_\alpha - v^\nu(A^{-1})^\mu_\nu\{p_\alpha,\phi^*_\mu\} = 0, \quad (4.13)$$
$$v^{n+1} - p_\alpha v^\alpha = 0.$$

With the substitutions $v^\alpha \to dx^\alpha$, $u_\alpha \to dp_\alpha$, and $v^{n+1} \to dS$ in (4.13), we get the following differential forms:

$$\eta^\alpha = dx^\alpha - dx^\nu(A^{-1})^\mu_\nu\{x^\alpha,\phi^*_\mu\},$$
$$\zeta_\alpha = dp_\alpha - dx^\nu(A^{-1})^\mu_\nu\{p_\alpha,\phi^*_\mu\}, \quad (4.14)$$
$$\theta = dS - p_\alpha dx^\alpha,$$

and besides

$$d\phi^*_\mu = \frac{\partial\phi^*_\mu}{\partial x^\alpha}dx^\alpha + \frac{\partial\phi^*_\mu}{\partial p_\alpha}dp_\alpha,$$
$$d\phi_{\mu'} = \frac{\partial\phi_{\mu'}}{\partial x^\alpha}dx^\alpha + \frac{\partial\phi_{\mu'}}{\partial p_\alpha}dp_\alpha. \quad (4.15)$$

The characteristic equations are now obtained by weakly putting these forms to zero; however, it is easily recognized that the last two forms, $d\phi^*_\mu$ and $d\phi_{\mu'}$, are not independent from the others. Indeed we have

$$\begin{aligned} d\phi^*_\mu &= \frac{\partial\phi^*_\mu}{\partial x^\alpha}(\eta^\alpha + dx^\nu(A^{-1})^\lambda_\nu\{x^\alpha,\phi^*_\lambda\}) \\ &\quad + \frac{\partial\phi^*_\mu}{\partial p_\alpha}(\zeta_\alpha + dx^\nu(A^{-1})^\lambda_\nu\{p_\alpha,\phi^*_\lambda\}) \\ &= \frac{\partial\phi^*_\mu}{\partial x^\alpha}\eta^\alpha + \frac{\partial\phi^*_\mu}{\partial p_\alpha}\zeta_\alpha + dx^\nu(A^{-1})^\lambda_\nu\{\phi^*_\mu,\phi^*_\lambda\} \\ &\approx \frac{\partial\phi^*_\mu}{\partial x^\alpha}\eta^\alpha + \frac{\partial\phi^*_\mu}{\partial p_\alpha}\zeta_\alpha. \end{aligned}$$

In conclusion we have a set of $(2n+3)-(m-2s+1)$ independent differential forms
[dim $Q^* = (2n+3)-(m-2s+1)$], and the characteristic system is given by

$$dx^\alpha = dx^\nu(A^{-1})^\mu_\nu\{x^\alpha,\phi^*_\mu\},$$
$$dp_\alpha = dx^\nu(A^{-1})^\mu_\nu\{p_\alpha,\phi^*_\mu\},$$
$$dS = p_\alpha dx^\alpha, \quad (4.16)$$
$$\phi^*_\mu = 0, \quad \phi_{\mu'} = 0.$$

Following the same discussion as in Sec. III, we have the characteristic manifold parametrized by
$2(n-m+s)+1$ arbitrary constants, where one of these is an unessential additive constant for $S$.

We can put the characteristic equations in a form analogous to that of Eqs. (3.40), if we solve the equations $\phi^*_\mu = 0$ in terms of $p_\mu$,

$$p_\mu = \psi_\mu(x^\alpha,p_{\mu'}) \quad (4.17)$$
$$(\mu = 0,1,\ldots,m-2s,\ \mu' = m-2s+1,\ldots,n).$$

Following the same arguments, we now get

$$\{p_\mu - \bar\psi_\mu, p_\nu - \bar\psi_\nu\} = O(\phi^2_{\mu'}), \quad (4.18)$$

where we have used Eq. (4.10), whereas, in the case of first-class constraints only, we would have gotten zero. It is easily verified that

$$\frac{\partial\phi^*_\mu}{\partial x^\alpha} = -A^{\cdot\nu}_\mu\{p_\alpha, p_\nu - \bar\psi_\nu\},$$
$$\frac{\partial\phi^*_\mu}{\partial p_\alpha} = +A^{\cdot\nu}_\mu\{x^\alpha, p_\nu - \bar\psi_\nu\} \quad (4.19)$$

when (4.17) holds; so we can write the characteristic system in the following form:

$$dx^\alpha = dx^\nu\{x^\alpha, p_\nu - \bar\psi_\nu\}, \quad (4.20)$$

$$dp_\alpha = dx^\nu \{ p_\alpha, p_\nu - \bar{\psi}_\nu \},$$
$$dS = p_\alpha \, dx^\alpha, \tag{4.21}$$
$$\phi^*_\mu = 0, \quad \phi_{\mu'} = 0. \tag{4.22}$$

In the form (4.20) we can more easily verify that the integrability conditions are satisfied, due to the property (4.18), using the Jacobi identity. As regards Eqs. (4.21), the same comment we made in Sec. III applies here as well. The point is rather that knowledge of the functions $S$ evaluated along the characteristics is not enough to recover a function $S$ of the coordinates $x^\alpha$ since, as we already know, such a function does not exist when second-class constraints are present.

When $m+1$ is even and $r = m+1$, by repeating the discussion we find $\dim \bar{Q} = 0$ and the associated space $Q^*$ will have dimension $2n+3$. So it will be spanned by any $(2n+3)$-dimensional basis, which can be chosen $(dx^\alpha, dS, dp_\alpha)$.

As a result the associated Pfaff system will be

$$dx^\alpha = 0,$$
$$dp_\alpha = 0, \quad \phi_\rho(x,p) = 0, \tag{4.23}$$
$$dS = 0,$$

which agrees with Eqs. (4.16) when no first-class contraints exist. As in the previous case the Hamilton–Jacobi function $S = S(x^\alpha)$ does not exist.

## V. THE HAMILTON–JACOBI FUNCTION

In the last section we saw how to get a characteristic system when second-class constraints are present, and we stressed that in such a case a function $S$ satisfying only the constraint equations (if we think of them as PDE in $S$) does not exist.[18] So we cannot really speak of a Hamilton–Jacobi method in such cases. Nevertheless, as we will see in the present section, the theory developed in Sec. III can again be useful.

We will consider two situations: the first is when the first-class constraints form a first-class subset, that is, when the Poisson bracket between any two of them is a first-class constraint. In such a case we will say that they are in weak involution among themselves. This can always be achieved as an application of general theorems.[16]

In this situation let us consider this set as defining our dynamical system, neglecting for the moment the second-class constraints. This set

$$\phi_\mu(x,p) = 0 \quad (\mu = 0,\ldots,f) \tag{5.1}$$

can be considered as a system of PDE in $S = S(x)$, following the theory developed in Sec. III.

Let us suppose that we have found a complete integral of the system (5.1),

$$S = \phi(x^\alpha, c_k) + c \quad (\alpha = 0,\ldots,n, \; k = f+1,\ldots,n), \tag{5.2}$$

where $c_k$ and $c$ are $n-f+1$ arbitrary constants.

In order to find the solutions of the characteristic equations, we put, as in Eqs. (3.41′),

$$\frac{\partial \phi(x,c)}{\partial c_h} - b^h = 0, \tag{5.3}$$

where $b^h$ are new $n-f$ constants. By adding to Eqs. (5.3) Eqs. (5.2) and

$$p_\alpha - f_\alpha(x,c) = 0, \tag{5.4}$$

where

$$f_\alpha(x,c) = \frac{\partial \phi(x,c)}{\partial x^\alpha}, \tag{5.5}$$

we get $2n+3-(f+1)$ equations for $x^\alpha$, $p_\alpha$, and $S$, which can be solved for $p_\alpha$ and $n-f$ $x^\alpha$, regarding $f+1$ of the coordinates $x^\alpha$ as parameters. We thus obtain an integral manifold for any given value of the $2(n-f)+1$ constants $c_k$, $b^k$, and $c$, as in Sec. III (the substitution of $f$ in place of $m$ must be performed).

The solution of the original characteristic system (4.16) can be derived from this solution by adding to the solution just found the second-class constraints $\phi_{\mu'}(x,p) = 0$, *which turn out to be constant and so have the meaning of some restriction on the constants* $c_k$ *and* $b^k$. The final number of the constants will be $2(n-f)+1-2s$.

In order to show this, let us write the characteristic vectors of the exterior differential system (5.1). These are, according to Sec. III [see Eq. (3.3)]:

$$v_\mu = -\{\phi_\mu, \cdot\}_{nh}. \tag{5.6}$$

The hypothesis made on $\phi_\mu$ now has the important consequence that $v_\mu$ are weakly equal to the characteristic vectors of our original system, Eq. (4.5), i.e., when they are calculated on the submanifold defined by *all* the constraints $\phi_\mu$ and $\phi_{\mu'}$:

$$v_\mu = -\{\phi_\mu, \cdot\}_{nh} \approx -\{\phi_\mu, \cdot\}^*_{nh}. \tag{5.7}$$

In fact, using the explicit expression of the Dirac brackets [see Eq. (4.6)], we have

$$\{\phi_\mu, \cdot\}^*_{nh} = \{\phi_\mu, \cdot\}_{nh} - \{\phi_\mu, \phi_{\mu'}\}_{nh} (C^{-1})^{\mu'\nu'} \{\phi_{\nu'}, \cdot\}_{nh},$$

where the last term on the rhs is weakly zero when $\phi_\mu = 0$ *and* $\phi_{\mu'} = 0$. This is because $\phi_\mu$ are assumed to be first class.

This fact has the consequence that, when all the constraints are satisfied, we have

$$v_\mu(\phi_{\mu'}) \approx 0, \tag{5.8}$$

that is, the characteristic vectors of the integral manifold we have found are also tangent to our original submanifold defined by the equations $\phi_\mu = 0$ and $\phi_{\mu'} = 0$.

This demonstrates that the restrictions imposed on the solution (5.3), (5.4) by $\phi_{\mu'}(x,p) = 0$ are indeed restrictions on the constants $c_k$ and $b^k$, and that in this way a solution of the original system can be found.

Another situation that may occur is when the constraints $\phi_\mu$ and $\phi_{\mu'}$ admit a maximal subset of first-class constraints (among themselves). We know from a general theorem that this always takes place, at least locally. Indeed we know that it is always possible to locally substitute the set $\phi_\mu$ and $\phi_{\mu'}$ with a new set such that $f+s+1$ of them are first class among themselves and the remaining $s$ are all second class.[16]

In general it is very difficult to find such a maximal subset, and doing so could be equivalent in almost all cases to completely solving the dynamics.

Nevertheless, it may turn out to be possible, or the con-

straints may already satisfy this condition.

If this is the case, let us call this subset

$$\psi_\lambda(x,p)=0 \quad (\lambda=0,1,\ldots,f+s). \tag{5.9}$$

The $\psi_\lambda$ include $\phi_\mu$, $\mu=0,1,\ldots,f$, and half of $\phi_{\mu'}$. Equations (5.9) alone define a Hamilton–Jacobi function $S(x^\alpha)$. So let us for the moment put aside the remaining constraints $\phi_{\mu'}$, $\mu'=f+s+1,\ldots,f+2s+1$, which we will call

$$\phi_{\mu'}(x,p)=\phi_{f+s+l}(x,p)$$
$$(\mu'=f+s+1,\ldots,f+2s,\ l=1,\ldots,s). \tag{5.10}$$

Let us suppose we have found a complete integral for the set of first-class (5.9):

$$S=\Psi(x,c_k)+c[k=1,\ldots,n-(f+s)], \tag{5.11}$$

where $c_k$ and $c$ are $n-f-s+1$ arbitrary constants.

We want to show that a solution of our original system (4.16) can be obtained by imposing the conditions (5.10) on the characteristic surface obtained from (5.11) by means of the equations

$$\frac{\partial\Psi(x,c)}{\partial c_h}-b^h=0, \tag{5.12}$$

where $b^h$ are new $n-f-s$ constants, and

$$p_\alpha-f_\alpha(x^\beta,c_k)=0 \tag{5.13}$$

with

$$f_\alpha(x^\beta,c_k)=\frac{\partial\Psi(x,c)}{\partial x^\alpha}.$$

The solution given by these $2n+3-(f+s+1)$ equations can be expressed by solving $p_\alpha$ and $n-(f+s)$ of $x^\alpha$ in terms of the remaining $f+s+1$ coordinates regarded as parameters and the constants $c_k$ and $b^k$. The integral manifold so obtained has dimension $f+s+1$ for any given value of the $2(n-f-s)+1$ constants $c_k$, $b^k$, and $c$ in $\mathbb{R}^{2n+3}$.

In order to show this, let us write the characteristic vectors of the system (5.9), which are now

$$v_\lambda=-\{\psi_\lambda,\cdot\}_{nh} \quad (\lambda=0,\ldots,f+s). \tag{5.14}$$

The situation is now quite different from the previous case. The analogous equation (5.7) does not hold any more.

If we require the integral manifold we have found to be an integral manifold of the system $\phi_{f+s+l}(x,p)=0$ as well, its dimensions will diminish from $f+s+1$ to $f+1$. In order to get this result, we must require that the characteristic space spanned by the vectors $v_\lambda$ of Eq. (5.14) be restricted by the requirement

$$v(\phi_{f+s+l})\approx 0 \quad (l=1,\ldots,s), \tag{5.15}$$

where $v$ is a generic vector given by

$$v=\lambda^\lambda v_\lambda \quad (\lambda=0,\ldots,f+s). \tag{5.16}$$

Equations (5.15) will give

$$\lambda^\lambda v_\lambda(\phi_{f+s+l})=-\lambda^\lambda\{\psi_\lambda,\phi_{f+s+l}\}=0. \tag{5.17}$$

Since the rank of the matrix $\|\{\psi_\lambda,\phi_{f+s+l}\}\|$ is $s$ (in a weak sense), we can solve Eqs. (5.17) for $s$ of $\lambda^\lambda$ in terms of the remaining $f+1$:

$$\lambda^{f+\mu}=-\lambda^\mu\{\phi_\mu,\phi_{f+s+l}\}(C^{-1})^{lm}$$
$$(l,m=1,\ldots,s,\ \mu=0,\ldots,f), \tag{5.18}$$

where

$$C_{lm}=\{\phi_{f+s+l},\phi_{f+m}\}.$$

In Eqs. (5.18) we have used the notation of Sec. IV regarding the labelling of the indices.

By substituting (5.18) in (5.16) we get

$$v=\lambda^\mu[v_\mu-\{\phi_\mu,\phi_{f+s+l}\}(C^{-1})^{lm}v_m]$$
$$=-\lambda^\mu[\{\phi_\mu,\cdot\}_{nh}-\{\phi_\mu,\phi_{f+s+l}\}(C^{-1})^{lm}\{\phi_{f+m},\cdot\}_{nh}]. \tag{5.19}$$

(The nonhomogeneous Poisson bracket between two constraints coincides with the usual Poisson bracket, since they do not depend on $S$.)

By adding to the lhs side of Eq. (5.19) terms which are zero when all the constraints are satisfied, we get

$$v\approx-\lambda^\mu[\{\phi_\mu,\cdot\}_{nh}-\{\phi_\mu,\phi_{f+s+l}\}_{nh}(C^{-1})^{lm}\{\phi_{f+m},\cdot\}_{nh}$$
$$+\{\phi_\mu,\phi_{f+l}\}_{nh}(\tilde{C}^{-1})^{lm}\{\phi_{f+s+m},\cdot\}_{nh}$$
$$-\{\phi_\mu,\phi_{f+l}\}_{nh}(C^{-1})^{lm}\{\phi_{f+s+m},\phi_{f+s+n}\}_{nh}$$
$$\times(\tilde{C}^{-1})^{np}\{\phi_{f+p},\cdot\}_{nh}]$$

$(l,m,n,p=1,\ldots,s)$, where $\tilde{C}_{lm}=C_{ml}$. We recognize the structure of the Dirac brackets:

$$v\approx-\lambda^\mu\{\phi_\mu,\cdot\}^*_{nh}. \tag{5.20}$$

In this way we recover the characteristic vectors (4.5):

$$v_\mu=-\{\phi_\mu,\cdot\}^*_{nh}. \tag{5.21}$$

In practice, to get this result when working on the solution (5.12), (5.13), it is only necessary to impose on this solution the restriction

$$\phi_{f+s+l}(x,p)=0, \tag{5.22}$$

where now, contrary to the first case considered in this section, Eq. (5.22) has the meaning of a restriction on the coordinates [which in the number of $f+s+1$ can be used to parametrize the solution (5.12) and (5.13)] and *not* on the constants $c_k, b^k$, which in this case are in the correct number right from the start.

The two situations described are two examples of the use of the Hamilton–Jacobi approach when second-class constraints are present. The method could be extended to intermediate cases where a number of constraints between $f+1$ and $f+s+1$ are known to be a first-class subset (among themselves).

## VI. PARAMETRIC FORMS OF THE CHARACTERISTIC SYSTEM

Until now we have discussed a parameter-free approach which in our opinion is the main feature of the Hamilton–Jacobi theory. We wish to mention here the possibility of a multiparameter approach. We can try to reformulate Eqs. (3.37) by defining $m+1$ parameters $\omega^\rho$ by the equations

$$d\omega^\rho=dx^\sigma(A^{-1})^\rho_\sigma \quad (\rho,\sigma=0,\ldots,m). \tag{6.1}$$

If this is possible the characteristic equations (3.37) become

$$dx^\alpha=\{x^\alpha,\phi_\rho\}\,d\omega^\rho,$$
$$dp_\alpha=\{p_\alpha,\phi_\rho\}\,d\omega^\rho, \quad \phi_\rho=0. \tag{6.2}$$
$$dS=p_\alpha\{x^\alpha,\phi_\rho\}\,d\omega^\rho,$$

If we check the integrability conditions of the Mayer system deduced from Eqs. (6.2) using the Jacobi identity, we find that these conditions boil down to requiring

$$\{\phi_\rho,\phi_\sigma\} = O(\phi^2). \tag{6.3}$$

Thus a multiparametric formulation such as (6.2) is possible only if the constraints $\phi_\rho$ ($\rho = 0,...,m$) satisfy the condition (6.3). The same remains true also for the characteristic system for a set of first- and second-class constraints [Eqs. (4.16)]: the system

$$\begin{aligned} dx^\alpha &= \{x^\alpha,\phi^*_\mu\}d\omega^\mu, \\ dp_\alpha &= \{p_\alpha,\phi^*_\mu\}d\omega^\mu, \\ dS &= p_\alpha\{x^\alpha,\phi^*_\mu\}d\omega^\mu, \\ \phi^*_\mu &= 0, \quad \phi_{\mu'} = 0, \end{aligned} \tag{6.4}$$

is integrable only if

$$\{\phi^*_\mu,\phi^*_\nu\} = O(\phi^2). \tag{6.5}$$

Besides (6.3) there is another case where one can use a multiparametric formulation. Namely, if the constraints $\phi_\rho$ (or $\phi_\mu$) form a closed Lie algebra under the Poisson (or Dirac) brackets, i.e.,

$$\{\phi_\rho,\phi_\sigma\} = C^\tau_{\rho\sigma}\phi_\tau, \tag{6.6}$$

where $C^\tau_{\rho\sigma}$ denote the structure constants, then the equations

$$\begin{aligned} dx^\alpha &= \{x^\alpha,\hat\phi_\rho\}\,d\tau^\rho, \\ dp_\alpha &= \{p_\alpha,\hat\phi_\rho\}\,d\tau^\rho \end{aligned} \tag{6.7}$$

with

$$\hat\phi_\rho = B^\sigma_\rho(\tau)\phi_\sigma, \tag{6.8}$$

where the functions $B^\sigma_\rho(\tau)$ satisfy

$$\frac{\partial}{\partial\tau^\tau}B^\sigma_\rho - \frac{\partial}{\partial\tau^\rho}B^\sigma_\tau = B^\xi_\tau B^\eta_\rho C^\sigma_{\xi\eta}, \tag{6.9}$$

are integrable. The functions $B^\sigma_\rho(\tau)$ are determined entirely by the structure constants [see Ref. 19]. In fact it turns out that

$$\begin{aligned} &\frac{\partial^2x^\alpha}{\partial\tau^\sigma\partial\tau^\rho} - \frac{\partial^2x^\alpha}{\partial\tau^\rho\partial\tau^\sigma} \\ &\quad = \{\{x^\alpha,\hat\phi_\rho\},\hat\phi_\sigma\} + \{x^\alpha,\phi_\tau\}\frac{\partial B^\tau_\rho}{\partial\tau^\sigma} - (\rho\leftrightarrow\sigma) \\ &\quad = -B^\eta_\rho B^\xi_\sigma C^\tau_{\eta\xi}\{\phi_\tau,x^\alpha\} \\ &\qquad + \{x^\alpha,\phi_\tau\}\left(\frac{\partial B^\tau_\rho}{\partial\tau^\sigma} - \frac{\partial B^\tau_\sigma}{\partial\tau^\rho}\right) = 0 \end{aligned}$$

and analogously for $p_\alpha$. The calculation goes in the same way even when second-class constraints are present.

Finally, let us consider a one-parameter formulation, by recovering the equations of motion derived by Dirac. This can be achieved by choosing a particular vector field of the characteristic "manifold"

$$v = -\lambda^\mu\{\phi_\mu,\cdot\}^* \quad (\mu = 0,...,m-r). \tag{6.10}$$

The corresponding Pfaff system is given by

$$\begin{aligned} dx^\alpha &= \lambda^\mu\{x^\alpha,\phi_\mu\}^*d\tau, \\ dp_\alpha &= \lambda^\mu\{p_\alpha,\phi_\mu\}^*d\tau, \\ &(\alpha = 0,...,n,\ \mu = 0,...,m-r), \end{aligned} \tag{6.11}$$

and

$$dS = \lambda^\mu p_\alpha\{x^\alpha,\phi_\mu\}^*d\tau, \quad \phi^*_\mu = 0,\ \phi_{\mu'} = 0, \tag{6.12}$$

where $\{\ ,\ \}^*$ is the usual Dirac bracket.

Let us conclude this section by considering the particular case where

$$\begin{aligned} \phi_\mu &= \begin{cases} \phi_0 = p_0 + H_i(x^i,p_i), \\ \phi_A = \phi_A(x^i,p_i), \end{cases} \\ &(i = 1,...,n,\ A = 1,...,m-r); \end{aligned} \tag{6.13}$$

Eqs. (6.11) become the equations of motion for a constrained system with a nonvanishing canonical Hamiltonian (for simplicity we will choose $\lambda_0 = 1$):

$$\begin{aligned} dx^i &= \{x^i,H_c + \lambda^A\phi_A\}^*\,d\tau, \\ dp_i &= \{p_i,H_c + \lambda^A\phi_A\}^*\,d\tau, \\ dx^0 &= d\tau, \quad \phi_A = 0, \\ dp_0 &= 0, \quad \phi_{\mu'} = 0, \end{aligned} \tag{6.14}$$

where we made use of the fact that $\phi_A$ and the second-class constraints $\phi_{\mu'}$ do not depend on $x^0$ and $p_0$.

## APPENDIX A

When first- and second-class constraints are present, that is, when the rank $r$ introduced in Sec. II is different from zero and $< m+1$, the problem of classifying the constraints $\phi_\rho$ into first class and second class can be solved in the following way.

We can start with the choice of a minor of maximum rank in the matrix $\|\{\phi_\rho,\phi_\sigma\}\|$. Since it is always possible to choose such a minor as a principal—hence antisymmetric—minor, let us suppose it to be $\|\{\phi_{\mu'},\phi_{\nu'}\}\|$, where $\mu'$, $\nu' = m-r+1,...,m$. The remaining functions $\phi_\mu$ ($\mu = 0,1,...,m-r$) are in general not first class, but from them we can construct a set of first-class constraints. To this end, we can observe that any row of the matrix $\|\{\phi_\rho,\phi_\sigma\}\|$ not belonging to the chosen minor must be a linear combination of the last $r$ rows. Explicitly this means

$$\{\phi_\mu,\phi_{\mu'}\} \approx A_{\mu\nu'}\{\phi_{\nu'},\phi_{\mu'}\},$$

which can be equivalently written

$$\{\phi_\mu - A_{\mu\nu'}\phi_{\nu'},\phi_{\mu'}\} \approx 0.$$

From this we see that we can define[3]

$$\phi^*_\mu = \phi_\mu - \{\phi_\mu,\phi_{\mu'}\}(C^{-1})^{\mu'\nu'}\phi_{\nu'}, \tag{A1}$$

which have the property $\{\phi^*_\mu,\phi_{\mu'}\} \approx 0$, with

$$C_{\mu'\nu'} = \{\phi_{\mu'},\phi_{\nu'}\}. \tag{A2}$$

We may now show that $\phi^*_\mu$ are first-class objects. Indeed let us choose $\phi^*_\mu$ and $\phi_{\mu'}$ as a new set of constraints.

We can verify that, by doing so, the rank $r$ does not change. In fact, if we put

$$\phi'_\mu = \phi^*_\mu, \quad \phi'_{\mu'} = \phi_{\mu'}, \tag{A3}$$

we have

$$\phi'_\rho = B_{\rho\sigma}\phi_0, \tag{A4}$$

where the matrix $B$ has rank $= m + 1$, since $\det B = 1$.

Owing to this, the rank of the matrix

$$\{\phi'_\rho, \phi'_\sigma\} \approx B_{\rho\lambda} B_{\sigma\tau} \{\phi_\lambda, \phi_\tau\} \tag{A5}$$

is again $r$. Then we necessarily have $\{\phi^*_\mu, \phi^*_\nu\} \approx 0$.

## APPENDIX B

Under the hypotheses already postulated in Sec. II, that is, that the functions $\phi_\rho$ are differentiable and that the rank of $x \to (\phi_0(x), \ldots, \phi_m(x))$ is equal to $m + 1$ for every $x$ belonging to the subset $V$ of the manifold $X$ (in our case $X = \mathbb{R}^{2n+3}$), defined by $\phi_\rho(x) = 0$, the subset $V$ is a submanifold of $X$ of dimension $N - m$ ($N = \dim X$) (Ref. 13, p. 228).

If $j$ denotes the inclusion mapping

$$j: V \to X, \tag{B1}$$

then the pullback $j^*$ and the differential mapping $j_*$ allow us to relate structures defined on $V$ and on $X$. For instance, if $f$ is a function $f: X \to \mathbb{R}$, then

$$j^* f: V \to \mathbb{R}, \tag{B2}$$

where

$$(j^* f)(x) = f(x), \quad \forall x \in V. \tag{B3}$$

It is convenient to introduce the following notation: If

$$f(x) = 0, \quad \omega_x = 0, \quad \forall x \in V, \tag{B4}$$

where $\omega_x$ is a differential form on $X$, we will write

$$f \approx 0, \ \omega \approx 0, \tag{B5}$$

and we will say that $f$, $\omega$ are weakly equal to zero.

Note that $f \approx 0$ is equivalent to $j^* f = 0$, when $f$ is a 0-form, but this is no longer true for a form of arbitrary degree.

In fact, it is possible to prove that if $\alpha \in \Lambda(T^*(X))$ is a differential form of degree $r$, $0 < r \leqslant N$, such that

$$j^* \alpha = 0, \tag{B6}$$

then

$$\alpha \approx \sum_\rho a^\rho \wedge d\phi_\rho, \tag{B7}$$

where $a^\rho$ is a differential form of degree $r - 1$.

The proof of the previous statement lies in the following. Let $\phi_\rho$ ($\rho = 1, \ldots, m$), $\psi_k$ ($k = m + 1, \cdots$) be a set of local coordinates for the manifold $X$, as is always possible under the stated hypotheses. A differential form of degree $r$ can be written as

$$\alpha = \sum_{\substack{(\rho)(k) \\ i=0}}^{r} \omega_{\rho_1 \cdots \rho_i k_{i+1} \cdots k_r} \, d\phi_{\rho_1} \wedge \cdots \wedge d\phi_{\rho_i} \wedge d\psi_{k_{i+1}} \wedge \cdots \wedge d\psi_{k_r}, \tag{B8}$$

then

$$j^* \alpha = 0 \tag{B9}$$

implies

$$j^* \omega_{k_1 \cdots k_r} = 0 \Leftrightarrow \omega_{k_1 \cdots k_r} \approx 0, \tag{B10}$$

where use is made of the fact that $j^*$ is an algebra homomorphism and

$$j^* \, d\phi_\rho = 0 \tag{B11}$$

due to $j^* \phi_\rho = 0$.

In conclusion, using Eqs. (B8) and (B11), we get

$$\alpha \approx \sum_\rho a^\rho \wedge d\phi_\rho, \tag{B12}$$

where $a^\rho$ is an $(r - 1)$ form.

The converse statement, that Eq. (B7) implies Eq. (B6), can be easily proved.

## APPENDIX C

In this appendix we want to analyze the problem posed by the presence of the 0-forms $\phi_\rho$ in the exterior differential system $A$.

Stated more precisely, let us consider an exterior differential system $A$ in the manifold $X$ (in our case $\mathbb{R}^{2n+3}$):

$$A = \begin{cases} \phi_\rho = 0 & (\rho = 0, \ldots, m), \\ \omega^\alpha = 0 & (\alpha = 1, \ldots, a), \end{cases} \tag{C1}$$

where $\omega^\alpha$ are differential forms of degree greater than zero.

An integral manifold of $A$ is a pair $(M, f)$, where $M$ is a submanifold $X$ and $f$ a differentiable mapping,

$$f: M \to X,$$

such that

$$\begin{cases} f^* \phi_\rho = 0, \\ f^* \omega^\alpha = 0. \end{cases} \tag{C2}$$

Let $g$ be the mapping

$$g: M \to V$$

and $j$ be the inclusion mapping (B1)

$$j: V \to X$$

so that

$$f = j \circ g. \tag{C3}$$

We know that $g$ too is a differentiable mapping.[20]

Now, finding an integral manifold $(M, f)$ of $A$ is equivalent to finding an integral manifold $(M, g)$ of the following system in $V$:

$$B = \{ j^* \omega^\alpha = 0 \}. \tag{C4}$$

Indeed, if $(M, f)$ is an integral manifold of $A$ in $X$, from (C2), (C3) we get

$$f^* \omega^\alpha = g^*(j^* \omega^\alpha) = 0,$$

so $(M, g)$ is an integral manifold of $B$. Conversely, if $(M, g)$ is an integral manifold of $B$, this means

$$g^*(j^* \omega^\alpha) = 0,$$

that is, $f^* \omega^\alpha = 0$. So the second set of Eqs. (C2) is satisfied.

On the other hand, from the definition of $V$ we get $j^* \phi_\rho = 0$, which implies $g^*(j^* \phi_\rho) = f^*(\phi_\rho) = 0$, so the system (C2) is satisfied, and the pair $(M, j \circ g)$ is an integral manifold of $A$ in $X$.

As a first application of this result, let us now consider the problem of determining the characteristic vectors when 0-forms are present. If $\overline{A}$,

$$\overline{A} = \begin{cases} \phi_\rho = 0, & d\phi_\rho = 0, \\ \omega^\alpha = 0, & d\omega^\alpha = 0, \end{cases} \tag{C5}$$

and $\overline{B}$,

$$\overline{B} = \begin{cases} j^*\omega^\alpha = 0, \\ d(j^*\omega^\alpha) = j^* d\omega^\alpha = 0, \end{cases} \tag{C6}$$

are the closures of $A$ and $B$, then the differential part of $\overline{A}$: $d\phi_\rho$, $\omega^\alpha$, $d\omega^\alpha$, generates the ideal $I_A$,

$$I_A = \left\{ \omega = \sum_\alpha \xi_\alpha \wedge \omega^\alpha + \sum_\alpha \eta_\alpha \wedge d\omega^\alpha + \sum_\rho \zeta^\rho \wedge d\phi_\rho \right\}, \tag{C7}$$

while $\overline{B}$ generates the ideal $I_B$,

$$I_B = \left\{ \tilde{\omega} = \sum_\alpha \tilde{\xi}_\alpha \wedge j^*\omega^\alpha + \sum_\alpha \tilde{\eta}_\alpha \wedge j^* d\omega^\alpha \right\}, \tag{C8}$$

where $\xi_\alpha$, $\eta_\alpha$, and $\zeta^\rho$ are arbitrary forms on $X$ and $\tilde{\xi}_\alpha$, $\tilde{\eta}_\alpha$ are arbitrary forms on $V$.

From (C7) and (C8) we have

$$I_B = j^* I_A. \tag{C9}$$

Let us observe that Eq. (C9) implies that the pullback of a form of $I_A$ belongs to $I_B$; we may ask whether other forms $\alpha$ belonging to $\wedge(T^*(X))$ exist such that $j^*\alpha \in I_B$. We may in fact prove the following statement:

$$j^*\alpha \in I_B \Leftrightarrow \alpha \underset{\sim}{\in} I_A, \tag{C10}$$

where the notation

$$\alpha \underset{\sim}{\in} I_A \tag{C11}$$

means $\omega_x \in (I_A)_x$ for any $x \in V$ (we will say that the differential form $\alpha$ on $X$ weakly belongs to $I_A$).

In fact, if $j^*\alpha \in I_B$, using Eq. (C9), which tells us that a form belonging to $I_B$ can be written as the pullback of some form of $I_A$, we may write

$$j^*\alpha = j^*\left( \sum_\alpha \xi_\alpha \wedge \omega^\alpha + \sum_\alpha \eta_\alpha \wedge d\omega^\alpha + \sum_\rho \zeta^\rho \wedge d\phi_\rho \right), \tag{C12}$$

where the term in brackets is an element of $I_A$. Using the fact that $j$ is an algebra homomorphism, we have

$$j^*\alpha = \sum_\alpha j^*\xi_\alpha \wedge j^*\omega^\alpha + \sum_\alpha j^*\eta_\alpha \wedge j^* d\omega^\alpha, \tag{C13}$$

from which we get

$$j^*\left( \alpha - \sum_\alpha \xi_\alpha \wedge \omega^\alpha - \sum_\alpha \eta_\alpha \wedge d\omega^\alpha \right) = 0. \tag{C14}$$

Thus, using (B7), we have

$$\alpha \approx \sum_\alpha \xi_\alpha \wedge \omega^\alpha + \sum_\alpha \eta_\alpha \wedge d\omega^\alpha + \sum_\rho \zeta^\rho \wedge d\phi_\rho, \tag{C15}$$

that is $\alpha \underset{\sim}{\in} I_A$. The converse is easily proved in an analogous way.

From the results (C10) we see that we can work in the space $X$ by considering forms which weakly belong to the ideal $I_A$.

The characteristic system is constructed by starting from the set of characteristic vector fields, which should be obtained from the condition

$$i_v(I_B) \subset I_B, \quad v \in T(V). \tag{C16}$$

We now want to demonstrate that this is equivalent to analyzing the consequences of the condition

$$i_{\bar{v}}(I_A) \underset{\sim}{\subset} I_A, \quad \bar{v} \in T(X). \tag{C17}$$

Indeed, if $v$ is a characteristic vector field of $B$, i.e., such that $i_v(I_B) \subset I_B$, since $I_B = j^* I_A$, this means

$$i_v(j^* I_A) \subset j^* I_A$$

or

$$j^* i_{j_* v}(I_A) \subset j^* I_A. \tag{C18}$$

Let us consider an extension $\bar{v}$ and $v$, that is a vector field on $X$ such that $\bar{v}_x = (j_* v)_x$ for every $x \in jV$; thus $\bar{v}$ will satisfy

$$i_{\bar{v}}\, d\phi_\rho = 0. \tag{C19}$$

From Eq. (C18) we get

$$j^* i_{\bar{v}}(I_A) \subset I_B,$$

and finally using Eq. (C11), we can write

$$i_{\bar{v}}(I_A) \underset{\sim}{\subset} I_A,$$

which is the result (C17).

Let us observe that this result guarantees that $\bar{v}$ is actually an extension of $v$, since by applying it to the 1-forms $d\phi_\rho$ of $I_A$ we get

$$i_{\bar{v}}(d\phi_\rho) \approx 0$$

inasmuch as no 0-forms are present in $I_A$, and this is exactly the definition of $\bar{v}$.

As a second application let us study how the Frobenius theorem about the integrability of Pfaff systems must be stated, when zero forms are present.

Let us consider the system $A$ in the case where $\omega^\alpha$ are independent 1-forms $\theta^\alpha$:

$$A = \begin{cases} \phi_\rho = 0 \quad (\rho = 0,\ldots,m), \\ \theta^\alpha = 0 \quad (\alpha = 1,\ldots,a). \end{cases} \tag{C20}$$

We already know that this system has the same solutions as the system

$$B = \{ j^*\theta^\alpha = 0 \}.$$

The Frobenius theorem states that the necessary and sufficient condition for the integrability of the system $B$ is

$$d(j^*\theta^\alpha) \wedge (j^*\theta^1) \wedge \cdots \wedge (j^*\theta^a) = 0 \tag{C21}$$

for any $\alpha$. From this, using the properties of the pullback, we get

$$j^*(d\theta^\alpha \wedge \theta^1 \wedge \cdots \wedge \theta^a) = 0$$

and from the result (B7) we get

$$d\theta^\alpha \wedge \theta^1 \wedge \cdots \wedge \theta^a \approx \sum_\rho \lambda^\rho d\phi_\rho,$$

or

$$d\theta^\alpha \wedge \theta^1 \wedge \cdots \wedge \theta^a \wedge d\phi_0 \wedge \cdots \wedge d\phi_m \approx 0, \tag{C22}$$

which is the integrability condition we were looking for.

The dual version of (C22) can likewise be easily obtained. Indeed the dual form of the system $B$ is given by the system of vector fields

$$B' = \{ v_\lambda \}, \quad v_\lambda \in T(V), \tag{C23}$$

where $\lambda = a+1,\ldots,n$, if $n$ is the dimension of the manifold, and $v_\lambda$ are defined by

$$i_{v_\lambda}(j^*\theta^\alpha) = 0 \tag{C24}$$

for all values of $\lambda$ and $\alpha$. The dual form of the Frobenius theorem requires for $v_\lambda$:

$$[v_\lambda, v_\sigma] = c_{\lambda\sigma\tau} v_\tau. \tag{C25}$$

Equation (C23) suggests the following form for the dual of the system $A$:

$$A' = \begin{cases} \bar{v}_\lambda \\ \phi_\rho = 0, \end{cases} \tag{C26}$$

$\bar{v}_\lambda$ being an extension of $v_\lambda$ to $T(X)$. It is obvious from the definition (C26) that $A'$ has the same solutions as $A$.

Let us observe that $A'$ is constructed by simply requiring the vectors $\bar{v}_\lambda$ to be weakly incident to the forms $\theta^\alpha$ and $d\phi_\rho$. Indeed we have

$$\begin{aligned}(v(j^*f))(\rho) &= v_\rho(j^*f) = (j_* v)_{j\rho}(f) \\ &= \bar{v}_{j_\rho}(f) = (j^*(\bar{v}(f)))(\rho),\end{aligned} \tag{C27}$$

that is

$$j^*(\bar{v}(f)) = v(j^*\theta), \tag{C28}$$

for any function $f$ on $X$. This can be extended to any 1-form:

$$j^*(i_{\bar{v}}\theta) = i_v(j^*\theta). \tag{C29}$$

Applying this result to $\bar{v}_\lambda$, we get

$$j^*(\bar{v}_\lambda \theta^\alpha) = i_{v_\lambda}(j^*\theta^\alpha), \tag{C30}$$

and furthermore we have

$$j^*(\bar{v}_\lambda \phi_\rho) = v_\lambda(j^*\phi_\rho) \tag{C31}$$

due to the definition of $V$.

Using the result of Appendix B, we may write (C30) and (C31) as

$$\begin{aligned} i_{\bar{v}_\lambda}\theta^\alpha &\approx 0, \\ i_{\bar{v}_\lambda} d\phi_\rho &\approx 0. \end{aligned} \tag{C32}$$

In order to find the integrability condition for the system $A'$, from (C25) and using the property (C28) we get

$$[v_\lambda, v_\sigma](j^*f) = c_{\lambda\sigma\tau} v_\tau(j^*f)$$

or

$$j^*([\bar{v}_\lambda, \bar{v}_\sigma](f)) = j^*(c'_{\lambda\sigma\tau}\bar{v}_\tau(f)),$$

where

$$j^* c'_{\lambda\sigma\tau} = c_{\lambda\sigma\tau},$$

that is,

$$[\bar{v}_\lambda, \bar{v}_\sigma] \approx c'_{\lambda\sigma\tau}\bar{v}_\tau. \tag{C33}$$

Equation (C33) is the integrability condition, that is, the Frobenius theorem, for the system $A'$.

## APPENDIX D

The transformation properties of the Hamilton–Jacobi function $S(x)$ under canonical transformations are known[21]; in the case of an infinitesimal transformation they have been given in particular by Bergmann.[22]

Here we only recall the results.

Let us consider the situation described in Sec. III. If we know a complete integral of system $\phi_\mu(x,p) = 0$ given by (3.37):

$$\begin{aligned} &S(x) = \phi(x^\alpha, c_k) + c \\ &(\alpha = 0,1,\ldots,n,\ k = m+1,\ldots,n), \end{aligned} \tag{D1}$$

we may ask what is the function

$$S'(x') = \phi(x'^\alpha, c_k) + c' \tag{D2}$$

when we submit the system to the finite canonical transformation generated by the function $w(x,p)$, via the equations

$$\begin{aligned} p'_\alpha &= e^{w(x,p)}{}_* p_\alpha, \\ x'^\alpha &= e^{w(x,p)}{}_* x^\alpha, \end{aligned} \tag{D3}$$

where the $*$ operation is defined in Eq. (3.32).

More generally, we may consider the one-parameter canonical transformation

$$\begin{aligned} p'_\alpha(\tau) &= e^{\tau w(x,p)}{}_* p_\alpha, \\ x'^\alpha(\tau) &= e^{\tau w(x,p)}{}_* x^\alpha. \end{aligned} \tag{D4}$$

The answer is given by the solution of the Hamilton–Jacobi equation:

$$\frac{\partial S'(x',\tau)}{\partial \tau} + w\left(x', \frac{\partial S'(x',\tau)}{\partial x'}\right) = 0, \tag{D5}$$

with the boundary condition

$$S'(x',0) = S(x). \tag{D6}$$

In the case of an infinitesimal canonical transformation, by neglecting higher orders of $\tau$, we get[21]

$$\delta S(x) = \tau w\left(x, \frac{\partial S}{\partial x}\right). \tag{D7}$$

Equation (D7) shows in particular that, when the canonical transformation is generated by the constraints $\phi_\mu$, that is, for $w$ a given linear combination of $\phi_\mu$, the function $S$ is invariant.

As stressed by Bergmann,[22] this has the consequence that $S$ is form-invariant under the invariance group (group generated by the constraints). In particular, the Hamilton–Jacobi theory does not require the setting of gauge conditions, which is, by the way, one of the main reasons for the relevance ascribed by the authors to a Hamilton–Jacobi approach to constrained systems in particle dynamics.

[1] Ph. Droz-Vincent, Lett. Nuovo Cimento **1**, 839 (1969); **7**, 206 (1973); Phys. Scripta **2**, 129 (1970); Rep. Math. Phys. **8**, 79 (1975); Ann. Inst. H. Poincaré **27**, 407 (1977); Phys. Rev. D **19**, 702 (1979); "*N*-Body Relativistic System," Ann. Inst. H. Poincaré **32**, 377 (1980); L. Bel, "Meccanica Relativistica Predictiva," curso Depart. Fis. Teor., Barcelona Univ. UAB FT-34, 1977; Ann. Inst. H. Poincaré **14**, 189 (1971). L. Bel and J. Martin, Ann. Inst. H. Poincaré **22**, 173 (1975); **33**, 409 (1980). L. Bel and F. X. Fustero, Ann. Inst. H. Poincaré **25**, 411 (1976). I. T. Todorov, "Dynamics of Relativistic Point Particles as a Problem with Constraints," Comm. JINR, EZ-10125, Dubna, 1976; "Constraint Hamiltonian Dynamics of Directly Interacting Relativistic Point Particles," Lectures at the Summer School in Math. Phys. Bogazici Univ., Bebek, Istanbul, 1979, and at the 17th Winter School of Theoretical Physics, Karpacz, 1980; "Constraint Hamiltonian Approach to Relativistic Point Particles Dynamics," lectures at the ICTP, Trieste, 1980. V. V. Molotkov and I. T. Todorov, Comm. Math. Phys. **79**, 11 (1981). S. N. Sokolov, Dokl. Akad. Nauk. SSSR **233**, 575 (1977) [Sov. Phys. Dokl. **22**, 198 (1977)]; Teor. Mat. Fiz. **36**, 193 (1978) [Theor. Math. Phys. **36**, 682 (1978)]. M. Kalb and P. Van Alstine, "Invariant Singular Action for the Relativistic Two-Body Problem: A Hamiltonian Formulation," Yale Report COO-3075-146, 1976. T. Takabayasi, Prog. Theor. Phys. **57**, 331 (1977); Suppl. Prog. Theor. Phys. **67**, 1 (1979). K. Kamimura and T. Shimizu, Prog. Theor. Phys. **58**, 383 (1977). K. Kamimura, Nuovo

Cimento B **65**, 181 (1981). D. Dominici, J. Gomis, and G. Longhi, Nuovo Cimento B **48**, 152 (1978); Nuovo Cimento A **48**, 257 (1978); A **56**, 263 (1980). R. Giachetti and E. Sorace, Nuovo Cimento A **43**, 281 (1978); Nuovo Cimento B **56**, 263 (1980); Lett. Nuovo Cimento **26**, 1 (1979). A. Komar, Phys. Rev. D **18**, 1881, 1887, 3017 (1978); **19**, 2908 (1979); "Relativistic Action-at-a-distance and quasiseparability," **24**, 2330 (1981); F. Rohrlich, Ann. Phys. (N.Y.) **117**, 292 (1979); **130**, 350 (1980); Phys. Rev. D **23**, 1305 (1981). H. Sazdjian, Nucl. Phys. B **161**, 469 (1979); Ann. Phys. (N.Y.) **136**, 136 (1981). A. Kihlberg, R. Marnelius, and N. Mukunda, Phys. Rev. D **23**, 2201 (1981); N. Mukunda and E. C. G. Sudarshan, Phys. Rev. D **23**, 2210 (1981); E. C. G. Sudarshan, N. Mukunda, and J. N. Goldberg, Phys. Rev. D **23**, 2218, 2231 (1981). L. Lusanna, "A Model for $N$ Classical Relativistic Particles," Nuovo Cimento A **64**, 65 (1981); L. Lusanna, Nuovo Cimento B **65**, 135 (1981). J. Gomis, J. A. Lobo, and J. M. Pons, Nuovo Cimento B **64**, 316 (1981); Ann. Inst. H. Poincaré **35**, 17 (1981). A. P. Balachandran *et al.*, "Relativistic Particle Interactions: a Third World View," preprint, University di Napoli, "Unified Geometrical Approach to Relativistic Particle Dynamics," preprint, University of Texas at Austin, 1981; "Separability in Relativistic Hamiltonian Particle Dynamics," Phys. Rev. D **26**, 3492 (1982); V. M. Penafiel, S. Orenstein, and K. Rafanelli, "Canonical Formalism for $N$ Interacting Relativistic Particles," preprint, Queens College, 1981. J. Samuel, "Relativistic Particle Models with Separable Interactions," Phys. Rev. D **26**, 3482 (1982); "Constraints in Relativistic Hamiltonian Mechanics," **26**, 3475 (1982).

[2]See also the *Proceedings of the Workshop "Relativistic Action at a Distance: Classical and Quantum Aspects,"* edited by J. Llosa, *Lecture Notes in Physics*, Vol. 162 (Springer, Berlin, 1981).

[3]P. A. M. Dirac, Can. J. Math. **2**, 129 (1950); **3**, 1 (1951); Proc. R. Soc. London A **246**, 326 (1958); *Lectures on Quantum Mechanics*, Belfer Graduate School, Yeshiva University, 1964.

[4]S. Shanmugadhasan, J. Math. Phys. **14**, 677 (1973); E. C. G. Sudarshan and N. Mukunda, *Classical Dynamics: A Modern Perspective* (Wiley, New York, 1974); A. G. Hanson, T. Regge, and C. Teitelboim, *Contributi del Centro Linceo Interdiscipl. di Scienze Mat. Fis. e Loro Applic., No. 22* (Lincei, Rome, 1976).

[5]For application of the Hamilton–Jacobi theory to constrained systems see A. Komar, Phys. Rev. D **18** (1978). For other applications, especially field theory, see, for instance: P. G. Bergmann, Phys. Rev. **144**, 1078 (1966); H. A. Kastrup, Phys. Lett. B **70**, 195 (1977); K. Kuchar, "Canonical Methods of Quantization," Utah Preprint UT 84 112, University of Utah, 1980, and references quoted therein.

[6]I. T. Todorov in Ref. 1.

[7]A. Komar in Ref. 1.

[8]G. Longhi in Ref. 2.

[9]In order to guarantee the stability condition $\phi_\rho = 0$, when $\phi_\rho$ belong to the first-class set, it is usually assumed that

$$\{H_c,\phi_\rho\} = \sum_\sigma c_{\rho\sigma}\phi_\sigma \approx 0.$$

[10]This assumption includes the case where $\|\partial\phi_\rho/\partial p_\alpha\|$ has rank $m+1$ as a particular but more physical case. Indeed, if this last condition does not hold, there will necessarily exist a relation between the coordinates $x^\alpha$, which usually contradicts the physical intuition, apart from particular systems like the rigid rotator.[11] In our subsequent development of the theory, this will require an interchange of some momentum with the corresponding conjugate coordinate, via a canonical transformation. Thus we will maintain the hypothesis (2.9) for the sake of generality.

[11]See Hanson *et al.* in Ref. 4.

[12]Mathematical textbooks in which the theory of systems of partial differential equations (PDE) of the first order can be found, are, for instance: E. Goursat, *A Course in Mathematical Analysis* (Dover, New York, 1959), Vol. II, Part Two, F. G. Tricomi, *Equazioni a Derivate Parziali* (Cremonese, Rome, 1957); L. E. P. Eisenhart, *Continuous Groups of Transformations* (Dover, New York, 1961); G. F. D. Duff, *Partial Differential Equations* (Univ. of Toronto Press, Toronto, 1956); J. Dieudonné, *Treatise on Analysis* (Academic, New York, 1974), Vol. IV, Chap. XVIII; Y. Choquet-Bruhat, C. DeWitt-Morette, and M. Dillard-Bleick, *Analysis, Manifolds and Physics* (North-Holland, Amsterdam, 1977), Chap. IV, Sec. C; R. Hermann, *Interdisciplinary Mathematics* (Math. Sci. Press, Brookline, MA, 1977), esp. Vol. XIV, Chaps. XV and XVI.

[13]Y. Choquet-Bruhat *et al.* in Ref. 12.

[14]Since Eqs. (2.14) determine $v$ only on the constraint surface, for the sake of simplicity we have chosen as a particular solution that given by Eqs. (2.17) and (2.18) as true equalities in the whole space $R^{2n+3}$. It is possible to verify that this particular choice does not entail any loss of generality.

[15]Goursat in Ref. 12.

[16]J. A. Schouten and W. V. der Kulk, *Pfaff's Problem and Its Generalizations* (Clarendon, Oxford, 1949).

[17]See, for instance, Goursat and Tricomi in Ref. 12.

[18]This point requires a further word. We have seen that $S$ as a function of $x$ does not exist. However, as mentioned at the beginning of this section, it can exist in some reduced space. Let us clarify this point. If we perform a canonical transformation on the original phase space, such that $2s$ of the new canonical variables are algebraically equivalent to the second-class constraints,[4] the remaining canonical variables define the so-called reduced phase space. Clearly, in this space, the set of equations for the Hamilton–Jacobi function reduces to $\phi^*_\mu = 0$, with the above-mentioned set of $2s$ variables put to zero. Since $\phi^*_\mu$ are first class, a Hamilton–Jacobi function will exist in the reduced configuration space.

[19]Sudarshan and Mukunda in Ref. 4.

[20]Pham Man Quan, *Introduction à la géométrie des variétés différentiables* (Dunod, Paris, 1969).

[21]F. J. Testa, J. Math. Phys. **11**, 2698 (1970); C. Lanczos, *Variational Principles of Mechanics* (Univ. of Toronto Press, Toronto, 1962).

[22]P. G. Bergmann in Ref. 5.

# The equivalence problem for nonconservative mechanics

J.-E. Werth
*Departamento de Física, Universidade Federal da Paraíba, 58.000-João Pessoa-PB, Brazil*

(Received 19 September 1983; accepted for publication 9 March 1984)

A geometric study of the equivalence problem for nonconservative mechanical systems is presented. Three equivalence relations for mechanical systems arise naturally in this framework: $\mathscr{L}$-, $\mathscr{G}$-, and $\mathscr{H}$-equivalence. Necessary and sufficient conditions for the $\mathscr{H}$-equivalence of two systems are derived. The connection between $\mathscr{G}$-transformations, $\mathscr{G}$-equivalence, and canonical transformations is investigated. Furthermore, the relationship to (geometric) quantization is discussed.

PACS numbers: 03.20. + i

## I. INTRODUCTION

There is a natural one-to-one correspondence between hyper-regular Lagrangian and Hamiltonian mechanical systems. However, the relationship between the associated dynamical systems (equations of motion) is more complicated. It is well known that the classically observable trajectories of some mechanical system do not uniquely determine their Lagrangian; this results in an ambiguity for the Hamiltonian and may lead to inequivalent quantum and statistical theories.[1]

A treatment of this problem for nonconservative systems is presented from the differential geometric point of view. Two mechanical $\mathscr{L}$-systems are called $\mathscr{L}$-equivalent if the associated dynamical systems on the tangent bundle $TQ$ (velocity phase space) coincide; they are $\mathscr{H}$-equivalent if the dynamical systems associated to the corresponding $\mathscr{H}$-systems on the cotangent bundle $T^*Q$ (momentum phase space) coincide. A more general definition of a Legendre transformation based on 1-forms provides for the construction of a map $\delta$ from the set of hyper-regular dynamical systems on $T^*Q$ into the set of dynamical systems on $TQ$. This is most useful when applied to the equivalence problem: Proposition 4 essentially says that $\mathscr{H}$-equivalence implies $\mathscr{L}$-equivalence. The converse is not true (Proposition 7). Thus the solution curves for different dynamical systems on $T^*Q$ may coincide when projected to configuration space $Q$.

Finally, we study the equivalence problem within the realm of geometric quantization and quantum mechanics. In general, $\mathscr{L}$-equivalent mechanical systems may lead to inequivalent quantum theories in the sense that the energy spectra are different. This motivates the introduction of a third type of equivalence: Two mechanical systems are $\mathscr{G}$-equivalent if the solution curves on $T^*Q$ differ by a certain canonical transformation, whereas the projected solution curves on $Q$ are identical. It is demonstrated that in a (geometrical) Mackey-type quantization via vector field operators,[2] $\mathscr{G}$-equivalent mechanical systems lead to unitarily equivalent operators.

## II. MECHANICAL SYSTEMS

In this section we shall briefly describe the most important properties of mechanical systems which can be found in Ref. 3 and Ref. 4. The notation is essentially that of Ref. 4. Denote by $\mathscr{F}M$ the space of smooth real-valued functions on a manifold $M$. Take a Lagrangian $L\in\mathscr{F}(TQ)$. Then the map $FL: TQ\to T^*Q$ given by $FL(v_q)\,w_q = dL(w_q)^1_{v_q}$ is called the *Legendre transformation* associated to $L$. Here $(w_q)^1_{v_q} = d/dt|_{t=0}(v_q + tw_q)\in T_{v_q}TQ$ denotes the vertical lift of $w_q\in T_qQ$ with respect to $v_q\in T_qQ$. $L\in\mathscr{F}(TQ)$ is called hyper-regular if $FL$ is a diffeomorphism, and $\mathscr{F}_{hr}(TQ)$ denotes the set of all hyper-regular Lagrangians. $\omega_0 = -d\theta_0$ denotes the canonical symplectic structure on $T^*Q$, and for $L\in\mathscr{F}_{hr}(TQ)$, $\omega_L$ denotes the symplectic structure $(FL)^*\omega_0$ on $TQ$. Let $\mathfrak{X}M$ be the Lie algebra of smooth vector fields on $M$. There is a linear isomorphism $\omega_L^\flat: \mathfrak{X}(TQ)\to\Omega^1(TQ)$ given by $\omega_L^\flat(X) = i_X\omega_L$, i.e., $\omega_L^\flat(X)\,\xi_{v_q} = \omega_L(X_{v_q}, \xi_{v_q})$ for any $\xi_{v_q}\in T_{v_q}TQ$. A 1-form $\pi\in\Omega^1(TQ)$ is called semibasic if $\pi(\xi_{v_q}) = 0$ for any $\xi_{v_q}\in T_{v_q}TQ$ with $T\tau_Q\,\xi_{v_q} = 0$. $\Omega^1_{sb}(TQ)$ denotes the space of semibasic 1-forms on $TQ$. Finally, denote by $\mathscr{L}(TQ) = \mathscr{F}_{hr}(TQ)\times\Omega^1_{sb}(TQ)$ the set of mechanical $\mathscr{L}$-systems. Define $\lambda: \mathscr{L}(TQ)\to\mathfrak{X}(TQ)$ by $\lambda(L,\pi) = (\omega_L^\flat)^{-1}\times(dE_L + \pi)$, where $E_L\in\mathscr{F}(TQ)$ is the energy given by $E_L(v_q) = FL(v_q)\,v_q - L(v_q)$. One says that $\lambda(L,\pi)$ is the dynamical system associated to the mechanical system $(L,\pi)$. We also write $\lambda(L,\pi) = X_L^{dE_L+\pi} = X_{(L,\pi)}$. Observe that $X_{(L,\pi)}\in\mathfrak{X}(TQ)$ is a second-order equation on $Q$, i.e., $T\tau_Q\,X_{(L,\pi)}(v_q) = v_q$ for $v_q\in TQ$.

## III. GENERALIZED FORCES

A generalized force on $Q$ is a differentiable map $K: TQ\to TQ$ which leaves the fibers $T_qQ$ invariant. $\mathscr{K}(TQ)$ denotes the space of generalized forces on $Q$. Given $L\in\mathscr{F}_{hr}(TQ)$, define $F^2L_q: T_qQ\to L(T_qQ, T^*_qQ)$ by $(F^2L_q(v_q)\,w_q)\,x_q = d/dt|_{t=0}FL_q(v_q + tw_q)\,x_q$. Since $L$ is hyper-regular, $F^2L_q(v_q): T_qQ\to T^*_qQ$ is a linear isomorphism; we also write $(F^2L_q(v_q)w_q)x_q = F^2L_q(v_q)(w_q, x_q)$. Moreover, $L\in\mathscr{F}_{hr}(TQ)$ induces a bijection $\phi^L: \mathscr{K}(TQ)\to\Omega^1_{sb}(TQ)$ by $(\phi^L(K))\xi_{v_q} = -F^2L_q(v_q)(K(v_q), T\tau_Q\xi_{v_q})$ for $\xi_{v_q}\in T_{v_q}TQ$. Thus given a generalized force field $\pi\in\Omega^1_{sb}(TQ)$ and a hyper-regular Lagrangian $L\in\mathscr{F}_{hr}(TQ)$, we associate a generalized force $K^\pi_L\in\mathscr{K}(TQ)$ via $\phi^L(K^\pi_L) = \pi$. A vector field $X\in\mathfrak{X}(TQ)$ is called vertical if $T\tau_Q X_{v_q} = 0$ for any $v_q\in TQ$. $\mathfrak{X}_V(TQ)$ denotes the space of vertical vector fields on $TQ$. Any $K\in\mathscr{K}(TQ)$ induces a vertical vector field $X_K\in\mathfrak{X}_V(TQ)$, $X_K(v_q) = (K(v_q))^1_{v_q}$. We shall need the following result.

*Proposition 1*: Let $\pi\in\Omega^1_{sb}(TQ)$. Then $X_{K^\pi_L} = X^\pi_L$.

*Proof*: Write $X_{K^\pi_L} = X_K$. Then (see Ref. 4, p. 117)

 0022-2488/84/082453-03$02.50  

$$\omega_L(X_K,\xi_{v_q}) = \xi_{v_q}\theta_L(X_K) - X_K(v_q)\theta_L(\xi) + \theta_L(v_q)[X_K,\xi]$$

for any (local) vector field $\xi$, $\xi(v_q) = \xi_{v_q} \in T_{v_q}TQ$. Choose $\xi$ such that $[X_K,\xi] = 0$. Since $\theta_L(\eta) = 0$ for any $\eta \in \mathfrak{X}_V(TQ)$, it follows that $\omega_L(X_K,\xi_{v_q}) = -X_K(v_q)\theta_L(\xi) = -d/dt\,|_{t=0}\theta_{\gamma(t)}(\xi_{\gamma(t)})$, where $\gamma(t)$ denotes the integral curve of $X_K$ at $v_q \in T_qQ \subset TQ$. The flow of $X_K$ leaves the fibers $T_qQ \subset TQ$ invariant, and, since $[X_K,\xi] = 0$, transforms integral curves of $\xi$ onto integral curves of $\xi$. Thus $T\tau_Q\xi_{v_q} = T\tau_Q\,\xi_{w_q}$ if $v_q$ and $w_q$ are connected by an integral curve of $X_K$. Hence, for any $\xi_{v_q} \in T_{v_q}TQ$,

$$\begin{aligned}\omega_L(X_K,\xi_{v_q}) &= -\frac{d}{dt}\Big|_{t=0} FL(\gamma(t))(T\tau_Q\xi_{v_q}) \\ &= -\frac{d}{dt}\Big|_{t=0} FL(v_q + tK(v_q))(T\tau_Q\xi_{v_q}) \\ &= \pi(\xi_{v_q}).\end{aligned}$$

This proves the assertion.

## IV. LEGENDRE TRANSFORMATION BASED ON 1-FORMS

The above treatment applies equally well to the Hamiltonian description of mechanical systems. We shall now introduce a more general definition of a Legendre transformation, emphasizing the Hamiltonian point of view. Given a 1-form $\alpha \in \Omega^1(T^*Q)$, define $F\alpha : T^*Q \to T^{**}Q = \cup_{q\in Q} T_q^{**}Q$ by $(F\alpha)(\beta_q)\gamma_q = \alpha(\gamma_q)^1_{\beta_q}$; $\beta_q,\gamma_q \in T_q^*Q$. $\alpha \in \Omega^1(T^*Q)$ is called hyper-regular if $F\alpha$ is a diffeomorphism, and $\Omega^1_{hr}(T^*Q)$ denotes the set of hyper-regular 1-forms on $T^*Q$. Consider the exterior derivative $d: \mathscr{F}(T^*Q) \to \Omega^1(T^*Q)$. Then $\mathscr{F}_{hr}(T^*Q)$ denotes the set $d^{-1}(\Omega^1_{hr}(T^*Q)) \subset \mathscr{F}(T^*Q)$ of hyper-regular Hamiltonians. Moreover, let $\mathscr{H}(T^*Q) = \mathscr{F}_{hr}(T^*Q) \times \Omega^1_{sb}(T^*Q)$ denote the set of mechanical $\mathscr{H}$-systems.

*Proposition 2*: Suppose $(H,\varphi) \in \mathscr{H}(T^*Q)$. Then $dH + \varphi \in \Omega^1_{hr}(T^*Q)$, and $FH = F(dH + \varphi)$.

*Proof*: It is sufficient to prove that $FH = F(dH + \varphi)$. Now, $F(dH + \varphi)(\beta_q)\gamma_q = (dH + \varphi)(\gamma_q)^1_{\beta_q} = dH(\gamma_q)^1_{\beta_q} + \varphi(\gamma_q)^1_{\beta_q}$. Since $\varphi$ is semibasic and $(\gamma_q)^1_{\beta_q}$ is vertical, $\varphi(\gamma_q)^1_{\beta_q} = 0$. The result now follows.

There is a bijection $\chi$: $\mathscr{L}(TQ) \to \mathscr{H}(T^*Q)$ given by $\chi(L,\pi) = (E_L \circ (FL)^{-1}, (FL)_*\pi)$. One knows[4] that for $H = E_L \circ (FL)^{-1}$ the diagram

$$\begin{array}{ccc} & FL & \\ TQ & \longrightarrow & T^*Q \\ & {\scriptstyle\rho}\searrow \quad \swarrow{\scriptstyle FH} & \\ & T^{**}Q & \end{array} \qquad (*)$$

commutes, where $\rho$ denotes the natural bijection $TQ \cong T^{**}Q$. Consider the linear isomorphism $\omega_0^b : \mathfrak{X}(T^*Q) \to \Omega^1(T^*Q)$. Define $\theta: \mathscr{H}(T^*Q) \to \mathfrak{X}(T^*Q)$ by $\theta(H,\varphi) = (\omega_0^b)^{-1}(dH + \varphi) \equiv X_{(H,\varphi)} \equiv X^{dH+\varphi}$. Let $\mathfrak{X}_{hr}(T^*Q) = (\omega_0^b)^{-1}(\Omega^1_{hr}(T^*Q))$ denote the set of hyper-regular vector fields on $T^*Q$. By Proposition 2, $\mathrm{Im}\,\theta \subset \mathfrak{X}_{hr}(T^*Q)$. Now let $\delta : \mathfrak{X}_{hr}(T^*Q) \to \mathfrak{X}(TQ)$ be the association given by $X \to (\rho^{-1} \circ F((\omega_0^b)^{-1}X))_*X$. It is readily checked that, if $\chi(L,\pi) = (H,\varphi)$, then $\lambda(L,\pi) = (\rho^{-1} \circ FH)_*X_{(H,\varphi)}$. Thus we have the following fact.

*Proposition 3*: The diagram

$$\begin{array}{ccc} \mathscr{L}(TQ) & \xrightarrow{\chi} & \mathscr{H}(T^*Q) \\ \downarrow{\scriptstyle\lambda} & & \downarrow{\scriptstyle\theta} \\ \mathfrak{X}(TQ) & \xleftarrow{\delta} & \mathfrak{X}_{hr}(T^*Q) \end{array}$$

commutes.

## V. $\mathscr{L}$-, $\mathscr{G}$-, AND $\mathscr{H}$-EQUIVALENCE

We shall now turn to the discussion of the equivalence problem. Two mechanical $\mathscr{L}$-systems $(L,\pi),(L',\pi') \in \mathscr{L}(TQ)$ are called $\mathscr{L}$-equivalent if $X_{(L,\pi)} = X_{(L',\pi')}$. Furthermore, two $\mathscr{L}$-equivalent systems $(L,\pi)$ and $(L',\pi')$ are called $\mathscr{G}$-equivalent if $\omega_L = \omega_{L'}$; they are called $\mathscr{H}$-equivalent if $\theta_L = \theta_{L'}$. If we denote the corresponding equivalence classes of $(L,\pi) \in \mathscr{L}(TQ)$ by $(L,\pi)_{\mathscr{L}}$, $(L,\pi)_{\mathscr{G}}$, and $(L,\pi)_{\mathscr{H}}$, respectively, then $(L,\pi)_{\mathscr{H}} \subseteq (L,\pi)_{\mathscr{G}} \subseteq (L,\pi)_{\mathscr{L}}$. In other words, the $\mathscr{H}$-equivalence relation is finer than the $\mathscr{G}$-equivalence relation, and the $\mathscr{G}$-equivalence relation is finer than the $\mathscr{L}$-equivalence relation. The converse is not true, as we shall prove in Sec. VI. For a characterization of $\mathscr{H}$-equivalence we shall need the following fact.

*Lemma 1*: Let $L, L' \in \mathscr{F}_{hr}(TQ)$. Then $\theta_L = \theta_{L'}$ iff $FL = FL'$.

*Proof*: The "if" part is clear. For the converse, $(FL)^*\theta_0 = (FL')^*\theta_0$ implies $FL(v_q)(T\tau_Q\xi_{v_q}) = FL'(v_q)(T\tau_Q\xi_{v_q})$ for any $\xi_{v_q} \in TTQ$. Since $T\tau_Q$ is surjective, $FL = FL'$ results.

*Proposition 4*: Let $(L,\pi),(L',\pi') \in \mathscr{L}(TQ)$. Then $(L,\pi)$ and $(L',\pi')$ are $\mathscr{H}$-equivalent iff $X_{\chi(L,\pi)} = X_{\chi(L',\pi')}$.

*Proof*: If $(L,\pi)$ and $(L',\pi')$ are $\mathscr{H}$-equivalent, then $X_{(L,\pi)} = X_{(L',\pi')}$ and $\theta_L = \theta_{L'}$. By Lemma 1, $FL = FL'$. It follows $X_{\chi(L,\pi)} = (FL)_*X_{(L,\pi)} = (FL')_*X_{(L',\pi')} = X_{\chi(L',\pi')}$. Conversely, if $(\theta \circ \chi)(L,\pi) = (\theta \circ \chi)(L',\pi')$, then, by Proposition 3, $\lambda(L,\pi) = \lambda(L',\pi')$. Moreover, if we write $\chi(L,\pi) = (H,\varphi)$, $\chi(L',\pi') = (H',\varphi')$, then $dH + \varphi = dH' + \varphi'$. Hence, by Proposition 2, $FH = FH'$. Since $FH \circ FL = \rho$, we conclude that $FL = FL'$. Hence, $\theta_L = \theta_{L'}$ and the result follows.

We come now to the connection between $\mathscr{G}$-equivalence and canonical transformations. A diffeomorphism $\delta : T^*Q \to T^*Q$ is called *canonical transformation* if $\delta^*\omega_0 = \omega_0$. The following result is easy to see.

*Lemma 2*: Let $L, L' \in \mathscr{F}_{hr}(TQ)$. Then $\omega_L = \omega_{L'}$ iff $FL \circ (FL')^{-1}$ is a canonical transformation.

$X, Y \in \mathfrak{X}(T^*Q)$ are called *canonically equivalent* if there is a canonical transformation $\delta : T^*Q \to T^*Q$ such that $\delta_*X = Y$.

*Proposition 5*: If two mechanical systems $(L,\pi)$ and $(L',\pi')$ are $\mathscr{G}$-equivalent, then $X_{\chi(L,\pi)}$ and $X_{\chi(L',\pi')}$ are canonically equivalent.

*Proof*: Note that $X_{\chi(L,\pi)} = (FL)_*X_{(L,\pi)}$ for any $(L,\pi) \in \mathscr{L}(TQ)$. Since $X_{(L,\pi)} = X_{(L',\pi')}$, we obtain $X_{\chi(L,\pi)} = (FL)_*(FL')^{-1}X_{\chi(L',\pi')} = (FL \circ (FL')^{-1})_*X_{\chi(L',\pi')}$. By Lemma 2, $FL \circ (FL')^{-1}$ is a canonical transformation. This proves the assertion.

## VI. $\mathscr{G}$-TRANSFORMATIONS

Given $\alpha \in \Omega^1 Q$, define $\tilde{\alpha} \in \mathscr{F}(TQ)$ by $\tilde{\alpha}(v_q) := \alpha_q(v_q)$, and $\pi_\alpha \in \Omega^1_{sb}(TQ)$ by $\tau_\alpha(\xi_{v_q}) := d\alpha(v_q, T\tau_Q \xi_{v_q})$. Observe that $\pi_\alpha = 0$ if $\alpha$ is closed.

*Proposition 6*: For $\alpha \in \Omega^1 Q$, $(L,\pi) \in \mathscr{L}(TQ)$ we have (i) $(L+\tilde{\alpha}, \pi - \pi_\alpha) \in \mathscr{L}(TQ)$; (ii) $(L,\pi)_{\mathscr{L}} = (L+\tilde{\alpha}, \pi - \pi_\alpha)_{\mathscr{L}}$; (iii) $\theta_{L+\tilde{\alpha}} = \theta_L + \tau_Q^* \alpha$; and (iv) $\omega_{L+\tilde{\alpha}} = \omega_L - \tau_Q^* d\alpha$.

*Proof*: We have $F(L+\tilde{\alpha}) = FL + \alpha \circ \tau_Q$; thus (i) follows. (iii) is easily checked. (iv) follows from (iii). Now, by Proposition 1, $X_L^\pi$ and $X_{L+\tilde{\alpha}}^\pi$ are vertical vector fields. Since $\omega_{L+\tilde{\alpha}} = \omega_L - \tau_Q^* d\alpha$, we obtain (1) $X_L^\pi = X_{L+\tilde{\alpha}}^\pi$. Observe that for the second-order equation $X_L^{dE_L}$ we have $(\tau_Q^* d\alpha)(X_L^{dE_L}, \xi) = \pi_\alpha(\xi)$ for any $\xi \in \mathfrak{X}(TQ)$. Since $E_L = E_{L+\tilde{\alpha}}$, it is easy to check that (2) $X_L^{dE_L} = X_{L+\tilde{\alpha}}^{dE_{L+\tilde{\alpha}} - \pi_\alpha}$. If we combine (1) and (2), the assertion (ii) follows.

Thus given $(L,\pi) \in \mathscr{L}(TQ)$ and $\alpha \in \Omega^1 Q$, $\alpha$ not closed, then $(L,\pi)$ and $(L+\tilde{\alpha}, \pi - \pi_\alpha)$ are $\mathscr{L}$-equivalent, but not $\mathscr{G}$-equivalent; furthermore, if $\alpha$ is closed, but not zero, then $(L,\pi)$ and $(L+\tilde{\alpha}, \pi - \pi_\alpha)$ are $\mathscr{G}$-equivalent, but not $\mathscr{H}$-equivalent. Summing up these facts, we have the following result.

*Proposition 7*: For any $(L,\pi) \in \mathscr{L}(TQ)$, we have the following sequence of strong inclusions:

$$(L,\pi)_{\mathscr{H}} \subset (L,\pi)_{\mathscr{G}} \subset (L,\pi)_{\mathscr{L}}.$$

In other words, $(L,\pi)_{\mathscr{L}} \not\subset (L,\pi)_{\mathscr{G}} \not\subset (L,\pi)_{\mathscr{H}}$. Now take $(L,\pi)$, $(L',\pi') \in \mathscr{L}(TQ)$. Suppose that $(L,\pi)$ and $(L',\pi')$ are $\mathscr{L}$-equivalent, but not $\mathscr{H}$-equivalent. Then $\lambda(L,\pi) = \lambda(L',\pi')$, and, by Proposition 3, $\delta((\theta \circ \chi)(L,\pi)) = \delta((\theta \circ \chi)(L',\pi'))$. On the other hand, Proposition 4 tells us that $(\theta \circ \chi)(L,\pi) \neq (\theta \circ \chi)(L',\pi')$. Thus we have the following result.

*Proposition 8*: $\delta : \mathfrak{X}_{hr}(T^*Q) \to \mathfrak{X}(TQ)$ is not injective.

A transformation $\phi : \mathscr{L}(TQ) \to \mathscr{L}(TQ)$ is called $\mathscr{G}$-transformation if for any $(L,\pi) \in \mathscr{L}(TQ)$, $\phi(L,\pi)$ and $(L,\pi)$ are $\mathscr{G}$-equivalent. By Proposition 6, any closed 1-form $\alpha \in \Omega^1_c Q$ induces a $\mathscr{G}$-transformation $\phi_\alpha : \mathscr{L}(TQ) \to \mathscr{L}(TQ)$, $\phi_\alpha(L,\pi) = (L+\tilde{\alpha}, \pi)$. Furthermore, any $c \in R$ induces a $\mathscr{G}$-transformation $\phi_c$ defined by $\phi_c(L,\pi) = (L+c, \pi)$. Observe that all these transformations leave $\mathscr{F}_{hr}(TQ) \subset \mathscr{L}(TQ)$ invariant. Now any $\mathscr{G}$-transformation of conservative mechanics (i.e., $\pi = 0$) comes from an element of $R \times \Omega^1_c Q$ (see Ref. 4, p. 216). Because of this, $R \times \Omega^1_c Q$ is called the $\mathscr{G}$-group of conservative Lagrangian mechanics. We remark that the $\mathscr{G}$-group of nonconservative Lagrangian mechanics is bigger than $R \times \Omega^1_c Q$.[5]

## VII. QUANTIZATION

We finally discuss our results within the realm of geometric prequantization (see for example Refs. 6 and 7). Roughly speaking, prequantization gives a way of converting classical systems into operators. Results of this type were also developed by Koopman.[8]

Consider the linear space $C_0^\infty(T^*Q)$ of real-valued compactly supported smooth functions on $T^*Q$. Let $\Omega$ denote the volume $\Omega = \omega_0^n$ on $T^*Q$; here $n$ is the dimension of $Q$. For a pre-Hilbert structure on $C_0^\infty(T^*Q)$ define $\langle f | g \rangle = \int_\Omega fg$, $f,g \in C_0^\infty(T^*Q)$. The corresponding Hilbert space is denoted by $L^2(T^*Q,\Omega)$. Now any vector field $X \in \mathfrak{X}(T^*Q)$ induces a differential operator $L_X : C_0^\infty(T^*Q) \to C_0^\infty(T^*Q)$; here $L_X$ is the Lie derivative. We shall also write $L_X = X^{\mathrm{op}}$. If $X$ is complete and $L_X \omega_0 = 0$, then $X^{\mathrm{op}}$ is essentially skew adjoint on the dense invariant domain $C_0^\infty(T^*Q) \subset L^2(T^*Q,\Omega)$ (see for example Refs. 2 and 9). Any canonical transformation $\delta : T^*Q \to T^*Q$ induces a unitary operator $U_\delta : L^2(T^*Q,\Omega) \to L^2(T^*Q,\Omega)$ by $U_\delta f = f \circ \delta$, i.e., $U_\delta f = \delta^* f$.

*Proposition 9*: Suppose that $X, Y \in \mathfrak{X}(T^*Q)$ are canonically equivalent. Then $X^{\mathrm{op}}$ and $Y^{\mathrm{op}}$ are unitarily equivalent.

*Proof*: There is a canonical transformation $\delta : T^*Q \to T^*Q$ such that $\delta_* X = Y$. Then $(L_Y \circ U_\delta) f = L_{\delta^* X} \delta^* f = \delta^* L_X f = (U_\delta \circ L_X) f$ for any $f \in C_0^\infty(T^*Q)$. This proves the assertion.

Hence, in view of Proposition 5, we conclude that for $\mathscr{G}$-equivalent $(L,\pi)$ and $(L',\pi')$, $X^{\mathrm{op}}_{\chi(L,\pi)}$ and $X^{\mathrm{op}}_{\chi(L',\pi')}$ are unitarily equivalent. In particular, if $\phi : \mathscr{L}(TQ) \to \mathscr{L}(TQ)$ is a $\mathscr{G}$-transformation, then $X^{\mathrm{op}}_{(\chi \circ \phi)(L,\pi)}$ and $X^{\mathrm{op}}_{\chi(L,\pi)}$ are unitarily equivalent.

## VIII. CONCLUSION

We have generalized some results in the theory of conservative mechanical systems to the nonconservative case. Moreover, the relationship to (geometric) quantization was analyzed. We used mechanical systems which are autonomous and holonomic, but of a general type in terms of external forces. A general theorem which determines all $\mathscr{G}$-transformations of nonconservative Lagrangian mechanics is currently under investigation.

Finally, it should be noticed that the above treatment applies equally well to the description of time-dependent systems. We observe that the damped harmonic oscillator $(L,\pi) = (\frac{1}{2}m\dot{q}^2 - U(q), \gamma \dot{q}\, dq)$ is $\mathscr{L}$-equivalent to the conservative time-dependent system $\tilde{L} = \exp((\gamma/m)t)(\frac{1}{2}m\dot{q}^2 - U(q))$, $t \in R$. However, both systems are neither $\mathscr{H}$- nor $\mathscr{G}$-equivalent. Several authors (see Ref. 10 and references therein) used the nonautonomous system $\tilde{L}$ in order to quantize the dissipative system $(L,\pi)$. Thus, the consideration of nonautonomous systems also provides an interesting field of potential application of our formalism.


## ACKNOWLEDGMENT

Research for this paper was supported by CNPq.



[1] M. Henneaux and L. C. Shepley, "Lagrangians for spherically symmetric potentials," J. Math. Phys. **23**, 2101 (1982).

[2] H. D. Doebner and J.-E. Werth, "Global properties of systems quantized via bundles," J. Math. Phys. **20**, 1011 (1979).

[3] C. Godbillon, *Géométrie Différentielle et Mécanique Analytique* (Hermann, Paris, 1969).

[4] R. Abraham and J. E. Marsden, *Foundations of Mechanics* (Benjamin, New York, 1978).

[5] J.-E. Werth (in preparation).

[6] B. Kostant, "Quantization and unitary representations," *Lecture Notes in Mathematics, Vol. 170* (Springer, Berlin, 1970), p. 87.

[7] J.-E. Werth, "Integrability for representations appearing in geometric pre-quantization," Ann. Inst. H. Poincaré **XXXVI**, 189 (1982).

[8] B. Koopman, "Hamiltonian systems and transformations in Hilbert space," Proc. Nat. Acad. Sci. (USA) **17**, 315 (1931).

[9] E. Nelson, "Analytic vectors," Ann. Math. **70**, 572 1959.

[10] H. Dekker, "Classical and quantum mechanics of the damped harmonic oscillator," Phys. Rep. **80**, 1 (1981).

# On a pointlike relativistic massive and spinning particle

Andreas Bette
*University of Stockholm, Institute of Theoretical Physics, Vanadisvägen 9, S-113 36 Stockholm, Sweden*

(Received 18 August 1983; accepted for publication 15 March 1984)

A pointlike massive and spinning relativistic particle is described as a confined system of two massless directly interacting spinning constituents. The approach is Hamiltonian. The employed phase space is, thus, a symplectic vector space equipped with global canonical and Poincaré-covariant twistor coordinates. The Poincaré-invariant generator of the phase space motion does not represent the energy of the total system. Consequently, the evolution parameter cannot be identified with the time. The generating function, however, makes the position four-vector and the proper time of the composite massive and spinning system into dynamical variables, i.e., functions of the evolution parameter. The phase flow may thus be interpreted as a simple particle dynamics in Minkowski space. In analogy with the definition of Bakamjian and Thomas for the center of energy of a relativistic massive and spinning particle, a definition of the center of energy of a massless particle with nonvanishing helicity is presented.

PACS numbers: 03.30. + p, 03.20. + i, 12.35. − i

## I. INTRODUCTION

It has been shown by Souriau[1] and Penrose[2–4] that, at least from the formal point of view, there exists a classical limit of the concept "massless spinning particle." In this limit helicity of the massless particle is a continuous degree of freedom. Upon an appropriate quantization *á la* Dirac, the discreteness of the helicity is recovered in a natural way.[3] Nevertheless, due to the correspondence principle, the obtained relativistic classical limit has an interest of its own. It turns out (Tod[5]) that the phase space of such a classical massless object may serve as a building block for the construction of the irreducible relativistic phase space of a massive spinning particle (Souriau[1]). The construction (Tod[5]) is performed using the first class constraint technique introduced by Dirac. Our conclusion is that the massive spinning particle obtained by the reduction procedure may be regarded as a bound (confined) system of two directly interacting massless spinning constituents. In this paper we present the possible dynamics of such a system. The famous no-interaction[6,7] theorem does not apply in our approach because the motion-generating (Poincaré-invariant) function is not the generator of infinitesimal time translations, i.e., is not the total energy (which is not a Poincaré invariant) of the considered system of directly interacting constituents. In addition the covariant and at the same time canonical variables are not physical. The following notational conventions will be used in this paper: Latin lower case letters as super- or subscripts will be used to denote the rank of a four-tensor (the abstract index notation, Penrose[8]). The same letters will also be used to denote concrete indices. Capital Latin indices as super- or subscripts will be used to denote the rank of a spinor (in both abstract and concrete form). Primed (sub-) superscripts will denote the complex conjugated (dotted) spinors. An overlined spinor symbol denotes a spinor obtained by complex conjugation. The simple rules for the spinor algebra are described in, e.g., Ref. 9. Here $e^{ijkn}$ and $g^{ij}$ denote the fully antisymmetric alternating flat pseudotensor and the flat metric tensor (signature − 2), respectively; $\epsilon^{AB}$ and $\bar{\epsilon}^{A'B'}$ denote the antisymmetric "metric" spinor-tensors in the spinor space and in its complex conjugate, respectively; $\sigma_i^{AA'}$ and $(\sigma^i_{AA'})$ denote fixed connecting quantities defining a Poincaré-invariant homomorphism from the spinor-space onto the four-vector space (for details see Ref. 9). Square brackets around indices will denote antisymmetrization, parentheses around indices will denote symmetrization, and the summation convention will be assumed throughout. Curly brackets will denote Poisson brackets. We say that two functions commute if their Poisson bracket vanishes. We choose units such that the constant velocity of light $c = 1$ and Planck's constant $h/2\pi = 1$.

## II. THE MASSLESS CLASSICAL PARTICLE AND ITS PHASE SPACE

A massless particle is characterized by its momentum four-vector $P^i$, its angular momentum four-tensor $M^{ik} = -M^{ki}$, and its helicity $s$. Under the action of the Poincaré group the components of the pair $(P^i, M^{ik})$ transform covariantly, i.e., under the action of the Lorentz group, $P^i$ constitutes a four-vector, and $M^{ik}$ constitutes an antisymmetric four-tensor of second rank, whereas, an arbitrary change of the reference event (the origin) in Minkowski space by means of a four-translation $T^i$ leaves $P^i$ invariant but transforms $M^{ik}$ according to the formula

$$M'^{ik} = M^{ik} - T^{[i}P^{k]}. \tag{2.1}$$

Regardless of whether a particle possesses mass or not, it will be called pointlike if we have a splitting

$$M^{ik} = 2X^{[i}P^{k]} + S^{ik}, \tag{2.2}$$

with $X^i$ representing its position four-vector and with $S^{ik} = -S^{ki}$ representing its translationally invariant intrinsic spin four-tensor such that

$$S^{ik}P_k = 0. \tag{2.3}$$

A system for which such decomposition is not possible will be called extended. In this note mainly pointlike, albeit possibly composite systems, will be considered.

The masslessness of a system is expressed by the fact that

$$m^2 := P^iP_i = 0. \tag{2.4}$$

 0022-2488/84/082456-05$02.50  

Equations (2.2)–(2.4) imply that the Pauli–Lubanski four-vector

$$S_i := -\tfrac{1}{2}e_{ijkm}M^{jk}P^m = -\tfrac{1}{2}e_{ijkm}S^{jk}P^m \tag{2.5}$$

is proportional to $P^i$ (Penrose[3]). The Poincaré-invariant, i.e., under the action of the Poincaré group, scalar proportionality constant $s$ is called the helicity of the massless system. Equations (2.2)–(2.4) also imply that if $s \neq 0$ then $X^i$ may be taken as any point on a null hyperplane defined by $X^iP_i = d$, where $d$ is a translationally dependent Lorentz scalar. Note that $s$ may be found once we are given $S^{ik}$ which fulfilling (2.3), and $P^i$ fulfilling (2.4) but it does not work in the opposite direction, i.e., knowing $s$ and $P^i$ fulfill (2.4) is not sufficient for the construction of $S^{ik}$ fulfulling (2.3). This is to be compared with the massive case where $S^{ik}$, which fulfills (2.3), and the Pauli–Lubanski four-vector $S^i$ are equivalent objects, i.e., carry the same information.

It has been demonstrated by Penrose[3] that the quantities $P^i$, $M^{ik}$, $s$, and $d$ describing a massless system may be consistently expressed as functions of Poincaré-covariant (see below) spinorial objects $\pi_A$, and $\omega^A$:

$$P^i := \sigma^i_{AA'}\pi^{A'}\bar{\pi}^A, \tag{2.6}$$

$$M^{kn} := i\sigma^k_{AA'}\sigma^n_{BB'}(\omega^{(A}\bar{\pi}^{B)}\bar{\epsilon}^{A'B'} - \bar{\omega}^{(A'}\pi^{B')}\epsilon^{AB}), \tag{2.7}$$

$$s := 1/2(\omega^A\bar{\pi}_A + \pi_{A'}\bar{\omega}^{A'}), \tag{2.8}$$

$$d := i/2(\omega^A\bar{\pi}_A - \pi_{A'}\bar{\omega}^{A'}). \tag{2.9}$$

Under the action of the (universal covering group of the identity connected component of the) Poincaré group the components of the pair $(\omega^A, \pi_{A'})$ transform covariantly, i.e., under the action of the (universal covering group of the identity connected component of the) Lorentz group $\omega^A$ and $\pi_{A'}$ constitute two different (Weyl) spinors, whereas, an arbitrary change of the reference event (the origin) in Minkowski space by means of a four-translation $T^i$ leaves $\pi_{A'}$ invariant but transforms $\omega^A$ according to the formula

$$\omega'^A = \omega^A - iT^i\sigma_i^{AA'}\pi_{A'}, \tag{2.10}$$

$$\bar{\omega}'^{A'} = \bar{\omega}^{A'} + iT^i\sigma_i^{AA'}\bar{\pi}_A. \tag{2.11}$$

Using the definitions in (2.6)–(2.9) it is now a straightforward task to check that the Poincaré covariance of the pair $(\omega^A, \pi_{A'})$ implies the Poincaré covariance of the pair $(P^i, M^{ik})$. The pair $(\omega^A, \pi_{A'})$ regarded as an abstract vector in a four-dimensional Poincaré-invariant complex vector space is called a twistor $Z^\alpha$. Its twistor conjugate $\bar{Z}_\alpha$ is represented by the pair $(\bar{\pi}_A, \bar{\omega}^{A'})$. Let us now treat $(\omega^A, \pi_{A'})$ as Poincaré-covariant coordinates of a point in a four-dimensional complex vector space $T$ (twistor space, eight real dimensions). Assume that $T$ is a symplectic vector space equipped with global Poincaré-covariant canonically conjugate coordinates $-i\omega^A$ and $\bar{\pi}_B$ (or, equivalently, $-i\pi_{A'}$ and $\bar{\omega}^{B'}$) fulfilling the following Poincaré-covariant canonical Poisson algebra:

$$\{\omega^A, \bar{\pi}_B\} = i\delta^A_B, \quad \{\omega^A, \pi_{B'}\} = 0, \tag{2.12}$$

$$\{\pi_{A'}, \pi_{B'}\} = 0, \quad \{\omega^A, \omega^B\} = 0, \tag{2.13}$$

$$\{\omega^A, \bar{\omega}^B\} = 0, \quad \{\pi_{A'}, \bar{\pi}_B\} = 0. \tag{2.14}$$

Using the definitions in (2.6) and (2.7) it is a straightforward but tedious task to check that the canonical Poincaré-covariant Poisson bracket relations in (2.12)–(2.14) imply the Poincaré-covariant Poisson bracket realization of the Poincaré algebra (see, e.g. Hughston,[10] p.33):

$$\{P^i, P^j\} = 0, \tag{2.15}$$

$$\{M^{ij}, P^k\} = 2g^{k[i}P^{j]}, \tag{2.16}$$

$$\{M^{ij}, M^{kn}\} = 2(g^{n[j}M^{i]k} + g^{k[i}M^{j]n}). \tag{2.17}$$

Note that the canonical action of the (universal covering group of the identity connected component of the) Poincaré group on $T$ is not transitive. This is so because points on a surface $s = \text{const}$ cannot, by means of a Poincaré transformation, be mapped onto points on another surface $s = \text{another const}$. Any surface $s = \text{const}$ constitutes a helicity shell in $T$.

Let us mention in passing that the quantization, *á la* Dirac, of (2.12)–(2.14) turns the $s$-function in (2.8) (after the normal ordering of terms) into an helicity operator with correct discrete eigenvalues (Penrose[4]).

In analogy to the massive spinning case we now define an inertial observer dependent center of energy of a massless particle:

$$Q^i := (1/\epsilon)\tilde{M}^{ik}t_k + (t/\epsilon)P^i, \tag{2.18}$$

where $t$ is the proper time of the observer, $t^i$ is its four-velocity ($t^it_i = 1$), and where

$$\epsilon := P^it_i, \tag{2.19}$$

$$\tilde{M}^{ik} := i\sigma^k_{AA'}\sigma^n_{BB'}(\rho^{(A}\bar{\pi}^{B)}\bar{\epsilon}^{A'B'} - \bar{\rho}^{(A'}\pi^{B')}\epsilon^{AB}), \tag{2.20}$$

with $\rho^A$ being a $t^i$-dependent spinor defined by

$$\rho^A := \omega^A - (s/\epsilon)\pi_{A'}\sigma_i^{AA'}t^i. \tag{2.21}$$

Here $Q^i$ in (2.18) is constructed in such a way that the canonical Poincaré-covariant Poisson brackets in (2.12)–(2.14) imply the following (not Poincaré covariant but physical) Poisson bracket relations:

$$\{Q^i, Q^j\} = 0, \tag{2.22}$$

$$\{Q^i, P_j\} = \delta^i_j - (1/P^mt_m)P^it_j. \tag{2.23}$$

Note that $Q^i$ is not a position four-vector. A novel, and since long searched for, feature of the introduced center of energy of a massless and spinning particle is the commutation relation in (2.22). The definition in (2.18) (compare with certain twistorial considerations presented by Penrose in Ref. 11) is an analog of the relativistic center of energy of a massive and spinning particle. The latter has been introduced a long time ago by Bakamjian and Thomas.[12] It should be noted that in terms of twistor variables it is not possible to define a unique position four-vector $X^i$ for the massless spinning particle. We may choose $X^i$ as any point on the hyperplane $X^iP_i = d$. With any such $X^i$ at hand we may construct $S^{ik}$ using (2.2). If $s = 0$ then $S^{ik} = 0$ and there is a unique $X^i$ available. In this case ($s = 0$) the center of energy in (2.18) becomes a covariant position four-vector lying on the hyperplane $X^iP_i = d$.

## III. THE POINTLIKE MASSIVE CLASSICAL RELATIVISTIC PARTICLE WITH SPIN AND ITS PHASE SPACE

Let two massless particles be described by two twistors $Z^\alpha = (\omega^A, \pi_{A'})$ and $W^\alpha = (\lambda^A, \eta_{A'})$, respectively. Under the

action of the (universal covering group of the identity connected component of the) Poincaré group the components of each pair are supposed to transform covariantly, i.e., in a manner described in the previous section. The two pairs $(\omega^A, \pi_{A'})$, $(\lambda^A, \eta_{A'})$ represent global Poincaré-covariant coordinates of a point in an eight-dimensional complex symplectic vector space $T \times T$ which defines a phase space for the two massless particles. Coordinates of $(\omega^A, \pi_{A'})$ fulfill its own canonical Poincaré-covariant Poisson bracket algebra introduced in (2.12)–(2.14) and so do coordinates of $(\lambda^A, \eta_{A'})$.

Besides, coordinates of $(\omega^A, \pi_{A'})$ commute with those of $(\lambda^A, \eta_{A'})$. Using the canonical and covariant twistor coordinates, six real-valued Poincaré-invariant (scalar) functions may be constructed (Tod[13]) on $T \times T$ according to the following recipe:

$$e := Z^\alpha \overline{Z}_\alpha + W^\alpha \overline{W}_\alpha = (\omega^A \bar{\pi}_A + \bar{\omega}^{A'} \pi_{A'}) + (\lambda^A \bar{\eta}_A + \overline{\lambda}^{A'} \eta_{A'}); \tag{3.1}$$

$$k := Z^\alpha \overline{Z}_\alpha - W^\alpha \overline{W}_\alpha = (\omega^A \bar{\pi}_A + \bar{\omega}^{A'} \pi_{A'}) - (\lambda^A \bar{\eta}_A + \overline{\lambda}^{A'} \eta_{A'}); \tag{3.2}$$

$$\mathrm{Re} f \quad \text{and} \quad \mathrm{Im} f, \tag{3.3}$$

where

$$f := I_{\alpha\beta} Z^\beta W^\alpha = \pi^{A'} \eta_{A'} \neq 0 \tag{3.4}$$

($I_{\alpha\beta}$ is the so-called infinity twistor); and

$$\mathrm{Re}\, a \text{ and } \mathrm{Im}\, a, \tag{3.5}$$

where

$$a := Z^\alpha \overline{W}_\alpha = (\omega^A \bar{\eta}_A + \overline{\lambda}^{A'} \pi_{A'}). \tag{3.6}$$

Assume that the two massless particles form a massive and, in general, spinning pointlike system $R$.

From now on we thus treat $T \times T$ as (a square root of) a (reducible) phase space of a massive and in general spinning fundamental pointlike physical system. In accordance with this identification we put linear momentum four-vector of $R$ to be

$$p^i := \sigma^i_{AA'} (\pi^{A'} \bar{\pi}^A + \eta^{A'} \bar{\eta}^A) = P^i_1 + P^i_2, \tag{3.7}$$

which implies that the mass squared of $R$ is given by

$$m^2 := P^i P_i = 2 f \bar{f} \neq 0, \quad \text{so that } |f| = (1/\sqrt{2}) m. \tag{3.8}$$

We put position four-vector of $R$ (our original identification) to be

$$x^i := (i/m^2) \sigma^i_{AA'} [\bar{f} (\omega^A \eta^{A'} - \lambda^A \pi^{A'}) - f (\bar{\omega}^{A'} \bar{\eta}^A - \overline{\lambda}^{A'} \bar{\pi}^A)]. \tag{3.9}$$

It is easy to see that $x^i$ is a true position four-vector, i.e., it is real and covariant under the action of the Poincaré group. We put the Pauli–Lubański four-vector to be (Perjes[14])

$$W^i := \sigma^i_{AA'} [k/2 (\pi^{A'} \bar{\pi}^A - \eta^{A'} \bar{\eta}^A) + a \eta^{A'} \bar{\pi}^A + \bar{a} \pi^{A'} \bar{\eta}^A]. \tag{3.10}$$

Equations (3.7) and (3.10) imply that

$$W^i p_i = 0, \tag{3.11}$$

and also that the spin squared is given by (Perjes[14] and Tod[13])

$$w^2 = -(1/m^2) W^i W_i = \tfrac{1}{4} k^2 + a\bar{a}. \tag{3.12}$$

The canonical Poincaré-invariant twistor Poisson algebra on $T \times T$ implies the following physically meaningful Poincaré-covariant commutation relation:

$$\{x^i, p^j\} = g^{ij}. \tag{3.13}$$

Thus $x^i$ and $p^i$ are conjugate variable but, as we shall see in a moment, they are canonically conjugate if and only if the system is spinless.

Defining the total angular four-momentum

$$L^{ik} := 2 x^{[i} p^{k]} + (1/m^2) e^{iknm} W_n p_m, \tag{3.14}$$

we obtain an important identity

$$L^{kn} = i \sigma^k_{AA'} \sigma^n_{BB'} [(\omega^{(A} \bar{\pi}^{B)} \bar{\epsilon}^{A'B'} - \bar{\omega}^{(A'} \pi^{B')} \epsilon^{AB}) + (\lambda^{(A} \bar{\eta}^{B)} \bar{\epsilon}^{A'B'} - \overline{\lambda}^{(A'} \eta^{B')} \epsilon^{AB})] = M_1^{kn} + M_2^{kn}. \tag{3.15}$$

Using the canonical twistor algebra and the relations in (3.7) and (3.15) it is straightforward to prove that $p^k$ and $L^{ij}$ fulfill Poincaré Poisson bracket algebra (Hughston[10]) such as in (2.15)–(2.17). In addition we also obtain

$$\{x^i, x^j\} = -(1/m^4) e^{ijkl} W_k p_l, \tag{3.16}$$

$$\{L^{ij}, x^k\} = 2 g^{k[i} x^{j]}, \tag{3.17}$$

$$\{W^i, x^j\} = -(2/m^2) W^{[i} p^{j]}. \tag{3.18}$$

The Poisson brackets in (3.16) and (3.18) are quite remarkable and on the quantum level of this approach imply that the space of four positions of a spinning pointlike particle should not be associated with any Poincaré-invariant configuration space. A Poincaré-invariant eight-dimensional configuration space is provided by $\pi_{A'}$, and $\eta_{A'}$, which are not directly physical. Here $x^i$, $p^i$ and $W^i$ represent 11 independent variables. The Poincaré-invariant scalars, e.g., arctan ($\mathrm{Im} f/\mathrm{Re} f$), arctan ($\mathrm{Im}\, a/\mathrm{Re}\, a$), $e$ and $k$ represent four additional (not mutually commuting) independent variables. Their physical meaning is not yet clear. Perhaps at some later stage they can be identified with operators of internal quantum numbers. The $T \times T$ is 16 dimensional and up to now we have only identified $11 + 4$ variables associated with the massive spinning pointlike system. The sixteenth variable is provided by the angle of rotation of a spacelike two-plane orthogonal to $p^i$ and $W^i$. The two orthogonal unit four-vectors spanning this polarization two-plane may be chosen to be

$$E^j := (i/m\sqrt{a\bar{a}}) \sigma^j_{AA'} (a \eta^{A'} \bar{\pi}^A - \bar{a} \pi^{A'} \bar{\eta}^A), \quad \text{if } a \neq 0, \tag{3.19}$$

$$E'^j := (i/m) \sigma^j_{AA'} (\eta^{A'} \bar{\pi}^A - \pi^{A'} \bar{\eta}^A), \quad \text{if } a = 0, \tag{3.20}$$

$$F^j := (1/wm\sqrt{a\bar{a}}) \sigma^j_{AA'} [\tfrac{1}{2} k (a \eta^{A'} \bar{\pi}^A + \bar{a} \pi^{A'} \bar{\eta}^A) - a\bar{a} (\pi^{A'} \bar{\pi}^A - \eta^{A'} \bar{\eta}^{A'})] \quad \text{if } a \neq 0, \tag{3.21}$$

$$F'^j := (i/m) \sigma^j_{AA'} (\eta^{A'} \bar{\pi}^A + \pi^{A'} \bar{\eta}^A), \quad \text{if } a = 0. \tag{3.22}$$

Summarizing we conclude that the sixteen variables characterizing two massless twistor particles also describe a massive and spinning pointlike system. Both descriptions are reducible, i.e., all Poincaré invariants are dynamical variables. Whether they are constants of the canonical flow motion depends on the generating function (see below). Note that the possibility of defining a unique position four-vector of the massive particle in terms of twistors variables [see

(3.9)] and thereby reassuring its pointlike structure exists if and only if it is thought of as composed of two (not less or more) massless components. Any massive and spinning system consisting of more then two massless twistor particles is necessarily extended.

Even if very much of the above material is known to people working with twistors the suggested identification of the physical massive variables is original and applies to the noninteracting as well as to the interacting case. Note that when $k = a = 0$, i.e., when the system is spinless, the canonical twistor Poisson brackets imply

$$\{x^i, x^j\} = 0, \tag{3.23}$$

$$\{x^i, p^j\} = g^{ij}. \tag{3.24}$$

In other words $x^i$ and $p^j$ show up as canonically conjugate variables. Geometrically we have that the eight-dimensional Poincaré-invariant (spanned by $x^i$ and $p^j$) subspace in the 13-dimensional submanifold of $T \times T$ defined by $k = a = 0$ is again a symplectic vector space. On the quantum level this implies that the space of four-positions of a spinless pointlike system defines a Poincaré-invariant configuration space.

## IV. DYNAMICS AND KINEMATICS OF THE MASSIVE POINTLIKE SPINNING SYSTEM

In the general spinning case consider a Poincaré-invariant (scalar) function:

$$H(Z^\alpha, W^\beta, \bar{Z}_\alpha, \bar{W}_\beta) = p^i p_i - W^i W_i (1/m^2) = 2f\bar{f} + \tfrac{1}{4}k^2 + a\bar{a}. \tag{4.1}$$

(All variables in our approach are made dimensionless by the use of $h$, $c$, and some numerical standard value for the mass variable $m_0$. We also use units such that $h/2\pi = c = m_0 = 1$.) The $H$ function, which is a (flat) generalization of the relativistic super-Hamiltonian described by Misner *et al.*[5] induces a canonical flow in $T \times T$. For the massive composite pointlike particle we obtain ($u$ is a parameter along the flow)

$$\frac{de}{du} = \{e, H\} = 0, \tag{4.2}$$

$$\frac{df}{du} = \{f, H\} = 0, \tag{4.3}$$

$$\frac{da}{du} = \{a, H\} = 0, \tag{4.4}$$

$$\frac{dk}{du} = \{k, H\} = 0, \tag{4.5}$$

$$\frac{dp^i}{du} = \{p^i, H\} = 0, \tag{4.6}$$

$$\frac{dx^i}{du} = \{x^i, H\} = 2p^i, \tag{4.7}$$

$$\frac{dW^i}{du} = \{W^i, H\} = 0, \tag{4.8}$$

$$\frac{dE^i}{du} = \{E^i, H\} = 2w^2 F^i, \tag{4.9}$$

$$\frac{dF^i}{du} = \{F^i, H\} = -2w^2 E^i. \tag{4.10}$$

We conclude that the super-Hamiltonian in (4.1) describes the massive free-particle case where all the invariants (Poincaré scalars) are constants of the motion and the polarization plane rotates with a constant angular velocity proportional to $w^2$.

The Lagrangian corresponding to the super-Hamiltonian in (4.1) reads

$$L = i(Z'^\alpha \bar{Z}_\alpha + W'^\alpha \bar{W}_\alpha) - i(Z^\alpha \bar{Z}'_\alpha + W^\alpha \bar{W}'_\alpha) - 2f\bar{f} - \tfrac{1}{4}k^2 - a\bar{a}. \tag{4.11}$$

The prime denotes differentiation with respect to the evolution parameter $u$. From (4.7) we obtain that

$$d\tau = 2m\, du \Rightarrow u = (1/2m)\tau + (1/2m)\tau_0, \quad \tau_0 = \text{const}, \tag{4.12}$$

where $\tau$ denotes the proper time of the massive pointlike spinning system.

Now $H$, in (4.1) [or $L$ in (4.11)] define also dynamics of the two massless twistor constituents. For the four-momentum spinors we obtain the following solutions:

$$\bar{\pi}^A = \bar{\pi}_1^A e^{iu(1/2m)\tau} + \bar{\pi}_2^A e^{-iu(1/2m)\tau}, \tag{4.13}$$

$$\bar{\eta}^A = [(w/\bar{a}) - (k/2\bar{a})]\bar{\pi}_1^A e^{iu(1/2m)\tau} - [(w/\bar{a}) + (k/2\bar{a})]\bar{\pi}_2^A e^{-iu(1/2m)\tau}, \tag{4.14}$$

where $\bar{\pi}_1^A$ and $\bar{\pi}_2^A$ are constant spinors depending on the initial values of the momentum four-vectors of the massless constituents and on the values of the constant Poincaré invariants (scalars).

The solutions in (4.13) and (4.14) represent two massless particles rotating around each other with the velocity of light along the world line of the pointlike massive and spinning particle they also happen to define.

## V. CONCLUSIONS AND REMARKS

We have presented a twistorial model of a pointlike massive and spinning particle. Such a particle is described by means of sixteen (phase space) variables and the super-Hamiltonian above (or the super-Lagrangian). We have shown that according to this model a massive pointlike particle has a dynamical origin and may be regarded as a system composed of two massless directly interacting twistor particles with in general nonvanishing helicity.

The presented model differs from the one presented by Souriau,[1] Arens,[16] and Tod[5] by being reducible which means that the action of the Poincaré group on the phase space is not transitive and also by the fact that in order to have a full description we need to specify the super-Hamiltonian (or the super-Lagrangian). No constraints are needed. Instead the super-Hamiltonian is arranged in such a way that certain Poincaré-invariant (scalar) quantities remain constant along the canonical flow if it is so desired.

A massive spinning system constructed with the aid of three massless twistor particles in a way similar to the one described above is necessarily extended. It defines three distinct world lines and looks from "outside" as a confined system of three massive and, in general, spinning particles.

## ACKNOWLEDGMENTS

Discussions with Professor S. Hojman are gratefully acknowledged.

We also thank the Royal Swedish Academy of Sciences for the support from the Lindahls Fund Award.

[1]J. M. Souriau, *Structure des systeme dynamique* (Dunod Universite, Paris, 1970).
[2]R. Penrose, Int. J. Theor. Phys. **1**, 61 (1968).
[3]R. Penrose and M. A. H. MacCallum, Phys. Rep. **C6**(4), 241 (1972).
[4]R. Penrose, in *Quantum Gravity, an Oxford Symposium* edited by C. J. Isham, R. Penrose, and D. W. Sciama (Clarendon, Oxford, 1975).
[5]P. Tod, Rep. Math. Phys. **11**(3), 339 (1977).
[6]D. G. Currie, T. F. Jordan, and E. C. G. Sudarshan, Rev. Mod. Phys. **35**, 350 (1963).
[7]H. Leutwyler, Nuovo Cimento **37**, 556 (1965).
[8]R. Penrose, in *Batelle Rencontres*, edited by C. M. de Witt and J. A. Wheeler (Benjamin, New York, 1968), pp. 135–149.
[9]F. A. E. Pirani, in *Lectures on General Relativity*, Brandeis Summer Institute in Theoretical Physics (Prentice–Hall, Englewood Cliffs, NJ, 1964), Vol. 1.
[10]L. P. Hughston, "Twistors and Particles," in *Lecture Notes in Physics Number 97* (Springer-Verlag, Berlin, 1979).
[11]R. Penrose and R. S. Ward, in *General Relativity and Gravitation, One Hundred Years after the Birth of Albert Einstein*, edited by Alan Held (Plenum, New York, 1980), Vol. 2.
[12]B. Bakamjian and L. H. Thomas, Phys. Rev. **92**, 1300 (1953).
[13]P. Tod, Doctoral Dissertation at Mathematical Institute, University of Oxford, U.K., 1975.
[14]Z. Perjes, Phys. Rev. D **11**, 2031 (1975).
[15]C. W. Misner, K. S. Thorne, and J. A. Wheeler, *Gravitation* (Freeman, San Francisco, 1973).
[16]R. Arens, Commun. Math. Phys. **21**, 139 (1971).

# Electromagnetic effects in an elastic circular cylindrical waveguide

A. E. Green
*Mathematical Institute, Oxford University, Oxford, England*

P. M. Naghdi
*Department of Mechanical Engineering, University of California, Berkeley, California 94720*

(Received 20 September 1983; accepted for publication 23 December 1983)

The problem of wave propagation in an elastic, thin circular cylindrical tube, including the possibility of propagation of rotating waves, is examined with the use of a nonlinear electromagnetic membrane theory. Some related special cases, which include wave propagation in a class of magnetic polarized rigid tubes and rigid plates, are also briefly discussed.

PACS numbers: 03.40.Kf, 41.10.Hv

## I. INTRODUCTION

The propagation of electromagnetic waves in deformable elastic materials has been investigated by Adkins and Rivlin,[1] who examined the propagation of electromagnetic waves in twisted circular cylindrical rods. In particular, they demonstrated (in Ref. 1) that for a given mode of transmission of waves the relevant secular equation yields a pair of waves propagated with different speeds which may be combined to give a wave that rotates during propagation. Electromechanical devices using such a rotating wave are of importance in microwave transmission. The rotation is usually produced by introducing ferrites into the waveguide and applying a static magnetic field (for references see Adkins and Rivlin[1]). In another paper, Adkins and Rivlin[2] have studied the propagation of electromagnetic waves along a circular cylindrical waveguide containing a holohedral isotropic material under the influence of static electric and magnetic fields. Additional references on the subject are cited in Adkins and Rivlin.[2]

The rotating waves in these ferrite-filled waveguides are nonreciprocal in the sense that if the direction of propagation is reversed, the direction of rotation of the waves is unaltered. For the twisted elastic cylinder, however, the wave is reciprocal. Adkins and Rivlin[1] also indicate how the problem of wave propagation in a twisted elastic tube could be solved but did not complete the solution. When the wall of the tube is thin compared with the radius of the tube (say the inner radius), the problem may be examined with the use of the electromagnetic membrane theory developed by Green and Naghdi[3] as part of a more general electromagnetic theory of deformable shell-like bodies. This is considered in Secs. II–IV and yields types of wave propagation similar in character to those found by Adkins and Rivlin[1] for circular cylinders, including the possibility of rotating waves.

Utilizing the basic theory of Sec. II, in Sec. V we consider wave propagation in a class of magnetic, polarized rigid tubes. For this class of material, however, it is found that waves are propagated without rotation. This is because thermodynamical considerations of a fairly general class of magnetic, polarized elastic material (Green and Naghdi[3]) deny the part of the constitutive relations used by Adkins and Rivlin[2] which led to the possibility of a rotating wave. The modification of the constitutive relations required to achieve a rotational wave is briefly discussed.

Extensive studies have been made of linear piezoelectric plate problems by Tiersten[4] and by Bugdayci and Bogy[5] and other authors. The linear piezoelectric theory for elastic plates was also considered as a special case of a general electromagnetic theory of plates and shells by Green and Naghdi.[3] It may be noted that wave propagation in a plate finitely deformed in its plane and acting as a waveguide may be studied as a special case of the membrane theory of Secs. II and III or for the particular problem of a circular cylinder considered in Sec. IV, as a limiting case when the radius of the cylinder tends to infinity. Similarly, wave propagation in a rigid plate in the presence of static magnetic fields may also be examined using the theory of Secs. II and III. Detailed solutions for these plate problems, however, are not considered here.

## II. NOTATION AND BASIC EQUATIONS

We consider here the basic equations relevant to the motion of a two-dimensional material surface in the presence of electromagnetic effects. Consider a material surface embedded in a Euclidean three-space $\mathscr{E}^3$ and let convected coordinates $\theta^\alpha(\alpha = 1,2)$ be assigned to each particle (or material point) of the surface. Further, let the surface occupied by the material surface in the deformed and the undeformed reference configurations be referred to as $\mathscr{s}$ and $\mathscr{S}$, respectively, and let $\mathbf{R}$ denote the position vector, relative to a fixed origin, of a typical point of $\mathscr{S}$. Then, $\mathbf{R} = \mathbf{R}(\theta^\alpha)$ specifies the place occupied by the material point in the undeformed configuration which we take to be the initial configuration. Likewise, the position vector of $\mathscr{s}$, relative to the same fixed origin, is given by $\mathbf{r} = \mathbf{r}(\theta^\alpha, t)$ and this specifies the place occupied by the material point $\theta^\alpha$ in the deformed configuration.

Let $\mathbf{a}_\alpha$ denote the base vectors along the $\theta^\alpha$-curves on $\mathscr{s}$ and let $\mathbf{a}_3$ be the unit normal to $\mathscr{s}$. Then,

$$\mathbf{a}_\alpha = \mathbf{r}_{,\alpha}, \quad \mathbf{a}_\alpha \cdot \mathbf{a}_\beta = a_{\alpha\beta}, \quad \mathbf{a}^\alpha \cdot \mathbf{a}_\beta = \delta^\alpha_\beta, \quad a^{\alpha\beta} = \mathbf{a}^\alpha \cdot \mathbf{a}^\beta,$$
$$a^{1/2}\mathbf{a}_3 = \mathbf{a}_1 \times \mathbf{a}_2, \quad a = \det a_{\alpha\beta}, \quad a^{1/2} = [\mathbf{a}_1 \mathbf{a}_2 \mathbf{a}_3] > 0, \quad (2.1)$$

where a comma stands for partial differentiation with respect to $\theta^\alpha$, $a_{\alpha\beta}$ and $a^{\alpha\beta}$ are the components of the first fundamental form of $\mathscr{s}$ and its conjugate and $\delta^\alpha_\beta$ is the Kronecker symbol in two-space. The velocity vector is defined by

$$\mathbf{v} = \dot{\mathbf{r}}, \quad (2.2)$$

where a superposed dot denotes time differentiation with respect to $t$ holding $\theta^\alpha$ fixed.

 0022-2488/84/082461-06$02.50  

In the above equations and throughout the paper, all Greek indices take the values 1,2, Latin indices take the values 1,2,3 and the usual summation convention over a repeated index (one subscript and one superscipt) is employed. Also, the raising and the lowering of indices of the components of surface tensors is accomplished with the use of $a_{\alpha\beta}$ and $a^{\alpha\beta}$. Whenever possible, in what follows, we use capital letters to represent the duals of quantities associated with $\mathscr{S}$ in the reference surface $\mathscr{S}$. For example, the base vectors and the components of the metric tensor of the reference surface $\mathscr{S}$ will be designated by $\mathbf{A}_i$ and by $A_{\alpha\beta}$, $A^{\alpha\beta}$. We note that results similar to those in (2.1) hold also for the surface $\mathscr{S}$ and record the following formulas:

$$\mathbf{R}=\mathbf{R}(\theta^\alpha),\quad \mathbf{A}_\alpha=\mathbf{R}_{,\alpha},$$
$$\mathbf{A}_\alpha\cdot\mathbf{A}_\beta=A_{\alpha\beta},\quad \mathbf{A}^\alpha\cdot\mathbf{A}_\beta=\delta^\alpha_\beta,\quad A^{\alpha\beta}=\mathbf{A}^\alpha\cdot\mathbf{A}^\beta, \tag{2.3}$$
$$A^{1/2}\mathbf{A}_3=\mathbf{A}_1\times\mathbf{A}_2,\quad A=\det A_{\alpha\beta},\quad A^{1/2}=[\mathbf{A}_1\mathbf{A}_2\mathbf{A}_3]>0.$$

In what follows, we take the reference configuration of the membrane to be the initial configuration so that $\mathbf{R}=\mathbf{R}(\theta^\alpha)=\mathbf{r}(\theta^\alpha,0)$.

In the remainder of this section, we summarize the principal results from the nonlinear membrane theory of elastic shells in the presence of electromagnetic effects, as derived by direct approach by Green and Naghdi.[3] (For an account of the membrane theory by direct approach without electromagnetic effects, see Sec. 14 of Naghdi's monograph[6].) Let $c$ be a closed curve on $\mathscr{S}$ and let

$$\boldsymbol{\nu}=\nu_\alpha\mathbf{a}^\alpha=\nu^\alpha\mathbf{a}_\alpha \tag{2.4}$$

be the outward unit normal to $c$ lying in the surface. Further, let the vector field $\mathbf{n}$, which depends on $\boldsymbol{\nu}$, represent the contact force (or the curve force vector) per unit length of $c$. Then, it can be shown that

$$\mathbf{n}=\mathbf{N}\boldsymbol{\nu}=\mathbf{N}^\alpha\nu_\alpha, \tag{2.5}$$

with the second-order tensor $\mathbf{N}$ defined by

$$\mathbf{N}=\mathbf{N}^\alpha\otimes\mathbf{a}_\alpha=N^{i\alpha}\mathbf{a}_i\otimes\mathbf{a}_\alpha,\quad \mathbf{N}^\alpha=N^{i\alpha}\mathbf{a}_i, \tag{2.6}$$

where the symbol $\otimes$ denotes tensor product. The local equation for conservation of mass can be expressed as

$$\rho a^{1/2}=\rho_R A^{1/2}\quad\text{or}\quad \dot\rho+\rho\,\mathrm{div}_s\,\mathbf{v}=0, \tag{2.7}$$

where $\rho$ and $\rho_R$ denote the mass densities of the surface $\mathscr{S}$ and the reference surface $\mathscr{S}$, respectively, and where

$$\mathrm{div}_s\,\mathbf{v}=\mathbf{v}_{,\alpha}\cdot\mathbf{a}^\alpha. \tag{2.8}$$

The local field equations for conservation of momentum and moment of momentum can be written as

$$\mathrm{div}_s\,\mathbf{N}+\rho(\mathbf{f}+\mathbf{f}_e)=\rho\dot{\mathbf{v}}, \tag{2.9a}$$
$$\mathbf{a}_\alpha\times\mathbf{N}^\alpha+\rho\mathbf{c}_e=\mathbf{0}, \tag{2.9b}$$

where $\mathbf{f}$ is the assigned field while $\mathbf{f}_e$ and $\mathbf{c}_e$ are the assigned force vector and assigned (axial) couple vector due to the electromagnetic fields, all per unit mass of the surface $\mathscr{S}$. Also, the divergence of a tensor field in (2.9) is defined by

$$a^{1/2}\,\mathrm{div}_s\,\mathbf{N}=(a^{1/2}\mathbf{N}^\alpha)_{,\alpha}, \tag{2.10}$$

and $(\ )_{,\alpha}$ stands for partial differentiation with respect to $\theta^\alpha$.

Electromagnetic effects in the membrane are represented by vector fields and scalar fields defined on $\mathscr{S}$ as follows: the electric field vectors,

$$\mathbf{e}^*_N=e^*_{Ni}\mathbf{a}^i\quad\text{or}\quad \mathbf{E}_N=\mathbf{F}^T\mathbf{e}^*_N=E_{Ni}\mathbf{A}^i; \tag{2.11a}$$

the electric displacement field vectors,

$$\bar{\mathbf{d}}_N=\bar{\mathbf{d}}^i_N\mathbf{a}_i\quad\text{or}\quad \bar{\mathbf{D}}_N=\Gamma\mathbf{F}^{-1}\bar{\mathbf{d}}_N=\bar{D}^{\,i}_N\mathbf{A}_i; \tag{2.11b}$$

the magnetic field vectors,

$$\mathbf{h}^*_N=h^*_{Ni}\mathbf{a}^i\quad\text{or}\quad \mathbf{H}_N=\mathbf{F}^T\mathbf{h}^*_N=H_{Ni}\mathbf{A}^i; \tag{2.11c}$$

the magnetic induction field vectors,

$$\mathbf{b}_N=b^{\,i}_N\mathbf{a}_i\quad\text{or}\quad \mathbf{B}_N=\Gamma\mathbf{F}^{-1}\mathbf{b}_N=B^{\,i}_N\mathbf{A}_i; \tag{2.11d}$$

the current density field vectors,

$$\mathbf{j}^*_N=j^{*i}_N\mathbf{a}_i\quad\text{or}\quad \mathbf{J}_N=\Gamma\mathbf{F}^{-1}\mathbf{j}^*_N=J^{\,i}_N\mathbf{A}_i; \tag{2.11e}$$

and the free charge field scalars,

$$e_N\quad\text{or}\quad E_N=\Gamma e_N; \tag{2.11f}$$

all for $N=0,1,\ldots,L$, where

$$\mathbf{F}=\mathbf{a}_i\otimes\mathbf{A}^i,\quad \Gamma=a^{1/2}/A^{1/2}. \tag{2.12}$$

Interpretations of the above fields in terms of electromagnetic vectors and scalars in the three-dimensional theory are included in Appendix B of the paper of Green and Naghdi.[3] The field equations governing the fields (2.11) are consequences of appropriate conservation laws for a moving shell-like body and are analogs of the three-dimensional conservation laws associated with the names of Faraday, Ampere, and Gauss. Moreover, the form of these field equations depends on the choice of particular interpretation. Here we adopt the trigonometric representation according to case (b) in Eq. (6.20) of Green and Naghdi.[3] The appropriate field equations are then given by

$$\begin{aligned}
&A^{-1/2}(B^{\,\alpha}_M A^{1/2})_{,\alpha}=-(\pi M/h)B^{\,3}_M-\hat{B}^{\,3}_M,\\
&A^{-1/2}(\bar{D}^{\,\alpha}_M A^{1/2})_{,\alpha}=(\pi M/h)\bar{D}^{\,3}_M+E_M,\\
&\dot{B}^{\,3}_M=-\epsilon^{\alpha\beta}E_{M\beta,\alpha},\\
&\dot{B}^{\,\alpha}_M=-\epsilon^{\alpha B}\{E_{M3,\beta}-(\pi M/h)E_{M\beta}-\hat{E}_{M\beta}\},\\
&\dot{\bar{D}}^{\,3}_M=\epsilon^{\alpha\beta}H_{M\beta,\alpha}-J^{\,3}_M,\\
&\dot{\bar{D}}^{\,\alpha}_M=\epsilon^{\alpha\beta}\{H_{M3,\beta}+(\pi M/h)H_{M\beta}\}-J^{\,\alpha}_M,
\end{aligned} \tag{2.13}$$

where

$$\epsilon^{12}=-\epsilon^{21}=A^{-1/2},\quad \epsilon^{11}=\epsilon^{22}=0, \tag{2.14}$$

and where $\hat{B}^{\,3}_M,\hat{E}_{M\beta}$ represent, respectively, contributions from the magnetic induction field and the electric field on the major surfaces of the shell-like body.

In the absence of thermal effects, we may express the constitutive equations for a magnetic, polarized elastic membrane in terms of a strain energy $\psi$ per unit mass of $\mathscr{S}$. We may therefore assume that, in addition to dependence on the fields $E_{Mi}$ and $H_{Mi}$, $\psi$ is a function of $\mathbf{a}_\alpha$ and their reference values or is equivalently of the form

$$\rho_R\psi=\hat{\psi}(e_{\alpha\beta},E_{Mi},H_{Mi};A_{\alpha\beta}), \tag{2.15}$$

where the relative strain measure $e_{\alpha\beta}$ is defined by

$$e_{\alpha\beta}=\tfrac{1}{2}(a_{\alpha\beta}-A_{\alpha\beta}). \tag{2.16}$$

Then, from Eq. (9.5) of Ref. 3, we have

$$N^{\alpha\beta}+N^{\alpha\beta}_e=\frac{1}{2}\frac{\rho}{\rho_R}\left(\frac{\partial\hat{\psi}}{\partial e_{\alpha\beta}}+\frac{\partial\hat{\psi}}{\partial e_{\beta\alpha}}\right), \tag{2.17a}$$
$$N^{3\beta}+N^{3\beta}_e=0, \tag{2.17b}$$

and

$$\overline{D}^{\,i}_{M} = -\frac{\partial\hat{\psi}}{\partial E_{Mi}}, \quad B^{\,i}_{M} = -\frac{\partial\hat{\psi}}{\partial H_{Mi}}, \tag{2.18}$$

where $N_e^{i\beta}$ represent contributions to the contact force arising from electromagnetic effects.

## III. TRANSVERSELY ISOTROPIC MEMBRANE

The theory of Sec. II models the main features of the response of finitely deforming (three-dimensional) thin magnetic, polarized elastic shell-like body regarded as a membrane. We now assume that the shell-like body in its reference configuration is (i) of uniform thickness $h$ and normals to its middle surface meet the major surfaces of the shell at points which are equidistant from the middle surface, and (ii) has material symmetry with respect to reflection along normals to the middle surface but is isotropic (with a center of symmetry) in directions tangent to any point of the middle surface.

Restrictions on the form of the Helmholtz free energy function $\hat{\psi}$ in (2.15), along with some material symmetry restrictions for the case in which the middle surface in the reference configuration is a plane (i.e., the case of an initially flat plate), have been discussed in some detail in Sec. V of Green and Naghdi.[3] A similar discussion can be carried out for a curved membrane. We do not record here the resulting restrictions on the general form of $\hat{\psi}$ in (2.15) but limit our attention to the case in which $\hat{\psi}$ is quadratic in the electromagnetic variables $E_{Mi}, H_{Mi}$ with $M$ limited to the values 0,1 only. Since $E_{0\alpha}$ and $H_{03}$ are zero in the representation used, the relevant nonzero components of the electromagnetic variables are given by the following list:

$$E_{1\alpha}, H_{0\alpha}, H_{1\alpha}, E_{03}, E_{13}, H_{13}. \tag{3.1}$$

Then, $\hat{\psi}$ can be expressed in the reduced form

$$\begin{aligned} -\hat{\psi} = & -\phi_0 + \tfrac{1}{2}\phi_1 A^{\alpha\beta}E_{1\alpha}E_{1\beta} + \tfrac{1}{2}\phi_2 A^{\alpha\beta}H_{0\alpha}H_{0\beta} \\ & + \tfrac{1}{2}\phi_3 A^{\alpha\beta}H_{1\alpha}H_{1\beta} + \tfrac{1}{2}\phi_4 E_{03}^2 \\ & + \tfrac{1}{2}\phi_5 E_{13}^2 + \tfrac{1}{2}\phi_6 H_{13}^2 \\ & + \tfrac{1}{2}\phi_7 A^{\alpha\gamma}A^{\beta\delta}e_{\alpha\beta}E_{1\gamma}E_{1\delta} \\ & + \tfrac{1}{2}\phi_8 A^{\alpha\gamma}A^{\beta\delta}e_{\alpha\beta}H_{0\gamma}H_{0\delta} \\ & + \tfrac{1}{2}\phi_9 A^{\alpha\gamma}A^{\beta\delta}e_{\alpha\beta}H_{1\gamma}H_{1\delta}, \end{aligned} \tag{3.2}$$

where $\phi_0, \phi_1, \ldots, \phi_9$ are scalar functions of the strain invariants

$$J_1 = A^{\alpha\beta}e_{\alpha\beta}, \quad J_2 = A^{\alpha\gamma}A^{\beta\delta}e_{\alpha\beta}e_{\gamma\delta}. \tag{3.3}$$

With the help of (3.2), from (2.18) and (2.17a) we obtain

$$\begin{aligned} &\overline{D}^{\,\alpha}_{1} = \phi_1 A^{\alpha\beta}E_{1\beta} + \phi_7 A^{\alpha\gamma}A^{\beta\delta}e_{\gamma\delta}E_{1\beta}, \\ &B^{\,\alpha}_{0} = \phi_2 A^{\alpha\beta}H_{0\beta} + \phi_8 A^{\alpha\gamma}A^{\beta\delta}e_{\gamma\delta}H_{0\beta}, \\ &B^{\,\alpha}_{1} = \phi_3 A^{\alpha\beta}H_{1\beta} + \phi_9 A^{\alpha\gamma}A^{\beta\delta}e_{\gamma\delta}H_{1\beta}, \\ &\overline{D}^{\,3}_{0} = \phi_4 E_{03}, \quad \overline{D}^{\,3}_{1} = \phi_5 E_{13}, \quad B^{\,3}_{1} = \phi_6 H_{13}, \end{aligned} \tag{3.4}$$

and

$$\begin{aligned} & N^{\alpha\beta} + N_e^{\alpha\beta} \\ & = \left(\frac{A}{a}\right)^{1/2}\left(\frac{\partial\phi_0}{\partial J_1}A^{\alpha\beta} + 2\frac{\partial\phi_0}{\partial J_2}A^{\alpha\gamma}A^{\beta\delta}e_{\gamma\delta}\right) + R^{\alpha\beta}, \end{aligned} \tag{3.5}$$

where the components $R^{\alpha\beta}$ are quadratic functions in the electromagnetic variables (3.1).

In the absence of body forces, it follows from Appendices A and B in Green and Naghdi[3] that

$$\rho(\mathbf{f} + \mathbf{f}_e) = \operatorname{div}_s \mathbf{N}_e + [\mathbf{t} + \mathbf{t}_e]_+ + [\mathbf{t} + \mathbf{t}_e]_-, \tag{3.6}$$

where the last two terms on the right-hand side of (3.6) represent the values of the mechanical force $\mathbf{t}$ and the force $\mathbf{t}_e$ due to the electromagnetic fields on the major surfaces of the membrane.

In the rest of this paper, we assume that the effect of the electromagnetic variables on the motion of the membrane is *small* so that quadratic terms in the variables (3.1) may be omitted in the equations of motion (2.9a). Then, the effect of $R^{\alpha\beta}$ in (3.5) and $\mathbf{t}_e$ in (3.6) may be neglected and the equations of motion reduce to

$$\rho\dot{\mathbf{v}} = a^{-1/2}(a^{1/2}S^{\alpha\beta}\mathbf{a}_\beta)_{,\alpha} + \mathbf{t}_+ + \mathbf{t}_-, \tag{3.7}$$

where

$$S^{\alpha\beta} = \left(\frac{A}{a}\right)^{1/2}\left(\frac{\partial\phi_0}{\partial J_1}A^{\alpha\beta} + 2\frac{\partial\phi_0}{\partial J_2}A^{\alpha\gamma}A^{\beta\delta}e_{\gamma\delta}\right). \tag{3.8}$$

## IV. WAVE PROPAGATION IN A STRETCHED, TWISTED CIRCULAR TUBE

Adkins and Rivlin[1] have studied the propagation of electromagnetic waves in magnetic, polarized, elastic isotropic circular rods subjected to torsion. They have also briefly discussed wave propagation in a twisted hollow circular rod or a tube. The theory of Sec. III is particularly relevant to the problem of electromagnetic wave propagation in a *thin* magnetic, polarized, elastic circular tube.

Consider a circular cylindrical membrane of initial radius $c$. The membrane is subjected to a large extension along its axis and a large twist about its axis so that the radius of the cylinder is changed to $kc$, where $k$ is a constant. The deformed cylinder is held in equilibrium and is subjected to small electromagnetic fields. The major surfaces of the cylindrical membrane will be regarded as perfectly conducting so that

$$\hat{B}^{\,3}_{M} = 0, \quad \hat{E}_{M\beta} = 0, \tag{4.1}$$

in Eq. (2.13). Let a material point on the reference surface in the reference configuration of the membrane be referred to a system of rectangular Cartesian axes with base vectors $\mathbf{e}_i$ ($i = 1,2,3$) and with $z$ axis directed along the axis of the cylindrical membrane. Further, let $(c,\theta,z)$ denote the cylindrical polar coordinates of the material point on the reference surface. Then, the position vector $\mathbf{R}$,$\mathbf{r}$ of the reference surface of the shell-like body in both the reference configuration and the current configuration may conveniently be taken as

$$\begin{aligned} &\mathbf{R} = c(\cos\theta\mathbf{e}_1 + \sin\theta\mathbf{e}_2) + z\mathbf{e}_3, \\ &\mathbf{r} = kc[\cos(\theta + \tau z)\mathbf{e}_1 + \sin(\theta + \tau z)\mathbf{e}_2] + \lambda z\mathbf{e}_3, \end{aligned} \tag{4.2}$$

where the quantities $k,\lambda,\tau$ are constants. Specifying the convected coordinates $\theta^\alpha$ in the deformed configuration by $\theta^1 = c\theta$ and $\theta^2 = z$, it follows from (4.2) that in the reference configuration

$$A = A_{11} = A_{22} = A^{11} = A^{22} = 1, \quad A_{12} = A^{12} = 0, \tag{4.3}$$

and that the components of the strain tensor $e_{\alpha\beta}$ and the associated invariants are

$$e_{11} = \tfrac{1}{2}(k^2 - 1), \quad e_{12} = \tfrac{1}{2}k^2\tau c, \quad e_{22} = \tfrac{1}{2}(\lambda^2 - 1 + k^2\tau^2 c^2),$$
$$J_1 = e_{11} + e_{22}, \quad J_2 = e_{11}^2 + 2e_{12}^2 + e_{22}^2, \quad a = k^2\lambda^2. \tag{4.4}$$

Also, from (3.4), (3.8), (4.3), and (4.4), we have

$$\begin{aligned}
\overline{D}_1^1 &= \phi_1 E_{11} + \phi_7(e_{11}E_{11} + e_{12}E_{12}),\\
\overline{D}_1^2 &= \phi_1 E_{12} + \phi_7(e_{22}E_{12} + e_{12}E_{11}),\\
B_1^1 &= \phi_3 H_{11} + \phi_9(e_{11}H_{11} + e_{12}H_{12}),\\
B_1^2 &= \phi_3 H_{12} + \phi_9(e_{22}H_{12} + e_{12}H_{11}),\\
\overline{D}_0^3 &= \phi_4 E_{03},\\
\overline{D}_1^3 &= \phi_5 E_{13}, \quad B_1^3 = \phi_6 H_{13},\\
B_0^1 &= \phi_2 H_{01} + \phi_8(e_{11}H_{01} + e_{12}H_{02}),\\
B_0^2 &= \phi_2 H_{02} + \phi_8(e_{22}H_{02} + e_{12}H_{01}),
\end{aligned} \tag{4.5}$$

and

$$\begin{aligned}
S^{11} &= (k\lambda)^{-1}\left(\frac{\partial\phi_0}{\partial J_1} + 2e_{11}\frac{\partial\phi_0}{\partial J_2}\right),\\
S^{22} &= (k\lambda)^{-1}\left(\frac{\partial\phi_0}{\partial J_1} + 2e_{22}\frac{\partial\phi_0}{\partial J_2}\right),\\
S^{12} &= 2(k\lambda)^{-1}e_{12}\frac{\partial\phi_0}{\partial J_2}.
\end{aligned} \tag{4.6}$$

The tube is maintained in equilibrium by an internal pressure $P$ and by forces over its end cross sections, the equations of equilibrium being satisfied with $P$ given by

$$P = (k/c)(S^{11} + 2\tau c S^{12} + \tau^2 c^2 S^{22}). \tag{4.7}$$

We consider now propagation of waves along the cylindrical tube and assume that all the electromagnetic variables in (4.5) are proportional to $\exp i(n\theta - pz + \omega t)$, where $n$ is an integer. Then, after removing this factor and when the free charge and current are zero, from (2.13) and (4.5) we obtain

$$\begin{aligned}
&\omega\phi_4 E_{03} - pH_{01} - (n/c)H_{02} = 0,\\
&-pE_{03} + \omega[(\phi_2 + \phi_8 e_{11})H_{01} + \phi_8 e_{12}H_{02}] = 0,\\
&-(n/c)E_{03} + \omega[\phi_8 e_{12}H_{01} + (\phi_2 + \phi_8 e_{22})H_{02}] = 0,
\end{aligned} \tag{4.8}$$

and

$$\begin{aligned}
&(n/c)E_{12} + pE_{11} + \omega\phi_6 H_{13} = 0,\\
&\omega\phi_5 E_{13} - (n/c)H_{12} - pH_{11} = 0,\\
&ipE_{13} + (\pi/h)E_{12} - i\omega[(\phi_3 + \phi_9 e_{11})H_{11} + \phi_9 e_{12}H_{12}] = 0,\\
&i(n/c)E_{13} - (\pi/h)E_{11} - i\omega[\phi_9 e_{12}H_{11}\\
&\quad + (\phi_3 + \phi_9 e_{22})H_{12}] = 0,\\
&ipH_{13} - (\pi/h)H_{12} + i\omega[(\phi_1 + \phi_7 e_{11})E_{11}\\
&\quad + \phi_7 e_{12}E_{12}] = 0,\\
&i(n/c)H_{13} + (\pi/h)H_{11} + i\omega[\phi_7 e_{12}E_{11}\\
&\quad + (\phi_1 + \phi_7 e_{22})E_{12}] = 0,
\end{aligned} \tag{4.9}$$

where all coefficients in (4.8) and (4.9) are functions of $J_1, J_2$, and are constants for specified values of $\lambda, k, \tau$.

A nonzero solution of equations (4.8) exists if

$$\begin{aligned}
&p = \alpha n \pm \beta,\\
&\alpha(\phi_2 + \phi_8 e_{22}) = (\phi_8/c)e_{12},\\
&\beta(\phi_2 + \phi_8 e_{22})\\
&\quad = \{[\omega^2\phi_4(\phi_2 + \phi_8 e_{22}) - n^2/c^2]\\
&\qquad \times [(\phi_2 + \phi_8 e_{11})(\phi_2 + \phi_8 e_{22}) - \phi_8^2 e_{12}^2]\}^{1/2}.
\end{aligned} \tag{4.10}$$

In the special case when the cylindrical tube is undeformed and in its reference configuration, $e_{11} = e_{22} = e_{12} = 0$, $J_1 = J_2 = 0$, and the various coefficients in (4.10) have the values

$$\begin{aligned}
&\alpha = 0, \quad \beta = \{\phi_2\phi_4\omega^2 - n^2/c^2\}^{1/2},\\
&\phi_2 = \phi_2(0,0), \quad \phi_4 = \phi_4(0,0),
\end{aligned} \tag{4.11}$$

where $\phi_4$ and $\phi_2$ are positive constants which represent, respectively, the dielectric constant and magnetic permeability for the undeformed cylinder.

The wave corresponding to (4.10) is a transverse electric wave. Following the procedure of Adkins and Rivlin,[1] for wave propagation in a twisted solid cylinder, we may combine two solutions corresponding to

$$\begin{aligned}
&n = m > 0, \quad p = \alpha m + \beta,\\
&n = -m < 0, \quad p = -\alpha m + \beta.
\end{aligned} \tag{4.12}$$

If $\alpha$ is positive and $\beta$ real, we obtain a wave which is elliptically polarized at each point and is propagated without distortion or attenuation, but each of the electric fields is rotated through an angle $\alpha$ per unit length traveled along the cylinder. If

$$\beta = -i\gamma, \quad \gamma \text{ real and } > 0, \tag{4.13}$$

then with (4.12) we obtain a similar type wave propagation but now the wave is attenuated.

We now turn our attention to Eqs. (4.9) and consider only the case for which the cylinder is isotropic so that

$$\phi_1 = \phi_4 = \phi_5, \quad \phi_2 = \phi_3 = \phi_6, \quad \phi_8 = \phi_9. \tag{4.14}$$

Also, we suppose that the cylinder is twisted through a small angle $\tau z$ at each point and is held at the same radius and length as when it is undeformed. This is possible if there is an internal pressure $P$ given by (4.7) and when

$$\begin{aligned}
&e_{11} = 0, \quad e_{12} = \tfrac{1}{2}\tau c, \quad e_{22} = \tfrac{1}{2}\tau^2 c^2,\\
&J_1 = \tfrac{1}{2}\tau^2 c^2, \quad J_2 = \tfrac{1}{2}\tau^2 c^2 + \tfrac{1}{4}\tau^4 c^4.
\end{aligned} \tag{4.15}$$

A nonzero solution of (4.9) gives a quartic equation for $p$ and for the special case under discussion yields the values

$$p = \pm\left(\phi_1\phi_2\omega^2 - \frac{n^2}{c^2} - \frac{\pi^2}{h^2}\right)^{1/2}, \tag{4.16a}$$

$$p = \pm\left(\phi_1\phi_2\omega^2 - \frac{n^2}{c^2} - \frac{\pi^2}{h^2}\right)^{1/2} + \frac{\phi_1\phi_9 + \phi_2\phi_7}{2\phi_1\phi_2}n\tau, \tag{4.16b}$$

where in obtaining (4.16) terms involving $\tau^2$ and higher powers of $\tau$ have been neglected. The coefficients $\phi_1, \phi_2, \phi_7, \phi_9$ in (4.16) take constant values when $J_1 = J_2 = 0$. The values in (4.16a) correspond to wave propagation in the undeformed cylinder. The second set of values in (4.16b) are associated with wave propagation which is similar in form to that given by (4.10) but with different behavior for the electromagnetic vectors. Again, as in (4.12) two solutions from (4.16b) may be combined to obtain an elliptically polarized wave.

## V. WAVE PROPAGATION IN A RIGID TUBE

Wave propagation in a solid circular cylindrical waveguide in the presence of static electric and magnetic fields has been discussed by Adkins and Rivlin.[2] Here we consider the corresponding problem for a circular cylindrical tube

employing the membrane theory of Sec. II. As in Sec. III, we suppose that the membrane is transversely isotropic and we restrict the electromagnetic independent variables to the list (3.1). Then, adopting the usual tensor notations, we may write

$$\mathbf{E}_{11} = E_{1\alpha}E_{1\beta}\mathbf{A}^{\alpha}\otimes\mathbf{A}^{\beta}, \quad \mathbf{H}_{00} = H_{0\alpha}H_{0\beta}\mathbf{A}^{\alpha}\otimes\mathbf{A}^{\beta},$$
$$\mathbf{H}_{11} = H_{1\alpha}H_{1\beta}\mathbf{A}^{\alpha}\otimes\mathbf{A}^{\beta}, \quad \mathbf{H}_{01} = H_{0\alpha}H_{1\beta}\mathbf{A}^{\alpha}\otimes\mathbf{A}^{\beta}, \tag{5.1}$$
$$\mathbf{K}_{10} = E_{1\alpha}H_{0\beta}\mathbf{A}^{\alpha}\otimes\mathbf{A}^{\beta}, \quad \mathbf{K}_{11} = E_{1\alpha}H_{1\beta}\mathbf{A}^{\alpha}\otimes\mathbf{A}^{\beta},$$

and the Helmholtz free-energy function $\hat{\psi}$ in (2.15) now becomes a function of the tensors defined in (5.1) and of $E_{03}, E_{13}, H_{13}$. With the help of symmetry restrictions discussed in Sec. V of Green and Naghdi,[3] it follows that

$$-\hat{\psi} = \phi(I_1, I_2, \ldots, I_{16}), \tag{5.2}$$

where the invariants $I_1, \ldots, I_{16}$ are defined by

$$I_1 = \operatorname{tr}\mathbf{E}_{11}, \quad I_2 = \operatorname{tr}\mathbf{H}_{00}, \quad I_3 = \operatorname{tr}\mathbf{H}_{11},$$
$$I_4 = E_{03}^2, \quad I_5 = E_{13}^2, \quad I_6 = H_{13}^2,$$
$$I_7 = (\operatorname{tr}\mathbf{H}_{01})^2, \quad I_8 = (\operatorname{tr}\mathbf{K}_{10})^2, \quad I_9 = (\operatorname{tr}\mathbf{K}_{11})^2,$$
$$I_{10} = \operatorname{tr}\mathbf{H}_{01}\operatorname{tr}\mathbf{K}_{10}\operatorname{tr}\mathbf{K}_{11}, \quad I_{11} = E_{03}E_{13}\operatorname{tr}\mathbf{H}_{01}, \tag{5.3}$$
$$I_{12} = E_{03}E_{13}\operatorname{tr}\mathbf{K}_{10}\operatorname{tr}\mathbf{K}_{11}, \quad I_{13} = E_{03}H_{13}\operatorname{tr}\mathbf{K}_{10},$$
$$I_{14} = E_{03}H_{13}\operatorname{tr}\mathbf{H}_{01}\operatorname{tr}\mathbf{K}_{11}, \quad I_{15} = E_{13}H_{13}\operatorname{tr}\mathbf{K}_{11},$$
$$I_{16} = E_{13}H_{13}\operatorname{tr}\mathbf{H}_{01}\operatorname{tr}\mathbf{K}_{10}.$$

We consider now one example in which there is a uniform static magnetic field $H$ parallel to the axis of the tube and superposed on this are small electromagnetic waves. Then, recalling from Sec. IV the notation $\theta^1 = c\theta, \theta^2 = z$ and retaining only the linear terms in the constitutive equations, we have

$$\overline{D}_1^1 = \alpha_1 E_{11}, \quad \overline{D}_1^2 = \alpha_1 E_{12} + \alpha_8 \overline{H}^2 E_{12},$$
$$B_0^1 = \alpha_2 H_{01}, \quad B_0^2 = \alpha_2(\overline{H} + H_{02}),$$
$$B_1^1 = \alpha_3 H_{11}, \quad B_1^2 = \alpha_3 H_{12} + \alpha_7 \overline{H}^2 H_{12}, \tag{5.4}$$
$$\overline{D}_0^3 = a_4 E_{03}, \quad \overline{D}_1^3 = \alpha_5 E_{13}, \quad B_1^3 = \alpha_6 H_{13},$$

where

$$\overline{H} = h^{1/2}H, \quad \alpha_r = 2\frac{\partial\phi}{\partial I_r} \quad (r = 1,2,\ldots,8), \tag{5.5}$$

and the constant coefficients $\alpha_r$ are evaluated at the values

$$I_2 = \overline{H}^2, \quad I_r = 0 \quad (r \neq 2). \tag{5.6}$$

Assuming that the electromagnetic variables are proportional to $\exp i(n\theta - pz + \omega t)$, after removal of this factor and when the free charge and current are zero, from (2.13), (4.1), and (5.4) we obtain

$$\alpha_4\omega E_{03} - pH_{01} - (n/c)H_{02} = 0,$$
$$pE_{03} - \alpha_2\omega H_{01} = 0, \tag{5.7}$$
$$(n/c)E_{03} - \alpha_2\omega H_{02} = 0,$$

and

$$\alpha_5\omega E_{13} - (n/c)H_{12} - pH_{11} = 0,$$
$$(n/c)E_{12} + pE_{11} + \alpha_6\omega H_{13} = 0,$$
$$ipE_{13} + (\pi/h)E_{12} - i\omega\alpha_3 H_{11} = 0, \tag{5.8}$$
$$i(n/c)E_{13} - (\pi/h)E_{11} - i\omega(\alpha_3 + \alpha_7\overline{H}^2)H_{12} = 0,$$
$$ipH_{13} - (\pi/h)H_{12} + i\omega\alpha_1 E_{11} = 0,$$
$$i(n/c)H_{13} + (\pi/h)H_{11} + i\omega(\alpha_1 + \alpha_8\overline{H}^2)E_{12} = 0.$$

A nonzero solution of equations (5.7) is possible if

$$p^2 = \alpha_2\alpha_4\omega^2 - n^2/c^2. \tag{5.9}$$

Also, a nonzero solution of (5.8) yields a quartic equation for $p$. We record here the results only for the case of a cylinder which is isotropic in its reference state when the coefficients in (5.8) which depend only on $\overline{H}^2$ are related as follows:

$$\alpha_5 = \alpha_4 = \alpha_1, \quad \alpha_6 = \alpha_3 = \alpha_2. \tag{5.10}$$

Then, the values of $p^2$ obtained from (5.8) are

$$p^2 = \alpha_1\alpha_2\omega^2 - \frac{\alpha_1}{\alpha_1 + \alpha_8\overline{H}^2}\left(\frac{\pi^2}{h^2} + \frac{n^2}{c^2}\right), \tag{5.11a}$$

$$p^2 = \alpha_1\alpha_2\omega^2 - \frac{\alpha_2}{\alpha_2 + \alpha_7\overline{H}^2}\left(\frac{\pi^2}{h^2} + \frac{n^2}{c^2}\right). \tag{5.11b}$$

For the two values $n = \pm m$ each of the three expressions (5.9), (5.11a), and (5.11b) yield the same value for $p$. The combination of two such waves therefore yields a wave which is propagated without rotation. The development of Adkins and Rivlin[1] for a cylindrical waveguide of solid circular cross section, which is obtained from the three-dimensional theory, resulted in the possibility of wave propagation with rotation and one might expect some similar mode of propagation would be possible also in a circular tube. The reason for this lack of rotation lies in some differences which occur in the basic constitutive equations used by Adkins and Rivlin,[2] after allowing also for any differences due to the nature of the present two-dimensional treatment of the subject. In the constitutive equations employed by Adkins and Rivlin,[2] the terms which give rise to a rotating wave in the context of three-dimensional theory may be regarded as being of the form

$$\overline{\mathbf{D}} = \alpha\mathbf{E}\times\mathscr{H}, \quad \mathbf{B} = \beta\mathbf{H}\times\mathscr{H}, \tag{5.12}$$

where $\mathscr{H}$ is the static magnetic field. Corresponding terms do not appear in the present paper because they are ruled out by thermodynamical considerations (see Green and Naghdi[3]). Moreover, within the three-dimensional framework, the relations (5.12) do not appear to be admissible from a thermodynamical viewpoint. However, owing to the way in which Adkins and Rivlin[2] have formulated their constitutive equations by first removing a time factor $\exp(i\omega t)$, the crucial terms might also be interpreted as arising from expressions of the form

$$\overline{\mathbf{D}} = \alpha\dot{\mathbf{E}}\times\mathscr{H}, \quad \mathbf{B} = \beta\dot{\mathbf{H}}\times\mathscr{H}. \tag{5.13}$$

Such expressions as (5.13) are admissible from a thermodynamical viewpoint since $\overline{\mathbf{D}}\cdot\dot{\mathbf{E}} + \mathbf{B}\cdot\dot{\mathbf{H}} = 0$, although they are not a part of what is usually defined as a magnetic polarized elastic medium. The forms (5.13) appear to be in line with what other authors use for such wave propagation problems when they write a part of their constitutive equations in the forms

$$\overline{\mathbf{D}} = i\alpha\mathbf{E}\times\mathscr{H}, \quad \mathbf{B} = i\beta\mathbf{H}\times\mathscr{H}. \tag{5.14}$$

In order to obtain results appropriate for wave propagation with rotation in a cylindrical tube, it appears to be necessary to add to the constitutive equations (5.4) terms which are the two-dimensional counterpart of the relations (5.13) but we omit details here.

## ACKNOWLEDGMENTS

The work of one of us (P. M. N.) was supported by the U. S. Office of Naval Research under Contract N00014-75-

C-0148, Project NR 064-436 with the University of California, Berkeley (U. C. B.) During 1983, A.E.G. held a visiting appointment in U. C. B.

[1]J. E. Adkins and R. S. Rivlin, Philos. Trans. R. Soc. London Ser. A **255**, 389 (1963).

[2]J. E. Adkins and R. S. Rivlin, Int. J. Eng. Sci. **1**, 187 (1963).

[3]A. E. Green and P. M. Naghdi, Philos. Trans. R. Soc. London Ser. A **309**, 559 (1983).

[4]H. F. Tiersten, *Linear Piezo-electric Plate Vibrations* (Plenum, New York, 1969).

[5]N. Bugdayci and D. B. Bogy, Int. J. Solids Struct. **17**, 1159 and 1179 (1981).

[6]P. M. Naghdi, *Handbuch der Physik*, edited by C. Truesdell (Springer-Verlag, Berlin, 1972), Vol. VIa/2, p. 425.

# The squared eigenstates of the sine–Gordon eigenvalue problem

D. J. Kaup
*Department of Physics, Clarkson University, Potsdam, New York 13676*

(Received 15 February 1984; accepted for publication 30 March 1984)

An analysis of the squared sine–Gordon eigenvalue problem in laboratory coordinates is presented. It is shown that unlike the unsquared laboratory coordinate eigenvalue problem, the squared laboratory coordinate eigenvalue problem may be cast into the form of a standard eigenvalue problem, wherein an eigenvalue independent operator operating on an eigenfunction generates the eigenvalue. With this form, it becomes rather elementary to obtain the squared-eigenfunction expansion of the sine–Gordon potentials as well as to demonstrate closure.

PACS numbers: 03.40.Kf

## I. INTRODUCTION

The squared eigenfunctions of an operator are important in at least two aspects. First, they can be used in an expansion of the potentials. This is actually a superposition principle. For example, for the KdV equation

$$u_t - 6u_x u + u_{xxx} = 0\,, \tag{1}$$

the expansion for the $N$-soliton case is[1]

$$u(x) = -4\sum_{j=1}^{N} k_j \psi_j^2(x)\,, \tag{2}$$

where $k_j^2$ is the $j$th bound state eigenvalue and $\psi_j$ is the $j$th bound state eigenfunction of the Schrödinger equation $\psi_{xx} - (u + k^2)\psi = 0$. As Eq. (2) demonstrates, $u$ is simply a linear sum of terms, each term of which can be associated with a given soliton.[1] I remark that this superposition principle also provides a viewpoint[2] from which one can understand and explain some of the well-known features of soliton collisions. These squared eigenstates also occur when one is considering perturbations, where they provide a transformation from the potentials into scattering space, and back again.[3,4] Thus an understanding of these squared eigenstates and their closure is of value. The first such analysis[5] was for the squared eigenfunctions of the Zakharov–Shabat eigenvalue problem.[6] The Schrödinger equation has been similarly treated in Ref. 7.

The purpose of this paper is to present a similar analysis for the "sine–Gordon eigenvalue problem"[8] which is

$$iv_{1x} - \tfrac{1}{2}\zeta v_1 + p_+ v_2 + (1/8\zeta)(v_1 \cos\theta + v_2 \sin\theta) = 0\,, \tag{3a}$$

$$iv_{2x} + \tfrac{1}{2}\zeta v_2 - p_+ v_1 - (1/8\zeta)(v_2 \cos\theta - v_1 \sin\theta) = 0\,, \tag{3b}$$

where $v = (v_1, v_2)^T$ is the eigenfunction, the superscript "$T$" designates the transpose,

$$p_\pm = \tfrac{1}{4}(\theta_t \pm \theta_x)\,, \tag{4}$$

$\zeta$ is the eigenvalue, and $\theta$ is the sine–Gordon field which satisfies

$$\theta_{tt} - \theta_{xx} + \sin\theta = 0\,. \tag{5}$$

The inverse scattering for (3) has been solved and is presented in Refs. 9 and 10. In this paper I shall follow the notation in Ref. 10. Also I shall use the definitions of the Jost functions, scattering data, etc., exactly as given in Ref. 10.

What is different about the sine–Gordon eigenvalue problem, Eq. (3), is that one cannot reduce this problem to the canonical form of $Lv = \zeta v$, where $L$ is a $\zeta$-independent operator. Various results such as orthogonality, closure, etc., for the Zakharov–Shabat case[5] and for the Schrödinger case[7] depend very much on the fact that $L$ is $\zeta$-independent. So the obvious question is whether or not similar results can be obtained for the sine–Gordon eigenvalue problem where $L$ is not $\zeta$-dependent.

However, as I shall now demonstrate, quite surprisingly, the *squared* sine–Gordon eigenstates can be cast into the canonical form, although the unsquare problem, Eq. (3), cannot be. To see this, let $u$ and $v$ be any two solutions of (3). Then it follows that

$$i\,\partial_x V_1 = (4\zeta^2 \cos\theta - 1)V_2 - 4\zeta^2 V_3 \sin\theta\,, \tag{6a}$$

$$i\,\partial_x V_2 = \tfrac{1}{4}(\cos\theta - 1/4\zeta^2)V_1 + 2ip_- V_3\,, \tag{6b}$$

$$i\,\partial_x V_3 = -2ip_- V_2 - \tfrac{1}{4}V_1 \sin\theta\,, \tag{6c}$$

where

$$V_1 = u^T v\,, \tag{7a}$$

$$V_2 = (1/4\zeta)u^T(\sigma_1 \sin\theta + \sigma_3 \cos\theta)v\,, \tag{7b}$$

$$V_3 = (1/4\zeta)u^T(\sigma_1 \cos\theta - \sigma_3 \sin\theta)v\,, \tag{7c}$$

and $\sigma_j$ are the Pauli spin matrices.

A justification for the choices made in (7) is in order. Although almost any combination could have been chosen for $V_1$, $V_2$, and $V_3$, the specific combination given in (7) was chosen because these are the natural terms which arise when one considers perturbations. For example, the first-order change in the transmission coefficient $a$ is given by

$$\delta a = \int_{-\infty}^{\infty} dx\,[\,V_1 \delta p_+ + \tfrac{1}{2}iV_2 \delta\theta\,]\,. \tag{8}$$

[To obtain Eq. (8), one may proceed as in Ref. 6, but using Eq. (3) instead of the Zakharov–Shabat equation.] [Also in Eq. (8), $V_1$ and $V_2$ are to be constructed from the $u = \phi$ and the $v = \psi$ eigenstates of Eq. (3), as defined in Ref. 10. But for the following argument this is not an important point.] The important feature of Eq. (8) is that $V_1$ and $V_2$ appear as components of a matrix which transforms one from perturbations in the potentials to perturbations in the scattering data. Since the inverse scattering problem of (3) has been solved,[9,10] we know that these $V_1$'s and $V_2$'s must satisfy some form of a closure property. So it is natural to start with the form of $V_1$ and $V_2$ as obtained from (8), which then is Eqs. (7a) and (7b). Then Eq. (6) follows from (3) upon defining $V_3$ as in (7c).

 0022-2488/84/082467-05$02.50  

Equation (6) may now be transformed into the canonical form as follows. From (6c) one has that $V_3$ is given by

$$V_3 = V_3(-\infty) + I^-[(i/4)V_1 \sin\theta - 2p_- V_2], \tag{9}$$

where

$$I^-[u] = \int_{-\infty}^{x} u\,dx. \tag{10a}$$

We shall similarly have need later of

$$I^+[u] = \int_{x}^{\infty} u\,dx. \tag{10b}$$

As shown by Eq. (9), up to an additive constant, $V_3$ is obtained from $V_1$ and $V_2$ by an integral operator acting on them. Meanwhile from (6b)

$$(1/4\zeta^2)V_1 = -4i\,\partial_x V_2 + \cos\theta V_1 + 8ip_- V_3. \tag{11a}$$

Now divide (6a) by $4\zeta^2$, using (11a) to eliminate the $\partial_x[V_1/(4\zeta^2)]$ term. There results

$$(1/4\zeta^2)V_2 = -4\,\partial_x^2 V_2 + i\,\partial_x(\cos\theta V_1) + \cos\theta V_2 - \sin\theta V_3 + 8\partial_x(p_- V_3). \tag{11b}$$

Consequently for those solutions where $V_3(-\infty) = 0$, Eq. (11) gives us an eigenvalue problem in the canonical form where

$$L^A V = (1/4\zeta^2)V, \tag{12}$$

and the elements of $L^A$ are

$$L^A_{11} = \cos\theta - 2p_- I^- \sin\theta, \tag{13a}$$

$$L^A_{21} = -i\,\partial_x \cos\theta - (i/4)\sin\theta I^- \sin\theta + 2i\,\partial_x p_- I^- \sin\theta, \tag{13b}$$

$$L^A_{12} = -4i\,\partial_x - 16ip_- I^- p_-, \tag{13c}$$

$$L^A_{22} = \cos\theta - 4\,\partial_x^2 - 16\,\partial_x p_- I^- p_- + 2\sin\theta I^- p_-, \tag{13d}$$

where in (12), $V = (V_1, V_2)^T$.

## II. SQUARED EIGENSTATES

Once we have the form of (12), it is more or less obvious how one could proceed. I shall first follow Ref. 5, wherein I shall define the adjoint problem as well as the Jost functions of these two problems. Next I go back and consider the solution for $V_3(-\infty) \neq 0$. I can then show that it is linearly dependent on the Jost functions, from which one readily obtains the expansion of the sine–Gordon potentials, $p_\pm$ and $\sin\theta$, in the squared eigenstates. Then closure $(L_1)$ of the squared eigenstates is demonstrated using the method of Gerdjikov and Khristov.[11] Finally the inner products between the eigenstates and their adjoint eigenstates are given, which follow from the divergence theorem.

As one has probably already noted, by using the superscript "$A$," I am considering the operator $L^A$ to be the adjoint operator. Taking this point of view, then the original operator must be

$$L_{11} = \cos\theta - 2\sin\theta I^+ p_-, \tag{14a}$$

$$L_{21} = 4i\,\partial_x - 16ip_- I^+ p_-, \tag{14b}$$

$$L_{12} = i\cos\theta\,\partial_x - (i/4)\sin\theta I^+ \sin\theta - 2i\sin\theta I^+ p_-\,\partial_x, \tag{14c}$$

$$L_{22} = \cos\theta - 4\,\partial_x^2 + 2p_- I^+ \sin\theta + 16p_- I^+ p_-\,\partial_x, \tag{14d}$$

and its eigenvalue problem is

$$LW = (1/4\zeta^2)W, \tag{15}$$

where $W = (W_1, W_2)^T$. If one defines $W_3$ by

$$W_3 = W_3(+\infty) - I^+[2p_-(W_1 + i\,\partial_x W_2) + (i/4)\sin\theta W_2], \tag{16}$$

then one can represent the general solution of (15) by

$$W_1 = -\zeta u^T \sigma_3 u, \tag{17a}$$

$$W_2 = u^T v, \tag{17b}$$

$$W_3 = (1/4\zeta)u^T(\sigma_1\cos\theta - \sigma_3\sin\theta)v, \tag{17c}$$

where $u$ and $v$ are again any solution of (3).

Let us now be definite and choose the following Jost functions:

$$|1,\zeta\rangle \equiv \begin{bmatrix} W_1 \\ W_2 \end{bmatrix} (u = v = \psi), \tag{18a}$$

$$|2,\zeta\rangle \equiv \begin{bmatrix} W_1 \\ W_2 \end{bmatrix} (u = v = \bar{\psi}), \tag{18b}$$

$$|3,\zeta\rangle \equiv \begin{bmatrix} W_1 \\ W_2 \end{bmatrix} (u = \psi, v = \bar{\psi}), \tag{18c}$$

where $\psi$ and $\bar{\psi}$ are defined as in Ref. 10. The states $|1,\zeta\rangle$ and $|2,\zeta\rangle$ have $W_3(+\infty) = 0$, while $|3,\zeta\rangle$ has $W_3(+\infty)$ as being nonzero. Whence for $n = 1,2$, or 3,

$$L|n,\zeta\rangle = \frac{1}{4\zeta^2}|n,\zeta\rangle - \frac{1}{\zeta}\delta_3^n \begin{bmatrix} \frac{1}{4}\sin\theta \\ 2ip_- \end{bmatrix}. \tag{19}$$

Similarly for the adjoint states, define

$$\langle 1,\zeta| = \begin{bmatrix} V_1 \\ V_2 \end{bmatrix}^T (u = \phi = v), \tag{20a}$$

$$\langle 2,\zeta| = \begin{bmatrix} V_1 \\ V_2 \end{bmatrix}^T (u = \bar{\phi} = v), \tag{20b}$$

$$\langle 3,\zeta| = \begin{bmatrix} V_1 \\ V_2 \end{bmatrix}^T (u = \phi, v = \bar{\phi}), \tag{20c}$$

then for $n = 1$, 2, or 3

$$\langle n,\zeta|(L^A)^T = \frac{1}{4\zeta^2}\langle n,\zeta| + \frac{1}{\zeta}\delta_3^n \begin{bmatrix} 2ip_- \\ -\frac{1}{4}\sin\theta + 2(\partial_x p_-) \end{bmatrix}^T. \tag{21}$$

When the sine–Gordon field $\theta$ is real, one has the symmetries[9,10]

$$\bar{\psi}(\zeta) = i\sigma_2\psi(-\zeta), \tag{22a}$$

$$\bar{\phi}(\zeta) = i\sigma_2\phi(-\zeta), \tag{22b}$$

whence

$$|2,\zeta\rangle = |1,-\zeta\rangle, \tag{23a}$$

$$|3,-\zeta\rangle = -|3,\zeta\rangle, \tag{23b}$$

and

$$\langle 2,\zeta| = \langle 1,-\zeta|, \tag{23c}$$

$$\langle 3,-\zeta| = -\langle 3,\zeta|. \tag{23d}$$

Thus the states $|2,\zeta\rangle$ and $\langle 2,\zeta|$ are not independent states.

One may now rather easily obtain the expansion of the sine–Gordon potentials. First from the analytical properties of $\phi$ and $\psi$,[9,10] one may show that

$$|3,\zeta\rangle = -2i\zeta \oint \frac{d\zeta'}{2\pi}\left(\frac{b}{a}\right)(\zeta')\frac{|1,\zeta'\rangle}{\zeta'^2-\zeta^2}, \tag{24a}$$

and that

$$\langle 3,\zeta| = 2i\zeta \oint \frac{d\zeta'}{2\pi}\left(\frac{\bar{b}}{a}\right)(\zeta')\frac{\langle 1,\zeta'|}{\zeta'^2-\zeta^2}. \tag{24b}$$

The definitions of $b,\bar{b}$, and $a$ are as in Ref. 10, and the contour $C$ is to be taken as in Fig. 1. [Strictly speaking, Eqs. (24) are only valid when $p_-$ and $\sin\theta$ are on compact support. However, resolution of (24) into the contributions from the poles at the zeros of $a(\zeta')$ and at $\zeta'=\zeta$, as well as an integral along the real axis will always give the proper result, even when $p_-$ and $\sin\theta$ are not on compact support. Still they would be expected to satisfy a weaker criteria such as was used in Refs. 9 and 10. For simplicity of notation, we shall from here on assume compact support.] Now from (19) and (21), for $n=3$, (24) yields

$$\begin{bmatrix}\frac{1}{4}\sin\theta\\ 2ip_-\end{bmatrix} = -\frac{i}{2}\int \frac{d\zeta'}{2\pi\zeta'^2}\left(\frac{b}{a}\right)(\zeta')|1,\zeta'\rangle, \tag{25a}$$

$$\begin{bmatrix}2ip_-\\ -\frac{1}{4}\sin\theta + 2(\partial_x p_-)\end{bmatrix}^T = -\frac{i}{2}\oint \frac{d\zeta'}{2\pi\zeta'^2}\left(\frac{\bar{b}}{a}\right)(\zeta')\langle 1,\zeta'|. \tag{25b}$$

These when expanded will yield the appropriate expansion formulas. For example, expanding (25a) gives

$$\frac{1}{4}\sin\theta = \sum_k \frac{C_k}{2\zeta_k}\psi_k^T\sigma_3\psi_k + \int_{-\infty}^{\infty}\frac{i\,d\xi}{4\pi\xi}\rho(\xi)\psi^T\sigma_3\psi(\xi), \tag{26a}$$

$$p_- = \sum_k \frac{iC_k}{4\zeta_k^2}\psi_k^T\psi_k - \int_{-\infty}^{\infty}\frac{d\xi}{8\pi\xi^2}\rho(\xi)\psi^T\psi(\xi), \tag{26b}$$

with the notation as in Ref. 10.

## III. CLOSURE

The proof of closure can be easily done if we follow the method of Gerdjikov and Khristov.[11] In fact their method of proof can be directly applied to the sine–Gordon case upon a suitable redefinition of the quantities involved. The key aspect of their method is to define the two functions $G(x,y;\zeta)$ and $\bar{G}(x,y;\zeta)$. For the sine–Gordon case, these must be defined by

$$\begin{aligned}2\pi\zeta a^2(\zeta)G(x,y;\zeta) &= \psi\circ\psi(x,\zeta)\widetilde{\phi\circ\phi}(y,\zeta)\theta(x-y)\\ &+ [2\psi\circ\phi(x,\zeta)\widetilde{\psi\circ\phi}(y,\zeta)\\ &- \phi\circ\phi(x,\zeta)\widetilde{\psi\circ\psi}(y,\zeta)]\theta(y-x),\end{aligned} \tag{27a}$$

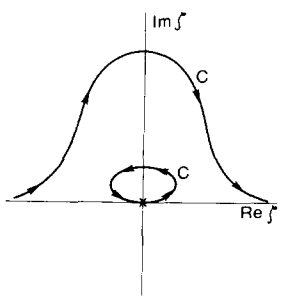



FIG. 1. The contour $C$ which consists of two parts, in the complex $\zeta$ plane. There is an essential singularity at $\zeta=0$. All zeros of $a(\zeta)$ are to lie between the two sections of $C$.

$$\begin{aligned}2\pi\zeta\bar{a}^2(\zeta)\bar{G}(x,y;\zeta) &= \bar{\psi}\circ\bar{\psi}(x,\zeta)\widetilde{\bar{\phi}\circ\bar{\phi}}(y,\zeta)\theta(x-y)\\ &+ [2\bar{\psi}\circ\bar{\phi}(x,\zeta)\widetilde{\bar{\psi}\circ\bar{\psi}}(y,\zeta)\\ &- \bar{\phi}\circ\bar{\phi}(x,\zeta)\widetilde{\bar{\psi}\circ\bar{\psi}}(y,\zeta)]\theta(y-x),\end{aligned} \tag{27b}$$

where the operations "$\circ$" and "$\bar{\circ}$" are defined by

$$u\circ v \equiv \begin{bmatrix}-\zeta v^T\sigma_3 u\\ v^T u\end{bmatrix}, \tag{28a}$$

$$u\circ v \equiv [v^T u, (1/4\zeta)v^T(\sigma_1\sin\theta + \sigma_3\cos\theta)u]. \tag{28b}$$

I shall comment that the functions $G$ and $\bar{G}$ are constructed to satisfy three conditions. First they are analytic in the appropriate half-plane. Second, the integral $\int_{-\infty}^{\infty} G(x,y;\zeta)h(y)dy$ is bounded, if $h(y)$ is integrable, for any $\zeta$ in the UHP (upper half-plane). (This is also the reason why the $x$ and $y$ dependence must depend on whether or not $x$ is greater than or less than $y$.) Similarly for $\bar{G}$, with $\zeta$ in the LHP. And third, they satisfy

$$\begin{aligned}G(x,y;\zeta) - \bar{G}(x,y;\zeta) &= \frac{1}{2\pi\zeta a^2(\zeta)}\psi\circ\psi(x,\zeta)\widetilde{\phi\circ\phi}(y,\zeta)\\ &- \frac{1}{2\pi\zeta\bar{a}^2(\zeta)}\bar{\psi}\circ\bar{\psi}(x,\zeta)\widetilde{\bar{\phi}\circ\bar{\phi}}(y,\zeta).\end{aligned} \tag{29}$$

Now one needs to only obtain certain asymptotic forms. For $|\zeta|\to\infty$ in the appropriate half-plane, one obtains[9,10]

$$\begin{aligned}2\pi G(x,y,\zeta) \to\ & e^{2ik(x-y)}\left[\begin{pmatrix}1\\0\end{pmatrix}(1,0) + O\left(\frac{1}{\zeta}\right)\right]\theta(x-y)\\ &+ e^{2ik(y-x)}\left[\begin{pmatrix}1\\0\end{pmatrix}(1,0) + O\left(\frac{1}{\zeta}\right)\right]\theta(y-x)\\ &+ O(1/\zeta^2),\end{aligned} \tag{30a}$$

$$2\pi\bar{G}(x,y;\zeta) = -2\pi G(x,y;-\zeta), \tag{30b}$$

while for $|\zeta|\to 0$ in the appropriate half-plane, one has

$$\begin{aligned}2\pi G(x,y;\zeta) \to\ & e^{2ik(x-y)}\left[\begin{pmatrix}0\\1/\zeta\end{pmatrix}\left(0,\frac{1}{4\zeta}\right) + O\left(\frac{1}{\zeta}\right)\right]\theta(x-y)\\ &+ e^{2ik(y-x)}\left[\begin{pmatrix}0\\1/\zeta\end{pmatrix}\left(0,\frac{1}{4\zeta}\right) + O\left(\frac{1}{\zeta}\right)\right]\theta(y-x)\\ &+ O(1),\end{aligned} \tag{31a}$$

$$2\pi\bar{G}(x,y;\zeta) = -2\pi G(x,y;-\zeta). \tag{31b}$$

The other asymptotic form required is the linear limit of $G$ and $\bar{G}$ (namely, their values when $\theta = 0 = \theta_t$). These limits are designated by $\Gamma$ and $\bar{\Gamma}$, and for the sine–Gordon case, I obtain

$$\begin{aligned}2\pi\Gamma(x,y;\zeta) = & e^{2ik(x-y)}\begin{bmatrix}1\\1/\zeta\end{bmatrix}\left[1,\frac{1}{4\zeta}\right]\theta(x-y)\\ &+ e^{2ik(y-x)}\begin{bmatrix}1\\-1/\zeta\end{bmatrix}\left[1,\frac{-1}{4\zeta}\right]\theta(y-x),\end{aligned} \tag{32a}$$

$$2\pi\bar{\Gamma}(x,y;\zeta) = -2\pi\Gamma(x,y;-\zeta). \tag{32b}$$

To complete the proof, one proceeds as follows.[11] First define

$$Q(x;\zeta)\equiv\int_{-\infty}^{\infty} G(x,y;\zeta)h(y)dy\,, \tag{33a}$$

$$\overline{Q}(x;\zeta)\equiv\int_{-\infty}^{\infty} \overline{G}(x,y;\zeta)h(y)dy\,, \tag{33b}$$

where $h$ is an arbitrary $L_1$ function, such that

$$\int_{-\infty}^{\infty}[|h_1(y)|+|h_2(y)|]dy<\infty\,. \tag{34}$$

It then follows that $Q$ and $\overline{Q}$ exist, and are analytic in the appropriate half-plane, except at the zeros of $a(\zeta)$ and $\bar{a}(\zeta)$. [See Eqs. (27).] Consider the function $Z(x)$ where

$$Z(x)=\left[\int_R-\int_D\right]Q(x,\zeta)d\zeta-\left[\int_R-\int_{\overline{D}}\right]\overline{Q}(x,\zeta)d\zeta\,, \tag{35}$$

where the contours $R$, $D$, and $\overline{D}$ are as defined in Fig. 2. Note that each contour consists of two separate segments. By Cauchy's theorem, one could express $Z(x)$ as a sum over the residues of $Q$ and $\overline{Q}$. I shall give the exact expression later. Instead for now I shall represent the result of this simply by

$$Z(x)=2\pi i\sum \mathrm{Res}[Q]+2\pi i\sum \mathrm{Res}[\overline{Q}]\,. \tag{36}$$

Now reorder the terms in (35), and subtract from it the identity

$$0=\left[\int_R-\int_D\right]d\zeta\int_{-\infty}^{\infty}dy\,\Gamma(x,y;\zeta)h(y)$$
$$-\left[\int_R-\int_{\overline{D}}\right]d\zeta\int_{-\infty}^{\infty}dy\,\overline{\Gamma}(x,y;\zeta)h(y)\,. \tag{37}$$

One then obtains

$$Z(x)=\int_R d\zeta\,[Q(x,\zeta)-\overline{Q}(x,\zeta)]$$
$$-\int_R d\zeta\int_{-\infty}^{\infty}dy[\Gamma(x,y;\zeta)-\overline{\Gamma}(x,y;\zeta)]h(y)$$
$$-\int_D d\zeta\int_{-\infty}^{\infty}dy[G(x,y;\zeta)-\Gamma(x,y;\zeta)]h(y)$$
$$+\int_{\overline{D}} d\zeta\int_{-\infty}^{\infty}dy[\overline{G}(x,y;\zeta)-\overline{\Gamma}(x,y;\zeta)]h(y)\,. \tag{38}$$

In the limit of $L\to\infty$ (see Fig. 2), the last two terms in (38) vanish. Then due to the symmetry in the $R$ contour (since $k=\frac{1}{2}\zeta-1/4\zeta$), the second term reduces to

$$\int_R d\zeta\int_{-\infty}^{\infty}dy(\Gamma-\overline{\Gamma})h(y)=\int_R\frac{d\zeta}{\pi}\int_{-\infty}^{\infty}e^{2ik(x-y)}h(y)dy\,. \tag{39}$$

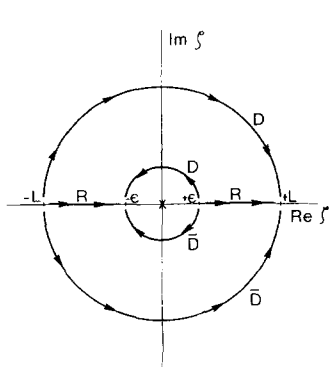



FIG. 2. The contours $D$, $\overline{D}$, and $R$ in the complex $\zeta$-plane. The constant $L$ is to be chosen larger than unity, and $\epsilon=1/(4L)$.

Due to (22), we have

$$\overline{G}(x,y;\zeta)=-G(x,y;-\zeta)\,, \tag{40}$$

whence for $L\to\infty$,

$$Z(x)=\int_R\frac{d\zeta}{\pi}\frac{\psi\circ\psi(x,\zeta)}{\zeta a^2(\zeta)}\int_{-\infty}^{\infty}dy\,\widetilde{\phi\circ\phi}(y,\zeta)h(y)$$
$$-\int_R\frac{d\zeta}{\pi}\int_{-\infty}^{\infty}e^{2ik(x-y)}h(x)dy\,. \tag{41}$$

Due to (40), the last sum in (36) is exactly equal to the first sum. If I would now evaluate (36), I find

$$Z(x)=\sum_k\frac{2\pi i}{\zeta_k(a_k')^2}$$
$$\times\left\{\begin{array}{l}\psi_k\circ\psi_k(x)\dot{h}_k+\left.\dfrac{\partial(\psi\circ\psi)(x)}{\partial\zeta}\right|_{\zeta=\zeta_k}h_k\\ -\left(\dfrac{a_k''}{a_k'}+\dfrac{1}{\zeta_k}\right)h_k\psi_k\circ\psi_k(x)\end{array}\right\}, \tag{42}$$

where

$$h_k\equiv\int_{-\infty}^{\infty}\widetilde{\phi_k\circ\phi_k}(y)h(y)dy\,, \tag{43a}$$

$$\dot{h}_k\equiv\int_{-\infty}^{\infty}\left[\frac{\partial\widetilde{\phi\circ\phi}(y)}{\partial\zeta}\right]_{\zeta_k}h(y)dy\,. \tag{43b}$$

From (41)–(43), for $L\to\infty$,

$$\int_C\frac{d\zeta}{\pi}\frac{\psi\circ\psi(x,\zeta)}{\zeta a^2(\zeta)}\int_{-\infty}^{\infty}\widetilde{\phi\circ\phi}(y,\zeta)h(y)dy$$
$$=\int_R\frac{d\zeta}{\pi}\int_{-\infty}^{\infty}e^{2ik(x-y)}h(y)dy\,. \tag{44}$$

Thus for any $h$ which is integrable, we have the closure relation

$$\int_C\frac{d\zeta}{\pi}\frac{\psi\circ\psi(x,\zeta)}{\zeta a^2(\zeta)}\widetilde{\phi\circ\phi}(y,\zeta)=\delta(x-y)I\,, \tag{45}$$

where $I$ is the identity $2\times2$ matrix. In terms of the notation in Eqs. (18) and (20), this may also be written as

$$\int_C\frac{d\zeta}{2\pi\zeta a^2(\zeta)}|x,\zeta\rangle\langle y,\zeta|=I\delta(x-y)\,, \tag{45'}$$

where it is now understood that $|\zeta\rangle=|1,\zeta\rangle$.

The last thing left to do is to give the orthogonality relations. From (12) and (15) one may obtain a divergence relationship, from which one may show for $\xi$ and $\xi'$ real that

$$\langle\xi|\xi'\rangle=2\pi\xi a^2(\xi)\delta(\xi'-\xi)\,, \tag{46a}$$

for the bound states

$$\langle\zeta_k|\xi\rangle=0=\langle\xi|\zeta_k\rangle\,, \tag{46b}$$

for "$P$-states"

$$\langle Pk|\xi\rangle=0=\langle\xi|Pk\rangle\,, \tag{46c}$$

while

$$\langle\zeta_k|\zeta_l\rangle=0\,, \tag{46d}$$

$$\langle\zeta_k|Pl\rangle=\langle Pk|\zeta_l\rangle=ia_k'^2\delta_l^k\,, \tag{46e}$$

$$\langle Pk|Pl\rangle=ia_k'^2\left(\frac{1}{\zeta_k}+\frac{a_k''}{a_k'}\right)\delta_l^k\,. \tag{46f}$$

The above "$P$-states" are defined by

$$|Pk\rangle = \left.\frac{\partial|\zeta\rangle}{\partial\zeta}\right|_{\zeta=\zeta_k}, \tag{47a}$$

$$\langle Pk| = \left.\frac{\partial\langle\zeta|}{\partial\zeta}\right|_{\zeta=\zeta_k}. \tag{47b}$$


[1]C. S. Gardner, J. M. Greene, M. D. Kruskal, and R. M. Miura, Comm. Pure Appl. Math. **27**, 97 (1974).

[2]D. J. Kaup, article in *Nonlinear Phenomena in Physics and Biology*, edited by Enns, Jones, Miura, and Rangnekar (Plenum, New York, 1981), pp. 95–123.

[3]D. J. Kaup, SIAM J. Appl. Math. **31**, 121 (1976).

[4]D. J. Kaup and A. C. Newell, Proc. R. Soc. London Ser. A **361**, 413 (1978).

[5]D. J. Kaup, J. Math. Anal. Appl. **54**, 849 (1976).

[6]V. E. Zakharov and A. B. Shabat, Zh. Eksp. Teor. Fiz. (USSR) **61**, 118 (1971) [Sov. Phys. JETP **34**, 62 (1972)].

[7]A. C. Newell, article in *Solitons*, edited by R. K. Bullough and P. J. Caudrey (Springer-Verlag, Berlin, 1980), pp. 170–285.

[8]M. J. Ablowitz, D. J. Kaup, A. C. Newell, and H. Segur, Phys. Rev. Lett. **30**, 1262 (1973).

[9]L. A. Takhtadzhyan and L. D. Faddeev, Teor. Mat. Fiz. (USSR) **21**, 160 (1974).

[10]D. J. Kaup, Stud. Appl. Math. **54**, 165 (1975).

[11]V. S. Gerdjikov and E. Kh. Khristov, Bulg. J. Phys. **7**, 28 (1980); **7**, 119 (1980).

# On the adiabatic stability of solitons and the matching of conservation laws[a)]

Pierre Lochak
*Laboratoire de Mathématiques de l'Ecole Normale Superieure, 45, rue d'Ulm, 75230 Paris, Cedex 05, France*

(Received 23 September 1983; accepted for publication 27 December 1983)

We derive a series of identities which generalize and simplify the results obtained for adiabatically modulated solitons in the case of perturbed specific integrable equations. It stresses the importance of the variational properties of the solitons, which make an adiabatic theorem plausible. A precise conjecture is made and its validity discussed from different points of view.

PACS numbers: 03.40.Kf

## I. INTRODUCTION

In a series of papers Kaup and Newell[1] and Karpman and Maslov[2] independently initiated the study of the perturbation theory for integrable equations. The example of a soliton obeying the KdV equation in a slowly changing medium was afterwards extensively studied (see Refs. 3, 4 and their bibliographies) and the production of a secondary soliton was finally pointed out in Ref. 5. This seems to ruin the validity of the adiabatic ansatz; however, adiabatically modulated solitons do have many features in common with the actual solution, a statement made precise through the study of the modified conservation laws. We show this in a more general and simple way than has yet been done, and next give arguments in favor of a possible adiabatic theorem for solitons, discussing the link with the periodic problem and Whitham's modulation theory.

## II. MATCHING THE CONSERVATION LAWS

### A. The Korteweg–deVries equation

We start with a typical example, namely

$$u_t + \alpha(t)uu_x + u_{xxx} = 0, \tag{1}$$

where we take $\alpha(t)$ to be a strictly positive $C^1$ function. When $\alpha$ is a constant, we denote (1) by $(\mathrm{KdV})_\alpha$—they are of course all equivalent by scaling. An equivalent way of writing (1) is found if one sets $q \equiv \alpha u$, $a(t) \equiv \exp(-\int^t \Gamma(s)ds) - \Gamma(s)$ any continuous function, which transforms (1) into

$$q_t + qq_x + q_{xxx} = -\Gamma(t)q. \tag{2}$$

Equation (2) models, for instance, the propagation of a wave governed by the KdV equation along a canal of varying depth.[1] We also mention that (2) includes the so-called cylindrical and spherical KdV equations.

$(\mathrm{KdV})_\alpha$ has an infinite sequence of integrals of the motion

$$H_{n,\alpha}[u] \equiv \int_{-\infty}^{+\infty} T_{n,\alpha}[u]dx, \tag{3}$$

$T_{n,\alpha}$ being a polynomial in $u,\ldots,u^{(n)}$—isobaric of weight $n$ with respect to the grading $\deg u \equiv 1$, $\deg \partial/\partial x \equiv \frac{1}{2}$; one has

$$H_{1,\alpha} \equiv M(u) = \int u, \quad H_{2,\alpha} \equiv E(u) = \int u^2,$$

$$H_{3,\alpha} = \int (\tfrac{1}{3}\alpha u^3 - u_x^2). \tag{4}$$

In order to normalize $T_{n,\alpha}$, we make the convention that it is obtained from $T_n \equiv T_{n,1}$ by changing $u$ into $\alpha u$ and then factoring out a power of $\alpha$, so that the resulting $T_{n,\alpha}$ is not dividable by $\alpha$. The first two integrals, $M(u)$ and $E(u)$, are conserved under (1), as is easily checked; in a more general way, the change of these integrals is given by the following:

*Lemma*: Let $u$ be a solution of (1), then

$$\forall n \geqslant 1, \quad \frac{d}{dt}H_{n,\alpha(t)}[u] = \dot{\alpha}(t)\cdot\int_{-\infty}^{+\infty} \frac{\partial T_{n,\alpha}}{\partial \alpha}[u]dx. \tag{5}$$

*Proof*: The variation of $H_{n,\alpha}$ is given by the right-hand side of (5), plus a term coming from the variation of $u$, $\alpha$ being held constant; this term, however, vanishes, because of the algebraic relations entailing the constancy of $H_{n,\alpha}$ for a constant $\alpha$. For instance, one has

$$\frac{d}{dt}M(u) = \frac{d}{dt}E(u) = 0; \quad \frac{d}{dt}H_{3,\alpha}[u] = \frac{1}{3}\dot{\alpha}\int_{-\infty}^{+\infty} u^3. \tag{6}$$

Next we construct the modulated soliton for (1); $(\mathrm{KdV})_\alpha$ has a soliton traveling at speed $c$,

$$S(x - ct,\alpha,c) \equiv 3c\alpha^{-1}\,\mathrm{sech}^2[\tfrac{1}{2}c^{1/2}(x - ct)], \tag{7}$$

the mass and energy of which are

$$M(s) = 12c^{1/2}\alpha^{-1}; \qquad E(s) = 24c^{3/2}\alpha^{-2}. \tag{8}$$

The modulated soliton is constructed as a traveling wave of constant *energy*. The choice to preserve (in some physical situation this integral may in fact represent the momentum) the energy rather than the mass is essential; it was found by Kaup and Newell in a direct way, as a byproduct of an inverse scattering calculation. We believe that the theorem below gives another justification for this and in the second part, we give yet another, more physical explanation, based on adiabatic invariance, which may have further consequences. We thus define

$$s(x,t,c_0,\delta) \equiv 3c(t)\alpha^{-1}(t)\,\mathrm{sech}^2\left[\tfrac{1}{2}c^{1/2}(t)\left(x - \int_0^t c(s)\,ds\right) - \delta\right]$$

$$= 3c_0\alpha^{1/3}(t)\,\mathrm{sech}^2\left[\frac{c_0^{1/2}}{2}\alpha^{2/3}(t)\right.$$

$$\left.\times\left(x - c_0\int_0^t \alpha^{4/3}(s)\,ds\right) - \delta\right], \tag{9}$$

[a)] This work was partially supported by A.T.P. (Action Thématique Programmée) "Application des Mathématiques Pures."

 0022-2488/84/082472-05$02.50  

choosing $c(t) \equiv c_0 \alpha^{4/3}(t)$ ($c_0 > 0$) to guarantee the constancy of the energy.

**Theorem:**

$$\forall n \geqslant 2 \qquad \frac{d}{dt} H_{n,\alpha(t)}[s] = \dot{\alpha}(t) \cdot \int_{-\infty}^{+\infty} \frac{\partial T_{n,\alpha}}{\partial \alpha}[s]\,dx, \quad (10)$$

where $s$ is the modulated soliton (9).

Recall that, by the lemma, the same relation holds true for the actual solution of (1), including the case $n = 1$. Only the conservation of mass is broken by $s$, since by (8), $M(s) = 12 c_0^{1/2} \alpha^{-1/3}(t)$ is not a constant. In fact, the balance of mass will be accounted for by the shelflike tail alluded to in the Introduction. We also point out that $\alpha$ is not assumed to be slowly varying [i.e., $\dot{\alpha}(t)$ small in (1) or $\Gamma(t)$ small in (2)].

*Proof of the theorem*: As was already shown in the pioneering paper of Lax,[6] one has

$$\forall n \geqslant 2 \qquad \frac{\partial H_{n,\alpha}}{\partial u}(u = s) = c_{n,\alpha} \frac{\partial E}{\partial u}(u = s). \quad (11)$$

This relation holds for any $n \geqslant 2$ ($\partial M / \partial u \equiv 1$) with a time-dependent—through $\alpha$—Lagrange multiplier $c_{n,\alpha}$; indeed, at any given time $t$, the modulated soliton has the shape $x$-dependence of the usual soliton for the corresponding values $\alpha = \alpha(t)$ and $c = c(t)$ and one needs only notice that (11) does not involve $t$ explicitly. Now

$$\frac{dH_{n,\alpha}}{dt} = \int \frac{\partial H_{n,\alpha}}{\partial u}(u = s)\dot{s}\,dx + \dot{\alpha} \int \frac{\partial T_{n,\alpha}}{\partial \alpha}(u = s)dx$$

$$= c_{n,\alpha} \int \frac{\partial E}{\partial u}(u = s)\dot{s}\,dx + \dot{\alpha} \int \frac{\partial T_{n,\alpha}}{\partial \alpha}(u = s)dx \quad (12)$$

and the first term on the right-hand side vanishes, due to the constancy of $E(s)$. The proof is finished.

*Amplification*: The situation is not very different with multisoliton states; indeed, start with a $N$-multisoliton $S_N(x,t,p_1,\ldots,p_{2N})$, containing $2N$ arbitrary parameters $(p_i)_{i=1}^{i=2N}$ which control the amplitudes, speeds, and phases of the soliton components. If the $p_i$'s are modulated so that

$$\frac{dH_{n,\alpha}}{dt}(s_N) = \dot{\alpha} \int \frac{\partial T_{n,\alpha}}{\partial \alpha}(s_N)dx \qquad \text{for} \qquad 1 < n \leqslant N+1, \quad (13)$$

one can write

$$\forall n > N+1 \qquad \frac{\partial H_{n,\alpha}}{\partial u}(u = s_N) = \sum_{k=2}^{N+1} C_{k,\alpha} \frac{\partial H_{k,\kappa}}{\partial u}(u = s_N) \quad (14)$$

for some multipliers $c_{k,\alpha}$, and (13) will still hold for $n > N+1$. Equations (13) yields $N$ equations for the modulation, whereas the multisoliton depends on $2N$ parameters, $N$ of which fix the spectrum of the operator $L = -\partial^2/\partial x^2 + s_N$. Hence, the modulation only determines the isospectral class. For a pure soliton, only a phase is left undetermined, and the modulated soliton should thus be a good approximation of the actual solution, up to translation. For similar problems, we refer to the papers by Benjamin and Mac Kean (Refs. 7 and 8; see also below, Sec. III).

In short, the modulation of the first integrals will be consistent with the whole hierarchy. Again, the mass poses a problem, because its gradient does not vanish at infinity. This suggests that, before generalizing the result of the theorem, we look at the periodic case for (1).

## B. The periodic problem for KdV

We refer to Mac Kean and Van Moerbecke[9] for the necessary material and, in order to use standard expressions, we change $u(x,t)$ into $-3u(\frac{1}{2}x,\frac{1}{2}t)$ so that (1) is transformed into

$$u_t = 3\alpha(t)uu_x - \tfrac{1}{2}u_{xxx}. \quad (15)$$

Also, we look for solutions with *primitive* period 1. If we set $q \equiv \alpha u$, (15) transforms into

$$q_t = 3qq_x - \tfrac{1}{2}q_{xxx} + (\dot{\alpha}/\alpha)q. \quad (16)$$

When $\dot{\alpha} \equiv 0$, solutions corresponding to solitons will have three *simple* eigenvalues, i.e., only one band of instability $(\lambda_1,\lambda_2)$ and the main difference with the previous case is that, apart from translation, there are now *two* degrees of freedom, $q$—which is an elliptic function—being determined, up to translation, by the values of $H_1 \equiv \int q$ and $H_2 \equiv \frac{1}{2}\int q^2$, subject only to $H_2 > \frac{1}{2}H_1^2$. This makes it possible, in the case of a nonconstant $\alpha$, to determine the modulated $q(x,t)$ such that

(i) for any fixed $t$, $q(x,t)$ as a function of $x$ has three simple eigenvalues.

(ii) $H_1(q) = \alpha(t)h_1$; $H_2(q) = \alpha^2(t)h_2$ with two constants $h_1$ and $h_2$.

For any $n$, we then have

$$\frac{\partial H_n}{\partial q} = a_n \frac{\partial H_1}{\partial q} + b_n \frac{\partial H_2}{\partial q} = a_n + b_n q. \quad (17)$$

Reverting to the original $u$, we find

$$H_n(q) = \alpha^{k_n} H_{n,\alpha}(u) \quad (18)$$

for some constants $k_n$. Hence

$$\frac{\partial H_n}{\partial q} = \alpha^{k_n - 1} \frac{\partial H_{n,\alpha}}{\partial u} \quad (19)$$

and

$$\frac{\partial H_{n,\alpha}}{\partial u} = \alpha^{1-k_n}(a_n + b_n \alpha u) = \alpha_n + \beta_n u, \quad (20)$$

yielding

$$\frac{dH_{n,\alpha}}{dt} = \int_0^1 \frac{\partial H_{n,\alpha}}{\partial u} u_t\,dx + \dot{\alpha} \int_0^1 \frac{\partial T_{n,\alpha}}{\partial \alpha}\,dx$$

$$= \lambda_{n,\alpha} \frac{d}{dt}\left(\int u\right) + \mu_{n,\alpha} \frac{1}{2} \frac{d}{dt}\left(\int u^2\right)$$

$$+ \dot{\alpha} \int_0^1 \frac{\partial T_{n,\alpha}}{\partial \alpha}\,dx = \dot{\alpha} \int_0^1 \frac{\partial T_{n,\alpha}}{\partial \alpha}\,dx \quad (21)$$

since $\int u \equiv h_1$ and $\frac{1}{2}\int u^2 \equiv h_2$.

Thus, the modulation is again consistent with the evolution of the $H_{n,\alpha}$ and $H_1$ is not exceptional any more. It is instructive to check in a direct way that the evolution of the simple spectrum is also correctly matched by the modulation: for a simple eigenvalue $\lambda$, if $q$ is a solution of (16), we have

$$\dot{\lambda} = \frac{\dot{\alpha}}{\alpha} \int_0^1 F^2 q\,dx, \quad (22)$$

where $F$ is a normalized eigenfunction for the eigenvalue $\lambda$:

$$F'' = (q - \lambda)F; \qquad \int_0^1 F^2 = 1. \quad (22\text{ bis})$$

Suppose $q$ is now the modulated function, according to (i) and (ii), with associated simple spectrum $(\lambda_0(t),\lambda_1(t),\lambda_2(t))$. Then (cf. Ref. 9)

$$F^2(x)=\frac{1}{2(\lambda-\lambda_1')}(-\sigma+2\lambda+q), \tag{23}$$

where $\sigma(t)=\lambda_0+\lambda_1+\lambda_2$ and $\lambda_1'$ is the zero of the derivative of the Hill discriminant between $\lambda_1$ and $\lambda_2$ ($\lambda_1'\in(\lambda_1,\lambda_2)$; $\Delta'(\lambda_1')=0$); hence (22) requires

$$\dot\lambda=\frac{\dot\alpha}{\alpha}\frac{1}{2(\lambda-\lambda_1')}\int_0^1(-\sigma+2\lambda+q)q\,dx$$

$$=\frac{\dot\alpha}{\alpha}\frac{1}{2(\lambda-\lambda_1')}[(-\sigma+2\lambda)H_1+2H_2]. \tag{24}$$

But we also have $F^2=\partial\lambda/\partial q$ and since

$$F^2(x)=\frac{1}{2(\lambda-\lambda_1')}\left[(-\sigma+2\lambda)\frac{\partial H_1}{\partial q}+\frac{\partial H_2}{\partial q}\right], \tag{25}$$

we obtain

$$\dot\lambda=\frac{1}{2(\lambda-\lambda_1')}[(-\sigma+2\lambda)\dot H_1+\dot H_2]. \tag{26}$$

Comparing (22) with (26), we find the conditions

$$H_1\sim\alpha(t);\qquad H_2\sim\alpha^2(t) \tag{27}$$

thus confirming the above result. This calculation raises two interesting questions:

(i) How does the discrepancy in the mass flux precisely appear when the period of the motion tends to infinity?

(ii) Does a slow variation of $\alpha$ prevent the opening up of the double spectrum? A positive answer would lead, via trace formula, to a proof that the modulated elliptic function is indeed, for slowly varying $\alpha$, a good approximation of the solution.

Finally, we briefly mention that $N$-periodic solutions can be treated in the same way; they depend on $2N+1$ parameters, $N+1$ of which determine the isospectral torus, which has dimension $N$. Again, the modulation only yields $N+1$ equations, to determine the isospectral class. For a nice geometric formulation of the modulation problem in the KdV (integrable) equation, we refer the reader to H. Flaschka *et al.*[10]

### C. Generalization

In this section, we spell out a generalization of the theorem in Sec. A. We keep on assuming that the Cauchy problem is solvable and work with sufficiently fast decreasing functions.

**Theorem:** Let

$$u_t=K_{\alpha(t)}(u,\ldots,u^{(n)}) \tag{28}$$

be a nonlinear equation such that, when $\alpha$ is fixed, the corresponding equation has a sequence of integrals of the motion

$$H_{n,\alpha}\equiv\int_{-\infty}^{+\infty}T_{n,\alpha}[u]dx. \tag{29}$$

We make the further assumption that one of the $H_{n,\alpha}$ (say $n=n_0$) is in fact independent of $\alpha$[we denote it by $E(u)$] and that there exist solitons $s(x-ct,\alpha,c)$ in the case when $\alpha$ is fixed, which satisfy the condition

$-\ \forall n\geqslant n_0$, $s$ is a critical point of $H_{n,\alpha}$ when $E$ is held fixed.

The conclusion is the following: start from the soliton profile $u(x,0)\equiv s(x,\alpha(0),c_0)$ and construct the modulated soliton, requiring that $E$ be a constant. Then, the relation

$$\frac{d}{dt}H_{n,\alpha}=\dot\alpha\int_{-\infty}^{+\infty}\frac{\partial T_{n,\alpha}}{\partial\alpha}dx \tag{30}$$

will hold for $n\geqslant n_0$, both for the actual solution and the modulated soliton.

*Proof*: It was already given in the general form in Sec. A.

*Amplification*: Multisolitons can be treated similarly; for details, see again Sec. A.

*Remarks*: 1. The simplest way of writing equations of type (28) is by scaling, in an integrable equation, on $u$, $x$, and $t$ (possibly simultaneously) and then making the scaling time dependent.

2. The condition of the theorem is satisfied at least for the KdV family; in fact, in the paper by Lax, there is indicated an easy way to check it. In particular, it also holds for the modified KdV equation.

*Examples*: By scaling on $x$ and $t$, one can study, for instance,

$$\varphi_{tt}-\varphi_{xx}=\alpha(t)\sin\varphi. \tag{31}$$

Here, we look briefly at the equation

$$v_t+\alpha^2(t)v^2v_x+v_{xxx}=0 \tag{32}$$

because, although it is obviously very similar to (1), there is an important point to be noticed. Equation (32) has again a hierarchy of conserved quantities $\tilde H_{n,\alpha}$ when $\alpha$ is constant, the first of which are

$$M(v)=\int_{-\infty}^{+\infty}v\,dx;\qquad E(v)=\int_{-\infty}^{+\infty}v^2\,dx;$$

$$\tilde H_{3,\alpha}=\int_{-\infty}^{+\infty}(\alpha^2v^4-6v_x^2)dx. \tag{33}$$

The soliton can be written

$$s(x,-ct,c,\alpha)=(6c)^{1/2}\alpha^{-1}\operatorname{sech}[c^{1/2}(x-ct)] \tag{34}$$

and

$$M(s)=\pi\sqrt{6}/\alpha\quad\text{independent of}\quad c;$$

$$E(s)=12\sqrt{c}/\alpha^2. \tag{35}$$

Hence, the modulated soliton is given by $c(t)=c_0\alpha^4(t)$;

$$s(x,t,c_0,\delta)=(6c_0)^{1/2}\alpha\operatorname{sech}\left[\alpha^2c_0^{1/2}\left(x-c_0\int_0^t\alpha^4(s)\,ds\right)-\delta\right]. \tag{36}$$

It satisfies relation (30) for $n\geqslant2$. The interesting point is the following: $M(v)$ and $E(v)$ are again both invariant under (32) and one again chooses to preserve $E(s)$ in the modulation; however, (32) and (1) are related via the Miura transform

$$u_t+\alpha uu_x+u_{xxx}=\left(2\alpha v+i\sqrt{6}\frac{\partial}{\partial x}\right)(v_t+v^2v_x+v_{xxx}),$$

$$u\equiv\alpha v^2+i\sqrt{6}v_x. \tag{37}$$

Under this transform, the energy $\int v^2\,dx$ corresponds to

the mass $\int u\,dx$. Why then in both cases does the soliton "choose" to preserve the former? A mathematical answer is given in the above theorem; a different, more physical explanation will be examined at the end of the second section.

## III. ADIABATIC STABILITY OF SOLITONS

We first explain, very briefly, why solitons are indeed privileged candidates for being adiabatic invariants. There exist at present two adiabatic theorems, one in classical mechanics, the other one in quantum mechanics. The first states the adiabatic invariance of the canonical momenta of completely integrable, finite-dimensional, Hamiltonian systems (which also includes the invariance of the phase volume); although it is used in physics, it is not rigorously demonstrated in more than one dimension for the configuration space (cf. V. Arnold, Ref. 11). The theorem in quantum mechanics states the invariance of the eigenstates associated with discrete eigenvalues in the spectrum of the Hamiltonian. It appeared in the late twenties (cf. M. Born and V. Fock, Ref. 12) and was neatly demonstrated by T. Kato (cf. Ref. 13).

The first one is thus dealing with finite-dimensional completely integrable systems, and the second with infinite-dimensional linear (hence with trivial action-angle variables) systems; they can be related to each other by considering the Hilbert space of "states" over the classical system, working with the Liouville operator (cf. Ref. 14).

Now, solitons embody the discrete part of the spectrum of infinite-dimensional (nonlinear) completely integrable systems; they are related to canonical momenta and should therefore be adiabatically invariant in some sense (see Ref. 15 for other arguments).

### A. The conjecture

Adiabaticity however, is often referred to in a rather loose way, and we feel it may be useful to give a precise formulation of an adiabatic theorem for solitons, modeled on the two existing ones (we again recall that the theorem in classical mechanics is far from being demonstrated, or even properly stated, in a general setting). First, one should work with time-dependent coefficients only; since a soliton is a local object, it experiences spatial gradients as time-dependent fields in the rest frame; in fact, experiments devised to check adiabatic stability in quantum mechanics always involve spatial gradients, rather than time-dependent potentials.

Also, from a purely mathematical point of view, the difficult thing is to find the right function space to work with, i.e., the right notion of distance between the adiabatically modulated soliton and the actual solution (we again refer to Benjamin and Mac Kean for a similar problem in a different kind of stability). The last issue is about the best way of modulating the soliton; we comment more on this later, but first, state our provisional theorem.

*Conjecture*: Let

$$u_t = K_{\alpha(t)}(u,u',\ldots,u^{(n)}) \tag{38}$$

be a nonlinear equation such that for any fixed $\alpha$, the corresponding equation is integrable with solitons $s(x-ct,c,\alpha)$ traveling at speed $c$; (38) preserves an "energy" $E(u)$.

Put $\alpha(t) = \tilde{\alpha}(t/\tau)$, where $\tilde{\alpha}$ is a nice ($C^1$) function, such that, $\tilde{\alpha}(z)\equiv 0$ for $z\leqslant 0$, $\tilde{\alpha}(z)\equiv 1$ for $z\geqslant 1$, $\tilde{\alpha}$ is monotonous.

Start with the initial data:

$$u(x,0) = s(x,c_0,\alpha(0)) \tag{39}$$

and construct the modulated soliton

$$s\equiv s(x,t,\alpha,c_0,\delta(t)) \tag{40}$$

which preserves energy. Then, there exists $\delta(t)$ such that

$$\lim_{\tau\to\infty} \sup_{0\leqslant t\leqslant\tau^{-1}} \|u(x,t) - s(x,t,c_0,\delta(t))\| = 0, \tag{41}$$

the norm $\|\cdot\|$ referring to some function space.

Thus, up to translation, the modulated soliton and the solution asymptotically ($\tau\to\infty$) coincide. We of course believe that the identities we displayed in part one should perhaps provide a tool for proving such a theorem. We again notice that they hold without any *slowness* in the variation of $\alpha$ being assumed, but one nice thing is that this would connect the stability of solitons with their extremal properties, giving flesh to the appealing picture that when the variation in $\alpha$ is not abrupt, the solution should follow the "bottom of the valley" represented by the modulated soliton. In this respect, the extension to multisoliton states may be misleading, because whereas solitons provide minima of constrained variational problems, multisolitons represent only critical points for such problems.

### B. The link with Whitham's modulation theory

We return to the problem of finding the correct constant of the motion to use, in order to display the modulated soliton. To this end, we use Whitham's approach to the modulation theory (see Ref. 16 for a thorough exposition), which, however, deals primarily with *periodic* waves. It does exhibit adiabatic invariance, namely under the form of "conservation of wave action," and since solitons are limits of periodic waves, there should exist a way to relate both theories; we now proceed to describe a first result along this line.

We start with the KdV and MKdV equations

$$u_t + 6uu_x + u_{xxx} = 0, \tag{42}$$

$$v_t + 6v^2v_x + v_{xxx} = 0. \tag{43}$$

These equations are both Lagrangian, if one considers "potentials" $\varphi$ and $\psi$ such that

$$u = \varphi_x; \qquad v = \psi_x \tag{44}$$

and Lagrangians

$$L[\varphi] = -\tfrac{1}{2}\varphi_t\varphi_x - \varphi_x^3 + \tfrac{1}{2}\varphi_{xx}^2, \tag{45}$$

$$L[\psi] = -\tfrac{1}{2}\psi_t\psi_x - \tfrac{1}{2}\psi_x^4 + \tfrac{1}{2}\psi_{xx}^2. \tag{46}$$

To emphasize the relevant parameters, we write the corresponding solitons in the form

$$s = a\,\mathrm{sech}^2[(a/2)^{1/2}(x - (\omega/k)t)], \tag{47}$$

$$\tilde{s} = a\,\mathrm{sech}[a(x - (\omega/k)t)], \tag{48}$$

$a$ being the amplitude, and $\omega$ and $k$ the frequency and wave number.

Following Whitham's prescription, one writes down the "averaged" Lagrangians

$$\mathcal{L}(\omega,k,a) = k\int_{-\infty}^{+\infty} L\,[s]; \qquad \tilde{\mathcal{L}}(\omega,k,a) \equiv \int_{-\infty}^{-\infty} \tilde{L}\,[\tilde{s}]. \tag{49}$$

It is worth noticing that $\mathcal{L}$ and $\tilde{\mathcal{L}}$ represent in fact the averaged Lagrangians for periodic trains of solitons having $k$ waves per unit length. Whitham's equations are expressed as

$$\mathcal{L}_a = 0; \qquad \frac{\partial \mathcal{L}_\omega}{\partial T} - \frac{\partial \mathcal{L}_k}{\partial X} = 0 \tag{50}$$

and similarly for $\tilde{\mathcal{L}}$. The first equation yields the "dispersion relation," namely $\omega = 2ak$ for the KdV equation and $\omega = ka^2$ for the MKdV equation.

The second equation is the conservation of wave action, which expresses a "dual" adiabatic invariance; in Whitham's words, " in the special case of a wave train [···] responding to changes of the medium in time, we have $\mathcal{L}_\omega = \text{const.}$" There is of course a dual statement if one exchanges space and time, which is of interest for the study of equations with space-dependent coefficients. We now compute $\mathcal{L}_\omega$; since $s = s(x - (\omega/k)t) = \varphi_x$, one has

$$-\tfrac{1}{2}\varphi_t\varphi_x = \frac{\omega}{2k}\varphi_x^2 = \frac{\omega}{2k}s^2, \tag{51}$$

and this is the only term in $L$ to give an $\omega$-dependent contribution

$$\mathcal{L}_\omega = \frac{\partial}{\partial\omega}\left(k\int_{-\infty}^{+\infty} - \frac{1}{2}\varphi_t\varphi_x\right)$$
$$= \frac{\partial}{\partial\omega}\left(\frac{\omega}{2}\int_{-\infty}^{+\infty} s^2\right) = \frac{1}{2}\int_{-\infty}^{+\infty} s^2 \sim a^{3/2}. \tag{52}$$

We thus recover the fact that when coefficients are slowly varying, modulation should keep $\mathcal{L}_\omega = \frac{1}{2}\int s^2$ fixed. The computation is the same for the MKdV equation, since we are concerned only with the $-\frac{1}{2}\psi_t\psi_x$ term, in the evaluation of $\tilde{\mathcal{L}}_\omega$. It is in fact obvious that this derivation can be extended to other equations since only the $t$-dependent part of the Lagrangian contributes. If one looks at Eq. (31) for instance, the above calculation will give the right answer again, namely that in the modulation, one should keep $\mathcal{L}_\omega = \int\varphi_x^2$ fixed.


[1]D. J. Kaup and A. C. Newell, Proc. R. Soc. London Ser. A **361**, 413–446 (1978).
[2]V. I. Karpman and E. Maslov, "Perturbation theory for solitons," Zh. Eksp. Teor. Fiz. **73**, 537–59 (1977) [Sov. Phys. JETP **46**, 281–91 (1977)]; "Structure of tails produced under the action of perturbations on solitons," Zh. Eksp. Teor. Fiz. **75**, 504–17 (1978) [Sov. Phys. JETP **48**, 252–9 (1978).]
[3]J. C. Fernandez, C. Froesche, and G. Reinisch, Phys. Scr. **20**, 545–551 (1979).
[4]K. Ko and H. H. Kuehl, Phys. Rev. Lett. **40**, 233 (1978).
[5]J. Wright, "Soliton production and solutions to perturbed KdV equations," Phys. Rev. A **21**, 335 (1980).
[6]P. D. Lax, Comm. Pure Appl. Math. **21**, 467–490 (1968).
[7]T. B. Benjamin, Proc. R. Soc. London Ser. A **328**, 153–183 (1972).
[8]H. P. Mac Kean, Comm. Pure Appl. Math. **30**, 347–353 (1977).
[9]H. P. Mac Kean and P. Van Moerbecke, Invent. Math. **30**, 217–274 (1975).
[10]H. Flaschka, M. G. Forrest, and D. W. Mac Laughlin, Comm. Pure Appl. Math. **33**, 739–784 (1980).
[11]V. Arnold, *Méthodes Mathématiques de la Mécanique Classique* (Mir, Moscow, 1976).
[12]M. Born and V. Fock, Z. Phys. **51** (1928).
[13]T. Kato, J. Phys. Soc. Jpn. **5** (1950).
[14]P. Lochak, Note aux Comptes rendus de l'Académie des Sciences, Octobre 1982.
[15]P. Lochak, in *Proceedings of the Perugia Symposium on "Wave-Particle Dualism"* (Reidel, Boston, in press).
[16]G. B. Whitham, *Linear and Non Linear Waves* (Wiley–Interscience, New York, 1976).

# Statistical interpretation of the Klein–Gordon equation in terms of the space-time Weyl curvature

E. Santamato[a)]
*Istituto di Fisica Sperimentale dell'Universita di Napoli, 80125 Napoli, Italy*

(Received 27 September 1983; accepted for publication 9 March 1984)

It is shown that a simple statistical interpretation of the Klein–Gordon equation can be obtained, provided a nontrivial vector transplantation law in space-time is assumed. The whole theory is constructed by using only classical probabilistic and dynamical concepts and applies equally well to both charged and neutral particles.

PACS numbers: 03.65.Bz, 03.65.Db

## I. INTRODUCTION

In the special and general relativity theory it is tacitly assumed that the vector lengths are invariant with respect to transplantation in space-time. However, it is well known from differential geometry that the notion of vector transplantation is of affine character and hence logically antecedent to the metric notion of length.[1] Therefore, the transplantation law of vectors can be prescribed quite independently of the metric tensor components.

In this work it is shown that, if one relaxes the assumption that the vector lengths do not change during a transplantation in space-time, a simple self-consistent statistical interpretation of the Klein–Gordon equation for a spinless relativistic quantum particle can be given.[2]

In particular, we assume that a quantum particle undergoes a relativistic random motion in space-time, similar to the one considered by Synge in his book.[3] This random motion can be described entirely using classical probabilistic concepts, i.e., no reference is made to "quantum interferences" between excluding alternatives. Probabilities are always positive definite and combine according to classical Laplace rules. Therefore, the proposed approach leads to a completely classical interpretation of the Klein–Gordon equation. Obviously, the particle wave function $\psi$ can no longer be interpreted as the probability amplitude of some event, since probability amplitudes have no role in the present approach. However, we will prove that the wave function yields all the information needed to describe the statistical behavior of a Gibbs ensemble of particles,[4] which is in accordance with the commonly accepted statistical interpretation of nonrelativistic quantum mechanics.[5] In this way, the statistical interpretation is extended to the relativistic domain.

Finally, the present statistical model permits solving some well-known difficulties connected with the Klein–Gordon equation as a one-particle equation. In particular, it is shown that a consistent statistical interpretation can be obtained in spite of the indefinite sign of the "spatial charge density" associated with the particle. Moreover, the same statistical interpretation also applies to neutral particles, where the notion of charge density becomes meaningless.

We consider the space-time Minkowsky metric of special relativity. Cartesian coordinates are used ($x^0 = ct, x^1 = x, x^2 = y, x^3 = z$; $c$ is the speed of light), so that the nonvanishing metric tensor components are $g_{00} = 1, g_{11} = g_{22} = g_{33} = -1$. The space-time where vectors change during a transplantation is curved, however, since the curvature tensor $R^{i}_{klm}$ is directly related to the change in the components of a vector, under a transplantation around an infinitesimal closed path.

Now, it is commonly accepted today that the metric properties of space-time cannot be prescribed *a priori*; rather, they are determined by the presence of matter. In turn, the motion of matter is affected by the metric properties of space-time. Then, it seems correct to admit that also the vector transplantation law should be determined by the presence of matter, and that the consequent space-time curvature acts on matter as a "guidance field." Empty space-time must have a Minkowsky metric, with constant metric tensor components (in Cartesian coordinates), and vectors in empty space-time must not change under a transplantation.

In the prsence of matter, we expect that the metric properties of space-time are altered. The metric tensor components are no longer constant and the vector lengths change during a transplantation. The Minkowsky metric is retained throughout this paper, however, since experiments show that gravitational effects, which are related to the departure of the actual metric from the empty-space Minkowsky metric, are negligible in elementary particles physics. But we suppose that the effects, related to the presence of a nontrivial transplantation law for vector lengths, may be relevant even if a single particle is considered.

Since the resulting theory is proved to be equivalent to the traditional relativistic quantum mechanics, based on the Klein–Gordon equation, we may say that the vector transplantation law accounts for the "quantum" effects, whereas the metric tensor accounts for the gravitational ones.

## II. THE PARTICLE RANDOM MOTION

In this work we assume that, during a transplantation from $x^i$ to $x^i + dx^i$ ($i = 0,1,2,3$) in space-time, the length $\|A\| = \sqrt{g_{ik} A^i A^k}$ of a (timelike) vector $A^i$ is changed according to the affine law

$$\delta\|A\| = \|A\|\phi_i \, dx^i, \tag{1}$$

where the quantities $\phi_i$ transform as the covariant compo-

[a)] Present temporary address: Department of Physics, University of California, Berkeley, CA 94720.

 0022-2488/84/082477-04$02.50  

nents of a vector and satisfy the integrability conditions

$$\partial_i \phi_k - \partial_k \phi_i = 0. \tag{2}$$

Here, $\partial_k \equiv \partial/\partial x^k$, and Einstein's summation convention for repeated indices is used.

Condition (2) is needed in order that the length $\|A\|$ of a vector is a well-determined function of its application point.

A transplantation law of the form (1) has been extensively used by Weyl in his unified theory of electromagnetism in the framework of general relativity.[6] For this reason, the space, supplemented by the transplantation law (1), is named the Weyl space with gauge field $\phi_i$. The gauge field components $\phi_i$ can be prescribed independently of the metric tensor components $g_{ik}$.

As previously noted, the space, endowed with the Weyl geometry, has a nonvanishing curvature tensor $R^{i}_{klm}$. The complete expression of $R^{i}_{klm}$ is quite intricate; however, in the present work, we need only the scalar curvature $R$, which is defined according to

$$R_{ik} = R^{l}_{ilk}, \quad R = g^{ik} R_{ik}. \tag{3}$$

A straightforward calculation shows that

$$R = (n-1)(n-2)\phi^{k}\phi_k - 2(n-1)\partial^k \phi_k, \tag{4}$$

a formula known as the Weyl formula.[7]

According to the statistical interpretation of quantum mechanics, the theory does not describe the motion of an individual particle, but only the statistical behavior of a relativistic Gibbs ensemble of particles.

Following Hakim,[8] we may think of a relativistic ensemble as formed by (i) a set of curves in space-time and (ii) an invariant measure given on this set. The set of curves yields the set of all possible trajectories in space-time, i.e., the path-space of the particle, while the invariant measure gives the probability that the particle moves along a given path.

A suitable path-space for the particle can be easily constructed as follows: Let $L(x,\dot{x})$ be a given Lagrangian of class $C_2$ in both arguments, which is homogeneous in the first degree in the variables $\dot{x}^i$ $(i = 0,1,2,3)$. The homogeneity of $L$ is required in order to have a theory which is independent of the choice of the parameter along the particle trajectory in space-time. This is an essential requirement for any workable relativistic theory.

Now, it is well known from analytical mechanics that starting from the Lagrangian $L(x,\dot{x})$, we can construct a one-parameter family of hypersurfaces $S(x) = \Sigma$ ($\Sigma$ being the parameter of the family) and a $(n-1)$-parameter congruence $K$ of curves in space-time, having the following properties.

(a) The quantities $x^i$ may be regarded as the components of the tangent vector to any curve $C$ of the congruence $K$.

(b) Each curve $C$ in $K$ is an *extremal*, i.e., along this curve the Euler–Lagrange equations associated to $L(x,\dot{x})$ are obeyed.[9]

(c) The family of hypersurfaces $S(x) = \Sigma$ is *geodesically equidistant* with respect to the Langrangian $L(x,\dot{x})$. By this we mean that, if a given curve $C$ in $K$ intersects any two hypersurfaces $S(x) = \Sigma_1$ and $S(x) = \Sigma_2$ of the family in the points $P_1$ and $P_2$, respectively, then the action integral $\int_{P_1 C}^{P_2} L(x,\dot{x}) d\tau = \Sigma_2 - \Sigma_1$, irrespective of the chosen extremal $C$.

In the sequel, it will be supposed that the family of geodesically equidistant hypersurfaces covers a region $G$ of space-time simply, and that $S(x)$ is of class $C^2$. Our considerations will be restricted to this region $G$.

The geometrical entity formed by the family of hypersurfaces $S(x) = \Sigma$ together with the congruence of extremals intersecting it is known as the Caratheódory *complete figure*, associated to the Langrangian $L(x,\dot{x})$.[10]

It is a well-known result of analytical mechanics that the family $S(x) = \Sigma$ is not completely arbitrary; rather, the function $S(x)$ must obey a first-order partial differential equation, the so-called Hamilton–Jacobi equation. We will deal with this equation later. For now, we remember only that the Caratheódory complete figure, and hence a congruence of extremals, can be uniquely associated to any solution of class $C^2$ of the Hamilton–Jacobi equation.

The complete figure is the fundamental geometric object of the present theory. Indeed, we now postulate that the path-space of a relativistic particle is given by the congruence of curves belonging to a family of hypersurfaces geodesically equidistant with respect to the homogeneous Lagrangian

$$L(x,\dot{x}) = \phi(x)\|\dot{x}\| + \sigma\lambda A_i(x)\dot{x}^i, \tag{5}$$

where

$$\phi(x) = \sqrt{1 + \gamma\lambda^2 R(x)}, \tag{6}$$

$R(x)$ is the scalar Weyl curvature (4), and the quantities $A_i(x)$ are the four-potential covariant components of an external electromagnetic field. In Eqs. (5) and (6), $\lambda = \hbar/mc$ is the reduced Compton length of the particle, $\sigma = e/\hbar c$ ($e$ is the charge in e.s.u.), and $\gamma$ is a numerical factor given by

$$\gamma = (n-2)/[4(n-1)] = \tfrac{1}{6}. \tag{7}$$

From Eqs. (5) and (6) we see that the Weyl curvature $R$ acts on the particle as an external "mesic" (i.e., scalar) field.

In the sequel, we will suppose that in the region $G$ of space-time $(1 + \gamma\lambda^2 R(x)) > 0$, so that $\phi(x)$ is real valued and positive. Using standard methods of analytical mechanics, we find the Hamilton–Jacobi equation corresponding to the Lagrangian (5) as

$$\lambda^2(\partial_i S - \sigma A_i)(\partial^i S - \sigma A^i) = 1 + \gamma\lambda^2 R. \tag{8}$$

In the same way, we can find the congruence of extremals belonging to the family of hypersurfaces $S(x) = \Sigma$ by solving the system of first-order differential equations

$$\dot{x}^i \equiv \frac{dx^i}{d\tau} = \lambda(\partial^i S - \sigma A^i), \tag{9}$$

where we used the special parameter $\tau$, which is related to the arc length $ds$ by

$$d\tau = \phi^{-1} ds = \phi^{-1}\sqrt{g_{ik}\, dx^i\, dx^k}. \tag{10}$$

This choice of the parameter turns out to be very convenient, and it will be used throughout this work. Finally, the Euler–Lagrange equations of motion read

$$\ddot{x}^i = \tfrac{1}{2}\gamma\lambda^2 \partial^i R + \sigma\dot{x}^k F^i_k, \tag{11}$$

where $F^i_k = \partial^i A_k - \partial_k A^i$ is the electromagnetic field tensor. From Eq. (11) we see that the particle is affected by a four-force which is proportional to the gradient of the Weyl curvature.

We now define a suitable measure on the congruence $K$ given by Eq. (9). Following Hakim,[8] we consider an arbitrary three-dimensional hypersurface $S_0$ intersecting all members of the congruence (without tangencies) in the region $G$. This $S_0$ may or may not coincide with a member of the geodesically equidistant family $S(x) = \Sigma$. Then, let $A$ be an open (Borel) subset of $S_0$ and let $C_A$ be the set of curves of the congruence intersecting $A$. Then $C_A$ generates a $\sigma$-field in the path-space and therefore, we can define a measure on this $\sigma$-field according to

$$\text{Meas}(C_A) = \int_{A \subset S_0} j^i \, d\Sigma_i, \tag{12}$$

where $d\Sigma_i$ is the "element of area" normal to the hypersurface $S_0$.

For consistency we require that the measure is independent of the parameter $\tau$ used along the curves of the congruence. This implies that the current components $j^i$ must be parameter independent. Moreover, the measure must be independent of the chosen hypersurface $S_0$, which implies the continuity equation

$$\partial_i j^i = 0. \tag{13}$$

Finally, the current $j^i$ must be parallel to the tangent vector along the particle trajectory, viz.,

$$j^i = \rho \dot{x}^i, \tag{14}$$

with $\rho > 0$. We note that since $j^i$ is parameter independent, the quantity $\rho$ must be dependent on the choice of the parameter along the particle trajectory. Then, in order to have a well-determined value of $\rho$, we will suppose that in Eq. (14) the parameter $\tau$, as defined in Eq. (10), is used.

Inserting Eq. (9) into Eqs. (13) and (14), we obtain

$$\partial_i [\rho(\partial^i S - \sigma A^i)] = 0. \tag{15}$$

We will refer to the scalar quantity $\rho(x)$ as the absolute density of the ensemble, representing the random motion of the particle. We stress that once the congruence of curves (9) is known, the quantity $\rho(x)$ is all that is needed to determine the underlying measure (12).

Up to this point the gauge vector $\phi_i$ in the transplantation law (1) is still undetermined and could be arbitrarily prescribed. However, as we previously noted, the vector $\phi_i$ must be related in some way to the presence of matter in space-time. This motivates our postulate that the gauge vector is related to the absolute density $\rho(x)$ by

$$\phi_i = -[1/(n-2)]\partial_i(\ln \rho). \tag{16}$$

We observe that $\phi_i$ satisfies the integrability conditions (2) and vanishes in empty space ($\rho \equiv 0$).

When relation (16) is inserted into the Weyl formula (4), the curvature $R$ may be expressed as a function of $\rho$ and its first and second partial derivatives, viz.,

$$R = \frac{1}{2\gamma}\left(\frac{\partial^i \partial_i \rho}{\rho} - \frac{\partial^i \rho \, \partial_i \rho}{2\rho^2}\right). \tag{17}$$

Therefore, the Hamilton–Jacobi equation (8) and the continuity equation (15) become a set of two coupled second-order differential equations for the unknown scalar functions $\rho(x)$ and $S(x)$.

Any solution $\{\rho, S\}$ of Eqs. (8) and (15) consistent with our postulate (16), with $\rho > 0$, yields a possible motion of the ensemble, i.e., a possible "state" of the particle. In fact, for any given solution $\{\rho, S\}$, the path-space of the particle is found by solving Eq. (9) and the underlying measure is found from Eq. (12) with $j^i = \lambda\rho \, \partial^i S$. Finally, the actual gauge vector of space-time is obtained from Eq. (16).

We notice that in view of property (b) of the congruence of particle trajectories, the particle motion obeys the Euler–Lagrange equation (11) with probability one.

The connection between the present theory and traditional relativistic quantum mechanics is easily found. Indeed, a direct calculation shows that the complex quantity $\psi$, defined as[11]

$$\psi = \sqrt{\rho}\, e^{iS}, \tag{18}$$

where $\{\rho, S\}$ is a solution of the couple of Eqs. (8) and (15), obeys the Klein–Gordon equation, viz.,

$$\lambda^2(\partial_i - \sigma A_i)(\partial^i - \sigma A^i)\psi + \psi = 0. \tag{19}$$

Conversely, if $\psi$ is a solution of Eq. (19), its square modulus $\rho = |\psi|^2$ and phase $S = \arg \psi$ solve the set of Eqs. (8) and (15). Notice that this correspondence is not one-to-one, since we may add to $S$ an arbitrary real-valued constant without affecting Eqs. (8) and (15). Then, two solutions $\psi$ and $\psi' = \psi e^{i\alpha}$ ($\alpha$ real) of Eq. (19), differing by a constant phase factor, correspond to the same random motion of the particle, and, therefore, they are physically equivalent. This is a well-known result of traditional quantum mechanics.

Relation (18) leads immediately to the required physical interpretation of the Klein–Gordon equation (19). The wave function constant-phase hypersurfaces form a family of hypersurfaces, which are geodesically equidistant with respect to the Langrangian (5). The congruence (9), belonging to this family, yields the set of all possible particle trajectories in space-time. The wave function square modulus yields the absolute density of the ensemble, representing the random motion of the particle. Finally, it should be stressed that, in view of Eqs. (9) and (14) we have $j^i = \lambda\rho(\partial^i S - \sigma A^i)$ and, therefore, $\rho$ is a parameter-invariant quantity. This is to be expected, indeed, since the actual gauge vector (16) and the actual transplantation law (1) cannot depend on the arbitrary choice of the parameter along the particle trajectories. This shows how useful an appropriate choice of the parameter can be.

## III. FINAL REMARKS

We have seen that the statistical behavior of the ensemble is determined by a congruence of curves in space-time, intersecting the family of hypersurfaces of constant wave function phase and by a measure defined on this congruence. However, we notice that the orientation of the curves forming the congruence is left undetermined by the theory, and, hence, it can be arbitrarily prescribed. In other words, if $\dot{x}^i = v^i(x)$ represents a member of the congruence, then $\dot{x}^i = -v^i(x)$ represents the same member.[12] Thus, the measure (12) is proportional to the probability that a curve of the congruence $K$ intersects the set $A$ on $S_0$, irrespective of its orientation, and therefore we can always choose the orientation of the current $j^i$ and the surface element $d\Sigma_i$ so that the measure (12) is positive definite.

The physical interpretation of the theory depends, however, strongly on the actual orientation of particle trajectories. This is due to the fact that physical observations are related to spacelike three-cuts of the particle trajectory in space-time, where a conventional ($t$-increasing) notion of evolution is assumed.[13] According to Feynman's interpretation,[14] we interpret the particles which cut the chosen spacelike hypersurface (i.e., the "present") pointing into the "future" as "particles" and the particles pointing into the "past" as "antiparticles." From Eq. (9), we see that if $\{\rho,S\}$ represents the random motion of the particle, say, with charge $e$, then $\{\rho,-S\}$ represents the motion of an antiparticle with opposite charge $-e$. Notice that the path-space as well as the underlying measure of the particle in its state $\{\rho,S\} \leftrightarrow \psi$ is the same as the corresponding antiparticle in the conjugate state $\{\rho,S\} \leftrightarrow \psi^*$. In other words, since no natural notion of evolution can be given in a completely covariant relativistic theory,[8] we can say that the particle and the antiparticle in conjugate states undergo the same random motion in space-time.[15]

We note that the present approach leads in a natural way to Feynman's interpretation of antiparticles. In particular, the quantity $Q = \int_A j_0\, dV$ extended to some volume $A$ of the three-dimensional physical space at the given time coordinate $t_0$ is proportional to the difference between the probability of observing the particle inside the volume $A$ at time $t_0$ as a true "particle," and the probability of observing it as an "antiparticle." The probability difference occurs [and hence $Q$ is not positive definite] since, in evaluating $Q$, it is assumed that the three-dimensional hyperplane element $d\Sigma \equiv dV$ has its normal oriented toward the "future." Thus, a conventional time-order on the particle trajectories has been introduced. Obviously, the total probability of finding the particle in either its "particle" or "antiparticle" state in the volume $A$ at time $t_0$ is the sum of the overmentioned probabilities. It is worth noting that the same statistical considerations apply both to charged and neutral particles.

In principle, the measure (12) may refer to a congruence of curves in space-time which may be either spacelike or timelike. However, causality requires that only timelike trajectories must be retained. Then, Eqs. (8) and (9) imply that $(1+\gamma\lambda^2 R) > 0$, and the Lagrangian (5) is positive along the particle motion. Throughout this work we have supposed that this is the case in the considered space-time region $G$.

We stress, however, that solutions of the Klein–Gordon equation (19) exist, which lead to a spacelike congruence of curves. In this case, although the geometric and probabilistic backgrounds of the theory may be well posed, the physical interpretation of the resulting spacelike stream of events is unclear. When the family of geodesically equidistant hypersurfaces $S(x) = \text{const}$ fails to cover the region $G$ of space-time simply, the present theory must be appropriately modified. This may happen, for example, in the presence of very strong external fields. The problem is of some interest, since it would correspond to "virtual" pair creation–annihilation processes, which are responsible, for instance, for "vacuum polarization." Although the present approach is not directly applicable to these pathological cases, it may be argued that after an appropriate definition of the particle path-space and of the underlying measure, a statistical interpretation of virtual processes in terms of a one-particle theory should be possible.

The theory may be formulated in curvilinear coordinates as long as all partial derivatives $\partial_i$ are to be replaced by covariant derivatives. Also, the generalization to the case of an external gravitational field is straightforward, provided we perform the substitution $R \to R - \hat{R}$, i.e., the curvature change from the Riemannian curvature $\hat{R}$, due to the actual transplantation law (1) for vector lengths. As expected, the covariant Klein–Gordon equation is obtained, with the Laplace operator replaced by the Laplace–Beltrami operator.

A possible connection between "quantum" effects and transplantation law for vector lengths has been demonstrated in this work. It is hoped that the proposed interpretation of the Klein–Gordon equation, where only classical probabilistic and geometric concepts are needed, will bring new insight to relativistic quantum mechanics.

## ACKNOWLEDGMENTS

We thank Dr. Carlos Perelman for enlightening conversations. This work was financially supported by Ministero Pubblica Istruzione, Italy.

[1] See, for example, L. Brillouin, *Tensors in Mechanics and Elasticity* (Academic, NY, 1964), Chap. V.

[2] A similar interpretation of the Schrödinger equation in terms of Weyl spatial curvature has been proposed by E. Santamato, Phys. Rev. D **29**, 216 (1984).

[3] J. L. Synge, *Relativity: The General Theory* (North-Holland, Amsterdam, 1960), Chap. IV.

[4] The Gibbs ensemble must not be confused with a system formed by $N$ noninteracting identical particles.

[5] See, for example, the review article on this argument by L. E. Ballantine, Rev. Mod. Phys. **42**, 358 (1970).

[6] H. Weyl, *Gravitation und Elektrizität* (Sitzber, Berlin, 1918), pp. 465–480. Reprinted in Lorentz, Einstein, and Minkowsky, *Das Relativitatprinzip* (Teubner, Leipzig, 1918).

[7] We use the standard notation $\phi^i = g^{ik}\phi_k$ and $\partial^i = g^{ik}\partial_k$. For a simple derivation of the Weyl formula see, for example, R. Adler, M. Bazin, and M. Schiffer, *Introduction to General Relativity* (McGraw-Hill, New York, 1965), p. 408.

[8] R. Hakim, J. Math. Phys. **8**, 1315 (1967).

[9] Due to the homogeneity of $L(x,\dot{x})$, the Euler–Lagrange equations are not independent.

[10] The complete figure associated to a homogeneous Lagrangian is extensively treated, for instance, in H. Rund, *The Hamilton–Jacobi Theory in the Calculus of Variations* (Van Nostrand, London, 1966), Chap. 3.

[11] This position is quite standard and has been used in both the "hydrodynamical" [E. Madelung, Z. Phys. **40**, 322 (1926); D. Bohm, Phys. Rev. **85**, 166 (1952)] and the "stochastic" [E. Nelson, Phys. Rev. **150**, 1079 (1966)] approach to nonrelativistic quantum mechanics.

[12] $p^i(x)$ is given by the right-hand side of Eq. (9).

[13] For instance, a conventional notion of evolution is given by the oriented family of hyperplanes $x^0 = ct = \text{const}$. For a more detailed discussion of this point, see Ref. 8.

[14] R. Feynman, Phys. Rev. **76**, 749 (1949).

[15] Obviously, when some notion of evolution is given the two motions are different, since particles and antiparticles move in opposite directions. However, covariance is lost.

# On the RHS description of resonances and virtual states

M. Gadella[a)]
*Facultad de Ciencias, Avenida de los Castros, s/n, Santander, Spain*

(Received 31 May 1983; accepted for publication 23 December 1983)

We discuss the role of the Møller wave operators in a rigged Hilbert space (RHS) model for resonances. We also construct a new RHS of analytic functions valid for the description of resonances and virtual states. These become generalized state vectors in a general physically realizable representation.

PACS numbers: 03.65.Db, 03.65.Nk

## I. INTRODUCTION

This paper is the third of a series trying to provide a mathematical rigorization for a model for resonances and virtual states in the context of the rigged Hilbert space (RHS) formulation of quantum mechanics.[1-3] It is divided into two sections. In the first one, we shall study the role of the Møller wave operators in this theory and in particular the precise meaning of the formulas $\Omega^{\pm}|E\rangle = |E^{\pm}\rangle$ appeared in Refs. 1–3 which has not been given until now. In the second section, we shall construct a new RHS of analytic functions and see how we can apply the previous formulation to it.

We start our study by assuming the existence of a free dynamics characterized by the Hamiltonian $H_0 = (-i\nabla)^2$ and a perturbed dynamics. For simplicity, we also assume that the perturbation is given by a potential $V$ which is zero outside a bounded region of the space, so that the dynamics can be characterized by perturbed Hamiltonian $H = H_0 + V$. Also, $H$ does not have a singular spectrum. We can decompose the Hilbert space of states $\mathscr{H}$ as

$$\mathscr{H} = \mathscr{H}_p \oplus \mathscr{H}_{ac}, \tag{1}$$

where $\mathscr{H}_p$ is the closure of the subspace spanned by all the eigenvalues of $H$ and $\mathscr{H}_{ac}$ its orthogonal subspace. The spectrum of the restriction of $H$ to $\mathscr{H}_{ac}$ precisely coincides with $\sigma_{ac}(H)$, the absolutely continuous spectrum of $H$. Hereafter, we shall call $H_1$ the restriction of $H$ to $\mathscr{H}_{ac}$.

Now, assume that the Møller wave operators exists and are complete.[4] The completeness condition requires that $\Omega^{+}\mathscr{H} = \Omega^{-}\mathscr{H} = \mathscr{H}_{ac}$.[5] Under these circumstances, $H_0$ is unitarily equivalent to $H_1$ which implies that $\sigma_{ac}(H_1) = \sigma(H_0) = R^{+}$.

The spectral theorem[6] says that there exists a unitary mapping $U$ and a Hilbert space $\mathscr{G}$ of functions such that $U\mathscr{H} = \mathscr{G}$ and $UH_0U^{-1}$ is the multiplication operator on $\mathscr{G}$. Here $\mathscr{G}$ is $L^2(R^{+}, dE)$. Since $H_1$ is unitarily equivalent to $H_0$, there must exist another unitary operator $V$ such that $V\mathscr{H} = L^2(R^{+}, dE)$ and $VH_1V^{-1}$ will be the multiplication operator on $V\mathscr{H}$. We recall that the multiplication operator $B$ on $L^2(R^{+}, dE)$ acts on every $f(E) \in \mathscr{D}(B)$ as $Bf(E) = Ef(E)$. Here $\mathscr{D}(B) = \{f(E) \in L^2(R^{+}, dE)$ such that $Ef(E) \in L^2(R^{+}, dE)\}$.

On the other hand, $H_1 = \Omega^{\pm} H_0 (\Omega^{\pm})^{-1}$ (Ref. 4). The equivalence between $H_0$ and $H_1$ can be given either by $\Omega^{+}$ or by $\Omega^{-}$. Both are unitary mappings from $\mathscr{H}$ onto $\mathscr{H}_{ac}$ because they are partial isometries[4] and because of the completeness condition. The operators $UH_0U^{-1}$ and $VH_1V^{-1}$ are equal on the space $L^2(R^{+}, dE)$. Therefore

$$UH_0U^{-1} = V\Omega^{\pm} H_0 (\Omega^{\pm})^{-1} V^{-1}. \tag{2}$$

Note also that $U$ and $V$ are not unique. In particular, $e^{i\alpha}U$ and $e^{i\beta}V$ play the same role, $\alpha$ and $\beta$ being any pair of real numbers. If we fix $U$ as the operator which diagonalizes $H_0$, i.e., $UH_0U^{-1}$ is the multiplication operator on $L^2(R^{+}, dE)$, we can define $V_{\pm} = U(\Omega^{\pm})^{-1}: \mathscr{H}_{ac} \to L^2(R^{+}, dE)$. Obviously $V_{\pm}$ diagonalizes $H_1$, are unitary, and $U = V_{\pm}\Omega^{\pm}$.

Throughout the next section, we shall extend the Møller wave operators to the dual spaces of the nuclear spaces of Hardy class functions defined in Refs. 2 and 3. In order to help the reader, we shall make our developments in full detail.

## II. ON THE EXTENSION OF THE MØLLER WAVE OPERATORS TO THE DUAL SPACES ON THE RHS OF HARDY CLASS FUNCTIONS

Consider the spaces $\Phi_{\pm}$ defined as in Ref. 2. We recall that $\Phi_{\pm}$ are the restrictions to $R^{+}$ of all the functions in $\Psi_{\pm}$ $\mathscr{S}(R) \cap H^2_{\pm}$, where (1) $\mathscr{S}(R)$ is the Schwartz space, i.e., the space of all infinitely differentiable functions on $R$ such that they and all their derivatives vanish at infinity faster than the inverse of any polynomial; and (2) $H^2_{+}$ and $H^2_{-}$ are, respectively, the functions of Hardy class on the upper and lower half of the complex plane $\mathbb{C}$.[7]

The $\Phi_{\pm} \in L^2(R^{+}, dE)$ are endowed with a nuclear topology which is finer than that inherited from $L^2(R^{+}, dE)$.

Consider $\Gamma_{\pm} = U^{-1}\Phi_{\mp} \subset \mathscr{H}$ and $\Gamma^{\pm} = V_{\pm}^{-1}\Phi_{\mp} \subset \mathscr{H}_{ac}$. Since $U = V_{\pm}\Omega^{\pm}$, it is obvious that $\Omega^{\pm}\Gamma_{\pm} = \Gamma^{\pm}$. Note that these are equalities between linear bijections.

*Proposition 1*: The $\Gamma_{\pm}$ are invariant under the action of $H_0$ and the $\Gamma^{\pm}$ are invariant under $H_1$.

*Proof*: First, note that

$$\Gamma^{\pm} = V_{\pm}^{-1}\Phi_{\mp} = \Omega^{\pm}U^{-1}\Phi_{\mp} = \Omega^{\pm}\Gamma_{\pm}. \tag{3}$$

The invariance of $\Gamma_{\pm}$ under $H_0$ can be proved as follows:

$$H_0\Gamma_{\pm} = H_0U^{-1}\Phi_{\mp} = U^{-1}[UH_0U^{-1}]\Phi_{\mp}, \tag{4}$$

$UH_0U^{-1}$ is the multiplication operator on $\Phi_{\pm}$. Then, taking into account that $E\Phi_{\pm} \subset \Phi^2_{\pm}$,

$$(4) = U^{-1}E\Phi_{\mp} \subset U^{-1}\Phi_{\mp} = \Gamma_{\pm}. \tag{5}$$

The second claim is also easily proven:

[a)] Present address: Facultad de Ciencias, Prado de la Magdalena, s/n. Valladolid, Spain.

 0022-2488/84/082481-05$02.50  

$$H_1\Gamma^{\pm} = H_1\Omega^{\pm}\Gamma_{\pm} = \Omega^{\pm}H_0\Gamma_{\pm} \subset \Omega^{\pm}\Gamma_{\pm} = \Gamma^{\pm}, \tag{6}$$

where the first identity is a consequence of (3) and the second one of $(\Omega^{\pm})^{-1}H_1\Omega^{\pm} = H_0$.

The next step consists of endowing $\Gamma_{\pm}$ and $\Gamma^{\pm}$ with respective topologies with the following conditions: $H_0$ must be continuous on $\Gamma_{\pm}$, $H_1$ on $\Gamma^{\pm}$, and $\Omega^{\pm}$ from $\Gamma_{\pm}$ onto $\Gamma^{\pm}$ with respect to them. We recall that we already have topologies on $\Phi_{\pm}$ defined as follows: on $\Psi_{\pm} = \mathscr{S}(R)\cap H^2_{\pm}$ we have topologies which are the restrictions to $\Psi_{\pm}$ of the topology on $\mathscr{S}(R)^2$. The Heaviside funtion $\theta \equiv \{\theta(E) = 0$ if $E \in R^-$: $\theta(E) = 1$ if $E \in R^+\}$ provides a bijection from $\Psi_{\pm}$ onto $\Phi_{\pm}$. Then the $\Phi_{\pm}$ have the final topologies given by $\Psi_{\pm}$ and $\theta$, respectively.[8] It is easily seen that these topologies confer on $\Phi_{\pm}$ the character of complete nuclear metrizable spaces. The proof uses the fact that if $\mathscr{V}$ is a locally convex nuclear space and if $h$ is a linear mapping from $\mathscr{V}$ into another vector space $\mathscr{W}$, then if we endow $\mathscr{W}$ with the final topology with respect to $\mathscr{V}$ and $h$, $\mathscr{W}$ is a nuclear locally convex space.[8,9] This is the result we need for the definition of the topologies on $\Gamma_{\pm}$ and $\Gamma^{\pm}$.

*Definition 1*: The $\Gamma_{\pm}$ have the final topologies with respect to $\Phi_{\mp}$ and $U^{-1}$. The $\Gamma^{\pm}$ have the final topologies with respect to $\Phi_{\mp}$ and $V_{\pm}^{-1}$.

Note that $U^{-1}$ and $V_{\pm}^{-1}$ are bijections so that $\Gamma_{\pm}$ and $\Gamma^{\pm}$ are complete nuclear metrizable spaces.

*Proposition 2*: The identity mappings from $\Gamma_{\pm}$ into $\mathscr{H}$ and $\Gamma^{\pm}$ into $\mathscr{H}_{ac}$ are continuous and nuclear. Also $H_0$ is continuous on $\Gamma_{\pm}$ and $H_1$ on $\Gamma^{\pm}$.

*Proof*: As a consequence of the definition of the topologies on $\Gamma_{\pm}$ and $\Gamma^{\pm}$ and the closed graph theorem,[10] $U$ and $V^{\pm}$ are continuous bijections. Denote by $i$ the injection from $\Phi_-$ into $L^2(R^+, dE)$ and $j$ that one from $\Gamma_+$ into $\mathscr{H}$. We already know that $i$ is continuous and nuclear:[11]

$$\begin{array}{ccc} \Phi_- & \xrightarrow{\ i\ } & L^2(R^+, dE) \\ \Big\downarrow{\scriptstyle U^{-1}} & & \Big\downarrow{\scriptstyle U^{-1}} \\ \Gamma_+ & \xrightarrow{\ i\ } & \mathscr{H} \end{array} \tag{7}$$

$j = U^{-1}iU$. Since $U$, $i$, and $U^{-1}$ are continuous, $j$ is also continuous. Since $U$ and $U^{-1}$ are continuous and $i$ nuclear, $j$ is also nuclear. The same arguments work for the other identity mappings.

Let $E$ be the multiplication operator on $L^2(R^+, dE)$. The restrictions of $E$ to $\Phi_{\pm}$ are continuous when $\Phi_{\pm}$ have their own topologies. We can write $H_1$ and $H_0$ as follows:

$$\begin{aligned} H_1 &= V_{\pm}^{-1}V_{\pm}, \\ H_0 &= U^{-1}\hat{E}U. \end{aligned} \tag{8}$$

The continuity of $U$, $V_{\pm}$, and their inverses give the desired result. Please, note that $U$ and $V_{\pm}$ are here looked as bijections between locally convex spaces and not as isometries between Hilbert spaces.

The above construction yields four new triplets:

$$\Gamma_{\pm} \subset \mathscr{H} \subset (\Gamma_{\pm})^{\times}, \quad \Gamma^{\pm} \subset \mathscr{H}_{ac} \subset (\Gamma^{\pm})^{\times}, \tag{9}$$

where the elements in $(\Gamma_{\pm})^{\times}$ and $(\Gamma^{\pm})^{\times}$ are the continuous antilinear functionals on $\Gamma_{\pm}$ and $\Gamma^{\pm}$, respectively. For that reason, we prefer calling these spaces the antiduals. For instance, $(\Gamma_+)^{\times}$ will be the antidual of $\Gamma_+$, etc. Now we are in the position of defining the action of the operators $\Omega^{\pm}$ on the antiduals.

*Definition 2*: Let $\xi_{\pm} \in (\Gamma_{\pm})^{\times}$. The action of $\Omega^{\pm}$ on $\xi_{\pm}$ is given by

$$\langle \Omega^{\pm} f_{\pm} | \Omega^{\pm} \xi_{\pm} \rangle = \langle f_{\pm} | \xi_{\pm} \rangle, \quad \forall f_{\pm} \in \Gamma_{\pm}. \tag{10}$$

*Proposition 3*: The $\Omega^{\pm}$ as defined above on $(\Gamma_{\pm})^{\times}$ are bicontinuous linear bijections onto $(\Gamma^{\pm})^{\times}$. The $\Omega^{\pm}$ are also bicontinuous bijections from $\Gamma_{\pm}$ onto $\Gamma^{\pm}$.

*Proof*: Consider the relation $\Gamma^{\pm} = \Omega^{\pm}\Gamma_{\pm}$. We endow $\Gamma^+$ and $\Gamma^-$ with the final topologies given by $\Gamma_+$ and $\Omega^+$ and $\Gamma_-$ and $\Omega^-$, respectively. The theorem on transitivity of final topologies says that the new topologies on $\Gamma^{\pm}$ are just the same as the old ones. Therefore, the $\Omega^{\pm}$ are continuous. The continuity of $(\Omega^{\pm})^{-1}$ follows from the closed graph theorem.[10]

By definition the transpose of $\Omega^{\pm}$, $(\Omega^{\pm})'$ is given by

$$\langle \Omega^{\pm} f_{\pm} | \xi_{\pm} \rangle = \langle f_{\pm} | (\Omega^{\pm})' \xi_{\pm} \rangle. \tag{11}$$

On the other hand,

$$\langle f_{\pm} | \Omega^{\pm} \xi_{\pm} \rangle = \langle (\Omega^{\pm})^{-1} f_{\pm} | \xi_{\pm} \rangle = \langle f_{\pm} | [(\Omega^{\pm})^{-1}]' \xi_{\pm} \rangle. \tag{12}$$

Thus $\Omega^{\pm} = [(\Omega^{\pm})^{-1}]'$ on $(\Gamma_{\pm})^{\times}$ $(\Omega^{\pm})^{-1}$ is bijective and bicontinuous from $\Gamma^{\pm}$ onto $\Gamma_{\pm}$. In Ref. 12, we can find that its transposes have the same property, provided we have endowed $(\Gamma_{\pm})^{\times}$ and $(\Gamma^{\pm})^{\times}$ with the corresponding strong topologies.[13]

*Definition 3*: The extension of $U$ and $V_{\pm}$ to the antiduals can be defined analogously:

$$\begin{aligned} &\langle Uf_{\pm} | U\xi_{\pm} \rangle = \langle f_{\pm} | \xi_{\pm} \rangle, \quad \forall f_{\pm} \in \Gamma_{\pm}, \quad \forall \xi_{\pm} \in (\Gamma_{\pm})^{\times}, \\ &\langle V_{\pm} f^{\pm} | V_{\pm} \xi^{\pm} \rangle = \langle f^{\pm} | \xi^{\pm} \rangle, \quad \forall f^{\pm} \in \Gamma^{\pm}, \quad \forall \xi^{\pm} \in (\Gamma^{\pm}). \end{aligned} \tag{13}$$

Note that $Uf_{\pm}$ and $V_{\pm}f^{\pm}$ belong to $\Phi_{\mp}$ and $U\xi_{\pm}$ and $V_{\pm}\xi^{\pm}$ belong to $(\Phi_{\mp})^{\times}$.

The corresponding extensions of $U^{-1}$ and $V_{\pm}^{-1}$ are also well defined with the aid of formulas similar to (13). The continuity properties of these extensions can be summarized as follows.

*Proposition 4*: We have that $U$ is a bicontinuous linear mapping from $(\Gamma_{\pm})^{\times}$ onto $(\Phi_{\mp})^{\times}$, and $V_{\pm}$ are bicontinuous linear mappings from $(\Gamma^{\pm})^{\times}$ onto $(\Phi_{\mp})^{\times}$. On the antiduals, it is also true that $U = V_{\pm}\Omega^{\pm}$. The inverses have the same properties.

*Proof*: It is as in Proposition 3.

Now, let $|\eta_{\pm}(E)\rangle$ be an antilinear functional on $\Phi_{\pm}$ such that $f^*_{\pm}(E) = \langle f_{\pm}(E') | \eta_{\pm}(E) \rangle$, $E \in R^+$, for all $f_{\pm}(E') \in \Phi_{\pm}$. It is well known that (1) $|\eta_{\pm}(E)\rangle \in (\Phi_{\pm})^{\times}$; and (2) $\hat{E}|\eta_{\pm}(E)\rangle = E|\eta_{\pm}(E)\rangle$ $(E \in R^+)$.

If $E\in R^-$, we can also find a continuous antilinear functional on $\Phi_\pm$. We know that any function in $\Phi_\pm$ can uniquely be extended into a function in $\Psi_\pm$. If $f_\pm(E')\in\Phi_\pm$ and we denote by the same symbol its unique extension into $\Psi_\pm$, we define $|\eta_\pm(E)\rangle$ $(E\in R^-)$ as $f^*_\pm(E)=\langle f_\pm(E')|\eta_\pm(E)\rangle$. We can also prove that (1) $|\eta_\pm(E)\rangle\in(\Phi_\pm)^\times$; and (2) $\hat{E}|\eta_\pm(E)\rangle=E|\eta_\pm(E)\rangle$ $(E\in R^-)$.

*Remark*: The relation $f^*_\pm(E)=\langle f_\pm(E')|\eta_\pm(E)\rangle$ is thus *everywhere* defined on $R$, since all the functions in $\Phi_\pm$ as well as their extensions into elements of $\Psi_\pm$ are continuous.

*Proposition 5*: $|\eta_+(E)\rangle=|\eta_+(E)\rangle\in\Phi^\times_+\cap\Phi^\times_-$.

*Proof*: Let $f\in\Psi_+\oplus\Psi_-$. For any point $E\in R$, there exists an antilinear functional $|\xi(E)\rangle\in(\Psi_+\oplus\Psi_-)^\times$, such that $[f(E')|\xi(E)\rangle=f^*(E)$. Let $\theta$ be the Heaviside function. A mapping $\theta^*$ from $(\Psi_+\oplus\Psi_-)^\times$ into $(\Phi_++\Phi_-)^\times$ can be defined as

$$\langle\theta f|\theta^*\xi\rangle=\langle f|\xi\rangle, \tag{14}$$

for all $f\in\Psi_+\oplus\Psi_-$ and all $\xi\in(\Psi_+\oplus\Psi_-)^\times$. Thus

$$\langle f(E')|\xi(E)\rangle=\langle\theta f(E')|\theta^*\xi(E)\rangle=f^*(E). \tag{15}$$

If $f(E')\in\Psi_\pm$, then $\theta f(E')\in\Phi_\pm$ and $\theta^*|\xi(E)\rangle=|\eta_\pm(E)\rangle$ so that $|\eta_+(E)\rangle=|\eta-(E)\rangle$. Hereafter, we call this functional $|\eta(E)\rangle$.

*Remark*: Actually, $|E_+\rangle$ and $|E_-\rangle$ are the restrictions to $\Phi_+$ and $\Phi_-$, respectively, of a functional $|\eta(E)\rangle\in(\Phi_++\Phi_-)^\times$. In the above proposition, we have identified $|\eta(E)\rangle$, $|\eta_+(E)\rangle$, and $|\eta_-(E)\rangle$.

Pick now $U^{-1}|\eta(E)\rangle\in(\Gamma_+)^\times\cap(\Gamma_-)^\times$. How does it operate? Let $f_\pm\in\Gamma_\pm$ and $Uf_\pm=f_\mp(E')\in\Phi_\mp$, then $\langle f_\pm|U^{-1}\eta(E)\rangle=\langle Uf_\pm|\eta(E)\rangle=\langle\Phi_\mp(E')|\eta(E)\rangle=f^*_+(E)$. Therefore, $U^{-1}|\eta(E)\rangle$ assigns the value on $E$ of its corresponding function in $\Psi_\mp$ to any function in $\Gamma_\pm$. For the sake of simplicity, we shall hereafter denote by $|E\rangle$ the functional $U^{-1}|\eta(E)\rangle$.

Now, it becomes clear that the formula $\Omega^\pm|E\rangle=|E^\pm\rangle$ means the $|E^\pm\rangle$ belong to $(\Gamma^\pm)^\times$, respectively. If $f^\pm\in\Gamma^\pm$, $\langle f^\pm|E^\pm\rangle=\langle\Omega^\pm f_\pm|\Omega^\pm E\rangle=\langle f_\pm|E\rangle=\langle f_\mp(E')|\eta(E)\rangle=f^*_\mp(E)$. When we apply $|E^\pm\rangle$ to any function $f^\pm\in\Gamma^\pm$, we obtain the complex conjugate of the value on $E$ of $\theta^{-1}V_\pm f^\pm\in\Psi$. This exactly is the same value we obtain by applying $|E\rangle$ to $(\Omega^\pm)^{-1}f^\pm=f_\pm$.

On the other hand, $H_0|E\rangle=E|E\rangle$ and $H_1|E^\pm\rangle=E|E^\pm\rangle$. This easily comes from $\hat{E}|\eta(E)\rangle=E|\eta(E)\rangle$. In fact, if we apply $U$ to both members of these equations, we obtain $U\hat{E}U^{-1}U|\eta(E)\rangle=EU|\eta(E)\rangle$ or $H_0U|\eta(E)\rangle=EU|\eta(E)\rangle$.

Thus $H_0|E\rangle=E|E\rangle$. Also, $H_1|E^\pm\rangle=H_1\Omega^\pm|E\rangle=\Omega^\pm H_0|E\rangle=\Omega^\pm E|E\rangle=E\Omega^\pm|E\rangle=E|E^\pm\rangle$.

Let $w$ be a point in $\mathbb{C}^+$ (upper half plane) or $\mathbb{C}^-$ (lower half plane). If $f_\pm(E')\in\Phi_\pm$, the functional $|\eta(w)\rangle$ such that $\langle f_\pm(E')|\eta(w)\rangle=f^*_\pm(w)$, $w\in\mathbb{C}^\mp$ is continuous in $\Phi_\pm$. We can also define $|w\rangle=U|\eta(w)\rangle$ and write $\Omega^\pm|w\rangle=|w^\pm\rangle$. Furthermore, $H_0|w\rangle=w|w\rangle$ and $H_1|w^\pm\rangle=w|w^\pm\rangle$.

Finally, we see that $S\Gamma_+$ is dense in $\mathcal{H}$, since $S$ is unitary (we have assumed completeness of $\Omega^\pm$). Therefore, if we assume that the incoming scattering states are in $\Gamma_+$, we make a negligible error if we consider all the outgoing states to be in $\Gamma_-$. Reciprocally, if we take the outgoing states in $\Gamma_-$, $S^{-1}\Gamma_-$ is dense in $\mathcal{H}$ and we can approach any incoming state by an element of $\Gamma_+$. Note that we do not require that $S\Gamma_+=\Gamma_-\Leftrightarrow\Gamma^+=\Gamma^-$. This would be too strong a condition, implying asymptotic completeness, with an obscure physical meaning and even unnecessary $(S=(\Omega^-)^{-1}\Omega^+)$.

*Remark*: All of above is also valid if we start with the spaces $\Phi_\pm$ as defined in Ref. 3. Those are nonmetrizable by contrast with the $\Phi_\pm$ as used up to here.[14]

## III. APPLICATION TO THE DESCRIPTION OF RESONANCES

The objective of the present section is double. First, we shall construct new triplets for the description of both resonances and virtual states and for which the above machinary remain valid. Then we use the results of Sec. I to gain more insight in the model for resonances described in Refs. 1–3.

For these new triplets, we choose $L^2(R^+,dE)$ as Hilbert space again. All we need is the spaces $\Phi_\pm$. We want the functions in $\Phi_\pm$ to be continuous, and they will be defined quite similarly as those of Refs. 2 and 3. There $\Phi_\pm$ will be useful for the description of resonances and virtual states as well as those spaces of Ref. 3. However, they have a structural advantage with respect to the latter: They are metrizable and, therefore, mathematically simpler. The idea is the following: For the description of virtual states all we need is being able to extend the functions in $\Phi_+$ or $\Phi_-$, as defined in Ref. 2 or in the first section of the present paper, into analytic functions on a strip contained in $\mathbb{C}^-$ or $\mathbb{C}^+$, respectively. In Ref. 3, $\Phi_\pm\subset\{\text{restrictions to } R^+ \text{ of functions in } H^2_\pm\}\cap\{\text{boundary values on } R^+ \text{ of entire analytic functions}\}$ so that these strips are in fact the whole half plane, which is obviously unnecessary.

Our new $\Phi_-$ will be constructed as follows: Let $S^{(\alpha)}(R)$ be the set of all infinitely differentiable functions $f(x)$ such that $e^{\alpha|x|}x^n(d^m/dx^m)f(x)$ is bounded for all $n, m\in N$.

Here, $\alpha$ is any positive real number. This is a complete nuclear locally convex space and also Fréchet.[15] Consider $S^{(\alpha)}(R^-)$, the closed subspace of all $S^{(\alpha)}(R)$ functions supported on $R^-$. If $\Psi_-$ is the space of the Fourier transforms of the members of $S^{(\alpha)}(R^-)$, then $\Psi_-$ has the required properties: Any function $f(E)\in\Psi_-$ belongs to $H^2_-$ after the Paley–Wiener theorem.[2,7] Furthermore, $f(E)$ can be analytically continued into the strip $\mathrm{Im}\,\omega<\alpha$. This result comes after the following.

**Theorem**[16]: Let $f\in L^2(R)$. Then $e^{b|x|}f(x)\in L^2(R)$ for all $b<a$ if and only if the Fourier tranform $\mathcal{F}(f)=\hat{f}$ has an analytic continuation into $|\mathrm{Im}\,z|<\alpha$ and

$$\int_{-\infty}^{\infty}|f(x)+iy|^2\,dx \tag{16}$$

is uniformly bounded for all $-a<y<a$.

As always, let $\Phi_-$ be the space of the restrictions to $R^+$ of all the functions in $\Psi_-$. Then $\Phi_-$ has the following properties.

(1) Any $\varphi(E)\in\Phi_-$ can be analytically continued into the region $\mathrm{Im}\,\omega<\alpha$. Also the extension of $\varphi(E)$ into $R$ belongs to

$H^2_- \cap \mathcal{S}(R)$ because of the Paley–Wiener theorem[2,7] and the fact $S^{(\alpha)}(R^-) \subset \mathcal{S}(R^-)$. We recall that the Fourier transform of a Schwartz function is also a Schwartz function.

(2) If $\varphi(E) \in \Phi_-$, then $E^n \varphi(E) \in \Phi_-$ for all $n \in N$. This is a consequence of $S^{(\alpha)}(R^-) \subset \mathcal{S}(R^-)$ and the properties of the Fourier transform.

(3) The space $\Phi_-$ can be endowed with a complete nuclear metrizable topology. To see it, note on the bijection existing betwen $\Phi_-$ and $S^{(\alpha)}(R^-)$. The Fourier transform $\mathcal{F}$ is a bijection between $S^{(\alpha)}(R^-)$ and $\Psi$. Then, the Heaviside function $\theta$ is a bijection between $\Psi_-$ and $\Phi_-$, since a function in $\Psi_-$ is fully determined by its values on $R^+$ (Ref. 17). Thus, $\Phi_- = \theta \circ \mathcal{F}[S^{(\alpha)}(R^-)]$ and we transport the topology on $S^{(\alpha)}(R)$ to $\Phi_-$ by means of $\theta \circ \mathcal{F}$. After the properties of the topology on $S^{(\alpha)}(R)$, those we have claimed for $\Phi_-$ follow.

(4) The space $\Phi_-$ is dense in $L^2(R^+)$. Furthermore, the canonical injection $\Phi_- \to L^2(R)$ is continuous. Obviously, any function in $\mathcal{S}(R^-)$ with compact support is in $S^{(\alpha)}(R^-)$. Thus, $S^{(\alpha)}(R^-)$ is dense in $L^2(R^-)$. Then, the procedure to show the denseness of $\Phi_-$ in $L^2(R^+)$ is indentical to the one followed for a similar problem in Ref. 2.

Proving the continuity of the canonical injection is not difficult. First of all, note that an equivalent family of seminorms on $S^{(\alpha)}(R^-)$ can be given by

$$p^2_{n,m}(f) = \int_{-\infty}^{0} \left| x^n \frac{d^m}{dx^m} f(x) \right|^2 e^{-\alpha x}\, dx, \quad f \in S^{(\alpha)}(R). \tag{17}$$

The proof of this assertion is a mere adaptation of the proof valid for the similar result on $(R)$ (Ref. 18). On the other hand, a set of seminorms for $\Phi_-$ can be given by

$$q_{n,m}(f(E)) = p_{n,m}[\mathcal{F}^{-1} \circ \theta^{-1}(f(E))]. \tag{18}$$

The only seminorm on $L^2(R^+)$ is its norm. Thus

$$\begin{aligned} \|f(E)\|^2 &= \int_{-\infty}^{0} |f(E)|^2\, dE \leqslant \int_{-\infty}^{\infty} |\theta^{-1} f(E)|^2\, dE \\ &= \int_{-\infty}^{\infty} |\mathcal{F}^{-1} \circ \theta^{-1} f(x)|^2\, dx \\ &= \int_{-\infty}^{0} |\mathcal{F}^{-1} \circ \theta^{-1} f(x)|^2\, dx, \end{aligned} \tag{19}$$

which yields the continuity of the canonical injection.[19]

After all of the above, we conclude that $\Phi_- \subset L^2(R^+) \subset \Phi^x_-$ is a RHS. If $\Phi_+$ is the space of all canonical conjugates of the elements in $\Phi_-$, we have a new RHS: $\Phi_+ \subset L^2(R^+) \subset \Phi^x_+$. Any function $f(E) \in \Phi_+$ can be analytically continued into the region $\operatorname{Im} \omega > \alpha$ and its extension into $R$ belongs to $H^2_+ \cap \mathcal{S}(R)$. With this exception, the functions in $\Phi_+$ satisfy the same properties as those of $\Phi_-$.

The results of Sec. I can be applied to these new spaces. Resonances and virtual states can be described either by using these RHS's or $\Psi_\pm \subset H^2_\pm \subset \Psi^x_\pm$, where the $\Psi_\pm$ are the spaces of the extensions into all of $R$ of functions in $\Phi_\pm$. Using these last RHS's instead of the others has some advantages we have already mentioned in Refs. 2 and 3. They can be summarized as follows

(1) The $\Psi_+$ and $\Psi_-$ have a trivial intersection. When using this description, the sets of incoming and outgoing states are disjoint. (The zero vector does not correspond to a physical state).

(2) When we operate, we need to make use of the values of $\varphi(E)$ on $R^-$ as well as those of $R^+$. These values are provided by the functions in $\Psi_\pm$.

(3) Virtual states and resonances are generalized vectors of different nature provided we use the representation $\Psi_\pm \subset H^2_\pm \subset \Psi^x_\pm$. Resonances are normalizable generalized eigenvectors of the multiplication operator. In the meantime, virtual states are normalizable generalized eigenvectors of it. This difference disappears when using the triplets $\Phi_\pm \subset L^2(R^+) \subset \Phi^x_\pm$.

However, it is important to remark that only the triplets $\Gamma_\pm \subset \mathcal{H} \subset \Gamma_\pm$ and $\Gamma^\pm \subset \mathcal{H}_{\text{ac}} \subset (\Gamma^\pm)^x$ have a straight interpretation in scattering theory and that the other two are representations of them. In the following, we are going to reformulate the results of Refs. 2 and 3 in terms of these more physical triplets. Our point of departure will be the spaces $\Phi_\mp$ has constructed in this section. The spaces $\Gamma_\pm$ and $\Gamma^\pm$ will change correspondingly.

Now assume a resonance is located at $Z_R = E_R - i\Gamma/2 (E_R .0; \Gamma > 0)$. If $f_+(E) \in \Phi_+, f_- = U^{-1} f_+(E)$, $\Omega^- f_- = f^-$, we have

$$\begin{aligned} f^*_+(Z_R) &= \langle f_+(E) | \eta(Z_R) \rangle = \langle f_- | Z_R \rangle \\ &= \langle \Omega^- f_- | \Omega^- Z_R \rangle = \langle f^- | Z_R^- \rangle, \end{aligned} \tag{20}$$

and also

$$f^*_-(Z^*_R) = \langle f^+ | Z^{*+}_R \rangle. \tag{21}$$

On the other hand,

$$\begin{aligned} &\langle \hat{E} f_+(E) | \eta(Z_R) \rangle \\ &\quad = \langle E f_+(E) | \eta(Z_R) \rangle = Z_R \langle f_+(E) | \eta(Z_R) \rangle \\ &\quad = Z_R f^*_+(Z_R), \end{aligned} \tag{22}$$

as proven in Ref. 1 and 2. Therefore

$$\begin{aligned} Z_R f^*_+(Z_R) &= \langle H_0 f_- | Z_R \rangle = \langle H_1 f^- | Z_R^- \rangle \\ &= Z_R \langle f^- | Z_R^- \rangle. \end{aligned} \tag{23}$$

This implies the validity of the formula $H_1 | Z_R^- \rangle = Z_R | Z_R^- \rangle$ in $(\Gamma^-)^x$. Analogous considerations yield $H_1 | Z^{*+}_R \rangle = Z^*_R | Z^{*+}_R \rangle$ in $(\Gamma^+)^x$.

*Proposition 4*:

$$\begin{aligned} &\text{For } t > 0, \quad e^{-itH_1} \Gamma^- \subset \Gamma^-; \\ &\text{for } t < 0, \quad e^{-itH_1} \Gamma \not\subset \Gamma^-; \\ &\text{for } t < 0, \quad e^{-itH_1} \Gamma^+ \subset \Gamma^+; \\ &\text{for } t > 0, \quad e^{-itH_1} \Gamma^+ \not\subset \Gamma^+. \end{aligned} \tag{24}$$

*Proof*: In Ref. 2, we have proven that

$$\begin{aligned} &\text{for } t > 0, \quad e^{-itE} \Phi_+ \subset \Phi_+; \quad \text{for } t < 0, \quad e^{-itE} \Phi_+ \not\subset \Phi_+; \\ &\text{for } t < 0, \quad e^{-itE} \Phi_- \subset \Phi_-; \quad \text{for } t > 0, \quad e^{-itE} \Phi_- \not\subset \Phi_-; \end{aligned} \tag{25}$$

the $\Phi_\pm$ being the spaces used in Sec. I. We can extend these results to the spaces $\Phi_\pm$ of Sec. II as well as those of Ref. 3. Then, after (25), (24) is straightforward. For instance, let $f^- \in \Gamma^-$ and $t > 0$. We want to show $e^{-itH_1} f^- \in \Gamma^-$.

Consider

$$U(\Omega^-)^{-1}e^{-itH_1}f^- = U[(\Omega^-)^{-1}e^{-itH_1}\Omega^-][(\Omega^-)^{-1}f^-] = Ue^{-itH_0}f^- = [Ue^{-itH_0}U^{-1}][Uf_-] = e^{-itE}f_+(E). \quad (26)$$

The last term in (26) is readily seen to be in $\Phi_+$, because $f_1(E)\in\Phi_+$ and (25). This, along with $\Gamma^- = \Omega^- U^{-1}\Phi_+$, proves our claim. With similar procedures, we obtain all relations (23).

Let us now study the behavior of the evolution operator $U(t) = e^{-itH_1}$ on the resonances. After Proposition 4, we can extend $U(t)$ into $(\Gamma^+)^\times$ if $t>0$ and also into $(\Gamma^-)^\times$ if $t<0$, in the following form: Consider $U^*(t) = e^{itH_1}$, the Hilbert space adjoint of $U(t)$. If $f^+\in\Gamma^+$ and $g^-\in\Gamma^-$,

$$\langle U^*(t)f^+|\varphi^+\rangle = \langle f^+|U^\times(t)\varphi^+\rangle, \quad t>0,$$
$$\langle U^*(t)g^-|\xi^-\rangle = \langle g^-|U^\times(t)\xi^-\rangle, \quad t<0, \quad (27)$$

where $\varphi^+\in(\Gamma^+)^\times$, $\xi^-\in(\Gamma^-)^\times$, and $U^\times(t)$ is the extension of $U(t)$ to the dual spaces. Definitions (27) are allowed by Proposition 4. In order to simplify the notation, we also call $e^{-itH_1}$ to $U^\times(t)$. If again, $Z_R$ is a point at which a resonance can be found, then

$$\langle f^-|e^{-itH_1}|Z_R^-\rangle = \langle e^{-itH_1}f^-|Z_R^-\rangle = \langle e^{itE}f_+(E)|\eta(Z_R)\rangle = e^{-itZ_R}f_+^*(E) = e^{-itZ_R}\langle f^-|Z_R^-\rangle, \quad t>0, \quad (28)$$

which yields

$$e^{-itH_1}|Z_R^-\rangle = e^{-itE_R}e^{-\Gamma t/2}|Z_R^-\rangle, \quad t>0. \quad (29)$$

After similar procedures, we can analogously obtain

$$e^{-itH_1}|Z_R^{*+}\rangle = e^{-itE_R}e^{\Gamma t/2}|Z_R^{*+}\rangle, \quad t<0, \quad (29')$$

and (29) and (29') follow from the part holding the symbol $\subset$ in Proposition 4. The other one says that (29) is not valid for $t<0$ and (29') is not valid for $t>0$. Equation (29) describes the decaying part of a resonance and (29') its growing part.[1,20]

Finally, if we have a virtual state placed at $E\in R^-$, $|E^-\rangle\in(\Gamma^-)^\times$ and $|E^+\rangle\in(\Gamma^+)^\times$ are equally good state vectors to describe them. The justification of it can be found in Ref. 3 and all the preceding.


[1]A. Böhm, J. Math. Phys. **22**, 2813 (1981).
[2]M. Gadella, J. Math. Phys. **24**, 1462 (1983).
[3]M. Gadella, J. Math. Phys. **24**, 2142 (1983).
[4]W. Amrein, J. Jauch, and B. Sinha, *Scattering Theory in Quantum Mechanics* (Benjamin, Reading, MA, 1977); M. Reed and B. Simon, *Scattering Theory* (Academic, New York, 1979), Sec. XI. 3; B. Simon, *Quantum Mechanics for Hamiltonians Defined as Quadratic Forms* (Princeton U. P., Princeton, 1971).
[5]This is equivalent to the condition of asymptotic completeness, since the singular spectrum of $H$ is emply. See Ref. 4.
[6]J. M. Jauch, *Foundations of Quantum Mechanics* (Addison–Wesley, Reading, MA, 1968); M. Reed and B. Simon *Functional Analysis* (Academic, New York, 1972).
[7]P. Duren, *Theory of $H_p$ Spaces*, Vol. 28 in *Pure and Applied Mathematics* (Academic, New York, 1970); K. Hoffman, *Banach Spaces of Analytic Functions* (Prentice–Hall, Englewood Cliffs, NJ, 1962); P. Koosis, *Introduction to $H_p$ Spaces, London Mathematical Society Lecture Note Series*, Vol. 40 (Cambridge U. P., Cambridge, 1980).
[8]Note that the final topology on C is the finest one for which C is locally convex and $h$ is continuous. In the present case, some authors prefer speaking about transferred topologies instead of final topologies since all we have is two spaces and a bijection between them. All these topics are studied, for instance, in the following book: J. Horvath, *Topological Vector Spaces and Distributions* (Addison–Wesley, Reading, MA, 1966).
[9]J. E. Roberts, Commun. Math. Phys. **3**, 98 (1966). Although Roberts' results refer to the initial topologies, some conclusions are valid for final ones.
[10]See p. 301 of the book mentioned in Ref. 8.
[11]The proof on the continuity and nulcearity of $i$ has not been clearly stated in Ref. 2. It is, however, simple. The canonical injection $\mathcal{S}(R^-)\xrightarrow{h}L^2(R^-)$ is continuous. Since $\mathcal{S}(R^-)$ is nuclear and its topology can be given by using the method of Ref. 9, it can be proven that $h$ is a nuclear mapping [see O. Melsheimer, J. Math. Phys. **15**, 902 (1974)]. Here $\mathcal{F}$ is a bicontinuous bijection, so that it preserves all the properties. Thus $s = \mathcal{F}h\mathcal{F}^{-1}$: $\Psi_+\to H^2_+$ is continuous and nuclear. On the other hand, if $f(E)\in H^2_+$,

$$\theta f(E) = \begin{cases} f(E), & \text{if } E>0, \\ \text{not defined}, & \text{if } E<0 \end{cases}$$

is a projection from $H^2_+$ into itself. It is, in addition, continuous from $H^2_+$ into $L^2(R^+)$. The diagram

$$\begin{array}{ccc} \Psi_+ & \xrightarrow{s} & H^2_+ \\ \theta\downarrow & & \downarrow\theta \\ \Phi_+ & \xrightarrow{i} & L^2(R^+) \end{array}$$

shows that $i = \theta s\theta^{-1}$ [note that $\theta$ is a bicontinuous bijection between $\Phi_+$ and $\Psi_+$, but not between $H^2_+$ and $L^2(R^+)$]. This proves that $i$ is continuous and since $s$ is nuclear, $i$ is also nuclear.
[12]V. K. Khoan, *Distributions, Analyse de Fourier Opérateurs aux Derivées Partielles* (Vuibert, Paris, 1972), Vol. 1, p. 74.
[13]For a definition of the strong topology, see, for instance, M. Reed and B. Simon, *Functional Analysis* (Academic, New York, 1972), p. 165.
[14]However the author is not sure whether $i$ and $j$ as in (7) are nuclear if we use the spaces as defined in Ref. 3. Continuity is, nevertheless, clearly assured.
[15]I. M. Gelfand and N. Ya. Vilenkin, *Generalized Functions* (Academic, New York, 1964), Vol. 4 and also see Vol. 2 of the same collection.
[16]M. Reed and B. Simon, *Fourier Analysis. Self Adjointness* (Academic, New York, 1975), p. 18.
[17]See Refs. 2 and 16 quoted therein. Also use the principle of analytic continuation.
[18]Reference 16, pp. 141–142.
[19]Reference 16, p. 129.
[20]A. Böhm, "The rigged Hilbert space and decaying states" in *Group Theoretical Methods in Physics*, Proceedings of the Austin Conference 1978, *Lecture Notes in Physics*, Vol. 94 (Springer-Verlag, Berlin, 1979); A. Böhm, Lett. Math. Phys.**3**, 455 (1979).

# A method of calculating the function of two noncommuting operators

P. Moretti
*Istituto di Ricerca sulle Onde Elettromagnetiche del Consiglio Nazionale delle Ricerche, Via Panciatichi 64, 50127 Firenze, Italy*

M. Mancini
*Istituto di Fisica Superiore dell'Università, 50139 Firenze, Italy*

(Received 23 August 1983; accepted for publication 30 March 1984)

A method is developed of evaluating a function of the sum of two noncommuting Hermitian operators. A physical example is given and a summation formula is obtained.

PACS numbers: 03.65.Db

## I. INTRODUCTION

In this paper we discuss a somewhat novel method of treating perturbatively a function of two noncommuting operators. Since this problem often appears in quantum mechanics, we intend to give a general exposition of this method which is based on the use of complex analysis.

Let us consider a Hermitian operator $A$ with an orthonormal set of eigenvectors $|a_j\rangle$ and (real) eigenvalues $a_j$,

$$A|a_j\rangle = a_j|a_j\rangle, \tag{1}$$

and a function $f(z)$, without singularities, expressible as a Taylor series

$$f(z) = \sum_n^{0\cdots\infty} c_n z^n, \quad c_n = f^{(n)}(0)/n! \tag{2}$$

where $f^{(n)}(0)$ denotes the $n$th derivative of $f(z)$ calculated at the point $z=0$. If $B$ is another Hermitian operator, noncommuting with $A$, we are interested in the evaluation of the operator $f(A+B)$ on the basis of the $|a_k\rangle$ vectors.

We start from the formula[1,2]

$$L = (2\pi i)^{-1}\int_{\gamma\infty} zG(z)dz, \quad \text{for } L \neq 0 \tag{3}$$

expressing a general Hermitian operator $L$ as a contour integral; $G(z)$ is the resolvent operator

$$G(z) = (z-L)^{-1}, \tag{4}$$

and $\gamma\infty$ is a closed contour enclosing all the nonzero eigenvalues of $L$. Since from Eq. (4) one has

$$LG = zG - 1,$$

it follows

$$L^2 = (2\pi i)^{-1}\int_{\gamma\infty} z^2 G(z)dz,$$

and so on, thus obtaining

$$L^m = (2\pi i)^{-1}\int_{\gamma\infty} z^m G(z)dz,$$

and, from Eq. (2),

$$f(L) = (2\pi i)^{-1}\int_{\gamma\infty} f(z)G(z)dz. \tag{5}$$

This formula gives a definition for $f(L)$, if the expression (2) does not hold.[3,4]

## II. THE CALCULATION OF THE MATRIX ELEMENT

Considering the operator $(A+B)$, we have

$$f(A+B) = (2\pi i)^{-1}\int_{\gamma\infty} f(z)G(z)dz \tag{6}$$

where $G(z) = [z-(A+B)]^{-1}$, and $\gamma\infty$ encloses all the eigenvalues of $(A+B)$ as well as the eigenvalues of $A$. If we pose

$$G_0(z) = (z-A)^{-1},$$

the following expansion holds:

$$G(z) = \sum_r^{0\cdots\infty} G_0(BG_0)^r. \tag{7}$$

The convergence of this expansion is discussed in Ref. 2. In this way we have, using the completeness relation for the vectors $|a_j\rangle$,

$$\begin{aligned}
\langle a_1|f(A+B)-f(A)|a_2\rangle &= (2\pi i)^{-1}\int_{\gamma\infty} f(z)\sum_r^{1\cdots\infty}\langle a_1|G_0(BG_0)^r|a_2\rangle dz \\
&= (2\pi i)^{-1}\sum_r^{1\cdots\infty}\int_{\gamma\infty} f(z)\Bigg\{\sum_{i_1,i_2,\ldots,i_r}\langle a_1|G_0B|a_{i_1}\rangle\langle a_{i_1}|G_0B|a_{i_2}\rangle \\
&\quad\times\langle a_{i_2}|G_0B\cdots G_0B|a_{i_r}\rangle\langle a_{i_r}|G_0|a_2\rangle\Bigg\}dz \\
&= (2\pi i)^{-1}\sum_r^{1\cdots\infty}\int_{\gamma\infty} f(z)\Bigg\{\sum_{i_1,i_2,\ldots,i_{r-1}}(z-a_1)^{-1}\langle a_1|B|a_{i_1}\rangle(z-a_{i_1})^{-1} \\
&\quad\times\langle a_{i_1}|B\cdots B|a_{i_{r-1}}\rangle(z-a_{i_{r-1}})^{-1}\langle a_{i_{r-1}}|B|a_2\rangle(z-a_2)^{-1}\Bigg\}dz \\
&= \sum_r^{1\cdots\infty}\sum_{\{i,r\}}\langle a_1|B|a_{i_1}\rangle\langle a_{i_1}|B\cdots B|a_{i_{r-1}}\rangle\langle a_{i_{r-1}}|B|a_2\rangle(2\pi i)^{-1}\int_{\gamma\infty} f_{\{i,r\}}(z)dz.
\end{aligned} \tag{8}$$

 0022-2488/84/082486-04$02.50  

Here the index $\{i,r\}$ denotes the set $(i_1,i_2,...,i_{r-1})$; if $N$ is the number of eigenvectors $|a_j\rangle$, the second sum is extended to the $N^{r-1}$ dispositions with repetitions of the indices $i_k$ taken $(r-1)$ at a time $(i_k = 1,2,...,N)$, and

$$f_{\{i,r\}}(z) = f(z)\left[(z-a_1)(z-a_{i_1})\cdots(z-a_{i_{r-1}})(z-a_2)\right]^{-1}. \tag{9}$$

Since one has[5]

$$(2\pi i)^{-1}\int_{\gamma\infty} f_{\{i,r\}}(z)dz = \sum [\text{residues of } f_{\{i,r\}}(z) \text{ inside } \gamma\infty],$$

if we denote as $R_{\{i,r\}}(a_{i_k})$ the residue at $z = a_{i_k}$ of the function $f_{\{i,r\}}(z)$, Eq. (8) is rewritten as

$$\langle a_1|f(A+B)-f(A)|a_2\rangle = \sum_{r}^{1\cdots\infty}\sum_{\{i,r\}}\langle a_1|B|a_{i_1}\rangle\langle a_{i_1}|B\cdots B|a_2\rangle \times\left[R_{\{i,r\}}(a_1)+\sum_{\nu}^{1\cdots r-1}R_{\{i,r\}}(a_{i_\nu})+R_{\{i,r\}}(a_2)\right]. \tag{10}$$

In calculating the residues, the expression (9) is to be examined in detail. The indices $i_k$ run through the whole set of the eigenvectors $|a_j\rangle$, and $|a_1\rangle,|a_2\rangle$, are fixed; it is therefore apparent that in general the $(r+1)$ factors in the denominator of $f_{\{i,r\}}(z)$ are not all different from each other. If the identical factors are grouped together, $(z-a_1)$ will appear $\lambda_1$ times, $(z-a_{i_k})\lambda_{i_k}$ times, $(z-a_2)\lambda_2$ times, each $\lambda$ depending on the choice of the set $\{i,r\}$, so that they are indicated as $\lambda\{i,r\}$. The allowed values for $\lambda_1\{i,r\}, \lambda_2\{i,r\}, \lambda_{i_k}\{i,r\}$ go from 0 to $(r+1)$, subject to the condition

$$\lambda_1\{i,r\}+\lambda_2\{i,r\}+\sum_{k}^{1\cdots r-1}\lambda_{i_k}\{i,r\} = r+1. \tag{11}$$

The function $f_{\{i,r\}}(z)$ is therefore rewritten as

$$f_{\{i,r\}}(z) = f(z)\left[(z-a_1)^{\lambda_1\{i,r\}}(z-a_{i_1})^{\lambda_{i_1}\{i,r\}}\cdots(z-a_{i_{r-1}})^{\lambda_{i_{r-1}}\{i,r\}}(z-a_2)^{\lambda_2\{i,r\}}\right]^{-1}. \tag{12}$$

In this way, a direct calculation[5] gives (the index $\{i,r\}$ on the $\lambda$ will be omitted for brevity)

$$R_{\{i,r\}}(a_{i_k}) = \left[(\lambda_{i_k}-1)!\right]^{-1}\sum_{\nu}^{0\cdots\lambda_{i_k}-1}\binom{\lambda_{i_k}-1}{\nu}f^{(\nu)}(a_{i_k}) + g_{i_k}^{(\lambda_{i_k}-\nu-1)}(a_{i_k}), \tag{13}$$

with

$$g_{i_k}(z) = [(z-a_1)^{\lambda_1}(z-a_{i_1})^{\lambda_{i_1}}\cdots(z-a_{i_{k-1}})^{\lambda_{i_{k-1}}} \times(z-a_{i_{k+1}})^{\lambda_{i_{k+1}}}\cdots(z-a_{i_{r-1}})^{\lambda_{i_{r-1}}} \times(z-a_2)^{\lambda_2}]^{-1},$$

where also the values 1,2, are allowed for $i_k$, and $\lambda_{i_k} > 0$. The use of these formulas will be displayed in a practical example.

## III. A PHYSICAL APPLICATION: THE DENSITY MATRIX

A useful application of the formula (10) is found in the calculation of the density matrix of two coupled systems[6]; if $\mathscr{H}_0$ is the Hamiltonian of two uncoupled systems and $\mathscr{H}_1$ is the coupling term, we have (posing $H_0 = \mathscr{H}_0/kT$, $H_1 = \mathscr{H}_1/kT$, where $k$ is the Boltzmann constant and $T$ is the absolute temperature with $a_j$ and $|a_j\rangle$ now the eigenvalues and the eigenvectors of $H_0$):

$$\langle a_1|\exp[-(H_0+H_1)]|a_2\rangle = \exp(-a_1)\delta_{12} + \langle a_1|H_1|a_2\rangle[\exp(-a_1) - \exp(-a_2)](a_1-a_2)^{-1}, \tag{14}$$

to the first order in $H_1$. If $|a_1\rangle$ and $|a_2\rangle$ are degenerate, we have $\lambda_1 = 2$, $\lambda_2 = 0$, so that, using Eq. (13) [the expression (14) is now undetermined] we have the result

$$\langle a_1|\exp[-(H_0+H_1)]|a_2\rangle = \exp(-a_1)[\delta_{12} - \langle a_1|H_1|a_2\rangle], \tag{15}$$

which, in this simple case, can be obtained carrying out in Eq. (14) the proper limit. Using these formulas it is possible to investigate the conditions under which the coupling term $H_1$ can be ignored in the evaluation of the density matrix.

As an example, we also calculate the term with $r = 2$. The set $\{i,r\}$ reduces to the index $i$; considering all the eigenvectors as nondegenerate, it is necessary to separate the cases when $\lambda_1 = \lambda_i = \lambda_2 = 1$ (i.e., $a_1 \neq a_i \neq a_2$), when $\lambda_1 = 2, \lambda_i = 0, \lambda_2 = 1$ (i.e., $a_i = a_1$), and $\lambda_1 = 1, \lambda_i = 0, \lambda_2 = 2$ (i.e., $a_i = a_2$). From the formula (10) we have for the term with $r = 2$,

$$\sum_{i\neq 1,2}\langle a_1|H_1|a_i\rangle\langle a_i|H_1|a_2\rangle \times[R_i(a_1)+R_i(a_i)+R_i(a_2)] + \langle a_1|H_1|a_1\rangle\langle a_1|H_1|a_2\rangle[R_1(a_1)+R_1(a_2)] + \langle a_1|H_1|a_2\rangle\langle a_2|H_1|a_2\rangle[R_2(a_1)+R_2(a_2)]. \tag{16}$$

All the residues can be calculated from Eq. (13) with the proper values of the $\lambda$:

$$R_i(a_1) = \exp(-a_1)[(a_1-a_2)(a_1-a_i)]^{-1},$$
$$R_i(a_i) = \exp(-a_i)[(a_i-a_1)(a_i-a_2)]^{-1},$$
$$R_i(a_2) = \exp(-a_2)[(a_2-a_1)(a_2-a_i)]^{-1},$$
$$R_1(a_1) = -\exp(-a_1)[(a_1-a_2)^{-1}+(a_1-a_2)^{-2}],$$
$$R_1(a_2) = \exp(-a_2)(a_2-a_1)^{-2},$$
$$R_2(a_1) = \exp(-a_1)(a_1-a_2)^{-2}$$
$$R_2(a_2) = -\exp(-a_2)[(a_2-a_1)^{-1}+(a_2-a_1)^{-2}].$$

The case with the only degeneracy $a_1 = a_2$ can be treated analogously, posing $\lambda_2 = 0$ (or $\lambda_1 = 0$); Eq. (16) becomes

$$\sum_{i\neq 1,2}\langle a_1|H_1|a_i\rangle\langle a_i|H_1|a_2\rangle[R_i(a_1)+R_i(a_i)] + \langle a_1|H_1|a_1\rangle\langle a_1|H_1|a_2\rangle R_1(a_1) + \langle a_1|H_1|a_2\rangle\langle a_2|H_1|a_2\rangle R_2(a_1). \tag{17}$$

The first two residues in this expression are calculated from Eq. (13) with $\lambda_1 = 2, \lambda_i = 1$, the third and the fourth with $\lambda_1 = 3$; we have

$$R_i(a_1) = -\exp(-a_1)[(a_1-a_i)^{-1}+(a_1-a_i)^{-2}],$$
$$R_i(a_i) = \exp(-a_i)(a_i-a_1)^{-2},$$
$$R_1(a_1) = R_2(a_1) = \tfrac{1}{2}\exp(-a_1).$$

When other degeneracies are present, other degrees of factorizations appear. As a limiting case, if all the eigenvalues of $H_0$ are degenerate ($H_0|a_j\rangle = a_1|a_j\rangle$), Eqs. (11) and (13) become

$$\lambda_1 = r+1,$$

and

$$R_{\{i,r\}}(a_1) = (2\pi i)^{-1}\int_{\gamma\infty}\exp(-z)(z-a_1)^{-(r+1)}dz$$
$$= [(-1)^r/r!]\exp(-a_1).$$

We have in this way

$$\langle a_1|\exp[-(H_0+H_1)]|a_2\rangle$$
$$= \exp(-a_1)\delta_{12} + \sum_r^{1\cdots\infty}\sum_{\{i,r\}}\langle a_1|H_1|a_{i_1}\rangle$$
$$\times\langle a_{i_1}|H_1\cdots H_1|a_2\rangle[(-1)^r/r!]\exp(-a_1)$$
$$= \sum_r^{0\cdots\infty}[(-1)^r/r!]\langle a_1|H_1^r|a_2\rangle\exp(-a_1)$$
$$= \langle a_1|\exp(-H_1)\exp(-H_0)|a_2\rangle$$
$$= \langle a_1|\exp(-H_0)\exp(-H_1)|a_2\rangle,$$

showing the obvious fact that, if the $H_0$ matrix is a multiple of the unit matrix, it commutes with $H_1$.

## IV. A SUMMATION FORMULA

We derive now an interesting summation formula by comparing the expression (10) and the expression obtained by the conventional perturbation method. First, the residue (13) of the function (9) at $z = a_{i_k}$ can be written, since one has

$$f^{\nu}(a_{i_k}) = \sum_n^{\nu\cdots\infty} c_n \nu!\binom{n}{\nu}a_{i_k}^{n-\nu},$$

in the following way:

$$R_{\{i,r\}}(a_{i_k}) = \sum_{\nu}^{0\cdots\lambda_{i_k}-1}\sum_n^{\nu\cdots\infty} c_n\binom{n}{\nu}[(\lambda_{i_k}-\nu-1)!]^{-1}$$
$$\times a_{i_k}^{n-\nu}g_{i_k}^{(\lambda_{i_k}-\nu-1)}(a_{i_k}).$$

Rearranging the summations,

$$R_{\{i,r\}}(a_{i_k}) = \sum_n^{0\cdots\infty} c_n R^{(n)}_{\{i,r\}}(a_{i_k}), \tag{18}$$

with

$$R^{(n)}_{\{i,r\}}(a_{i_k}) = \sum_{\nu}^{0\cdots\lambda_{i_k}-1}[(\lambda_{i_k}-\nu-1)!]^{-1}$$
$$\times g_{i_k}^{(\lambda_{i_k}-\nu-1)}(a_{i_k})\binom{n}{\nu}a_{i_k}^{n-\nu}\eta(n-\nu),$$

where $\eta(n-\nu)$ is the step function: $\eta(m)=1$ if $m\geqslant 0$, $\eta(m)=0$ if $m<0$. Moreover, one has

$$f(A+B) = \sum_n^{0\cdots\infty} c_n(A+B)^n,$$

and

$$(A+B)^n = \sum_r^{0\cdots n}\{A^{n-r},B^r\},$$

where $\{A^{n-r},B^r\}$ denotes the sum of $\binom{n}{r}$ addends, all different from each other since $A$ and $B$ do not commute, containing $(n-r)$ times the factor $A$ and $r$ times the factor $B$. Another way of writing our matrix element is therefore

$$\langle a_1|f(A+B)-f(A)|a_2\rangle$$
$$= \sum_n^{0\cdots\infty} c_n\sum_r^{0\cdots n}\langle a_1|\{A^{n-r},B^r\}|a_2\rangle. \tag{19}$$

After inspection we obtain [the term with $r=0$ gives merely $f(A)$]

$$\langle a_1|f(A+B)-f(A)|a_2\rangle$$
$$= \sum_n^{0\cdots\infty} c_n\sum_r^{1\cdots\infty}\sum_{\{i,r\}}\langle a_1|B|a_{i_1}\rangle$$
$$\times\langle a_{i_1}|B\cdots B|a_{i_{r-1}}\rangle\langle a_{i_{r-1}}|B|a_2\rangle$$
$$\times\sum_{(n_1+n_{i_1}+\cdots+n_{i_{r-1}}+n_2)=n-r} a_1^{n_1}a_{i_1}^{n_{i_1}}\cdots a_{i_{r-1}}^{n_{i_{r-1}}}a_2^{n_2}, \tag{20}$$

where the last sum is extended to the $\binom{n}{r}$ sets of $(r+1)$ positive integers $n_1,n_{i_1},\ldots,n_{i_{r-1}},n_2$, such that their sum is $(n-r)$; the sum over $r$ is extended to infinity since the terms with $r>n$ obviously vanish. Comparing the expressions (10) and (20) we have, using Eq. (18),

$$\sum_r^{1\cdots\infty}\sum_{\{i,r\}}\langle a_1|B|a_{i_1}\rangle\langle a_{i_1}|B\cdots B|a_2\rangle$$
$$\times\sum_n^{0\cdots\infty} c_n\Bigg[\sum_{(n_1+n_{i_1}+\cdots+n_2)=n-r} a_1^{n_1}a_{i_1}^{n_{i_1}}\cdots a_2^{n_2}$$
$$- R^{(n)}_{\{i,r\}}(a_1) - \sum_{\mu}^{1\cdots r-1} R^{(n)}_{\{i,r\}}(a_{i_\mu})$$
$$- R^{(n)}_{\{i,r\}}(a_2)\Bigg] = 0. \tag{21}$$

Since the operator $B$ and the form of the function $f(z)$ are arbitrary, the term in square brackets identically vanishes. We have in this way for a given choice $(3,4,\ldots,r+1)$ of the set $\{i,r\}$ (that becomes an index $r$)

$$\sum_{(n_1+n_2+\cdots+n_{r+1})=n-r} a_1^{n_1}a_2^{n_2}\cdots a_{r+1}^{n_{r+1}} = \sum_{\mu}^{1\cdots r+1} R_r^{(n)}(a_\mu). \tag{22}$$

If all the $a_\mu$ are different from each other, this equation becomes

$$\sum_{(n_1+n_2+\cdots+n_{r+1})=n-r} a_1^{n_1}a_2^{n_2}\cdots a_{r+1}^{n_{r+1}}$$
$$= a_1^n[(a_1-a_2)(a_1-a_3)\cdots(a_1-a_{r+1})]^{-1}$$
$$+ a_2^n[(a_2-a_1)(a_2-a_3)\cdots(a_2-a_{r+1})]^{-1}+\cdots$$
$$+ a_{r+1}^n[(a_{r+1}-a_1)(a_{r+1}-a_2)\cdots(a_{r+1}-a_r)]^{-1}.$$

For example, posing $(r+1)=n$,

$$a_1+a_2+\cdots+a_n$$
$$= a_1^n[(a_1-a_2)\cdots(a_1-a_n)]^{-1}$$
$$+ a_2^n[(a_2-a_1)\cdots(a_2-a_n)]^{-1}$$
$$+\cdots+a_n^n[(a_n-a_1)\cdots(a_n-a_{n-1})]^{-1},$$

and, posing $r = n$,

$$1 = a_1^n [(a_1 - a_2)(a_1 - a_3)\cdots(a_1 - a_n)(a_1 - a_{n+1})]^{-1} + \cdots + a_{n+1}^n [(a_{n+1} - a_1)(a_{n+1} - a_2) \cdots(a_{n+1} - a_{n-1})(a_{n+1} - a_n)]^{-1}.$$

If we choose $r > n$, the first member of Eq. (22) vanishes, so that

$$\sum_{\mu}^{1\cdots r+1} R_r^{(n)}(a_\mu) = 0.$$

If some $a_\mu$ are coincident, the general equations (22) and (18) are to be used. In the limiting case where all the $a_\mu$ coincide, the first member of Eq. (22) gives $\binom{n}{r}$ times $a_1^{n-r}$, exactly the same value of $R_r^{(n)}(a_1)$ given by Eq. (18) [only the term with $\nu = \lambda_1 - 1 = r$ is nonvanishing in Eqs. (13) and (18)].

We would like to point out that the aim of this paper is to give a contribution to the use of the complex analysis methods in quantum mechanics. These methods are generally more rapid than the usual algebraic methods, and are extensively used in many branches of theoretical physics [see, for examples, Refs. 3 and 7]. Moreover, they sometimes give interesting unexpected results, as shown in this last section.


[1]T. Kato, Progr. Theor. Phys. **4**, 154 (1949).
[2]A. Messiah, *Mécanique Quantique* (Dunod, Paris, 1964), Chap. XVI.
[3]E. Merzbacher, Am. J. Phys. **36**, 814 (1968).
[4]V. Smirnov, *Cours de Mathématiques Supérieures* (Mir, Moscow, 1972), Tome III, deuxième partie, Chap. IV.
[5]G. Sansone, *Lezioni sulla Teoria delle Funzioni di una Variabile Complessa* (Cedam, Padova, 1963), Chap. III.
[6]F. Hartmann-Boutron, Ann. Phys. **9**, 285 (1975).
[7]M. L.Goldberger and K. M. Watson, *Collision Theory* (Wiley, New York, 1964).

# On embedded eigenvalues for the one-dimensional Schrödinger equation

Maria Hoffmann-Ostenhof[a)]
*Institut für Theoretische Physik, Universität Wien, Boltzmanngasse 5, A-1090 Wien, Austria*

Thomas Hoffmann-Ostenhof
*Institut für Theoretische Chemie und Strahlenchemie, Universität Wien, Währingerstrasse 17, A-1090 Wien, Austria*

(Received 2 August 1983; accepted for publication 24 February 1984)

We consider the one-dimensional Schrödinger equation $(-d^2/dx^2 + V(x) - \lambda)u(x) = 0$ for $x \geqslant 0$, with Dirichlet boundary conditions, where $\lambda \geqslant 0$ and $V(x)$ is real valued, bounded, and vanishes for $x \to \infty$. Conditions on $V$ for the absence of non-negative eigenvalues are derived. We show for instance: Let $\lambda > 0$, $u \not\equiv 0$. If for some $c > 0$, $0 < \mu < m \leqslant 1$, $|V + cx^{-m}| \leqslant \mu\sqrt{\lambda}\, x^{-1}$ for $x$ large, then $u \notin L^2(0, \infty)$. We also construct explicit examples with $u \in L^2$, $\lambda \geqslant 0$ and $-c_-(1+x)^{-\alpha} \leqslant V(x) \leqslant -c_+(1+x)^{-\alpha}$, $c_- > c_+ > 0$, with $\alpha \in (0,1]$ for $\lambda > 0$, respectively, $\alpha \in (0,2]$ for $\lambda = 0$.

PACS numbers: 03.65.Ge

## I. INTRODUCTION

We consider the one-dimensional Schrödinger equation

$$\frac{-d^2u}{dx^2} + (V - \lambda)u = 0 \tag{1.1}$$

in $(0, \infty)$, where $\lambda \geqslant 0$ and $V$ is a real-valued function which vanishes for $x \to \infty$. In the following $d/dx$ will be denoted by $'$. For (1.1) and its generalizations to the $n$-dimensional case various conditions on $V$ are known which imply that (1.1) has no nontrivial $L^2$-solution. See Refs. 1–6 for results for the $n$-dimensional case and see, e.g., Refs. 7–11 for the one-dimensional case and for other references. In fact it has been shown that Refs. 1 and 6, roughly speaking, if $V = V_1 + V_2$ satisfies

$$V_2 = o(|x|^{-1}) \text{ and } \limsup_{x \to \infty} x \nabla V_1(x) \leqslant 0\,, \tag{1.2}$$

and if $(-\Delta + V - \lambda)u = 0$ with $\lambda > 0$ and $u \not\equiv 0$, then $u \notin L^2(R^n)$. Very sharp results of this kind have been obtained in the one-dimensional case, see, e.g., Refs. 8–10. For the case $\lambda = 0$, less seems to be known, see however Refs. 1 and 8.

In Sec. II, sufficient conditions on $V$ are given implying nonexistence of positive and zero eigenvalues $\lambda$ for (1.1). The proofs are very simple. We note that our results for positive eigenvalues are closely related to Weidmann's results.[7]

Several examples of (1.1) with embedded eigenvalues $\lambda > 0$ are found in the literature. See for instance Refs. 7, 11, 12, 13, and 9 for other references. The canonical example for a positive eigenvalue is that of von Neumann and Wigner[6,12] where $V(x) = -8x^{-1}\sin 2x + O(x^{-2})$ for $x$ large. $V$ clearly oscillates and decays slowly. However, oscillation of $V'$ rather than of $V$ can produce positive eigenvalues. This is already evident from Weidmann's results and his example for a positive eigenvalue.

In Sec. III, we shall give examples of this type (see also Eastham and Kalf[11]) for positive and zero eigenvalues. Further, we discuss the relation of the asymptotics of $V$ and the asymptotics of the eigenfunction $u$ in these examples.

## II. NONEXISTENCE OF NON-NEGATIVE EIGENVALUES

In this section we consider differential equations of the form

$$-u'' + (V - \lambda)u = 0 \quad \text{for } a \leqslant x < \infty \tag{2.1}$$

with $\lambda \geqslant 0$, $u \in L^2(a, \infty)$ with boundary condition $u(a) = 0$. $V(x)$ is assumed to be a

real-valued, measurable bounded function on $[a, \infty)$

$$\text{with } \lim_{x \to \infty} |V(x)| = 0\,. \tag{2.2}$$

As is well known[14,15] the operator $-d^2/dx^2 + V$ with Dirichlet boundary conditions admits a unique self-adjoint realization in $L^2(a, \infty)$ which we denote by $H$, where the domain $D(H)$ is given by

$$D(H) = \{u \in L^2(a, \infty) : u(a) = 0, u \text{ and } u' \text{ absolutely continuous on } (a, \infty),\ u' \in L^2(a, \infty),\ -u'' + Vu \in L^2(a, \infty)\}\,.$$

Let $\sigma_{ess}(H)$ denote the essential spectrum of $H$ (for a definition of $\sigma_{ess}$ see, e.g., Ref. 16) then $\sigma_{ess}(H) = [0, \infty)$.[7,14] Furthermore, if $u$ is an $L^2$-solution of $(H - \lambda)u = 0$, then $u \in D(H)$ and $u(x) \to 0$ for $x$ tending to infinity.[14]

**Theorem 2.1:** Let $V$ satisfy (2.2). Suppose that $V = V_1 + V_2$ and $V_1$ is absolutely continuous for $x \geqslant x_0$, for some $x_0 \geqslant a$. Let $f$ be an absolutely continuous function, bounded, positive, and monotonically increasing for $x \geqslant x_0$. Suppose

$$V_2^2 \leqslant \frac{f'}{f} V_1' - \left(\frac{f'}{f}\right)^2 (V_1 - \lambda_0) \text{ and } V_1 < \lambda_0\,, \tag{2.3}$$

a.e., for $x \geqslant x_0$ for some $\lambda_0 \geqslant 0$. If $\lambda_0 > 0$, then (2.1) has no nontrivial $L^2$-solution for $\lambda \geqslant \lambda_0$. If $\lambda_0 = 0$ and if either

$$V(x) \leqslant \frac{3}{4} x^{-2} \text{ or } \lim_{x \to \infty} \int_{x_0}^{x} V(y)dy = -\infty\,, \tag{2.4}$$

then (2.1) has no nontrivial $L^2$-solution for $\lambda \geqslant 0$.

a) Supported by "Fonds zur Förderung der wissenschaftlichen Forschung in Österreich" Projectnr.P4925.

 0022-2488/84/082490-04$02.50  

*Remark 2.1:* The assumption that $V$ is bounded and tends to zero for $x\to\infty$ can be somewhat weakened, however, the proof of Theorem 2.1 gets more involved.

By a special choice of the function $f$ in (2.3) we obtain the following:

*Corollary 2.1:* Assume in Theorem 2.1 instead of (2.3) that $V_1\leqslant 0$, $V_1'\geqslant 0$, and $V_2\in L^1(x_0,\infty)$. Then (2.1) has no nontrivial $L^2$-solution for $\lambda>0$.

*Proof of Theorem 2.1:* We consider first the case $\lambda_0>0$. Let $Hu=\lambda u$, for some $\lambda\geqslant\lambda_0$. Without loss $u$ can be chosen real. We assume that $u\in L^2$ and derive a contradiction. Since $V-\lambda_0<0$, Sturm's theorem asserts that $u$ oscillates near infinity (see, e.g., Wong[17]). Hence we can choose $x_0\leqslant x_1<x_2$ with $u'(x_1)=u'(x_2)=0$ with $x_2$ arbitrarily large. Obviously we have for $x\geqslant x_0$

$$(u'^2)'=(V_1-\lambda)(u^2)'+2V_2uu'\,. \tag{2.5}$$

Multiplying (2.5) by $f(V_1-\lambda)^{-1}$ and integrating from $x_1$ to $x_2$ leads to

$$(fu^2)(x_2)-(fu^2)(x_1)+\int_{x_1}^{x_2}\left(\frac{f}{(V_1-\lambda)}\right)'u'^2\,dx - \int_{x_1}^{x_2}u^2f'\,dx+2\int_{x_1}^{x_2}fV_2(V_1-\lambda)^{-1}uu'\,dx=0\,. \tag{2.6}$$

By the geometric arithmetic inequality

$$2\int_{x_1}^{x_2}fV_2(V_1-\lambda)^{-1}uu'\,dx \leqslant\int_{x_1}^{x_2}u^2f'\,dx+\int_{x_1}^{x_2}\frac{f^2}{f'}\left(\frac{V_2}{(V_1-\lambda)}\right)^2u'^2\,dx \tag{2.7}$$

and we get

$$(fu^2)(x_2)-(fu^2)(x_1) \geqslant-\int_{x_1}^{x_2}\left\{\left(\frac{f}{(V_1-\lambda)}\right)'+\frac{f^2}{f'}\left(\frac{V_2}{(V_1-\lambda)}\right)^2\right\}u'^2\,dx\,. \tag{2.8}$$

Choosing $x_2$ sufficiently large the left-hand side of (2.8) becomes negative, since $fu^2\to 0$ for $x\to\infty$. (Note that $u^2+u'^2>0$.) But the right-hand side of (2.8) is positive due to the condition (2.3) which leads to a contradiction.

For the case $\lambda_0=0$ the proof runs the same way as for $\lambda_0>0$. We just have to show that $u$ oscillates near infinity. But this is a consequence of condition (2.4) which can be seen, e.g., from Knowles,[8] respectively, from Wong.[17]

*Proof of Corollary 2.1:* If $V_2\not\equiv 0$ choose $f=\int_{x_0}^{x}|V_2(x')|dx'$. Though $f'$ might vanish it is an admissible choice for $f$ as is easily seen by following the proof of Theorem 2.1. Since $V_1\leqslant 0$, $V_1'\geqslant 0$, and $\int_{x_0}^{\infty}|V_2(x)|dx\to 0$ for $x_0\to\infty$ condition (2.3) holds for $x_0$ sufficiently large and the corollary is implied. For $V_2\equiv 0$ set $f=1-x^{-1}$ for instance.

*Remark 2.2*: (a) The corollary is covered by a result of Weidmann.[7] Particularly for the case $V=V_2$ the result is known from Wallach.[18]

(b) Theorem 2.1 has obvious extensions to the $n$-dimensional case if $V(x)=V(|x|), x\in R^n$. We suspect that there are also generalizations to the nonspherical case.

(c) Condition (2.4) can be replaced by other assumptions, and there is extensive literature on oscillation properties of eigenfunctions of Sturm Liouville problems, see, e.g., Ref. 17.

(d) As can be seen from inequality (2.8) we can replace in Theorem 2.1 the condition "$f$ bounded" by the following condition: Assume $u\in L^2, u'(x_n)=0$ for all $n$, where $x_n\to\infty$ for $n\to\infty$, then $(fu^2)(x_n)\to 0$ for $n\to\infty$.

With the aid of Remark 2.2(d) we obtain

*Corollary 2.3:* Let $V=-cx^{-m}+W$ for some $c>0$, $0<m\leqslant 1$, with $|W|<\mu\sqrt{\lambda_0}\,x^{-1}$ for some $0<\mu<m$, $\lambda_0>0$. Then (2.1) has no nontrivial $L^2$-solution for $\lambda\geqslant\lambda_0$.

*Proof*: Choose $V_1=-cx^{-m}$, $V_2=W$ and $f(x)=x^{\mu}$, $\mu<m$. Then it is easily verified that condition (2.3) is fulfilled. In order to apply Theorem 2.1 [see Remark 2.2(d)] we must show that, if $u\in L^2$ and $u'(x_n)=0$ for all $n$, where $x_n\to\infty$ for $n\to\infty$, then $x_n^{\mu}u^2(x_n)\to 0$ for $n\to\infty$. But this can be seen from the following: Let $u'(x_i)=0$, $i=1,2$. Multiplying (2.1) by $x^m u'(x)$ we obtain by partial integration

$$m\int_{x_1}^{x_2}x^{m-1}u'^2\,dx+\frac{\lambda m}{2}\int_{x_1}^{x_2}x^{m-1}u^2\,dx +\int_{x_1}^{x_2}Vx^m uu'\,dx=\frac{\lambda}{2}x^mu^2\Big|_{x_1}^{x_2}\,.$$

Since $|V|\leqslant dx^{-1}$ for some $d<\infty$, we have

$$\left|\int_{x_1}^{x_2}Vx^muu'\,dx\right| \leqslant\frac{d}{2}\left(\int_{x_1}^{x_2}x^{m-1}u^2\,dx+\int_{x_1}^{x_2}x^{m-1}u'^2\,dx\right)$$

leading to

$$\frac{\lambda}{2}x_2^m\,u^2(x_2)\leqslant\frac{\lambda}{2}x_1^m\,u^2(x_1)+b\left(\int_{x_1}^{\infty}x^{m-1}u^2\,dx+\int_{x_1}^{\infty}x^{m-1}u'^2\,dx\right)<\infty$$

for some $0<b<\infty$, where we used $m\leqslant 1$ and $u, u'\in L^2$. Since this holds for arbitrarily large $x_2$ with $u'(x_2)=0$, $x_n^{\,m}\,u^2(x_n)\leqslant C<\infty$ for all $n$ follows.

*Remark 2.3:* For $m=1$ this result is implicit in Corollary 3.4 of Knowles.[8]

## III. EXAMPLES

We start with a specific construction on which our examples are based.

*Lemma 3.1:* Let $a_0=b_0=1$ and $a_n, b_n>0$, $n>1$. Let

$$N_0=0,\quad N_n=\frac{1}{2}\sum_{k=0}^{n-1}\left(\frac{1}{a_k}+\frac{1}{b_k}\right),\quad n\geqslant 1, \tag{3.1}$$

$$M_0=1/2a_0,\quad M_n=1/2a_n+N_n,\quad n\geqslant 1.$$

Let $c_0=1$ and

$$c_n=c_{n-1}b_{n-1}/a_n\quad\text{for } n\geqslant 1\,. \tag{3.2}$$

Then

$$-u'' + W(x)u = 0 \quad \text{for } x \geqslant 0, \tag{3.3}$$

where

$$u(x) = \sum_{n=0}^{\infty} (-1)^n c_n \{\sin(a_n \pi(x - N_n))\chi_N^{(n)} + \cos(b_n \pi(x - M_n))\chi_M^{(n)}\} \tag{3.4}$$

with $\chi_N^{(n)}$, respectively, $\chi_M^{(n)}$ denoting the characteristic functions of the interval $[N_n, M_n)$, respectively, $[M_n, N_{n+1})$ and

$$W(x) = -\pi^2 \sum_{n=0}^{\infty} (a_n^2 \chi_N^{(n)} + b_n^2 \chi_M^{(n)}) . \tag{3.5}$$

*Proof*: A direct calculation.

*Remark 3.1:* If one chooses $W(x) = W(-x)$ and $u(x) = -u(-x)$ for $x < 0$, (3.3) holds for the whole real axis.

In the following $c_\pm$, $d_\pm$ will denote suitable positive constants.

**Theorem 3.1:** There is a potential $V(x)$ obeying

$$-c_-(x+1)^{-1} \leqslant V(x) \leqslant -c_+(x+1)^{-1} \quad \text{for } x \geqslant 0, \tag{3.6}$$

$c_- > c_+ > 0$, such that

$$-u'' + (V - \pi^2)u = 0 \quad \text{for } x \geqslant 0, \tag{3.7}$$

with $u(0) = 0$, $u \in L^2[0, \infty)$. In particular

$$d_-(x^2+1)^{-(1-\gamma)} \leqslant (u^2 + u'^2)(x) \leqslant d_+(x^2+1)^{-(1-\gamma)} \tag{3.8}$$

with $\gamma \in (0, \frac{1}{2})$ depending on $V$, and $d_- < d_+ < \infty$.

*Proof*: We use the construction of Lemma 3.1. We set

$$V(x) = W(x) + \pi^2 \tag{3.9}$$

and

$$a_n = 1 + n^{-1}, \quad b_n = 1 + \gamma n^{-1} \quad \text{for } n \geqslant 1, \gamma \in (0, \tfrac{1}{2}) . \tag{3.10}$$

Therefore, for $x < 1$, $V(x) = -\pi^2$ and for $x \geqslant 1$

$$V(x) = -\pi^2 \sum_{n=1}^{\infty} \{(2n^{-1} + n^{-2})\chi_N^{(n)} + (2n^{-1}\gamma + n^{-2}\gamma^{-2})\chi_M^{(n)}\} . \tag{3.11}$$

In order to verify (3.6) we note that for $x \in [N_n, N_{n+1})$

$$\frac{1}{2}(n-1) \leqslant \sum_{k=1}^{n-1} (1+k)^{-1} \leqslant N_n \leqslant x \leqslant N_{n+1} \leqslant \sum_{k=1}^{n} (1 + k^{-1}\gamma)^{-1} + 1 \leqslant (1+\delta)n , \tag{3.12}$$

where $\delta > 0$ arbitrarily small and $n$ sufficiently large. Following now immediately from (3.11) and (3.12) is (3.6). Finally we show (3.8). By (3.4) we have

$$(u^2 + u'^2)(x) = \sum_{n=0}^{\infty} c_n^2 \{(a_n^2\pi^2 + (1 - \pi^2 a_n^2) \times \sin^2(a_n \pi(x - N_n)))\chi_N^{(n)} + (b_n^2\pi^2 + (1 - \pi^2 b_n^2) \times \cos^2(b_n \pi(x - M_n)))\chi_M^{(n)}\} . \tag{3.13}$$

By our choice (3.10) of $a_n, b_n$, it is easily seen that for $n$ large

$$c_n^2 \leqslant (u^2 + u'^2)(x) \leqslant \pi^2 a_n^2 c_n^2 \quad \text{for } x \in [N_n, N_{n+1}) . \tag{3.14}$$

Noting that by (3.2) for suitable constants $k_+ > k_- > 0$,

$$c_n^2 = (1 + n^{-1})^{-2} \prod_{k=1}^{n-1} (k+\gamma)^2 (k+1)^{-2} \leqslant k_+ (\Gamma(n+\gamma))^2 (\Gamma(n+1))^{-2}$$

and using that

$$\Gamma(x) = x^{x - 1/2} e^{-x} (2\pi)^{1/2} (1 + O(1/x))$$

for large $x$, we obtain for $n$ large

$$c_n^2 \leqslant c_+ n^{2(\gamma - 1)} . \tag{3.15}$$

We use now (3.15) and (3.12) to obtain from (3.14) the desired result.

*Remark 3.2:* By working out $\int_0^\infty u^2\, dx$ according to (3.4) we could have gotten directly conditions on $a_n, b_n$ which ensure $u \in L^2$. We included the asymptotics of $u$ since we consider the relation of the asymptotics of $V$ and $u$ interesting. See Ref. 19 for a discussion.

*Remark 3.3:* As can be seen from Corollary 2.3 one cannot choose $c_+$ and $c_-$ in (3.6) arbitrarily close.

Similar to the preceding result one can show the following:

**Theorem 3.1′:** Let $0 < \alpha < 1$, then there exists a $V(x)$ so that

$$-c_-(x+1)^{-\alpha} \leqslant V(x) \leqslant -c_+(x+1)^{-\alpha}$$

and

$$-u'' + Vu = \pi^2 u \quad \text{for } x \geqslant 0,$$

where

$$d_- \exp\{-m(1+\delta)x^{1-\alpha}\} \leqslant (u^2 + u'^2)(x) \leqslant d_+ \exp\{-m(1-\delta)x^{1-\alpha}\}$$

with $\delta > 0$ arbiitrarily small, $m$ a suitable positive constant.

*Proof*: Choosing $a_n = 1 + n^{-\alpha}$, $b_n = 1 + \gamma n^{-\alpha}$, $\gamma \in (0,1)$ and again setting $V(x) = W(x) + \pi^2$ in our construction the result above can be verified. We omit details.

Note the difference in the asymptotic behavior of the eigenfunctions $u$ in Theorem 3.1 and Theorem 3.1′.

We continue with examples for zero eigenvalues.

**Theorem 3.2:** There is a potential $V(x)$ defined on $[0, \infty)$ with

$$-c_-(x+1)^{-2} \leqslant V(x) \leqslant -c_+(x+1)^{-2} \tag{3.16}$$

so that $-u'' + Vu = 0$ on $[0, \infty)$. $u^2 + u'^2$ decays polynomially.

*Proof*: We again use the construction of Lemma 3.1. Set $V = W$ and choose

$$a_n = e^{-\alpha n}, \quad b_n = \gamma e^{-\alpha n} \quad \text{for } n \geqslant 1, \alpha, \gamma > 0 \text{ and } \gamma < e^{-\alpha} . \tag{3.17}$$

It is easily seen that for $x \in [N_n, N_{n+1})$ and $n$ large

$$x \leqslant N_{n+1} \leqslant 1 + \frac{1}{2}\left(1 + \frac{1}{\gamma}\right) \sum_{k=1}^{n} e^{\alpha k} \leqslant 1 + \frac{1}{2\alpha}\left(1 + \frac{1}{\gamma}\right) \int_1^{n+1} e^{\alpha x}\, dx \leqslant C_1(\alpha,\gamma) e^{\alpha n} , \tag{3.18}$$

with $C_1(\alpha,\gamma) > 0$ depending on $\alpha$ and $\gamma$. Since

$$V(x) \geqslant -\pi^2 e^{-2\alpha n} \quad \text{for } x \in [N_n, M_{n+1}) , \tag{3.19}$$

$$V(x) \geqslant -\pi^2 C_1(\alpha,\gamma)^2 x^{-2}$$

for large $x$. On the other hand we have

$$x \geqslant N_n \geqslant 1 + \frac{1}{2}\left(1 + \frac{1}{\gamma}\right)\int_0^{n-1} e^{\alpha x}\,dx \geqslant C_2(\alpha,\gamma)e^{\alpha n}. \tag{3.20}$$

But $V(x) \leqslant -\pi^2\gamma^2 e^{-2\alpha n}$ for $x \in [N_n, N_{n+1})$ and therefore,

$$V(x) \leqslant -\pi^2\gamma^2 C_2^2(\alpha,\gamma)^2 x^{-2} \tag{3.21}$$

for $x$ large. The boundedness of $V$ together with (3.19) and (3.21) imply (3.16).

It remains to show that $u \in L^2(0,\infty)$ and that $u^2 + u'^2$ decays polynomially. Let us consider now (3.13). The choice (3.17) of $a_n, b_n$ implies that for large $n$

$$\pi^2 c_n^2 b_n^2 \leqslant u^2 + u'^2 \leqslant c_n^2 \quad \text{for } x \in [N_n, N_{n+1}). \tag{3.22}$$

But

$$c_n^2 = a_n^{-2}\prod_{k=1}^{n-1}\left(\frac{b_k}{a_k}\right)^2 = \exp\left(2n(\alpha + \ln\gamma)\right). \tag{3.23}$$

Note that because of (3.17) $\alpha + \ln\gamma < 0$. Since by (3.4)

$$\int_0^\infty u^2\,dx \leqslant \sum_{n=0}^{\infty} c_n^2, \tag{3.24}$$

the quotient criterium implies $u \in L^2(0,\infty)$. Finally we apply (3.18) and (3.20) to (3.23) and combine the resulting inequalities with (3.22). Therefore, we get

$$d_- x^{-2(1+\mu)} \leqslant u^2 + u'^2 \leqslant d_+ x^{-2\mu}$$

for large $x$ with $1 + \alpha^{-1}\ln\gamma = -\mu$.

*Remark 3.4:* Knowles[8] proves that if $-(m^2 + 2m + \frac{1}{4}) \leqslant V(x)x^2 \leqslant -(m^2 - 2m + \frac{1}{4})$, with some $m > 0$, for $x$ large then $-u'' + Vu = 0$ implies $u \notin L^2$.

Similar to the preceding result one can show the following:

**Theorem 3.2':** Let $0 < \epsilon < 1$, then there is a $V(x)$ such that

$$-c_-(x+1)^{-2(1-\epsilon)} \leqslant V(x) \leqslant -c_+(x+1)^{-2(1-\epsilon)}$$

and $-u'' + Vu = 0$ for $x \geqslant 0$, where $d_- \exp(-mx^\epsilon) \leqslant u^2 + u'^2 \leqslant d_+ \exp(-(1-\delta)mx^\epsilon)$. Here $c_\pm, d_\pm, m$ are suitable positive constants and $\delta$ can be chosen arbitrarily small.

*Proof:* Choose $a_n = n^{-\alpha}$, $b_n = n^{-\alpha}\gamma$ for $n \geqslant 1, 0 < \gamma < 1$ with $\alpha \equiv \epsilon^{-1} - 1 > 0$. We omit details.

## ACKNOWLEDGMENTS

We thank J. Weidmann for a helpful discussion and H. Grosse for a useful remark which helped to improve an earlier version of Theorem 2.1. Further we thank E. Harrell for bringing Eastham and Kalf's book to our attention.

[1] S. Agmon, J. Analyse Math. **23**, 1 (1970).
[2] T. Kato, Commun. Pure Appl. Math. **12**, 403 (1959).
[3] B. Simon, Commun. Pure Appl. Math. **22**, 531 (1969).
[4] R. Froese, I. Herbst, M. Hoffmann-Ostenhof, and T. Hoffmann-Ostenhof, J. Analyse Math. **41**, 272 (1982).
[5] R. Froese and I. Herbst, Commun. Math. Phys. **87**, 429 (1982).
[6] M. Reed and B. Simon, *Methods of Modern Mathematical Physics IV: Analysis of Operators* (Academic, New York, 1978).
[7] J. Weidmann, Math. Z. **98**, 268 (1967).
[8] I. Knowles, Proc. R. Soc. Edinburgh Sec. A **84**, 197 (1979).
[9] M. S. P. Eastham, *Lecture Notes in Mathematics, No. 564* (Springer, Berlin, 1976), p. 72.
[10] F. V. Atkinson and W. N. Everitt, Proc. R. Soc. Edinburgh Sec. A **80**, 57 (1978).
[11] M. S. P. Eastham and H. Kalf, *Schrödinger Operators with Continuous Spectrum* (Pitman, London, 1983).
[12] J. von Neumann and E. Wigner, Z. Phys. **30**, 465 (1929).
[13] C. Thurlow, Proc. R. Soc. Edinburgh Sec. A **84**, 197 (1979).
[14] B. Simon, Bull. Am. Math. Soc. **7**, 447 (1982).
[15] T. Kato, *Perturbation Theory for Linear Operators* (Springer, Berlin, 1976).
[16] M. Reed and B. Simon, *Methods of Modern Mathematical Physics I: Functational Analysis* (Academic, New York, 1972).
[17] J. S. W. Wong, Trans. Am. Math. Soc. **144**, 197 (1969).
[18] S. Wallach, Am. J. Math. **70**, 833 (1948).
[19] R. Froese, I. Herbst, M. Hoffmann-Ostenhof, and T. Hoffmann-Ostenhof, Proc. R. Soc. Edinburgh Sec. A. **95**, 25 (1983).

# On the inverse scattering transform of multidimensional nonlinear equations related to first-order systems in the plane

A. S. Fokas and M. J. Ablowitz
*Department of Mathematics and Computer Science, Clarkson College of Technology, Potsdam, New York 13676*

(Received 17 February 1983; accepted for publication 30 December 1983)

The inverse problem associated with a rather general system of $n$ first-order equations in the plane is linearized. When the system is hyperbolic, this is achieved by utilizing a Riemann–Hilbert problem; similarly, a "$\bar{\partial}$" (DBAR) problem is used when the system is elliptic. The above result can be employed to linearize the initial value problem associated with a variety of physically significant equations in $2+1$, i.e., two spatial and one temporal dimensions. Concrete results are given for the $n$-wave interaction in $2+1$ and for various forms of the Davey–Stewartson equations. Lump solutions (solitons in $2+1$) of the latter equation are given a definitive spectral characterization and are obtained through a linear system of algebraic equations.

PACS numbers: 03.65.Nk

## I. INTRODUCTION

This paper is the seventh in a series of articles dedicated to the study of (a) solving inverse problems in the plane for appropriately decaying potentials [i.e., given suitable scattering data reconstruct the potential $q(x,y)$]; (b) solving the initial value problem of certain evolution equations in $2+1$, i.e., two spatial and one temporal dimensions [i.e., given $q(x,y,0)$ and assuming it is sufficiently decaying for large $x, y$, find $q(x,y,t)$]. Our program of study began with a certain nonlinear integro-differential evolution equation, the so-called Benjamin–Ono equation (BO).[1] We have found that the BO equation has many features similar to multidimensional problems; in this sense it acts as a pivot from $1+1$ to $2+1$. The inverse problem associated with the "time"-dependent Schrödinger equation

$$i\psi_y + \psi_{xx} + (q+\lambda)\psi = 0\,, \tag{1.1}$$

as well as the initial value problem of the related Kadomtsev–Petviashvili, KPI, equation

$$(q_t + 6qq_x + q_{xxx})_x = 3q_{yy}, \tag{1.2}$$

were considered by the authors in Refs. 2 and 3. The inverse problem associated with

$$-\psi_y + \psi_{xx} + (q+\lambda)\psi = 0\,, \tag{1.3}$$

and the related KPII equation

$$(q_t + 6qq_x + q_{xxx})_x = -3q_{yy} \tag{1.4}$$

were considered by Bar Yaacov and the authors in Ref. 4. (We note that both KPI and KPII arise naturally in physical contexts. For example, in water waves the sign of $q_{yy}$ depends on the relevant magnitude of gravity and surface tension.) The inverse problem associated with

$$\psi_x = \lambda B\psi + q\psi + J\psi_y \tag{1.5}$$

was considered in Refs. 5 and 6. In Eq. (1.5), $B$ and $J$ are constant $n\times n$ diagonal matrices and $q(x,y,t)$ is an $n\times n$ off-diagonal matrix containing the potentials (or field variables). The hyperbolic (i.e., $J$ real) and aspects of the elliptic (i.e., $J$ imaginary) versions of (1.5) were considered in a rather compact form in Ref. 5. The elliptic case with lumps (i.e., decaying solitons in $2+1$) was considered in Ref. 6. We point out that the manifestation of lumps is one of the novel features of the inverse scattering transform (IST) in $2+1$. Also in Ref. 6 a method for solving inverse problems in the plane as well as solving the initial value problems of the related nonlinear equations was announced. Equation (1.5) can be used to solve various physical nonlinear equations in $2+1$. Among them are the $n$-wave interaction,[7] the so-called Davey–Stewartson (DS) equation[8] (which is the long wave limit of the Benney–Roskes equation[9]), and the modified KP equation (MKP). Actually, in analogy with KP there exist DSI and DSII as well as MKPI and MKPII.

In this paper we investigate, in more detail, the inverse problem associated with (1.5) as well as the initial value problems of some related nonlinear multidimensional equations. In particular in Sec. II we outline the essential steps of the generalized IST we use for solving multidimensional problems. In Secs. III and IV we consider the hyperbolic and elliptic versions of (1.5), respectively. The existence or not of lumps for the related nonlinear equations, is equivalent to the existence or not of homogeneous solutions of a certain linear integral equation characterizing suitable Jost eigenfunctions. In the elliptic case we investigate these homogeneous modes and show how they generate lumps. Pure lumps can be computed through a linear system of algebraic equations (for the hyperbolic case we do not expect lumps). In Sec. V we show how the results of Secs. III–IV can be used for solving the initial value problem of certain nonlinear equations. Considering a specific nonlinear equation implies a specific time-dependent part of the Lax pair associated with the given equation. In Sec. V we indicate how this time-dependent part of the Lax pair uniquely determines the evolution of the underlying scattering data. Once this evolution has been determined the results of Secs. III–IV can be directly used for the exact integration of the given equation. Concrete results are given for the $n$-wave interaction, for DSI and for DSII. MKPI and MKPII can be treated in an exactly analogous manner.

A review of our work on multidimensional IST as well

 0022-2488/84/082494-12$02.50  

as some related aspects of one-dimensional problems can be found in Ref. 10.

We note that in Refs. 1–6 we make some assumptions about the compactness of certain operators, the existence or not of eigenvalues, etc. Some of these assumptions can be justified by assuming that a certain norm of $q$ is sufficiently small. A rigorous investigation of these and other questions (e.g., uncovering some constraints on the scattering data) is presently under investigation.

## II. THE GENERAL FRAMEWORK

Prior to our investigation of multidimensional IST, in connection with equations in 2 + 1 it was established that:

(i) Many physically important multidimensional nonlinear equations are related to pairs of linear systems, the so-called Lax pairs[11] (i.e., to pairs of linear systems the solvability condition of which implies the given equation). In particular KP, DS, MKP, and $n$-wave interaction were related to Lax pairs with time-independent parts given by (1.1), (1.3), and (1.5).

(ii) The above multidimensional equations have interesting algebraic properties. In particular, KP possesses infinitely many symmetries and conserved quantities[12,13] (see also Ref. 14).

(iii) There exist several direct methods for finding *particular solutions* of the above equations. Such methods are (a) appropriate use of Bäcklund transformation[15]; (b) the bilinear approach of Hirota[16]; (c) the more general $\tau$-function approach of Sato, Miwa, Jimbo, Date, and Kasiwara related to Kac–Moody algebras[17]; (d) the "dressing method" of Zakharov–Shabat[18] [this method exploits the fact that there exist linear integral equations of the Gel'fand–Levitan–Marchenko (GLM) type which have the property that their solutions also solve certain nonlinear equations in 2 + 1]; (e) the extension of the "Riemann–Hilbert (RH) method" of Zakharov, Shabat, and Mikhailov,[19] proposed by Manakov[20] (this method uses certain nonlocal RH boundary value problems to directly obtain solutions of several nonlinear equations in 2 + 1); (f) the direct linearizing method proposed by the authors[21,2] and extended recently by Santini and the authors[22] (see also Ref. 23). (This method is closely related to the perturbation approach of Rosales[24] and exploits the fact that there exist rather general linear integral equations, involving arbitrary measure contour, which can be used to linearize certain nonlinear equations in 2 + 1.) For the interrelations of (d)–(f), see Ref. 22.

In spite of the above progress, the question of finding a suitable method for solving *the initial value problems* of multidimensional equations was essentially open. In this regard we mention that some interesting results had been obtained about KPI[25,26] and about the three-wave interaction.[27] However, it was not clear from this work how a viable unified scheme to handle the above and other equations could be obtained. In particular in Manakov's treatment of KPI the usual IST had to be supplemented with solving an additional pair of GLM-type equations in scattering space; also the lump solutions (algebraically decaying solitons) were excluded. Similarly Kaup's treatment of the three-wave interaction exploits crucially the existence of characteristic coordinates. Furthermore, what is perhaps more important, the IST had been considered so far, only within the framework of RH problems (local in 1 + 1, nonlocal in KPI). However, this framework seems in general inadequate for handling other multidimensional problems (e.g., KPII).

We have recently developed a method for both solving inverse problems in the plane as well as solving initial value problems of certain nonlinear equations in 2 + 1. This method stems from the work of the authors on BO, KPI, hyperbolic systems, and elliptic systems and from the work of the authors and Bar Yaacov on KPII. We would like to emphasize that the investigation of KPII was of crucial importance in the development of this method, since it was the first case the inadequacy of the RH formulation of the IST was discovered. From the treatment of KPII became clear that the RH problem had to be replaced by the so-called "$\bar\partial$" (DBAR) problem.

The main steps of this method can be summarized as follows:

(i) Define an eigenfunction $\mu(x, y,\kappa)$, which is bounded for all complex values of the "spectral parameter" $\kappa$ and which is appropriately normalized ($\mu \to 1$ as $\kappa \to \infty$). This eigenfunction is usually defined in terms of a Fredholm linear integral equation of the second type, and it may have different representations in different sections of the complex $\kappa$ plane. To derive this integral equation, one regards the terms involving $q(x, y)$ as a "forcing" and uses a suitable Fourier transform or a Green's function formulation. The above integral equation may have homogeneous solutions. These homogeneous modes (corresponding to discrete eigenfunctions) are rather important because they give rise to lumps (i.e., 2 + 1 decaying solitons).

(ii) Compute $\partial\mu/\partial\bar\kappa$. This is in general expressed in terms of some other bounded eigenfunction, which we call $N(x, y,l,\kappa)$, and appropriate scattering data. We note that in some problems (e.g., BO, KPI) $\mu(x, y,\kappa)$ is a *sectionally meromorphic function of* $\kappa$, i.e., it is holomorphic, modulo poles, in regions of the complex $\kappa$ plane separated by certain contours and it has a jump across these contours. In these cases $\partial\mu/\partial\bar\kappa$, which measures the "departure of $\mu(x, y,\kappa)$ from holomorphicity," will be zero everywhere except on the pole locations and on the above contours.

(iii) Employ a suitable "symmetry" relationship between $N$ and $\mu$ to express $\partial\mu/\partial\bar\kappa$ in terms of $\mu$ and appropriate scattering data. If $\mu$ has homogeneous solutions, then one needs also to establish a relationship between $\mu$ and these homogeneous modes. We have so far encountered two types of symmetry conditions: "discrete" (KPII, elliptic systems) and "differential" (BO,KPI). The relationship between $\partial\mu/\partial\bar\kappa$, $\mu$, and scattering data, which we call the scattering equation, is the central equation associated with the inverse problem of a given equation. This scattering equation defines, in general, a "$\bar\partial$" problem, i.e., given $\partial\mu/\partial\bar\kappa$, find $\mu$. In the case that $\mu$ is sectionally meromorphic this "$\bar\partial$" problem degenerates to a RH problem.

(iv) Use the following extension of Cauchy's formula[28],

$$\mu(x, y,\kappa) = \frac{1}{2\pi i}\int\int_R \frac{[\partial\mu(x, y, z)/\partial\bar z]\, dz\wedge d\bar z}{z-\kappa}$$

$$+\frac{1}{2\pi i}\int_C \frac{\mu(x,y,z)}{z-\kappa}\,dz \tag{2.1}$$

(where $R$ and $C$ are appropriate region contour in the $z$ plane, respectively) to solve the "$\bar\partial$" problem. Its solution is found, in general, in terms of a linear integral equation for $\mu(x,y,\kappa)$. Equation (2.1) is uniquely defined in terms of the above-mentioned scattering data. If $\mu$ has a homogeneous solution at $\kappa_l$, then $\partial\mu/\partial\bar\kappa$ has a $\delta$ function at $\kappa=\kappa_l$; thus pure lump solutions are always found in closed form as the solution of a linear system of algebraic equations.

(v) Calculate the potential $q(x,y)$ directly from the solution of the inverse problem [typically given by integrals over $\mu(x,y,\kappa)$ and the scattering data]. The above discussion summarizes the steps needed for the solution of an inverse problem.

(vi) In order to solve the initial value problem of some related nonlinear evolution equation, one needs only to find the evolution of the scattering data. This can be achieved in a simple manner by employing the time-dependent part of the Lax pair of the given equation. Furthermore, the initial scattering data can always be expressed in terms of the initial data $q(x,y,0)$. Thus Eq. (2.1) [and hence the formula for $q(x,y,t)$] is uniquely defined in terms of the initial data.

At this point we would like to make two remarks: (i) Our motivation for using the "$\bar\partial$" problem came from the significant work of Beals and Coifman[29] on the IST of nonlinear equations in 1 + 1 related to first-order systems on the line. In their treatment, the inverse problem is formulated in terms of an RH problem with respect to certain rays in the complex plane. However, they indicate [in particular, see Ref. 29(c)] that the RH problem should be viewed as a special case of a "$\bar\partial$" problem. (ii) RH problems have also been useful in connection with two other recent significant discoveries: (a) the integration of the Ernst equation (the static, axisymmetric reduction of the vacuum Einstein's equations), in particular with regards to the constructive proof of the Geroch conjecture[30]; (b) the integration of the self-dual Yang–Mills (SDYM) equations in four-dimensional Euclidean space, in particular with respect to the Atiyah–Ward construction.[31] The SDYM equations, as a result of their special structure, although defining a four-dimensional model, have many properties similar to those of two-dimensional problems. Motivated from the above discussion we expect that the "$\bar\partial$" problem will also be useful for the exact integration of various models in both the fields of relativity as well as that of particle physics. Furthermore, we also expect these ideas to be useful for multidimensional differential–difference and pure difference equations.

## III. HYPERBOLIC SYSTEMS IN THE PLANE

We first consider the inverse problem associated with the hyperbolic system

$$\mu_x = i\kappa\hat J\mu + q\mu + J\mu_y, \quad \hat J f \doteq Jf - fJ, \tag{3.1}$$

where $\mu(x,y,\kappa)$ is an $n$th-order matrix, $J$ is a constant real diagonal matrix with elements $J_1 > J_2 \cdots > J_n$, and $q(x,y)$ is an $n$th-order off-diagonal matrix containing the potentials $q_{ij}(x,y)$. We assume that $q_{ij}(x,y)\to 0$ rapidly enough for large $x,y$. Equation (3.1) is obtained from the well-known[32] equation

$$\psi_x = i\lambda B\psi + q\psi + J\psi_y, \tag{3.2}$$

by taking $B=0$ [this is without loss of generality since all the equations solvable by (3.2) are independent of $B$ [32,14]] and using $\psi=\mu\exp[i\kappa(Jx+Iy)]$ ($I$ denotes the unit diagonal matrix).

### A. Bounded eigenfunctions

Let $\pi_0\mu$, $\pi_+\mu$, $\pi_-\mu$ denote the diagonal, strictly upper diagonal, and strictly lower diagonal parts of the matrix $\mu$. A solution of (3.1) bounded for all complex values of $\kappa=\kappa_R+i\kappa_I$ and tending to $I$ as $\kappa\to\infty$ is given by

$$\mu(x,y,\kappa)=\begin{cases}\mu^+(x,y,\kappa), & \kappa_I>0,\\ \mu^-(x,y,\kappa), & \kappa_I\leqslant 0,\end{cases} \tag{3.3}$$

where $\mu^\pm(x,y,\kappa)$ satisfy the following linear integral equations:

$$\begin{aligned}\mu^\pm(x,y,\kappa) = I &+ \frac{1}{2\pi}\int_{-\infty}^{x} d\xi\, E_{x-\xi,J}e^{i\kappa(x-\xi)\hat J}\\ &\times(\pi_0+\pi_\pm)q(\xi,\eta)\mu^\pm(\xi,\eta,\kappa)\\ &-\frac{1}{2\pi}\int_{x}^{\infty} d\xi\, E_{x-\xi,J}e^{i\kappa(x-\xi)\hat J}\\ &\times\pi_\mp q(\xi,\eta)\mu^\pm(\xi,\eta,\kappa).\end{aligned} \tag{3.4}^\pm$$

In Eqs. $(3.4)^\pm$ the linear operator $E_{x-\xi,J}$ is defined by

$$\begin{aligned}&(E_{x-\xi,J}f(\cdot))(x-\xi,y)\\ &\doteq\int_{-\infty}^{\infty}d\eta\int_{-\infty}^{\infty}dm\, e^{im(x-\xi)J+im(y-\eta)}f(\eta)\\ &=f(y+(x-\xi)J),\end{aligned} \tag{3.5}$$

where $f(y+(x-\xi)J)$ denotes the matrix obtained from $f(x)$ by evaluating the $l$th row of the matrix $f(x)$ at $y+(x-\xi)J_l$. Also from the definition of $J$ it follows that

$$e^{\hat J}f = e^{J}fe^{-J}. \tag{3.6}$$

Equations $(3.4)^\pm$ can be derived by using a Green's function approach similar to the one used in Refs. 3 and 4. Alternatively, one may use Fourier transforms in the $y$ direction. For this it is more convenient to work with Eq. (3.2). Let

$$\begin{aligned}\tilde\psi(x,m)&\doteq\int_{-\infty}^{\infty}\psi(x,\eta)e^{-im\eta}\,d\eta, \quad \psi(x,y)\\ &=\frac{1}{2\pi}\int_{-\infty}^{\infty}\tilde\psi(x,m)e^{imy}\,dm.\end{aligned} \tag{3.7}$$

Then taking the $y$-Fourier transform of (3.2) with $\lambda=0$ it follows that

$$\begin{aligned}\tilde\psi(x,m) = e^{imJx}A(m) &+ \int_{-\infty}^{x}d\xi\int_{-\infty}^{\infty}d\eta\\ &\times e^{imJ(x-\xi)}q(\xi,\eta)\psi(\xi,\eta)e^{-im\eta},\end{aligned}$$

where $A(m)$ is an arbitrary function of $m$. Hence

$$\begin{aligned}\psi(x,y) = &\frac{1}{2\pi}\int_{-\infty}^{\infty}dm\, e^{imJx+imy}A(m)\\ &+\frac{1}{2\pi}\int_{-\infty}^{x}d\xi\int_{-\infty}^{\infty}d\eta\int_{-\infty}^{\infty}dm\end{aligned}$$

$$\times e^{imJ(x-\xi)+im(y-\eta)}q(\xi,\eta)\psi(\xi,\eta)\,. \tag{3.8}$$

Using $\psi=\mu\exp[i\kappa(Jx+y)]$ it follows that a solution of (3.1) is given by

$$\mu(x,y,\kappa)=\frac{1}{2\pi}\int_{-\infty}^{\infty}dm\,e^{imJx+imy+i\kappa\hat{J}x}A(m)$$
$$+\frac{1}{2\pi}\int_{-\infty}^{x}d\xi\int_{-\infty}^{\infty}d\eta\int_{-\infty}^{\infty}dm$$
$$\times e^{imJ(x-\xi)+im(y-\eta)}e^{i\kappa\hat{J}(x-\xi)}q(\xi,\eta)\mu(\xi,\eta,\kappa)\,. \tag{3.9}$$

The first integral of (3.9) represents a "homogeneous" solution of (3.1), while the second a particular one. Equations $(3.4)^{\pm}$ can now be easily derived by properly choosing $A(m)$, so that boundness for all complex values of $\kappa$ is achieved. In particular,

$$\int_{-\infty}^{x}d\xi\,e^{i\kappa(x-\xi)\hat{J}}f(\xi)$$
$$=\int_{-\infty}^{x}d\xi\,e^{i\kappa(x-\xi)\hat{J}}(\pi_0+\pi_++\pi_-)f(\xi)$$
$$=\int_{-\infty}^{x}d\xi\,e^{i\kappa(x-\xi)\hat{J}}(\pi_0+\pi_+)f(\xi)$$
$$-\int_{x}^{\infty}d\xi\,e^{i\kappa(x-\xi)\hat{J}}\pi_-f(\xi)$$
$$+\int_{-\infty}^{\infty}dl\,e^{i\kappa(x-\xi)\hat{J}}\pi_-f(\xi)\,.$$

Note that the first two integrals are bounded for $\kappa_I>0$, while the third integral can be absorbed in the homogeneous solution by choosing

$$A(m)=-2\pi\int_{-\infty}^{\infty}d\xi\int_{-\infty}^{\infty}d\eta$$
$$\times e^{-imJ\xi-im\eta-i\kappa\xi\hat{J}}\pi_-\mu^+(\xi,\eta,\kappa)+\delta(m)I\,. \tag{3.10}$$

Equation $(3.4)^-$ is derived in a similar way.

### B. Scattering equation and scattering data

*Assuming that the linear integral equations $(3.4)^{\pm}$ have no homogeneous solutions*, it follows that $\mu^+$ and $\mu^-$ are holomorphic functions of $\kappa$, for $\kappa_I>0$ and $\kappa_I<0$, respectively. Hence the function $\mu(x,y,\kappa)$ defined by (3.3) is a *sectionally holomorphic function of $\kappa$ having a jump across $\kappa_I=0$.* Thus $\partial\mu/\partial\bar{\kappa}=0$ for all $\kappa$ with $\kappa_I\neq0$ and $\partial\mu/\partial\bar{\kappa}=\mu^+(x,y,\kappa)-\mu^-(x,y,\kappa)$ for $\kappa=\kappa_R$. Rather than following in detail the method of Sec. 2, i.e., computing $\mu^+-\mu^-$ in terms of some other bounded eigenfunction $N$ and then establishing a "symmetry" condition relating $\mu$ and $N$, we find it more convenient to obtain directly a scattering equation:

$$\mu^+(x,y,\kappa)-\mu^-(x,y,\kappa)=\int_{-\infty}^{\infty}dl\,\mu^-(x,y,l)e^{ilJx+ily}f(l,\kappa)$$
$$\times e^{-i\kappa Jx-i\kappa y},\quad \kappa=\kappa_R\,. \tag{3.11}$$

In Eq. (3.11) the scattering data $f(l,\kappa)$ satisfies

$$f(l,\kappa)-\int_{-\infty}^{\infty}dm\,T_+(l,m)f(m,\kappa)$$
$$=T_+(l,\kappa)-T_-(l,\kappa),\quad l,\kappa\ \text{real}, \tag{3.12}$$

where

$$T_{\pm}(l,\kappa)\doteq\frac{1}{2\pi}\int_{-\infty}^{\infty}d\xi\int_{-\infty}^{\infty}d\eta\,e^{-ilJ\xi-il\eta}\pi_{\pm}q(\xi,\eta)$$
$$\times\mu^{\mp}(\xi,\eta,\kappa)e^{i\kappa\xi J+i\kappa\eta},\quad l,\kappa\ \text{real}\,. \tag{3.13}^{\pm}$$

Before indicating how the above equations can be derived, we note the remarkable fact that Eq. (3.12) can be solved in *closed form*. This is because its kernel is strictly upper triangular.

*Example*: Suppose that $n=2$, i.e.,

$$q(x,y)=\begin{pmatrix}0 & q_1(x,y)\\ q_2(x,y) & 0\end{pmatrix},\quad T_+(l,\kappa)=\begin{pmatrix}0 & T_{+12}(l,\kappa)\\ 0 & 0\end{pmatrix},$$
$$T_-(l,\kappa)=\begin{pmatrix}0 & 0\\ T_{-21}(l,\kappa) & 0\end{pmatrix}, \tag{3.13}$$

where

$$T_{+12}(l,\kappa)\doteq\frac{1}{2\pi}\int_{-\infty}^{\infty}d\xi\int_{-\infty}^{\infty}d\eta$$
$$\times\{q(\xi,\eta)\mu^-(\xi,\eta,\kappa)\}_{12}e^{i\kappa\xi J_2+i\kappa\eta-il\xi J_1-il\eta},$$
$$T_{-21}(l,\kappa)\doteq\frac{1}{2\pi}\int_{-\infty}^{\infty}d\xi\int_{-\infty}^{\infty}d\eta$$
$$\times\{q(\xi,\eta)\mu^+(\xi,\eta,\kappa)\}_{21}e^{i\kappa\xi J_1+i\kappa\eta-il\xi J_2-il\eta}.$$

Then Eq. (3.12) implies

$$f_{22}(l,\kappa)=0\,,\quad f_{21}(l,\kappa)=-T_{+12}(l,\kappa)\,,\quad f_{12}(l,\kappa)=T_{-12}(l,\kappa)\,,$$
$$f_{11}(l,\kappa)=-\int_{-\infty}^{\infty}dm\,T_{+12}(l,m)T_{-21}(m,\kappa)\,.$$

Equation (3.12) can be solved in a similar manner for any $n$.

To derive Eq. (3.11), let

$$\Delta(x,y,\kappa)\doteq\mu^+(x,y,\kappa)-\mu^-(x,y,\kappa)\,,\quad \kappa\ \text{real}\,.$$

Then Eqs. $(3.4)^{\pm}$ imply that $\Delta$ satisfies

$$\Delta(x,y,\kappa)=\frac{1}{2\pi}\int_{-\infty}^{\infty}d\xi\,E_{x-\xi,J}e^{i\kappa(x-\xi)\hat{J}}$$
$$\times(\pi_+q(\xi,\eta)\mu^-(\xi,\eta,\kappa)-\pi_-q(\xi,\eta)\mu^+(\xi,\eta,\kappa))$$
$$+\frac{1}{2\pi}\int_{-\infty}^{x}d\xi\,E_{x-\xi,J}$$
$$\times e^{i\kappa(x-\xi)\hat{J}}q(\xi,\eta)\Delta(\xi,\eta,\kappa)\,. \tag{3.14}$$

One can show that if

$$\Delta(x,y,\kappa)=\int_{-\infty}^{\infty}dl\,\mu^-(x,y,l)e^{ilJx+ily}f(l,\kappa)e^{-i\kappa Bx-i\kappa y}, \tag{3.15}$$

where $\mu^-(x,y,\kappa)$ solves $(3.4)^-$ and $f(l,\kappa)$ solves (3.12), then $\Delta(x,y,\kappa)$ also solves (3.14). The details are given in Appendix A.

### C. The solution of the inverse problem

Equation (3.11) is the central equation associated with the inverse problem of (3.1). It defines a nonlocal RH problem in the complex $\kappa$ plane for the sectionally holomorphic

matrix function $\mu(x, y,\kappa)$. By taking the "minus" projection of (3.1) it follows that $\mu^-(x, y,\kappa)$ solves the following linear integral equation:

$$\mu^-(x, y,\kappa) + \frac{1}{2\pi i}\int_{-\infty}^{\infty} dl \int_{-\infty}^{\infty} dv \times \frac{\mu^-(x, y,l)e^{ilJx}f(l,\nu)e^{-i\nu Jx + i(l-\nu)y}}{\nu - \kappa + i0} = I . \tag{3.16}$$

Alternatively, Eq. (3.16) can be derived by using (2.1).

Equation (3.16) uniquely defines $\mu^-(x, y,\kappa)$ in terms of the scattering data $f(l,\nu)$. Once $\mu^-(x, y,\kappa)$ is found the potential $q(x, y)$ is easily obtained:

$$q(x, y) = -\frac{1}{2\pi}\hat{J}\int_{-\infty}^{\infty} dl \int_{-\infty}^{\infty} dv \times \mu^-(x, y,l)e^{ilJx}f(l,\nu)e^{-i\nu Jx + i(l-\nu)y} . \tag{3.17}$$

To derive Eq. (3.17), note that if one seeks an asymptotic expansion of $\mu(x, y,\kappa)$ for large $\kappa$ in the form

$$\mu(x, y,\kappa) = I + \mu_1(x, y)/\kappa + O(1/\kappa^2) , \tag{3.18}$$

one obtains [by using (3.1)]

$$q = -i\hat{J}\mu_1 . \tag{3.19}$$

On the other hand, large $\kappa$ asymptotics of (3.16) imply that

$$\mu_1(x, y) = \frac{1}{2\pi i}\int_{-\infty}^{\infty} dl \int_{-\infty}^{\infty} dv\, e^{ilJx + ily}f(l,\nu)e^{-i\nu Jx - i\nu y} . \tag{3.20}$$

Comparing (3.19) and (3.20), Eq. (3.17) follows.

Equations (3.12), (3.16), and (3.17) provide formally the complete solution of the inverse problem of (3.1): given the scattering data $T_+(l,\kappa)$, $T_-(l,\kappa)$ use Eq. (3.12) to obtain $f(l,\kappa)$ in closed form. Then Eq. (3.16) yields $\mu^-(x, y,\kappa)$, and finally Eq. (3.17) produces $q(x, y)$.

### D. Direct linearization

The essence of the "direct methods" (d)–(f) of Sec. II is the existence of linear integral equations [such as (3.16)], the solutions of which are related [through some formulas such as (3.17)] to the solution of inverse problems. Establishing such a relationship between linear integral equations and inverse problems is straightforward since in this case the IST formalism may be bypassed. However, as was pointed out in Sec. II, purely direct methods are not suitable for solving initial value problems since, given $q(x, y,0)$, it is not clear how the measure-contour–scattering data can be chosen.

In what follows we apply the direct linearizing method to the solution of the inverse problem of (3.1) [see Sec. II(f)]. In this way some of the results of Secs. IIIB and IIIC, namely, Eqs. (3.16) and (3.17), are also verified.

*Proposition 3.1:* Let $\mu(x, y,\kappa)$ be a solution of the linear integral equation in $\kappa$,

$$\mu(x, y,\kappa) + \frac{1}{2\pi i} \times \int\int_C \frac{d\zeta(l,\nu)\,\mu(x, y,l)e^{ilJx + ily}f(l,\nu)e^{-i\nu Jx - i\nu y}}{\nu - \kappa} = I , \tag{3.21}$$

where the measure-contour matrix $C$, $d\zeta(l,\nu)$, and $f(l,\nu)$ are essentially arbitrary. Assume that the homogeneous integral equation corresponding to (3.21) has only the zero solution. Then

$$q(x, y) \doteqdot -\frac{1}{2\pi}\hat{J}\int\int_C d\zeta(l,\nu)\mu(x, y,l) \times e^{ilJx + ily}f(l,\nu)e^{-i\nu Jx - i\nu y} \tag{3.22}$$

solves Eq. (3.1).

To prove the above result, let us first introduce some notation: Define the linear operators $L_\kappa$ and $P_{x, y, f}$ through

$$(L_\kappa g)(x, y) \doteqdot \frac{d}{dx}g - i\kappa\hat{J}g - J\frac{d}{dy}g , \tag{3.23}$$

$$(P_{x, y, f}g)(\kappa) \doteqdot -\frac{1}{2\pi i}\int\int_C \times \frac{d\zeta(l,\nu)\,g(l)e^{ilJx + ily}f(l,\nu)e^{-i\nu Jx - i\nu y}}{\nu - \kappa} . \tag{3.24}$$

By direct computations one may verify that

$$[L_\kappa, P_{x, y, f}]g(x, y,\kappa) = -\frac{1}{2\pi}\hat{J}\int\int_C d\zeta(l,\nu)\,g(x, y,l) \times e^{ilJx + ily}f(l,\nu)e^{-i\nu Jx - i\nu y} . \tag{3.25}$$

Equation (3.21) can be written as

$$\mu(x, y,\kappa) = I + (P_{x, y, f}\mu)(\kappa) . \tag{3.26}$$

Applying the operator $L_\kappa - qI$ on (3.26) and using (3.25), it follows that

$$(L_\kappa - q)\mu(x, y,\kappa) = P_{x, y, f}\{(L_\kappa - q)\mu(x, y,\kappa)\} . \tag{3.27}$$

Hence, assuming that the homogeneous version of (3.21) has only the zero solution, it follows that $(L_\kappa - q)\mu = 0$, i.e., Eq. (3.1).

## IV. ELLIPTIC SYSTEMS IN THE PLANE

We now consider the inverse problem associated with the elliptic system

$$\mu_x = i\kappa\hat{J}\mu + q\mu - iJ\mu_y , \tag{4.1}$$

where $\mu(x, y,\kappa)$ is an $n$th-order matrix, $J$ is a constant real diagonal matrix with all its entries different from each other, and $q(x, y)$ is an $n$th-order off-diagonal matrix containing the potentials $q_{ij}(x, y)$. We assume that $q_{ij}(x, y)\to 0$ rapidly enough for large $x$ and $y$. Equation (4.1) is obtained from the well-known[32] equation [i.e., (3.2) with $B = 0$, and $J$ replaced by $iJ$]

$$\psi_x = q\psi - iJ\psi_y , \tag{4.2}$$

through the transformation $\psi = \mu\exp[i\kappa Jx - \kappa y]$.

### A. Bounded eigenfunctions and their relations

For the solution of the inverse problem associated with (4.1) we follow very closely the steps outlined in Sec. II. We first consider step (i), i.e., we introduce an eigenfunction $\mu(x, y,\kappa)$ which solves (4.1), is bounded for all complex values of $\kappa$, and tends to $I$ as $\kappa\to\infty$. Such an eigenfunction satisfies the following Fredholm linear integral equation:

$$\mu(x, y,\kappa) = I + (G_{\kappa_R,\kappa_I,q}\mu(\cdot,\cdot,\kappa))(x, y) . \tag{4.3}$$

The operator $G_{\kappa_R,\kappa_I,q}$ is a linear integral matrix operator de-

fined by the following: Let $f(x,y)$ be some $n\times n$ matrix, then the $ij$th entry of the operator $G_{\kappa_R,\kappa_I,q}$ applying on $f(x,y)$ is

$$\{(G_{\kappa_R,\kappa_I,q}f)(x,y)\}_{ij}$$
$$\doteq\frac{1}{2\pi}\Big(\int_{-\infty}^{x}d\xi\int_{-\infty}^{C_{ij}\kappa_I}dm\int_{-\infty}^{\infty}d\eta - \int_{x}^{\infty}d\xi\int_{C_{ij}\kappa_I}^{\infty}dm\int_{-\infty}^{\infty}d\eta\Big)$$
$$\times\{e^{(mJ+i\kappa\hat{J})(x-\xi)+im(y-\eta)}q(\xi,\eta)f(\xi,\eta)\}_{ij}\,,\quad J_i>0\,, \tag{4.4}$$

where $C_{ij}\doteq(J_i-J_j)/J_i$ and for $J_i<0$ the integrals with respect to $dm$ are replaced by $\int_{C_{ij}\kappa_I}^{\infty}dm$ and $\int_{-\infty}^{C_{ij}\kappa_I}$, respectively. $f$ is a matrix and $\{f\}_{ij}$ denotes its $ij$th entry. Hence

$$\{e^{(mJ+i\kappa\hat{J})(x-\xi)+im(y-\eta)}q(\xi,\eta)f(\xi,\eta)\}_{ij}$$
$$=e^{[mJ_i+i\kappa(J_i-J_j)](x-\xi)+im(y-\eta)}\{q(\xi,\eta)f(\xi,\eta)\}_{ij}\,. \tag{4.5}$$

Sometimes it will be convenient to work with the column vectors of the matrix $\mu$. Letting $\mu=(\mu^1,\ldots,\mu^j,\ldots,\mu^n)$, it follows from (4.3) that $\mu^j$ satisfies

$$\mu^j(x,y)=\mathbf{I}^j+(\mathbf{g}^j_{\kappa_R,\kappa_I,q}\,\mu^j(\cdot,\cdot,\kappa))(x,y)\,, \tag{4.6$^j$}$$

where $\mathbf{I}^j$ denotes the $j$th unit vector, and

$$\{(\mathbf{g}^j_{\kappa_R,\kappa_I,q}\mathbf{f})(x,y)\}_l$$
$$\doteq\frac{1}{2\pi}\Big(\int_{-\infty}^{x}d\xi\int_{-\infty}^{C_{lj}\kappa_I}dm\int_{-\infty}^{\infty}d\eta - \int_{x}^{\infty}d\xi\int_{C_{lj}\kappa_I}^{\infty}dm\int_{-\infty}^{\infty}d\eta\Big)$$
$$\times e^{[mJ_l+i\kappa(J_l-J_j)](x-\xi)+im(y-\eta)}\{q(\xi,\eta)\mathbf{f}(\xi,\eta)\}_l\,,$$
$$J_l>0 \tag{4.7}$$

[for $J_l<0$ the integrals with respect to $dm$ must be altered just as in (4.4)]; if $\mathbf{f}$ is a vector, $\{\mathbf{f}\}_l$ denotes its $l$ th entry.

Equation (4.3) can be derived in a similar way as Eq. (3.4). Comparing (4.3) to (3.4) it follows that (a) Equation (4.3), in contrast to (3.4), has no jump across $\kappa_I=0$. (b) Equation (4.3) *depends explicitly on* $\kappa_I$. We emphasize this dependence by writing $G_{\kappa_R,\kappa_I,q}$ instead of $G_{\kappa,q}$. However, for simplicity of notation, we still write $\mu(x,y,\kappa)$ instead of $\mu(x,y,\kappa_R,\kappa_I)$. From the above it follows that the solution $\mu(x,y,\kappa)$, although bounded everywhere (i.e., for all complex values of $\kappa$), is analytic nowhere with respect to $\kappa$, since $\partial\mu/\partial\bar{\kappa}\neq 0$.

The "departure from holomorphicity" of $\mu(x,y,\kappa)$ is measured by $\partial\mu/\partial\bar{\kappa}$. Hence we are led to step (ii) of Sec. 2, namely, compute $\partial\mu/\partial\bar{\kappa}$. Differentiating (4.3) with respect to $\bar{\kappa}$ it follows that for all $\kappa$ which are not eigenvalues of (4.3)

$$\frac{\partial\mu(x,y,\kappa)}{\partial\bar{\kappa}}=\Omega(x,y,\kappa_R,\kappa_I) + \left(G_{\kappa_R,\kappa_I,q}\frac{\partial\mu(\cdot,\cdot,\kappa)}{\partial\bar{\kappa}}\right)(x,y)\,,\quad \kappa\in R^*\,, \tag{4.8}$$

where the matrix $\Omega$ is defined by

$$\{\Omega\}_{ii}=0\,,\quad \{\Omega\}_{ij}=T_{ij}(\kappa_R,\kappa_I)e^{\theta_{ij}(x,y,\kappa_R,\kappa_I)}\,,\quad i\neq j\,,$$
$$\theta_{ij}(x,y,\kappa_R,\kappa_I)=iC_{ij}(J_i\kappa_R x+\kappa_I y)\,, \tag{4.9}$$

and $T_{ij}$ is given by

$$T_{ij}(\kappa_R,\kappa_I)\doteq\frac{i}{4\pi}\operatorname{sgn}(J_i)C_{ij}\int_{-\infty}^{\infty}d\xi\int_{-\infty}^{\infty}d\eta \times\{q(\xi,\eta)\mu(\xi,\eta,\kappa)\}_{ij}e^{-\theta_{ij}(\xi,\eta,\kappa_R,\kappa_I)}\,. \tag{4.10}$$

[$R^*$ denotes the set of points in the complex $\kappa$ plane for which (4.3) has a homogeneous solution.]

Equation (4.8) motivates the introduction of another bounded eigenfunction, which we call $N(x,y,\kappa_R,\kappa_I)$:

$$N(x,y,\kappa_R,\kappa_I)=w(x,y,\kappa_R,\kappa_I) + (G_{\kappa_R,\kappa_I,q}N(\cdot,\cdot,\kappa_R,\kappa_I))(x,y)\,, \tag{4.11}$$

where $w$ is the matrix $\Omega$ with $T_{ij}=1$, i.e.,

$$\{w\}_{ii}=0\,,\quad w_{ij}=e^{\theta_{ij}(x,y,\kappa_R,\kappa_I)}\,,\quad i\neq j\,. \tag{4.12}$$

To show that $N(x,y,\kappa_R,\kappa_I)$ also solves (4.1) one needs only to show that the $w$ above satisfies the "homogeneous" version of (4.1), i.e.,

$$w_x=(i\kappa_R-\kappa_I)\hat{J}w-iJw_y\,;$$

this is straightforward.

Equations (4.9), (4.12) imply

$$\Omega=\sum_{\substack{i,j=1\\ i\neq j}}^{n}T_{ij}w^{ij}\,, \tag{4.13}$$

where $w^{ij}$ is a matrix with zeros everywhere, except at its $ij$th entry,

$$\{w^{ij}\}_{lv}=0\,,\quad i\neq l\ \text{and/or}\ j\neq v\,,\quad \{w^{ij}\}_{ij}=e^{\theta_{ij}}\,. \tag{4.14}$$

Hence, Eq. (4.8) implies

$$\frac{\partial\mu}{\partial\bar{\kappa}}=\sum_{\substack{i,j=1\\ i\neq j}}^{n}T_{ij}N^{ij}\,, \tag{4.15}$$

where the matrix $N^{ij}$ is defined by

$$N^{ij}(x,y)=w^{ij}(x,y)+(G_{\kappa_R,\kappa_I,q}N^{ij}(\cdot,\cdot,\kappa_R,\kappa_I))(x,y)\,. \tag{4.16}$$

Step (iii) of the method of Sec. II consists of finding a relationship between $\mu(x,y,\kappa)$ and $N^{ij}(x,y,\kappa_R,\kappa_I)$. This crucial relationship is as follows:

$$N^{ij}(x,y,\kappa_R,\kappa_I)=\mu(x,y,\kappa_R+i(J_j/J_i)\kappa_I)w^{ij}(x,y,\kappa_R,\kappa_I)\,. \tag{4.17}$$

To derive (4.17), first note that the matrix $N^{ij}$ has zeros everywhere except on the $j$th column. Let $\mathbf{N}^{ij}$ denote the $j$th column of $N^{ij}$ (its only nontrivial column). Then $\mathbf{N}^{ij}$ satisfies

$$\mathbf{N}^{ij}(x,y,\kappa_R,\kappa_I)=\mathbf{I}^i e^{\theta_{ij}(x,y,\kappa_R,\kappa_I)} + (\mathbf{g}^j_{\kappa_R,\kappa_I,q}\mathbf{N}^{ij}(\cdot,\cdot,\kappa_R,\kappa_I))(x,y)\,, \tag{4.18}$$

where $\mathbf{g}^j_{\kappa_R,\kappa_I,q}$ is defined by (4.7). It turns out that the operator $\mathbf{g}^j_{\kappa_R,\kappa_I,q}$ possesses the following property:

$$(\mathbf{g}^j_{\kappa_R,\kappa_I,q}e^{\theta_{ij}(\cdot,\cdot,\kappa_R,\kappa_I)-\theta_{ij}(x,y,\kappa_R,\kappa_I)}\mathbf{f}(\cdot,\cdot))(x,y)$$
$$=(\mathbf{g}^i_{\kappa_R,(J_j/J_i)\kappa_I,q}f(\cdot,\cdot))(x,y)\,. \tag{4.19}$$

The above result can be easily established by a simple change of variables: The left-hand side involves the exponential of

$$[mJ_l+i\kappa(J_l-J_i)-i\kappa_R(J_i-J_j)](x-\xi) + i(m-C_{ij}\kappa_I)(y-\eta)\,.$$

Simplifying the above and then letting $m = \hat{m} + C_{ij}\kappa_I$, it becomes

$$[\hat{m}J_l + i(\kappa_R + i\kappa_I J_j/J_i)(J_l - J_i)](x-\xi) + i\hat{m}(y-\eta)\,.$$

The above is just the exponential involved in the right-hand side. Also the $\int_{-\infty}^{C_{ij}\kappa_I} dm$ transforms to $\int_{-\infty}^{C_{li}(J_j/J_i)\kappa_I} d\hat{m}$. This completes the proof of (4.19).

Using (4.19) in Eq. (4.18) and comparing it with (4.6), it follows

$$\mathbf{N}^{ij}(x,y,\kappa_R,\kappa_I) = \boldsymbol{\mu}^i(x,y,\kappa_R + i(J_j/J_i)\kappa_I)e^{\theta_{ij}(x,y,\kappa_R,\kappa_I)} \tag{4.20}$$

or equivalently Eq. (4.17).

Equations (4.15) and (4.20) imply

$$\frac{\partial\mu(x,y,\kappa)}{\partial\bar{\kappa}} = \sum_{\substack{i,j=1\\ i\neq j}}^{n} \mu\left(x,y,\kappa_R + i\frac{J_j}{J_i}\kappa_I\right)\Omega^{ij}(x,y,\kappa_R,\kappa_I)\,,$$
$$\kappa\notin R^*\,, \tag{4.21}$$

where the matrix $\Omega^{ij}$ has a nonzero element only at its $ij$th entry,

$$\{\Omega^{ij}(x,y,\kappa_R,\kappa_I)\}_{vl} = \begin{cases} 0, & v\neq i \text{ and/or } l\neq j\,,\\ T_{ij}(\kappa_R,\kappa_I)e^{\theta_{ij}(x,y,\kappa_R,\kappa_I)}, & v=i,\ l=j\,. \end{cases} \tag{4.22}$$

Equation (4.21) is the basic equation needed for the solution of the inverse problem of (4.1). However, this equation is complete only if Eq. (4.3) has no homogeneous solutions. In the sequel we are going to obtain the additional relationship needed in the case that homogeneous modes exist. The following proposition is valid:

*Proposition 4.1*: Suppose that Eq. $(4.6)^i$ has a homogeneous solution $\boldsymbol{\phi}^i(x,y)$ at $\kappa_R^i\kappa_I^i$. Then Eq. $(4.6)^j, j=1,\ldots,n, i\neq j$, also has a homogeneous solution $\boldsymbol{\phi}^i(x,y)\exp[-\hat{\theta}_{ji}(x,y,\kappa_R^i,\kappa_I^i)]$ at the position $\kappa_R^i, \kappa_I^i J_i/J_j$, where $\hat{\theta}_{ji}(x,y,\kappa_R^i,\kappa_I^i) = \theta_{ji} + i\sigma_{ji}$ and $\sigma_{ji}$ is a constant depending on $\kappa^i$.

The above result follows directly from Eq. (4.19): Since $\boldsymbol{\phi}^i(x,y)$ is a homogeneous solution of $(4.6)^i$, then

$$\boldsymbol{\phi}^i(x,y) = (\mathbf{g}^i_{\kappa_R^i,\kappa_I^i,q}\boldsymbol{\phi}^i(\cdot,\cdot))(x,y)\,. \tag{4.23}$$

Using (4.19), it follows that

$$\boldsymbol{\phi}^i(x,y)e^{-\theta_{ji}(x,y,\kappa_R^i,\kappa_I^i)} = (\mathbf{g}^j_{\kappa_R^i,\kappa_I^i(J_i/J_j),q}\boldsymbol{\phi}^i(\cdot,\cdot)e^{-\theta_{ji}(\cdot,\cdot,\kappa_R^i,\kappa_I^i)})(x,y)\,, \tag{4.24}$$

which implies Proposition 4.1.

The homogeneous solutions $\boldsymbol{\phi}^i$ give rise to lumps for the corresponding nonlinear equations in a way similar to that found in BO and KPI.[33] Suppose that $(4.6)^i$ has homogeneous solutions $\boldsymbol{\phi}^i_{l_i}$ at $\kappa^i_{l_i}$, where $l_i = 1,2,\ldots,\Lambda_i$. Then all $\boldsymbol{\mu}^j$, $j=1,2,\ldots,n, j\neq i$, will have singularities at $\kappa^j_{R_{l_i}}, \kappa^j_{I_{l_i}}J_i/J_j$. Assuming these singularities are simple poles, we obtain the following representation for the vector $\boldsymbol{\mu}^i$:

$$\boldsymbol{\mu}^i(x,y,\kappa) = \tilde{\boldsymbol{\mu}}^i(x,y,\kappa) + \sum_{l_1=1}^{\Lambda_i}\frac{\boldsymbol{\phi}^1_{l_1}\exp[\hat{\theta}_{i1}(x,y,\kappa^1_{R_{l_1}},\kappa^1_{I_{l_1}})]}{\kappa - (\kappa^1_{R_{l_1}} + i\kappa^1_{I_{l_1}}J_1/J_i)} + \cdots + \sum_{l_i=1}^{\Lambda_i}\frac{\boldsymbol{\phi}^i_{l_i}}{\kappa-\kappa^i_{l_i}} + \cdots + \sum_{l_n=1}^{\Lambda_n}\frac{\boldsymbol{\phi}^n_{l_n}\exp[\hat{\theta}_{in}(x,y,\kappa^n_{R_{l_n}},\kappa^n_{I_{l_n}})]}{\kappa - (\kappa^n_{R_{l_n}} + i\kappa^n_{I_{l_n}}J_n/J_i)}\,, \tag{4.25a}$$

where

$$\hat{\theta}_{ij}(x,y,\kappa^i_{R_{l_i}},\kappa^i_{I_{l_i}}) = \theta_{ij}(x,y,\kappa^i_{R_{l_i}},\kappa^i_{I_{l_i}}) + (\sigma_{ij})^i_{l_i}\,. \tag{4.25b}$$

The above representation takes the concise form

$$\boldsymbol{\mu}^i(x,y,\kappa) = \tilde{\boldsymbol{\mu}}^i(x,y,\kappa) + \mathbf{S}^i(x,y,\kappa)\,, \tag{4.26}$$

where $\tilde{\boldsymbol{\mu}}^i$ is the nonsingular part of $\boldsymbol{\mu}^i$ and $\mathbf{S}^i$ represents the singular portion of $\boldsymbol{\mu}^i$:

$$\mathbf{S}^i(x,y,\kappa) = \sum_{l_1=1}^{\Lambda_1}\cdots\sum_{l_n=1}^{\Lambda_n}\sum_{j=1}^{n}\frac{\boldsymbol{\phi}^j_{l_j}\exp[-\hat{\theta}_{ij}(x,y,\kappa^j_{R_{l_j}},\kappa^j_{I_{l_j}})]}{\kappa - (\kappa^j_{R_{l_j}} + i\kappa^j_{I_{l_j}}J_j/J_i)}\,. \tag{4.27}$$

Equations (4.21), (4.26), and (4.27) imply

$$\frac{\partial\mu(x,y,\kappa)}{\partial\bar{\kappa}} = \begin{cases} \displaystyle\sum_{\substack{i,j=1\\ i\neq j}}^{n}\mu\left(x,y,\kappa_R + i\frac{J_j}{J_i}\kappa_I\right)\Omega^{ij}(x,y,\kappa_R,\kappa_I)\,,\\ \quad \kappa\notin R^*\,,\\ \Delta(x,y,\kappa)\,,\quad \kappa\notin R^*\,, \end{cases} \tag{4.28}$$

where $R^*$ is the set of eigenvalues of (4.3) and $\Delta$ is the matrix $\Delta(x,y,\kappa) = (\boldsymbol{\Delta}^1(x,y,\kappa),\ldots,\boldsymbol{\Delta}^n(x,y,\kappa))$ with

$$\boldsymbol{\Delta}^i(x,y,\kappa) = \pi\sum_{l_1=1}^{\Lambda_1}\cdots\sum_{l_n=1}^{\Lambda_n}\sum_{j=1}^{n}\boldsymbol{\phi}^j_{l_j}\exp[-\hat{\theta}_{ij}(x,y,\kappa^j_{R_{l_j}},\kappa^j_{I_{l_j}})] \times\delta\left[\kappa - \left(\kappa^j_{R_{l_j}} + i\kappa^j_{I_{l_j}}\frac{J_j}{J_i}\right)\right]. \tag{4.29}$$

To solve the "$\bar{\partial}$" problem associated with (4.28) one has to further establish a relationship between $\boldsymbol{\mu}^i$ and $\boldsymbol{\phi}^i$. This relationship is as follows:

$$\lim_{\kappa\to\kappa^i_{l_i}}\left[\boldsymbol{\mu}^i(x,y,\kappa) - \frac{\boldsymbol{\phi}^i_{l_i}(x,y)}{\kappa-\kappa^i_{l_i}}\right] = \boldsymbol{\phi}^i_{l_i}(-iJ_i x + y + \gamma^i_{l_i})\,,\quad l_i = 1,2,\ldots,\Lambda_i\,, \tag{4.30}$$

where $\boldsymbol{\phi}^i_{l_i}$ is normalized by $\lim[\boldsymbol{\phi}^i_{l_i}(J_i x + iy)]\to i$ for large $x, y$, and the constant $\gamma^i_{l_i}$ can be fixed from asymptotics. Equation (4.30) is valid provided that

$$1 - \frac{\operatorname{sgn}(J_i)}{2i\pi}\int_{-\infty}^{\infty}d\xi\int_{-\infty}^{\infty}d\eta\,\{q(\xi,\eta)\boldsymbol{\phi}^i_{l_i}(\xi,\eta)\}_i = 0\,,$$
$$\int_{-\infty}^{\infty}d\xi\int_{-\infty}^{\infty}d\eta\exp[-\theta_{li}(\xi,\eta,\kappa^i_{R_{l_i}},\kappa^i_{I_{l_i}})] \times\{q(\xi,\eta)\boldsymbol{\phi}^i_{l_i}(\xi,\eta)\}_l = 0\,,\quad l=1,2,\ldots,n,\ l\neq i\,. \tag{4.31}$$

To prove the above introduce

$$\hat{\boldsymbol{\mu}}^i_{l_i}(x,y,\kappa) \doteq \boldsymbol{\mu}^i(x,y,\kappa) - \boldsymbol{\phi}^i_{l_i}(x,y)/(\kappa-\kappa^i_{l_i})\,. \tag{4.32}$$

Then $\hat{\boldsymbol{\mu}}^i_{l_i}(x,y,\kappa^i_{l_i})$ satisfies

$$\hat{\mu}^i_{l_i} - \mathbf{g}^i_{\kappa^i_{R_{l_i}},\kappa^i_{I_{l_i}},q^{\hat{\mu}^i_{l_i}}} = \mathbf{I}^i + \frac{\partial}{\partial\kappa}(\mathbf{g}^i_{\kappa_R,\kappa_I,q}\boldsymbol{\phi}^i_{l_i})\Big|_{\kappa=\kappa^i_{l_i}}. \tag{4.33}$$

The left-hand side of (4.33) has, by definition, a homogeneous solution $\boldsymbol{\phi}^i_{l_i}$. Hence (4.33) is solvable iff the right-hand side satisfies a certain orthogonality condition. The derivative with respect to $\kappa$ appearing in (4.33) gives two contributions, one from the integrand and one from the boundary of integration. These contributions are most easily evaluated by introducing the change of variables $m = \hat{m} + 2\kappa_I$. Then the integrand involves the exponential of

$$(\hat{m}J_l + \bar{i}\bar{\kappa}J_l - i\kappa J_i)(x-\xi) + i(\hat{m} + 2\kappa_I)(y-\eta).$$

This yields a contribution proportional to $-iJ_i(x-\xi) + (y-\eta)$ and hence a particular solution $-\boldsymbol{\phi}^i_{l_i}(iJ_i x - y)$. Hence $\hat{\boldsymbol{\mu}}^i_{l_i}$ equals this particular solution plus the homogeneous solution $\gamma^i_{l_i}\boldsymbol{\phi}^i_{l_i}$, where the constant $\gamma^i_{l_i}$ can be fixed from asymptotics. The contribution from the boundary of integration yields the orthogonality conditions (4.31).

### B. The inverse problem

The basic equations formally expressing the solution of the inverse problem of (4.1) are the matrix equation

$$\mu(x,y,\kappa) - (T_{x,y,\Omega}\,\mu(x,y,\cdot))(\kappa) - S(x,y,\kappa) = I, \tag{4.34}$$

and the following set of vector equations:

$$(-ixJ_1 + y + \gamma^1_{l_1})\boldsymbol{\phi}^1_{l_1} - \hat{\mathbf{T}}^1_{l_1}(x,y) - \hat{\mathbf{S}}^1_{l_1}(x,y) = \mathbf{I}^1,$$
$$l_1 = 1,2,\ldots,\Lambda_1, \tag{4.35}$$

$$\cdots$$

$$(-ixJ_n + y + \gamma^n_{l_n})\boldsymbol{\phi}^n_{l_n} - \hat{\mathbf{T}}^n_{l_n}(x,y) - \hat{\mathbf{S}}^n_{l_n}(x,y) = \mathbf{I}^n,$$
$$l_n = 1,2,\ldots,\Lambda_n. \tag{4.35}$$

In (4.34) the linear operator $T_{x,y,\Omega}$ is defined by

$$(T_{x,y,\Omega}f(\cdot))(\kappa) \doteq \frac{1}{2\pi i}\int_{R_\infty}\int \sum_{\substack{i,j=1\\ i\neq j}}^{n} f(z_R + i(J_j/J_i)z_I)$$
$$\times\Omega^{ij}(x,y,z_R,z_I)\,dz\wedge d\bar{z}(z-\kappa)^{-1} \tag{4.36}$$

($R_\infty$ is the entire $z$-complex plane, $dz\wedge d\bar{z} = -2i dz_R\,dz_I$), and the matrix $S(x,y,\kappa)$ has columns $\mathbf{S}^i$, $i = 1,2,\ldots,n$, which are defined by (4.27). In (4.35) the vectors $\hat{\mathbf{T}}^i_{l_i}, \hat{\mathbf{S}}^i_{l_i}$ are defined by

$$\hat{\mathbf{T}}^i_{l_i}(x,y) \doteq \{(T_{x,y,\Omega}\,\mu(x,y,\cdot))(\kappa^i_{l_i})\}_i, \tag{4.37}$$

$$\hat{\mathbf{S}}^i_{l_i} \doteq \sum_{l_1=1}^{\Lambda_1}\cdots\sum_{l_i=1}^{\Lambda_i}\cdots\sum_{l_n=1}^{\Lambda_n}\sum_{j=1}^{n}\frac{\boldsymbol{\phi}^j_{l_j}e^{-\hat{\theta}_{ij}(x,y,\kappa^i_{k_{l_j}},\kappa^j_{l_j})}}{\kappa^i_{l_i} - (\kappa^j_{l_j} + i\kappa^j_{l_j}J_j/J_i)},$$

where $\{f\}_i$ denotes the $i$th column of the matrix $f$ and $\Sigma^{\Lambda_i}_{l_i=1}$ denotes the sum from $l_i = 1$ to $l_i = \Lambda_i$ unless any of the denominators vanishes. Equations (4.35) can be written in the concise form

$$\Phi(x,y)(-ixJ + yI + \Gamma) - \hat{T}(x,y) - \hat{S}(x,y) = I, \tag{4.38}$$

where the columns of $\Phi, \hat{T}, \hat{S}$ are $\boldsymbol{\Phi}^i_{l_i}, \hat{\mathbf{T}}^i_{l_i}, \hat{\mathbf{S}}^i_{l_i}$, respectively, and $\Gamma$ is a diagonal matrix with elements $\gamma^i_{l_i}$.

The linear integral equations (4.34), $(4.35)^i_{l_i}$ define the functions

$$\mu(x,y,\kappa),\quad \{\boldsymbol{\phi}^i_{l_i}\}_{l_i=1}^{\Lambda_i},\quad i = 1,2,\ldots,n, \tag{4.39}$$

in terms of the *scattering data*

$$\{\kappa^i_{l_i},\gamma^i_{l_i},(\sigma_{ij})^i_{l_i}\}_{l_i=1}^{\Lambda_i},\quad i = 1,2,\ldots,n,$$
$$T_{ij}(\kappa_R,\kappa_I),\quad i,j = 1,2,\ldots,n,\ i\neq j. \tag{4.40}$$

Given the scattering data (4.40), Eqs. (4.34), $(4.35)^i_{l_i}$ yield the functions $\mu(x,y,\kappa), \boldsymbol{\phi}^i_{l_i}$. Then the potential $q(x,y)$ can be reconstructed from

$$q(x,y) = \hat{J}\left\{\frac{1}{2\pi}\int_{R_\infty}\int \sum_{\substack{i,j=1\\ i\neq j}}^{n}\mu\left(x,y,z_R + i\frac{J_j}{J_i}z_I\right)\right.$$
$$\times\Omega^{ij}(x,y,z_R,z_I)\,dz\wedge d\bar{z} - i\Bigg(\sum_{l_1=1}^{\Lambda_1}\cdots\sum_{l_n=1}^{\Lambda_n}\sum_{j=1}^{n}\boldsymbol{\phi}^j_{l_j}$$
$$\times\exp\left[-\hat{\theta}_{ij}(x,y,\kappa^j_{R_{l_j}},\kappa^j_{I_{l_j}})\right],\ldots,\sum_{l_1=1}^{\Lambda_1},\ldots,$$
$$\left.\times\sum_{l_n=1}^{\Lambda_n}\sum_{j=1}^{n}\boldsymbol{\phi}^j_{l_j}\exp\left[-\theta_{nj}(x,y,\kappa^j_{R_{l_j}},\kappa^j_{i_{l_j}})\right]\Bigg)\right\}. \tag{4.41}$$

The above equations can be obtained as follows: Eq. (4.34) is simply obtained by substituting (4.28) in (2.1). To derive Eqs. (4.38) take the limit of the $i$th column of (4.34) as $\kappa\to\kappa^i_{l_i}$. Equation (4.41) can be derived from (4.34) in the same way that (3.17) was derived.

We note that pure lump solutions correspond to $\Omega^{ij} = 0$ and hence are obtained via the *linear system of algebraic equations* $(4.35)^i_{l_i}$ with $\Omega^{ij} = 0$.

*Example*: Consider the case $n = 2$ and assume that $\Lambda_1 = \Lambda_2 = 1$. Let

$$\kappa^1_{l_1} = \kappa_1,\quad \kappa^2_{l_2} = \kappa_2,\quad \gamma^1_{l_1} = \gamma_1,\quad \gamma^2_{l_2} = \gamma_2. \tag{4.42}$$

Then (4.35) and (4.41) yield

$$(-iJ_1x + y + \gamma_1)\boldsymbol{\phi}^1(x,y) = \mathbf{I}^1 + \frac{\boldsymbol{\phi}^2(x,y)\exp\left[-\hat{\theta}_{12}(x,y,\kappa_{R_2},\kappa_{I_2})\right]}{\kappa_1 - (\kappa_{R_2} + i(J_2/J_1)\kappa_{I_2})}, \tag{4.43}$$

$$(-iJ_2x + y + \gamma_2)\boldsymbol{\phi}^2(x,y) = \mathbf{I}^2 + \frac{\boldsymbol{\phi}^1(x,y)\exp\left[-\hat{\theta}_{21}(x,y,\kappa_{R_1},\kappa_{I_1})\right]}{\kappa_2 - (\kappa_{R_1} + i(J_1/J_2)\kappa_{I_1})}, \tag{4.44}$$

$$q(x,y) = -i\hat{J}\left\{\boldsymbol{\phi}^1 + \boldsymbol{\phi}^2 e^{-\hat{\theta}_{12}(x,y,\kappa_{R_2},\kappa_{I_2})},\right.$$
$$\left.\times\boldsymbol{\phi}^2 + \boldsymbol{\phi}^1\exp\left[-\hat{\theta}_{21}(x,y,\kappa_{R_1},\kappa_{I_1})\right]\right\}. \tag{4.45}$$

Equations (4.43) and (4.44) imply

$$\boldsymbol{\phi}^1(x,y) = \frac{1}{\pi(x,y)}\begin{pmatrix} -iJ_2x + y + \gamma_2 \\ (1/\alpha)\exp\left[-\hat{\theta}_{12}(x,y,\kappa_{R_2},\kappa_{I_2})\right]\end{pmatrix},$$
$$\boldsymbol{\phi}^2(x,y) = \frac{1}{\pi(x,y)}\begin{pmatrix} (1/\beta)\exp\left[-\hat{\theta}_{21}(x,y,\kappa_{R_1},\kappa_{I_1})\right] \\ -iJ_1x + y + \gamma_1\end{pmatrix}, \tag{4.46}$$

where

$$\pi(x,y) \doteq (-iJ_1x + y + \gamma_1)(-iJ_2x + y + \gamma_2) - e^{-(\hat{\theta}_{12}+\hat{\theta}_{21})}/\alpha\beta,$$

$$\alpha \doteq \kappa_1 - (\kappa_{R_2} + i(J_2/J_1)\kappa_{I_2}), \quad \beta \doteq \kappa_2 - (\kappa_{R_1} + i(J_1/J_2)\kappa_{I_1}).$$

Hence (4.45) yields

$$q(x,y) = \frac{-i}{\pi(x,y)} \times \begin{pmatrix} 0 & (J_1 - J_2) \times \exp[-\hat{\theta}_{21}(x,y,\kappa_{R_1},\kappa_{I_1})]J_2^* \\ (J_2 - J_1) \times \exp[-\theta_{12}(x,y,\kappa_{R_2},\kappa_{I_2})]J_1^* & 0 \end{pmatrix}, \tag{4.47}$$

where

$$J_1^* = 1/\alpha - iJ_1x + y + \gamma_1, \quad J_2^* = 1/\beta - iJ_2x + y + \gamma_2.$$

It can be verified that $\phi^1$, $\phi^2$, and $q$ as defined from (4.46) and (4.47) satisfy the eigenvalue equation (4.1) [where $\kappa = \kappa_1$ for the first column of the matrix equation (4.1) and $\kappa = \kappa_2$ for the second column].

### C. A direct linearization

For completeness we now state the analog of Proposition 3.1.

*Proposition 4.2:* Let $\mu(x,y,\kappa)$ be a solution of the linear integral equation in $\kappa$,

$$\mu(x,y,\kappa) = I + \int_{R_\infty} \int \sum_{\substack{i,j=1 \\ i\neq j}}^{n} d\xi_{ij}(z_R,z_I)\,\mu(x,y,z_R + i(J_j/J_i)z_I)w^{ij}(x,y,z_R,z_I)(z-\kappa)^{-1}, \tag{4.48}$$

where the measures $d\xi_{ij}(z_R,z_I)$ are arbitrary. Assume that the homogeneous integral equation corresponding to (4.48) has only the trivial solution. Then

$$q(x,y) \doteq \frac{1}{2\pi}\hat{J}\int_{R_\infty}\int \sum_{\substack{i,j=1 \\ i\neq j}}^{n} d\xi_{ij}(z_R,z_I) \times \mu\left(x,y,z_R + i\frac{J_j}{J_i}z_I\right)w^{ij}(x,y,z_R,z_I) \tag{4.49}$$

solves Eq. (4.1).

The proof of Proposition 4.2 is similar to that of Proposition 3.1 and is therefore omitted.

## V. ON THE INITIAL VALUE PROBLEM OF CERTAIN EVALUATION EQUATIONS

Suppose that the potentials $q$ of (3.1) or (4.1) also solve some evolution equation, i.e., $q = q(x,y,t)$. The scope of this section is to establish how the results of Secs. 3 and 4 can be used to solve the initial value problem of this evolution equation. First consider the hyperbolic case: $q(x,y,t)$ is expressed [see (3.17)] in terms of $\mu^-(x,y,t,\kappa)$ and the scattering data $f(\kappa,l,t)$. Also $\mu^-(x,y,t,\kappa)$ is uniquely determined [see (3.16)] in terms of $f(\kappa,l,t)$. Furthermore, given $q(x,y,0)$ Eqs. (3.12) and $(3.13)^{\pm}$ imply $f(\kappa,l,0)$. Hence, to compute $q(x,y,t)$ one only needs to find the evolution of the scattering data $f(\kappa,l,t)$. Similarly for the elliptic case: $q(x,y,t)$ is uniquely determined from the scattering data (4.40) [through Eqs. (4.34), (4.35), and (4.41)]. Furthermore, given $q(x,y,0)$, one can compute the initial scattering data [for example, $T_{ij}(\kappa_R,K_I,0)$ can be computed from (4.10)].

A specific evolution equation for $q(x,y,t)$ implies a specific Lax pair. In what follows we shall show how the time-dependent part of a Lax pair uniquely determines the evolution of the scattering data. We consider three different types of Lax pairs; these Lax pairs contain, as special cases, the $n$-wave interaction in 2 + 1, the DSI and the DSII equations.

### A. Lax pairs containing the $n$-wave interaction in 2 + 1

Consider the Lax pair

$$\psi_x = J\psi_y + q\psi, \quad \psi_t = A_1\psi + A_2\psi_y \tag{5.1}$$

and assume that

$$A_1 \to 0, \quad A_2 \to A_{20} \quad \text{for large } x,y, \tag{5.2}$$

where $J, A_{20}$ are constant real diagonal matrices. Hence for large $x,y$, $\psi \to \exp[i\kappa(Jx + y + A_{20}t)]$. However, the IST was based on an eigenfunction normalized to $I$ at $\infty$. Thus, let $\psi = \mu\exp[i\kappa(Jx + y + A_{20}t)]$ and (5.1) yield (3.1) and

$$\mu_t = i\kappa(A_2\mu - \mu A_{20}) + A_1\mu + A_2\mu_y. \tag{5.3}$$

Recall that the RH problem associated with the inverse problem of (3.1) is given by Eq. (3.11), i.e.,

$$\mu^+(x,y,\kappa) - \mu^-(x,y,\kappa) = \int_{-\infty}^{\infty} dl\,\mu^-(x,y,l)\hat{f}(x,y,t,l,\kappa), \tag{5.4}$$

where

$$\hat{f}(x,y,t,l,\kappa) = e^{il\hat{J}x + ily}f(l,\kappa,t)e^{-i\kappa Jx - i\kappa y}. \tag{5.5}$$

The evolution of the scattering data $f(l,\kappa,y)$ associated with (5.1) is given by

$$f(l,\kappa,t) = e^{ilA_{20}t}f(l,\kappa,0)e^{-i\kappa tA_{20}}. \tag{5.6}$$

To prove (5.6), apply the operator

$$L_\kappa \doteq \partial_t - i\kappa A_2 - A_1 - A_2\partial_y \tag{5.7}$$

to Eq. (5.4). Then use

$$L_\kappa\mu^{\pm} = -i\kappa\mu^{\pm}A_{20}, \quad L_\kappa = L_l + i(l-\kappa)A_2, \quad \hat{f}_y = i(l-\kappa)\hat{f},$$

to obtain

$$\hat{f}_t = ilA_{20}\hat{f} - i\kappa\hat{f}A_{20}. \tag{5.8}$$

Equation (5.8) implies (5.6).

### B. Lax pairs containing DSI

Consider the Lax pair

$$\psi_x = q\psi + J\psi_y, \quad \psi_t = A_1\psi + A_2\psi_y + A_3\psi_{yy}, \tag{5.9}$$

and assume that

$$A_1 \to 0, \quad A_2 \to 0, \quad A_3 \to A_{30} \quad \text{for large } x,y, \tag{5.10}$$

where $J$ is a real constant diagonal matrix and $A_{30}$ is a purely imaginary diagonal matrix. Letting $\psi = \mu\exp[i\kappa(Jx + y + i\kappa A_{30}t)]$, Eqs. (5.9) yield (3.1) and

$$\mu_t = A_3\mu_{yy} + A_2\mu_y + A_1\mu + (i\kappa)^2(A_3\mu - \mu A_{30}) + 2i\kappa A_3\mu_y + i\kappa A_2\mu . \quad (5.11)$$

The evolution of the scattering data $f(l,\kappa,0)$ [of (3.16) and (3.17)] associated with (5.9) is given by

$$f(l,\kappa,0) = e^{(il)^2 tA_{30}} f(l,\kappa,0) e^{-(i\kappa)^2 tA_{30}} . \quad (5.12)$$

Equation (5.12) can be derived in a similar way as Eq. (5.6).

### C. Lax pairs containing DSII

Consider the Lax pair

$$\psi_x = q\psi - iJ\psi_y , \qquad \psi_t = A_1\psi + A_2\psi_y + A_3\psi_{yy} , \quad (5.13)$$

and assume (5.10) ($J$ and $A_{30}$ are real and purely imaginary diagonal matrices, respectively). Letting $\psi = \mu \exp[\kappa(iJ_x - y + \kappa A_{30}t)]$, Eqs. (5.13) yield (4.1) and

$$\mu_t = A_3\mu_{yy} + A_2\mu_y + A_1\mu + \kappa^2(A_3\mu - \mu A_{30}) - 2\kappa A_3\mu_y - \kappa A_2\mu . \quad (5.14)$$

The evolution of the scattering data $\Omega^{ij}$ [of (4.41)] associated with (5.13) is given by

$$\Omega^{ij}(x,y,\kappa_R,\kappa_I,t) = e^{\hat{\kappa}^2 A_{30}t}\Omega^{ij}(x,y,\kappa_R,\kappa_I,0)e^{-\kappa^2 A_{30}t} , \qquad \hat{\kappa} \doteqdot \kappa_R + i(J_j/J_i)\kappa_I . \quad (5.15)$$

To derive Eq. (5.15) define the linear operator $L_\kappa$,

$$L_\kappa \doteqdot \partial_t - A_3\partial_y^2 - A_2\partial_y - A_1 - \kappa^2 A_3 + 2\kappa A_3\partial_y + \kappa A_2 .$$

Then (5.14) yields

$$L_\kappa\,\mu(\kappa) + \kappa^2\mu A_{30} = 0 , \quad L_\kappa \frac{\partial\mu}{\partial\bar{\kappa}} + \kappa^2 \frac{\partial\mu}{\partial\bar{\kappa}} A_{30} = 0 , \qquad \mu(\kappa) \doteqdot \mu(x,y,\kappa) . \quad (5.16)$$

Applying the operator $L_\kappa$ to (4.21) and using (5.16), Eq. (5.15) follows in a straightforward way. Similarly one can establish that (see Appendix B)

$$\frac{\partial}{\partial t}\gamma^i_{I_i} = -2\kappa^i_{I_i}A_{30_i} ,$$

$$\frac{\partial}{\partial t}(\sigma_{ji})^i_{I_i} = A_{30_j}\left[2i\kappa^i_{R_{I_i}}\kappa^i_{I_{I_i}}\left(1 - \frac{J_i}{J_j}\right) - \left(1 - \frac{J_i}{J_j}\right)^2(\kappa^i_{I_i})^2\right] . \quad (5.17)$$

### D. Some physically interesting equations

The $n$-wave resonant interactions and variants of the so-called DS equation are contained as special cases in the Lax pairs (5.1), (5.9), and (5.13).

#### 1. The n-wave interaction

The $n$-wave interaction equations are

$$q_{ij_t} = \alpha_{ij}q_{ij_x} + \beta_{ij}q_{ij_y} + \sum_{\substack{\kappa=1 \\ \kappa\neq j}}^{n} (\alpha_{i\kappa} - \alpha_{\kappa j})q_{i\kappa}q_{\kappa j} , \qquad i,j,\kappa = 1,\ldots,n , \quad (5.18)$$

where the real constants $\alpha_{ij}$ and $\beta_{ij}$ are related to the $x$ and $y$ components of the underlying group velocities. Equations (5.18) are the compatibility conditions of the Lax pair (5.1) with $A_1$ and $A_2$ defined by

$$\{A_1\}_{ij} = \alpha_{ij}q_{ij} , \quad \{A_1\}_{ii} = 0 , \quad A_2 = \mathrm{diag}(C_1,\ldots,C_n) , \quad (5.19)$$

where $J_i, C_i$ are given in terms of $\alpha_{ij}, \beta_{ij}$ via

$$\alpha_{ij} = \frac{C_i - C_j}{J_i - J_j} , \quad \beta_{ij} = C_i - J_i\alpha_{ij} . \quad (5.20)$$

Hence, the initial value problem of (5.18) can be solved via (3.16) and (3.17), where the initial scattering data $f(l,\kappa,0)$ can be calculated from (3.12), $(3.13)^{\pm}$, and $f(l,\kappa,t)$ is found from (5.6), where

$$A_{20} = \mathrm{diag}(C_1,\ldots,C_n) . \quad (5.21)$$

#### 2. The DS equation

Consider the Lax pair

$$\psi_x = q\psi + B\psi_y , \qquad \psi_t = A_1\psi + A_2\psi_y + A_3\psi_{yy} , \quad (5.22)$$

where

$$B = \mathrm{diag}(b_1,b_2) , \quad A_3 = \mathrm{diag}(e_1,e_2) , \quad A_2 = \gamma_{12}q , \quad \gamma_{12} = \gamma_{21} \doteqdot (e_1 - e_2)/(b_1 - b_2) , \quad (5.23)$$

$$q = \begin{pmatrix} 0 & q_{12} \\ q_{21} & 0 \end{pmatrix} , \quad A_1 = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} .$$

The compatibility condition of (5.22) yields some nonlinear equations for $q_{12}, q_{21}$. However, in contrast to the usual cases, $A_{ij}$ depend nonlocally on $q_{12}, q_{21}$. Actually one finds

$$q_{ij_t} = \frac{e_i - e_j}{(b_i - b_j)^2}q_{ij_{xx}} + \frac{2(b_ie_j - b_je_i)}{(b_i - b_j)^2}q_{ij_{xy}} + \frac{e_ib_j^2 - b_i^2e_j}{(b_i - b_j)^2}q_{ij_{yy}} + (A_{ii} - A_{jj})q_{ij} , \quad (5.24)$$

where $A_{ij}$ are defined in terms of $q_{ij}$ through the equations

$$A_{ij} = \frac{\gamma_{ij}}{b_i - b_j}q_{ij_x} + \left(\frac{2e_i - b_i\gamma_{ij}}{b_i - b_j}\right)q_{ij_y} , \quad i \neq j , \quad (5.25a)$$

$$A_{11_x} - b_1A_{11_y} = \frac{\gamma_{12}}{b_2 - b_1}(q_{12}q_{21})_x + \frac{2e_1 - b_1\gamma_{12}}{b_2 - b_1}(q_{12}q_{21})_y \quad (5.25b)$$

and $A_{22}$ satisfies an equation obtained from (5.25b) by $1 \leftrightarrow 2$.

Variants of the DS equation can be obtained by the following "reduction" of (5.22). Let

$$b_2 = -b_1 , \quad e_2 = -e_1 , \quad e_1 = i , \qquad q_{21} = \sigma\bar{q}_{12} \doteqdot \sigma\bar{Q}\,(\sigma = \pm 1) , \quad b_1^2 = \pm 1 . \quad (5.26)$$

Using (5.26) in (5.22)–(5.25b) and letting $\phi = i(A_{11} - A_{22}) + (\sigma_2/b_1^2)|q|^2$ one is lead to the following (the overbar denotes complex conjugate).

*DSI*: We call DSI the following set of equations:

$$iQ_t + \tfrac{1}{2}(Q_{xx} + Q_{yy}) = -\sigma|Q|^2Q + \phi Q , \quad \sigma = \pm 1 , \qquad \phi_{xx} - \phi_{yy} = 2\sigma(|Q|^2)_{xx} . \quad (5.27)$$

Equations (5.27) are the compatibility conditions of (5.9) with

$$J_1 = 1 , \quad J_2 = -1 , \quad A_2 = iq , \quad A_3 = \mathrm{diag}(i,-i) , \quad (5.28)$$

$$A_1=\begin{pmatrix}A_{11} & A_{12}\\ A_{21} & A_{22}\end{pmatrix},\quad q=\begin{pmatrix}0 & Q\\ \sigma\bar{Q} & 0\end{pmatrix},$$

where the entries of $A_1$ are defined by

$$\begin{aligned}&A_{12}=\tfrac{1}{2}i(Q_x+Q_y),\quad A_{21}=\tfrac{1}{2}i\sigma(-\bar{Q}_x+\bar{Q}_y),\\ &A_{11_x}-A_{11_y}=-\tfrac{1}{2}i\sigma[(Q\bar{Q})_x+(Q\bar{Q})_y],\\ &A_{22_x}+A_{22_y}=\tfrac{1}{2}i\sigma[(Q\bar{Q})_x-(Q\bar{Q})_y].\end{aligned}\tag{5.29}$$

Hence, assuming that $A_{11},A_{22}$ tend to zero for large $x,y$ (we note in this context that the question of the precise conditions under which this occurs is under investigation), the initial value problem of (5.27) can be solved via (3.16), (3.17). In these equations $J=\mathrm{diag}(1,-1)$; furthermore, the initial scattering data $f(l,\kappa,0)$ is determined from (3.12), $(3.13)^{\pm}$, and $f(l,\kappa,t)$ is found from (5.12), where

$$A_{30}=\mathrm{diag}(i,-i).\tag{5.30}$$

*DSII*: We call DSII the following set of equations:

$$\begin{aligned}&iQ_t+\tfrac{1}{2}(Q_{yy}-Q_{xx})=\sigma|Q|^2Q+\phi Q,\\ &\phi_{xx}+\phi_{yy}=-2\sigma(Q^2)_{xx}.\end{aligned}\tag{5.31}$$

Equations (5.31) are the compatibility conditions of (5.13) with

$$J_1=1,\quad J_2=-1,\quad A_2=-q,\quad A_3=\mathrm{diag}(i,-i),\tag{5.32}$$

and $A_1,q$ are defined by (5.28), where

$$\begin{aligned}&A_{12}=-\tfrac{1}{2}(iQ_x+Q_y),\quad A_{21}=-\tfrac{1}{2}\sigma(-i\bar{Q}_x+\bar{Q}_y),\\ &A_{11_x}+iA_{11_y}=\tfrac{1}{2}\sigma[i(Q\bar{Q})_x+(Q\bar{Q})_y],\\ &A_{22_x}-iA_{22_y}=\tfrac{1}{2}\sigma[-i(Q\bar{Q})_x+(Q\bar{Q})_y].\end{aligned}\tag{5.33}$$

Hence (again, as before we assume that $A_{11},A_{22}$ tend to zero for large $x$, $y$) the initial value problem of (5.31) can be solved via (4.34), (4.35), (4.41). In these equations $J=\mathrm{diag}(1,-1)$; furthermore, the evolution of the scattering data (4.40) is determined from Eqs. (5.15) and (5.17), where $A_{30}=\mathrm{diag}(i,-i)$.


## ACKNOWLEDGMENTS

This work was partially supported by the Air Force Office of Scientific Research under Grant No. 78-3674-D, the Office of Naval Research under Grant No. N00014-76-C-0867, and the National Science Foundation under Grant No. MCS-8202117. We appreciate useful discussions with D. Bar Yaacov.


## APPENDIX A

In this appendix we derive Eq. (3.12). Substituting $\Delta$ from (3.15) to (3.14), it follows that (3.11) is valid iff

$$\begin{aligned}&\int_{-\infty}^{\infty}\mu^-(l)e^{iJx}f(l,\kappa)e^{i(l-\kappa)y}\,dl-\frac{1}{2\pi}\int_{-\infty}^{\infty}d\xi\,Ee^{i\kappa(x-\xi)J}\\ &\times(\pi_+q\mu^-(\kappa)-\pi_-q\mu^+(\kappa))e^{i\kappa Jx}-\frac{1}{2\pi}\int_{-\infty}^{\infty}d\xi\\ &\times\int_{-\infty}^{\infty}dl\,Ee^{i\kappa(x-\xi)J}q\mu^-(l)e^{iJ\xi}f(l,\kappa)e^{i(l-\kappa)\eta}=0,\end{aligned}\tag{A1}$$

where $\mu^{\pm}(l)\doteq\mu^{\pm}(x,y,l)$ and $E\doteq E_{x-\xi,J}$. On the other hand, multiplying Eq. $(3.4)^-$ from the right by $\exp[ilJx]f(l,\kappa)\exp[i(l-\kappa)y]$ and integrating with respect to $l$, we find

$$\begin{aligned}&\int_{-\infty}^{\infty}\mu^-(l)e^{iJx}f(l,\kappa)e^{i(l-\kappa)y}\,dl-\frac{1}{2\pi}\int_{-\infty}^{\infty}d\xi\\ &\times\int_{-\infty}^{\infty}dl\,Ee^{il(x-\xi)J}(\pi_+q\mu^-(l))e^{il\xi J}f(l,\kappa)e^{i(l-\kappa)y}\\ &-\int_{-\infty}^{\infty}e^{iJx}f(l,\kappa)e^{i(l-\kappa)y}\,dl-\frac{1}{2\pi}\int_{-\infty}^{x}d\xi\\ &\times\int_{-\infty}^{\infty}dl\,Ee^{il(x-\xi)J}q\mu^-(\kappa)e^{il\xi J}f(l,\kappa)e^{i(l-\kappa)y}=0.\end{aligned}\tag{A2}$$

Letting $m=\tilde{m}-\kappa$ in the third integral of (A1) and $m=\tilde{m}-l$ in the fourth integral of (A2), it follows that these two terms are equal. Hence, subtracting (A1) − (A2), we find

$$\begin{aligned}&\int_{-\infty}^{\infty}dl\,e^{iJx}f(l,\kappa)e^{i(l-\kappa)y}\,dl\\ &-\frac{1}{2\pi}\int_{-\infty}^{\infty}d\xi\int_{-\infty}^{\infty}dl\int_{-\infty}^{\infty}d\eta\int_{-\infty}^{\infty}dm\\ &\times e^{i(m+l)(x-\xi)J}(\pi_+q\mu^-(\kappa))e^{il\xi J}f(l,\kappa)e^{i(l-\kappa)y+im(y-\eta)}\\ &-\frac{1}{2\pi}\int_{-\infty}^{\infty}d\xi\int_{-\infty}^{\infty}d\eta\int_{-\infty}^{\infty}dm\,e^{i(m+\kappa)(x-\xi)J}\\ &\times(\pi_+q\mu^-(\kappa)-\pi_-q\mu^+(\kappa))e^{i\kappa\xi J+im(y-\eta)}=0.\end{aligned}\tag{A3}$$

Letting $l=\bar{m}$, $m+l=\bar{m}$, and $m+\kappa=\bar{m}$, in the first, second, and third integrals of (A3), we conclude that (3.11) is valid iff $f(\bar{m},\kappa)$ satisfies (3.12).

## APPENDIX B

In this appendix we derive Eqs. (5.17). To simplify the calculations, we consider (5.14) for large $x$, $y$:

$$\mu_t\sim A_{30}(\mu_{yy}-2\kappa\mu_y)+\kappa^2(A_{30}-\mu A_{30}).\tag{B1}$$

Denoting $\{\mu^i\}_j$ by $\mu_j^i$, it follows that $\mu_j^i$ satisfies

$$\mu_{j_t}^i=A_{30_j}(\mu_{j_{yy}}^i-2\kappa\mu_{j_y}^i)+\kappa^2(A_{30_j}-A_{30_i})\mu_j^i.\tag{B2}$$

Similarly $\phi_j^i$, the homogeneous solution of (4.23) satisfies (B2) with $\kappa$ replaced by $\kappa^i$. Using the equations that $\mu_j^i$, $\phi_j^i$ satisfy and

$$\tilde{\mu}^i\doteq\mu^i-\phi^i/(\kappa-\kappa^i),\tag{B3}$$

we find an equation for $\tilde{\mu}_j^i$. Taking the limit of this equation as $\kappa\to\kappa^i$ and using (4.30), we find

$$\frac{\partial}{\partial t}\gamma^i=-2\kappa^iA_{30_i}.\tag{B4}$$

To derive the evolution equation of $(\sigma_{ij})^i$, we first let

$$\boldsymbol{\mu}^j=\tilde{\boldsymbol{\mu}}^j+\frac{\boldsymbol{\phi}^ie^{-\hat{\theta}_{ji}}}{\kappa-(\kappa_R^i+i\kappa_I^iJ_i/J_j)}.\tag{B5}$$

Since the evolutions of $\boldsymbol{\mu}^j$, $\boldsymbol{\phi}^i$, and $\tilde{\boldsymbol{\mu}}^j$ are known ($\tilde{\boldsymbol{\mu}}^j$ evolves like $\boldsymbol{\mu}^j$), we find the evolution of $\hat{\theta}_{ji}$, which implies

$$\frac{\partial}{\partial t}(\sigma_{ji})^i=iA_{30_j}\left[2i\kappa_I^i\kappa\kappa_R^i\left(1-\frac{J_i}{J_j}\right)-\left(1+\frac{J_i}{J_j}\right)^2(\kappa_I^i)^2\right].\tag{B6}$$


[1] A. S. Fokas and M. J. Ablowitz, "The Inverse Scattering Transform for the Benjamin Ono—A Pivot to Multidimensional Problems," Stud. Appl. Math. **68**, 1 (1983).

[2] A. S. Fokas and M. J. Ablowitz, "On the Inverse Scattering and Direct

Linearizing Transforms for the KP," Phys. Lett. A **94**, 67 (1983).
[3]A. S. Fokas and M. J. Ablowitz, "On the Inverse Scattering of the Time Dependent Schrödinger Equation and the Associated KPI Equation," I.N.S. #22, Nov. 1982, Stud. Appl. Math. (to be published).
[4]M. J. Ablowitz, D. Bar Yaacov, and A. S. Fokas, "On the IST for KP," Stud. Appl. Math. **69**, 135 (1983).
[5]A. S. Fokas, "On the Inverse Scattering of First Order Systems in the Plane Related to Nonlinear Multidimensional Equations," Phys. Rev. Lett. **51**, 3–6 (1983).
[6]A. S. Fokas and M. J. Ablowitz, "Method of Solution for a Class of Multidimensional Nonlinear Evolution Equations," Phys. Rev. Lett. **51**, 7–10 (1983).
[7]M. J. Ablowitz and H. Segur, *Solitons and the Inverse Scattering Transform* (SIAM, Philadelphia, 1981), pp. 1–425.
[8]A. Davey and K. Stewartson, Proc. R. Soc. London Ser. A **338**, 101 (1974).
[9]D. J. Benney and G. J. Roskes, Stud. Appl. Math. **48**, 377 (1969).
[10](a) A. S. Fokas and M. J. Ablowitz, "On the Inverse Scattering Transform for Multidimensional (2 + 1) Problems," lectures given at School and Workshop on Nonlinear Phenomena" OAXTEPEC, Mexico, 29 Nov.–17 Dec. 1982; (b) M. J. Ablowitz and A. S. Fokas, "Comments on the Inverse Scattering Transform and Related Nonlinear Evolution Equations," *ibid.* 29 Nov.–17 Dec. 1982.
[11]P. D. Lax, Commun. Pure Appl. Math. **21**, 467 (1968).
[12]W. Oevel and B. Fuchssteiner, Phys. Lett. A **88**, 323 (1982).
[13]H. H. Chen, J. C. Lee, and J. E. Lin, preprint, *Nonlinear Waves* (to be published).
[14]B. G. Konopelchenko, Phys. Lett. A **86**, 346 (1981); Commun. Math. Phys. (1982); J. Phys. A **15** (1982).
[15]H. H. Chen, J. Math. Phys. **16**, 2382 (1975); D. V. Chudnovski, J. Math. Phys. **20**, 2416 (1979).
[16]J. Satsuma, J. Phys. Soc. Jpn. **40**, 286 (1976); J. Satsuma and M. J. Ablowitz, J. Math. Phys. **20**, 1496 (1979); A. Nakamura, Phys. Rev. Lett. **46**, 751 (1981); Phys. Lett A **88**, 55 (1982); J. Math. Phys. **22**, 2456 (1981); J. Phys. Soc. Jpn. **50**, 2469 (1981); **51**, 19 (1981).
[17]M. Kashiwara and T. Miwa, Proc. Jpn. Acad. A **57**, 342 (1981); E. Date, M. Kashiwara, and T. Miwa, Proc. Jpn. Acad. A **57**, 387 (1981); M. Jimbo, E. Date, M. Kashiwara, and T. Miwa, J. Phys. Soc. Jpn. **50**, 3806 (1981); RIMS Preprints 359, 360.
[18]V. E. Zakharov and P. B. Shabat, Funct. Anal. Appl. **8**, 226 (1974).
[19](a) V. E. Zakharov and P. B. Shabat, Funct. Anal. Appl. **13**, 166 (1979); (b) V. E. Zakharov and A. V. Mikhailov, Zh. Eksp. Teor. Fiz. **74**, 1053 (1978) [Sov. Phys. JETP **47**, 1017 (1978)].
[20]S. V. Manakov (private communication).
[21]A. S. Fokas and M. J. Ablowitz, Phys. Rev. Lett. **47**, 1096 (1981).
[22]P. Santini, M. J. Ablowitz, and A. S. Fokas, "The Direct Linearization and the Dressing Method," preprint.
[23]M. J. Ablowitz, A. S. Fokas, and R. L. Anderson, "The Direct Linearizing Transform and the Benjamin-Ono Equation," Phys. Lett. A **93**, 375 (1983).
[24]R. Rosales, Stud. Appl. Math. **59**, 117 (1978).
[25](a) V. E. Zakharov and S. V. Manakov, Sov. Sci. Rev. Phys. Rev. **1**, 133 (1979); (b) S. V. Manakov, Physica D **3**, 420 (1981).
[26]H. Segur, "Comments on IS for the Kadomtsev-Petviashvili Equation," in *Mathematical Methods in Hydrodynamics and Integrability in Dynamical Systems (La Jolla Institute, 1981)*, AIP Conference Proceedings, No. 88, edited by M. Tabor and Y. M. Treve (Am. Inst. Physics, New York, 1982), pp. 211–228.
[27]J. Kaup, Physica D **1**, 45 (1980); also H. Cornile, J. Math. Phys. **20**, 1653 (1979); L. P. Niznik, Ukrainian Math. J. **24**, 110 (1972).
[28]L. Hörmander, *Complex Analysis in Several Variables* (Van Nostrand, Princeton, NJ, 1966).
[29]R. Beals and R. R. Coifman, "Scattering, Transformations Spectrals, et Nonlinear Equations of Evolution" Seminaire Goulaoric-Meyer-Schwartz, exp. 22, Ecole Polytechnique, Palaiseau, 1980–1981; (b) R. Beals and R. R. Coifman, "Scattering and Inverse Scattering for First Order Systems," preprint; (c) R. Beals and R. R. Coifman, "Scattering, Transformations Spectrals, and Nonlinear Equations of Evolution" Seminaire Goulaoric-Meyer-Schwartz, exp. 21, Ecole Polytechnique, Palaiseau, 1981–1982.
[30]I. Hauser and F. J. Ernst, J. Math. Phys. **21**, 1126, 1418 (1980); **22**, 1051 (1981); see also V. A. Belinskii and V. E. Zakharov, Zh. Eksp. Teor. Fiz. **75**, 1955 (1978) [Sov. Phys. JETP **48**, 985 (1978)]; C. Cosgrove, J. Math. Phys. **22**, 2624 (1981).
[31]M. F. Atiyah and R. S. Ward, Commun. Math. Phys. **55**, 117 (1977); M. F. Atiyah, N. J. Hitchin, V. G. Drinfield, and Yu. I. Manin, Phys. Lett. A **65**, 185 (1978).
[32]M. J. Ablowitz and R. Haberman, Phys. Rev. Lett. **35**, 1185 (1975).
[33]In both cases of BO as well as KPI the analog of Eq. (4.3) depends analytically on $\kappa$ and hence the corresponding eigenfunctions could only admit poles in $\kappa$. In contrast to the above $\mu(x, y, \kappa)$ does not depend analytically on $\kappa$. However, we still assume that the singularities of $\mu(x, y, \kappa)$ take the form of poles in $\kappa$. This must be investigated further to establish whether any other functional form is possible.

# Null infinity from a quasi-Newtonian view

Jeffrey Winicour
*Physics Department, University of Pittsburgh, Pittsburgh, Pennsylvania 15260*

(Received 6 December 1983; accepted for publication 9 March 1984)

A scheme is elaborated which generates unique characteristic initial value data for a general relativistic fluid in terms of initial value data for a Newtonian fluid. The chief ingredient is the demand that the evolution of the general relativistic fluid yield the evolution of the Newtonian fluid as a parameter $\lambda$, corresponding to $1/c$, goes to zero. The resulting gravitational null data is examined in detail to the first three nontrivial orders in $\lambda$. To this order, the data is asymptotically flat at future null infinity and appears to properly incorporate an outgoing radiation condition. The implications of these results for the calculation of gravitational radiation from quasi-Newtonian systems is discussed.

PACS numbers: 04.20. − q, 02.60. + y

## I. INTRODUCTION

This work is a follow-up of a null-cone approach to post-Newtonian calculations proposed in an earlier paper[1] (hereafter referred to as Paper I). The ideas of that paper seemed especially attractive for formulating radiative corrections. Null cones of a $\lambda$-parameter family of space-times ($1/\lambda$ representing the velocity of light in these space-times) were arranged to yield absolute-time slices in the $\lambda = 0$ Newtonian limit. In this way, for example, a natural geometric formulation of the quadrupole radiation formula might be possible. The $\ddot{Q}$-terms on one side of the equation would refer to the Newtonian quadrupole dependence on absolute time while, on the other side, the radiation terms would represent the leading-order general-relativistic energy flux reaching null infinity along the corresponding null cones. Thus, radiation calculations could be based upon the exact curved-space null cones, instead of the customary quasi-Minkowskiian formulation of approximation schemes for post-Newtonian hydrodynamics.[2–9] In the limit $\lambda = 0$, these exact null cones reduce to absolute-time slices on which to express Newtonian properties.

The approach is based upon the characteristic initial value formulation of general relativity with interior fluid source. Here the initial data on a null cone, for the gravitational field, consist of the shear (or astigmatism) of the null rays. In Paper I, the details were worked out to the leading nontrivial order in $\lambda$. The condition that data on an initial cone yield a Newtonian space-time (in the sense of Cartan), for $\lambda = 0$, turned out to be remarkably simple. Essentially, the gravitational potential emerges in a second role as a potential for the shear of these null cones. The formalism and leading-order results are recapitulated in Sec. II.

Futamase and Schutz[8,9] have recently developed a complementary post-Newtonian approximation scheme based upon the space-like, rather than null, initial value problem. They adopt the usual harmonic coordinate formalism which reduces Einstein's equation to a Minkowskiian wave equation. The initial value formulation avoids the divergent retarded integrals characteristic of previous schemes which attempted to rule out incoming radiation by boundary conditions at the false past null infinity of the Minkowskiian "light cones." Instead, their boundary conditions are the Cauchy data on a space-like hypersurface extending to spatial infinity. In this way, they have constructed a mathematically well-defined approximation scheme and have greatly refined the methodology of the harmonic coordinate approach. Certain aspects of their scheme are still not as ideal as might be desired.

(i) For a given Newtonian system, the post-Newtonian corrections are not unique but depend upon the choice of gravitational Cauchy data. They circumvent this by introducing an ensemble average over all possible Cauchy data, the physical justification being the random incoming radiation, impinging on any realistic system.

(ii) Because their scheme is based upon the false Minkowskiian light cones, it is not applicable to properties of future null infinity such as radiation fields and energy flux. Nevertheless, their work represents an important achievement and has provided a mathematically well-defined derivation of the radiation-reaction corrections to Newtonian equations of motion.[10]

The null approach was formulated with the specific intent of studying radiation fields, from a post-Newtonian viewpoint. Initial data on an outgoing null cone lead instantaneously to radiation fields at future null infinity. This radiation can be studied without even having to evolve Einstein's equation. The crucial physical issue is the selection of initial null data. In the presence of matter, it is known that the mathematically simplest initial data, choosing the shear of the null cone to vanish, introduce unphysical incoming radiation. What is the proper $\lambda$-dependent null data for studying post-Newtonian corrections? Paper I gives a procedure for determining such null data by the requirement that the $\lambda$-dependent system have a Newtonian limit in some neighborhood $u_0 \leqslant u < u_1$ of the initial null cone $u_0$. In Sec. III, we formalize that procedure in terms of an iteration scheme for determining $O(\lambda^n)$ data from $O(\lambda^{n-1})$ data. At each order, the data are given by the solution of a Poisson equation.

A similar role for Poisson equations appears in the impressive work of Persides[11] who carries out a $\lambda$-expansion using a space-like but asymptotically null coordinate system. Persides' formalism is not based upon an initial value formulation and comparison with our work is not straightforward particularly with respect to boundary conditions. A chief difference appears to be that Persides sets to zero certain otherwise arbitrary source terms for his Poisson equations.

 0022-2488/84/082506-09$02.50  

As a result, the asymptotic shear of his asymptotically null hypersurfaces vanishes to leading order in the $\lambda$-expansion. In our formalism, this would correspond to the presence of incoming radiation.

The initial null data determined by the Poisson equation are unique provided the asymptotic behavior of the source is consistent with the boundary condition that the solution vanish at infinity.[12] This property of the asymptotic behavior is established in Sec. IV, up through the first three nontrivial orders in $\lambda$. To this order, the data are unique and are consistent with asymptotic flatness at null infinity, as discussed in Sec. V.

The following heuristic model exhibits the mechanism by which unique null data can be determined by the requirement of a Newtonian limit and reveals the sense in which these data contain no incoming radiation. For the $\lambda$-dependent flat space-time metric

$$ds^2 = du^2 + 2\lambda\, du\, dr - \lambda^2 r^2(d\theta^2 + \sin^2\theta\, d\phi^2) \quad (1.1)$$

in null polar coordinates, consider the wave equation

$$\frac{2\lambda}{r}\frac{\partial}{\partial r}\left(r\frac{\partial F}{\partial u}\right) - \nabla^2 F = S. \quad (1.2)$$

Here $S$ is a smooth, compact, $\lambda$-independent external source for the $\lambda$-dependent field $F(u, r, \theta, \phi, \lambda)$. We require that $F$ vanish at future null infinity and be uniformly differentiable with respect to $\lambda$ and $x^\alpha$, for $0 \leqslant \lambda \leqslant \lambda_1$, in some future neighborhood of the initial null cone $u_0$. Expanding $F$ as a series in $\lambda$, $F = \Sigma F^{(n)}\lambda^n$, the boundary condition at infinity implies that $F^{(0)}$ is uniquely determined from $\nabla^2 F^{(0)} = -S$. Furthermore, for $n > 0$, each succeeding term is uniquely determined from

$$\nabla^2 F^{(n)} = \frac{2}{r}\frac{\partial}{\partial r}\left(r\frac{\partial F^{(n-1)}}{\partial u}\right).$$

Thus, under the above conditions, (1.2) determines unique null data $F(u_0, x^i; \lambda) = \Sigma F^{(n)}(u_0, x^i)\lambda^n$. (In fact, in this oversimplified model, this procedure yields a unique series solution whose successive terms contain the monopole, dipole, quadrupole, ... radiation from the source.) To identify the manner in which uniqueness of the solution corresponds to the exclusion of incoming radiation, we examine the behavior of $F$ in the region exterior to the source. For simplicity, consider the spherically symmetric solution $F = (1/r)[f(u) + g(u + 2\lambda r)]$, which contains outgoing and incoming waves described by $f$ and $g$, respectively. We have $\partial F/\partial\lambda = 2\dot{g}(u + 2\lambda r)$, which must vanish at null infinity, uniformly for $0 \leqslant \lambda \leqslant \lambda_1$, on the initial null cone. By choosing $\lambda = r_0/r$, this implies that $g(u_0 + 2r_0) = 0$ so that the initial data for incoming waves must effectively vanish. Thus, by the use of stringent analytic conditions, it is possible to eliminate incoming waves from the data while keeping the outgoing waves intact. Of course, all incoming waves from past null infinity are not eliminated by this mechanism. Data on the initial cone can only determine $F$ to the future of $u_0$ and are ignorant of imploding-exploding waves which radiate to future null infinity in the past of $u_0$. That is an advantage of this technique in the curved space case, where the complete exclusion of incoming waves in terms of past null infinity is not only irrelevant physically but also impossible using existing techniques.

## II. NEWTONIAN GRAVITY ON THE NULL CONE

We present here the formalism developed in Paper I. We adhere to the conventions of that paper with one major exception: we base the ð-operator on the unit sphere metric $q_{AB}$, as noted in Sec. II B below.

### A. The $\lambda$-dependent system

We consider a $\lambda$-parameter family of space-times such that physical space-time, with the velocity of light equal to 1, is characterized by $\lambda = 1$ and Cartan's version of Newtonian space-time, with infinite velocity of light, is obtained in the limit $\lambda \to 0$. By utilizing the available diffeomorphism freedom, these space-times are described on a common manifold in such a way that they share a common family of null cones whose apices trace out a time-like geodesic. In the null coordinate system based upon this family of cones, with $x^0 = u$, $x^1 = r$, $x^A = (x^2, x^3) = (\rho, \phi)$, the $\lambda$-dependent family of metrics takes the form

$$ds^2 = (e^{2\lambda^2\beta}V/r - \lambda^4 r^2 h^{AB}U_A U_B)du^2 + 2\lambda e^{2\lambda^2\beta}\, du\, dr + 2\lambda^3 r^2 U_A\, du\, dx^A - \lambda^2 r^2 h_{AB}\, dx^A\, dx^B, \quad (2.1)$$

where $h^{AB}h_{BC} = \delta^A{}_C$ and $\det(h_{AB}) = \sin^2\theta$. In addition, we set $V = r + \lambda^2 W$ and $h_{AB} = q_{AB} + \lambda^2\gamma_{AB}$. These choices of $\lambda$-factors ensure that (2.1) reduces to (1.1) in the flat-space case.

As matter source for this system, we take the $\lambda$-dependent ideal fluid stress tensor

$$T_{\mu\nu} = (\rho + \lambda^2 p)w_\mu w_\nu - \lambda^2 p g_{\mu\nu},$$

where the four-velocity $w_\mu$ has the form $w_\mu = t_\mu + \lambda^2 v_\mu$, with $t_\mu = (u + \lambda r)_{,\mu}$. Einstein's equation is then equivalent to the matter evolution equation $T^\nu_{\mu;\nu} = 0$ and the following six components of Einstein's equation:

$$-8\pi\lambda^{-2}T_{11} = -4\beta_{,1}/r + \lambda^2 c^{AB}c_{AB}/4, \quad (2.2a)$$

$$-8\pi\lambda^{-2}T_{1A} = -(r^4 e^{-2\lambda^2\beta}h_{AB}U^B_{,1})_{,1}/2r^2 - 2\beta_{,A}/r + \beta_{,1A} - c_{AB}{}^{:B}/2, \quad (2.2b)$$

$$-8\pi r^2(T - g^{AB}T_{AB}) = 2\lambda^{-2}e^{-2\lambda^2\beta}V_{,1} + \lambda^{-2}\mathscr{R} + 2\beta^{:A}_{A} + 2\lambda^2\beta^{:A}\beta_{:A} - e^{-2\lambda^2\beta}r^{-2}(r^4 U^A)_{,1:A} + \lambda^2 r^4 e^{-4\lambda^2\beta}h_{AB}U^A{}_{,1}U^B_{,1}/2, \quad (2.2c)$$

$$8\pi\lambda^{-2}T_{AB}m^A m^B = e^{-2\lambda^2\beta}m^A m^B[r(rh_{AB,0})_{,1}/\lambda - (rVc_{AB})_{,1}/2 - 2e^{2\lambda^2\beta}(\beta_{:AB} + \lambda^2\beta_{:A}\beta_{:B}) + r^2 h_{AC}U^C{}_{,1:B} - \lambda^2 r^4 e^{-2\lambda^2\beta}h_{AC}h_{BD}U^C_{,1}U^D_{,1}/2 + 2rU_{A:B} + \lambda^2 r^2 c_{AB}U^D{}_{:D}/2 + \lambda^2 r^2 c_{AB:D}U^D - \lambda^2 r^2 c^D_A(U_{B:D} - U_{D:B})]. \quad (2.2d)$$

Here capital Latin indices are raised with $h^{AB} = 2m^{(A}\bar{m}^{B)}$, $\mathscr{R}$ is the Ricci scalar of the two-geometry defined by $h_{AB}$, a colon represents the associated covariant derivative, and we have introduced the shear tensor $c_{AB} = \gamma_{AB,1} = h_{AB,1}/\lambda^2$.

The unconstrained data on an initial null cone consist of the matter data $\rho, v_1, v_A$ (with $p$ determined by an equation of state) and the gravitational data $h_{AB}$ (or, equivalently, $c_{AB}$).

In terms of these data, radial integration of the "hypersurface" equations (2.2a)–(2.2c) determines the entire initial metric and radial integration of the gravitational "evolution" equation (2.2d) determines the time derivative of the gravitational data. The matter evolution equation determines the time derivative of the matter data. All integration constants are fixed by smoothness conditions which require that $\beta$, $v_A$, $W$, and $c_{AB}$ vanish at the origin.

### B. The Newtonian limit

In paper I, the necessary and sufficient conditions were found for the above $\lambda$-dependent family to yield, in the limit $\lambda \to 0$, a Cartan version of Newtonian space-time with absolute-time slices $t$, a Euclidean metric $e^{\mu\nu}$, and a connection $\Gamma^{\rho}{}_{\mu\nu} = t_{,\mu}\, t_{,\nu} e^{\rho\sigma} \Phi^*{}_{,\sigma}$, where $\nabla^2 \Phi^* = 4\pi\rho$. (Here $\Phi^*$ is the Newtonian potential appropriate to a reference frame with freely falling origin.) It is convenient to state these conditions using the $\eth$-operator for the unit-sphere metric $q^{AB} = 2q^{(A}\bar{q}^{B)}$. For the sake of definiteness, we set $q^A = (1, i/\sin\theta)/\sqrt{2}$, the essential point being that $q^A{}_{,0} = q^A{}_{,1} = 0$. The conventions are $f_{,A} q^A = \eth f/\sqrt{2}$, for scalars $f$, and $(\bar{\eth}\,\eth - \eth\,\bar{\eth})\eta = 2s\eta$, for spin-weight $s$-quantities $\eta$. We define the shear-potential $\alpha$, up to $l = 0$ and $l = 1$ harmonics, by $c_{AB} q^A q^B = \eth^2 \alpha$, then for general-relativistic data on an initial null cone $u_0$ to go over to Newtonian data on the corresponding absolute-time slice $t_0$, the shear potential must be related to the Newtonian potential by

$$\eth^2 [r^2(\alpha + \bar{\alpha})]_{,1} = -4\eth^2 \Phi^*,$$

in the limit $\lambda \to 0$. Furthermore, for the first-time derivative of $\alpha$, at $u = u_0$, to be a smooth function of $\lambda$, $\eth^2(\alpha - \bar{\alpha})$ must vanish, for $\lambda = 0$, so that the initial shear must be pure electric. We incorporate these conditions by setting, for $\lambda = 0$,

$$(r^2 \alpha)_{,1} = -2\Phi^*. \tag{2.3}$$

This last equation fixes the $l = 0$ and $l = 1$ parts of $\alpha$, which leads to some calculational convenience even though these parts are not of physical significance. The potential $\Phi^*$ is related to the usual Newtonian potential $\Phi$ by

$$\Phi^* = \Phi + a + r a_m Y_{1m}, \tag{2.4}$$

where we use the summation convention for spherical harmonic indices. Here $a$ and $a_m$ depend only on time. They are chosen so that $\Phi^*$ and its spatial gradients vanish along the free-fall (or geodesic) world-line $r = 0$. We assume that the fluid has compact spatial support so that

$$\Phi \triangleq \phi_{lm} Y_{lm}/r^{l+1}, \tag{2.5}$$

where "$\triangleq$" stands for "equals, in the exterior of the fluid" and where, for example, $\phi_{00} = -m(4\pi)^{1/2}$, with $m$ being the mass of the Newtonian system.

Near the origin, $\Phi^* = O(r^2)$. Thus the integral of (2.3),

$$\alpha = -2r^{-2} \int_0^r \Phi^*(r')dr', \tag{2.6}$$

leads, for $\lambda = 0$, to a smooth $\alpha$ which vanishes at the origin. Because integrals of this type arise often, we introduce the shorthand

$$\int f = \int_0^r f(r')dr' \quad \text{and} \int^{\infty} f = \int_0^{\infty} f(r')dr',$$

for functions $f$ with the appropriate behavior to make these integrals well defined.

When (2.3) is satisfied for null cones in some neighborhood of the initial null cone, $u_0$, Euler's equations for the fluid follow from the $\lambda = 0$ limit of Einstein's equation. Thus, in these circumstances, a Newtonian limit exists in this neighborhood. However, as demonstrated in paper I, this requirement on the $u$-dependence of the $\lambda = 0$ limit of $\alpha$ determines, via the gravitational evolution equation, the initial $\lambda$-dependence of $\alpha$, at $u = u_0$. In Sec. III, we investigate the details of this interdependence.

### C. The weak field limit

There is an important distinction between a Newtonian limit, with its connection and degenerate space-time metric, and the weak field limit, with its background Minkowski metric. In addition to a Newtonian limit, there is also a weak field limit associated in a natural way with the $\lambda$-dependent system (2.1). This is obtained by setting $u = \lambda\tilde{u}$ and $g_{\mu\nu} = \lambda^2 \tilde{g}_{\mu\nu}$, so that

$$d\tilde{s}^2 = (e^{2\lambda^2\beta} V/r - \lambda^4 r^2 h^{AB} U_A U_B) d\tilde{u}^2 + 2e^{2\lambda^2\beta} d\tilde{u}\, dr + 2\lambda^2 r^2 U^A d\tilde{u}\, dx^A - r^2 h_{AB}\, dx^A\, dx^B.$$

The Einstein tensor for this system then satisfies

$$\tilde{G}_{\mu\nu} = -8\pi\lambda^2 [(\rho + \lambda^2 p)\tilde{w}_\mu \tilde{w}_\nu - \lambda^2 p \tilde{g}_{\mu\nu}],$$

where $\tilde{g}_{\mu\nu}\tilde{w}^u \tilde{w}^\nu = 1$. Now $\tilde{\rho} = \lambda^2 \rho$ and $\tilde{p} = \lambda^2 p$ go to zero as $\lambda \to 0$.

The $\tilde{g}_{\mu\nu}$-system is often taken as the starting point for post-Newtonian approximations, with the new dynamical time scale $\tilde{u}$ then introduced to formulate the Newtonian or "slow motion" limit.[13] The $\tilde{g}_{\mu\nu}$-system is more convenient for discussing asymptotic radiative properties because Penrose's compactification of null infinity[14] is then directly applicable in the $\lambda = 0$ limit. Also, the $\tilde{g}_{\mu\nu}$-system is essential for carrying out numerical investigations of post-Newtonian effects.[15,16] However, it should be borne in mind that it is the $g_{\mu\nu}$-system which is appropriate for investigating the existence of a Newtonian limit.

## III. DETERMINATION OF THE NEWTONIAN INITIAL SHEAR

Given initial data for a Newtonian space-time, we now present a formal method for determining unique initial null data for a $\lambda$-dependent family of Einstein space-times. Our hypothesis is that $\lambda = 0$ represents a Newtonian space-time, $\lambda = 1$ represents an Einstein space-time with the physical value of the velocity of light, and that the $\lambda$-dependent system (2.1) extrapolates between them analytically. In accord, we expand all quantities in the form

$$f(x^\alpha; \lambda) = \sum_{n=0}^{\infty} f^{(n)}(x^\alpha)\lambda^n. \tag{3.1}$$

Note that we assume the existence of a family of null cones for $0 < \lambda \leqslant 1$, in some neighborhood $u_0 \leqslant u < u_1$ of the initial null cone. For a matter distribution with a sufficiently large quadrupole moment the bending of light would lead to caustics that destroy this null cone property. However, this bending effect goes to zero as the velocity of light goes to

infinity in the Newtonian limit so that such a family of cones should exist for some range $0<\lambda\leqslant\lambda_1$. For $\lambda_1\leqslant 1$, the physical space-time would be excluded from the system. Presumably this would show up as a divergence of the $\lambda$-series for the shear. No post-Newtonian approximation scheme might be valid for such a system. In the space-like approach, the false Minkowski light cones would not reveal the gravitational lens effects. Strangely, the predominance of observable physical systems fall into the class $\lambda_1>1$ (for some choice of apex). It is precisely this circumstance that gives physical importance to the post-Newtonian approach.

We now discuss the iteration scheme.

Gravitational quantities such as $\beta$ and $W$, which are scalar fields on the two-spaces of constant $u$ and $r$, are expanded in the manner of (3.1). For the vector field $U_A$ and tensor field $c_{AB}$, we introduce scalar potentials $U_A q^A=\eth Z/\sqrt{2}$ and $c_{AB}q^A q^B=\eth^2\alpha$, with $Z$ and $\alpha$ having expansions as in (3.1). Although the $l=0$ part of $Z^{(n)}$ and the $l=0$ and $l=1$ parts of $\alpha^{(n)}$ are physically irrelevant, it will be convenient to give them specific values in the $n=0$ case. To give a unique expansion for the complex *dyad* vector, we first fix its phase by requiring

$$m^A|_{r=0}=q^A \quad \text{and} \quad m^A{}_{,1}\bar{m}_A=0.$$

Then, using $h_{AB}=q_{AB}+\lambda^2\gamma_{AB}$, it follows that we can set

$$m^A=(1+\lambda^4 Q)q^A+\lambda^2 P\bar{q}^A,$$

where the complex scalars $Q$ and $P$ have expansions as in (3.1).

As initial matter data for the $\lambda$-dependent system, we choose the initial Newtonian data, so that $\rho(x^\alpha;\lambda)=\rho(x^\alpha)$ and $v^\beta(x^\alpha;\lambda)=v^\beta(x^\alpha)$, for $u=u_0$. In addition we fix the geodesic world-line traced out by the origin by requiring that it pass through the Newtonian center-of-mass, at $u=u_0$, with the center-of-mass velocity. Then, given the initial choice of angular coordinates, the initial data $\rho(u_0,r,x^A)$ and $v^\alpha(u_0,r,x^A)$ are functionally unique. Furthermore, there is no remaining gauge freedom in the $\lambda$-dependent system. As we shall see, this arrangement leads formally to unique gravitational initial data when the Newtonian limit is preserved under evolution.

There seems to be no other natural way to eliminate diffeomorphism freedom and the resulting simplicity of a uniquely determined $\lambda$-dependent system for each choice of Newtonian initial data has great technical and conceptual advantages. However, it would be of interest to investigate the transformation properties under, say, change of position or initial velocity of the origin, to check the consequences of the Galilean symmetries of the Newtonian theory.

We note that, for $u>u_0$, $\rho$ and $v^\alpha$ become $\lambda$-dependent and therefore so does the pressure, for which we assume a $\lambda$-independent equation of state $p(\rho)$. Also, except in the presence of special symmetries, the center-of-mass world-line will veer away from the origin since the center-of-mass does not, in general, follow a geodesic (free-fall trajectory) of the Newton–Cartan theory.

We now apply the $\lambda$-expansion to formulate the iteration scheme which determines the initial gravitational data. After applying $\eth^{-1}$ to (2.2b) and $\eth^{-2}$ to (2.2d), the $n$th-order parts of Eqs. (2.2a), (2.2b), and (2.2d) take the form

$$-4r\beta^{(n)}{}_{,1}=J_\beta^{(n)}, \tag{3.2}$$

$$(r^4 Z^{(n)}_{,1})_{,1}=2r^4(\beta^{(n)}/r^2)_{,1}-(2+\eth\bar{\eth})r^2\alpha^{(n)}+J_Z^{(n)}, \tag{3.3}$$

$$[r^2(\alpha^{(n)}-Z^{(n)})]_{,1}=-2\beta^{(n)}+J_\alpha^{(n)}. \tag{3.4}$$

Here $J_\beta{}^{(n)}$, $J_Z{}^{(n)}$, and $J_\alpha{}^{(n)}$ are determined by order $(n-2)$ hypersurface fields except for an $\alpha_{,0}^{(n-1)}$ term in $J_\alpha{}^{(n)}$ [see (3.5) below]. Because of the use of the inverse $\eth$ operator, the $l=0$ part of $J_Z{}^{(n)}$ and the $l=0$ and $l=1$ parts of $J_\alpha{}^{(n)}$ are at this stage arbitrary, corresponding to the physical irrelevence of the corresponding parts of the potentials $Z^{(n)}$ and $\alpha^{(n)}$. They may be fixed at some later stage for calculational ease. In the expansion of the evolution equation (2.2d), the time derivative of $\alpha$ appears as an order $(n-1)$ term in $J^{(n)}$; more specifically

$$J_\alpha^{(n)}=2r\int\frac{(r^2\alpha_{,0}^{(n-1)})_{,1}}{r}+\text{order }(n-2)\text{ hypersurface terms.} \tag{3.5}$$

A Poisson equation for the $n$th-order gravitational data $\alpha^{(n)}$ can now be obtained in the following way. To Eq. (3.2), add the equation obtained from applying $\partial/\partial r$ to (3.3) and the equation obtained by applying $[-2+(\partial/\partial r)r^2(\partial/\partial r)]$ to (3.4). There results

$$r^2\nabla^2(r^2\alpha^{(n)})_{,1}=J_\beta^{(n)}+(r^2J^{(n)}_{\alpha,1})_{,1}-2J_\alpha^{(n)}+J^{(n)}_{Z,1}, \tag{3.6}$$

where $\nabla^2$ is the Laplacian, i.e.,

$$r^2\nabla^2=\frac{\partial}{\partial r}r^2\frac{\partial}{\partial r}+\eth\bar{\eth}.$$

This Poisson equation uniquely determines $\alpha^{(n)}$ in terms of order $(n-1)$ quantities provided the boundary conditions that $(r^2\alpha^{(n)})_{,1}$ vanish at infinity, modulo $l=0$ and $l=1$ harmonics, can be consistently applied. When desired, ambiguity with respect to $l=0$ and $l=1$ harmonics may be eliminated by considering the equation obtained by applying $(\eth\bar{\eth}+2)\eth\bar{\eth}$ to (3.6). Since this operator commutes with $\nabla^2$ and annihilates $l=0$ and $l=1$ terms, the resulting equation may be solved for $(\eth\bar{\eth}+2)\eth\bar{\eth}\alpha^{(n)}$ in standard fashion.

The iteration scheme begins at $n=0$ with specification of the Newtonian data $\rho(u_0,r,x^A)$ and $v^\alpha(u_0,r,x^A)$. This determines the initial value of $\alpha^{(0)}$ via (2.3). Assuming that the initial value of $\alpha^{(n)}$ has been determined, $\alpha^{(n+1)}$ is found by first integrating (3.2) and (3.3) to find $\beta^{(n)}$ and $Z^{(n)}$ and integrating the $n$th-order version of (2.2c) to find $W^{(n)}$. Then $\alpha^{(n)}{}_{,0}$ is the only remaining quantity needed to determine $\alpha^{(n+1)}$ via (3.6). But at this stage the iteration scheme has determined $\alpha^{(n)}$ completely in terms of the initial Newtonian data $\rho$ and $v^\alpha$ so that $\alpha^{(n)}{}_{,0}$ is ultimately determined by the time derivatives of the Newtonian $\rho$ and $v^\alpha$, which, in turn, are determined by Euler's equations for the Newtonian fluid. Thus, formally, the expansion can be continued to arbitrary order. The details involved in this procedure will become evident in the ensuing calculations leading up to a Poisson equation for $\alpha^{(2)}$.

### A. Formulas for $n=0$

From paper I, we have the following results:

$$r^2\alpha^{(0)}=-2\int\Phi^*, \tag{3.7}$$

$$\beta^{(0)} = 2\pi \int \rho^{(0)} r, \tag{3.8}$$

$$r^2 Z^{(0)} = r^2 \alpha^{(0)} + 2 \int \beta^{(0)}, \tag{3.9}$$

$$W^{(0)}/2r = \Phi^* - \beta^{(0)}. \tag{3.10}$$

Here (3.7) is just a rewriting of (2.3); and (3.8)–(3.10) result from inserting (3.7) into the hypersurface equations. Note that all $l = 0$ and $l = 1$ freedom has been removed. Equations (3.7)–(3.10) must hold for $u \geqslant u_0$. (For $u = u_0$, we have $\rho^{(0)} = \rho$.)

In working out higher $n$-order quantities, some additional relationships are useful. We define (the time-dependent quantity)

$$m\langle r, x^A \rangle = 4\pi \int \rho^{(0)} r^2.$$

Then, using $\rho r = (\partial/\partial r)(1/r) \int \rho r^2 + (1/r^2) \int \rho r^2$, we obtain

$$2\beta^{(0)} = \frac{m\langle r, x^A \rangle}{r} + \int \frac{m\langle r, x^A \rangle}{r^2}. \tag{3.11}$$

By parts integration, we also have

$$\int \beta^{(0)} = r\beta^{(0)} - \frac{m\langle r, x^A \rangle}{2}. \tag{3.12}$$

By applying $\eth\bar{\eth}$ to (3.7) and reexpressing the integral in terms of the Poisson equation for $\Phi^*$, we obtain

$$\eth\bar{\eth}\alpha^{(0)} = 2\Phi^*{}_{,1} - 2m\langle r, x^A \rangle / r^2. \tag{3.13}$$

To determine the dyad quantity $P = m^A q_A / \lambda^2$, we carry out the dyad expansion of $h_{AB,1} q^A q^B = \lambda^2 \eth^2 \alpha$. We find

$$2P^{(0)} = -\eth^2 \int \alpha^{(0)}, \tag{3.14}$$

where, according to (3.13),

$$\eth\bar{\eth} \int \alpha^{(0)} = 2\Phi^* - 2 \int \frac{m\langle r, x^A \rangle}{r^2}. \tag{3.15}$$

### B. Formulas for $n = 1$

For $u = u_0$, the Poisson equation (3.6) takes the explicit form

$$r^2 \nabla^2 (r^2 \alpha^{(1)})_{,1} = 16\pi (\rho^{(0)} v^{(0)} r^2)_{,1} - 16\pi \rho^{(0)} r^2 v_1{}^{(0)} - 4(r^3 \Phi^*{}_{,0})_{,1}/r, \tag{3.16}$$

where we have introduced a potential for the fluid's angular momentum density by $\rho v_A q^A = \eth(\rho\nu)/\sqrt{2}$. (For $u = u_0$, we have $\nu^{(0)} = \nu$ and $v_1{}^{(0)} = v_1$.) The term involving $\Phi^*{}_{,0}$ in (3.16) arises from rewriting the $\alpha^{(0)}{}_{,0}$ terms, indicated in (3.5) and (3.6), using $(r^2 \alpha^{(0)})_{,1} = -2\Phi^*$. We can determine $\Phi^*{}_{,0}$ from the time derivative of the Newtonian Poisson equation

$$\nabla^2 \Phi^*{}_{,0} = 4\pi \rho^{(0)}_{,0}.$$

Using the Newtonian mass conservation equation, this becomes

$$r^2 \nabla^2 \Phi^*{}_{,0} = 4\pi (r^2 \rho v_1)_{,1} + 2\pi \eth\bar{\eth}[\rho(\nu + \bar{\nu})], \tag{3.17}$$

for $u = u_0$, thus determining $\Phi^*{}_{,0}$ from the initial matter data. In this way, the determination of the order $n = 1$ gravitational data reduces to solving the Poisson equations (3.17) and (3.16), in that order.

The order $n = 1$ hypersurface equations, for $u = u_0$, take the form

$$\beta^{(1)} = 4\pi \int \rho v_1 r, \tag{3.18}$$

$$(r^4 Z^{(1)}_{,1})_{,1} = 2r^4 (\beta^{(1)}/r^2)_{,1} - (2 + \eth\bar{\eth}) r^2 \alpha^{(1)} + 16\pi \rho r^2 \nu, \tag{3.19}$$

$$W^{(1)}_{,1} = (2 - \eth\bar{\eth})\beta^{(1)} + \eth\bar{\eth}[r^4 (Z^{(1)} + \bar{Z}^{(1)})]_{,1}/4r^2 + \eth^2\bar{\eth}^2 \int \frac{(\alpha^{(1)} + \bar{\alpha}^{(1)})}{4}. \tag{3.20}$$

In general, the data $\alpha^{(1)}$ will not be real, as in the $n = 0$ case, due to the presence of the angular momentum source term in (3.16), i.e.,

$$r^2 \nabla^2 (r^2 \operatorname{Im} \alpha^{(1)})_{,1} = 16\pi (\rho^{(0)} r^2 \operatorname{Im} \nu^{(0)})_{,1}.$$

Thus the shear at this order will have both "electric" and "magnetic" parts. Similarly $Z^{(1)}$ also has, in general, an imaginary part. However, the equation

$$[r^2(\alpha^{(1)} - Z^{(1)})]_{,1} = 2r\left(r \int \alpha^{(0)}_{,0}\right)_{,1} - 2\beta^{(1)}, \tag{3.21}$$

which results from the $\lambda$-expansion of (2.2d), implies that $\operatorname{Im}(\alpha^{(1)} - Z^{(1)}) = 0$.

### C. Formulas for $n = 2$

According to (3.6), the gravitational data are obtained from

$$r^2 \nabla^2 (r^2 \alpha^{(2)})_{,1} = J^{(2)}_{\beta} + (r^2 J^{(2)}_{\alpha,1})_{,1} - 2J^{(2)}_{\alpha} + J^{(2)}_{Z,1}. \tag{3.22}$$

Calculation of the source terms becomes lengthy, although straightforward, at this state. We find, for $u = u_0$,[17]

$$J^{(2)}_{\beta} = -8\pi r^2 (\rho v_1{}^2 + p) - \tfrac{1}{2} r^2 (\eth^2 \alpha^{(0)}) \bar{\eth}^2 \alpha^{(0)}, \tag{3.23}$$

$$J^{(2)}_{Z} = \eth^{-1}\{16\pi (2)^{1/2} r^2 \rho v_1 v_A q^A + 2(\beta^{(0)} r^4 \eth Z^{(0)}_{,1})_{,1} + [r^4 (\eth^2 \alpha^{(0)}) \bar{\eth} Z^{(0)}]_{,1} - r^2 \bar{P}^{(0)} \eth^3 \alpha^{(0)} - 2r^2 (\eth \bar{P}^{(0)}) \eth^2 \alpha^{(0)} + r^2 P^{(0)} \eth\, \bar{\eth}^2 \alpha^{(0)}\}, \tag{3.24}$$

$$J^{(2)}_{\alpha} = 2r \int \frac{(r^2 \alpha^{(1)}_{,0})_{,1}}{r} + \eth^{-2}\{-16\pi \rho (v_A q^A)^2 - 2(\eth\beta^{(0)})^2 - 4\beta^{(0)} \eth^2 \beta^{(0)} - (rW^{(0)} \eth^2 \alpha^{(0)})_{,1} - \tfrac{1}{2} r^4 (\eth Z^{(0)}_{,1})^2 + r^2 (\bar{\eth}\, \eth^2 \alpha^{(0)}) \eth Z^{(0)} - 4P^{(0)} \eth\, \bar{\eth}\Phi^* - 2(\eth P^{(0)}) \bar{\eth}\Phi^* + 2(\bar{\eth} P^{(0)}) \eth\Phi^*\}. \tag{3.25}$$

The only term occurring in (3.23)–(3.25) which has not been previously reduced in terms of matter data is $\alpha^{(1)}{}_{,0}$. It is determined by taking the time derivative of (3.16). In the process, it is necessary to use the Euler equations of the Newtonian fluid to reduce the expressions $(\rho\nu)_{,0}$, $(\rho v_1)_{,0}$, and $\Phi^*{}_{,00}$ to hypersurface quantities. We omit the details here as it turns that the $\alpha^{(1)}{}_{,0}$ term plays no essential role in our remaining considerations.

The inverse $\eth$ operator appearing in (3.24) and (3.25) cannot be explicitly removed, as was possible in the $n = 0$ and $n = 1$ cases. However, taking note of the identity $\eth^2\bar{\eth}^2\alpha = \eth\,\bar{\eth}(\eth\,\bar{\eth} + 2)\alpha$, an equation equivalent to (3.22), free of inverse $\eth$ operators, may be obtained in the manner suggested in the discussion following (3.6). Alternatively, (3.22)

may be used as it stands by simply assigning a unique definition to $\eth^{-1}$. For instance, define $\eth^{-1}$ of a spin-weight $s$ function $\eta$ to be that spin-weight $(s-1)$ function $f$ which contains no harmonics with $l<s$ and satisfies $\eth f=\eta$.

## IV. ASYMPTOTIC BEHAVIOR

In the previous section, we have presented an iteration scheme which determines Einsteinian characteristic initial data from Newtonian initial data by means of the Poisson equation (3.6). The source terms of this Poisson equation reduce to two types: (i) matter terms which vanish outside the compact support of the density $\rho$, and (ii) gravitational terms which extend to infinity. We now investigate, for $n=0$, 1, and 2, whether the asymptotic behavior of these gravitational terms is compatible with the existence of a unique solution to the Poisson equation. Actually, since only the $l\geqslant 2$ parts of the potential $\alpha$ for the gravitational data are of physical relevance, our interest lies in the projected Poisson equation

$$\nabla^2(r^2\mathscr{P}\alpha^{(n)})_{,1}=\mathscr{P}S^{(n)}, \tag{4.1}$$

where $\mathscr{P}$ is the operator which projects out $l=0$ and $l=1$ harmonics and

$$r^2S^{(n)}=\mathscr{P}\left[J_\beta^{(n)}+r^2((r^2J_\alpha^{(n)})_{,1}/r^2)_{,1}+J_{Z,1}^{(n)}\right]. \tag{4.2}$$

Here (4.1) follows from (3.6) in view of the identity $[\mathscr{P}, \nabla^2]=0$.

For $n=0$, no asymptotic difficulties arise but we need the details of the asymptotic behavior in order to investigate the $n=1$ and $n=2$ cases. According to (2.4)–(2.6),

$$\alpha^{(0)}=-a_mY_{1m}-2\left(\frac{a}{r}\right)-\left(\frac{2}{r^2}\right)\int\Phi, \tag{4.3}$$

where, in the region exterior to the matter,

$$\Phi\triangleq-(m/r)+(d_mY_{1m}/r^2)+\mathscr{P}\phi_{lm}Y_{lm}/r^{l+1}. \tag{4.4}$$

Here, as is evident from (2.4), $a=-\Phi_0$ (the value of the standard Newtonian potential at the origin), $a_m$ determines the components of $\nabla\Phi$ at the origin, and $d_m$ determines the Newtonian dipole moments. The choice of initial center-of-mass coordinates would imply that $d_m$ and $d_{m,0}$ vanish, for $u=u_0$. However, in this section, we relax this center-of-mass condition to demonstrate that it has no substantial effect on the behavior of $\mathscr{P}\alpha^{(1)}$ or $\mathscr{P}\alpha^{(2)}$. From (4.3) and (4.4), $\alpha^{(0)}$, in the exterior region, has the form

$$\alpha^{(0)}\triangleq-a_mY_{1m}-\frac{2a}{r}+\frac{2m\ln r}{r^2}+\frac{2d_mY_{1m}}{r^3}+\frac{c^{(0)}}{r^2}+\mathscr{P}\frac{2\phi_{lm}Y_{lm}}{lr^{l+2}}. \tag{4.5}$$

The "integration constant" $c^{(0)}$ is determined by the matter distribution by using (3.13):

$$\bar{\eth}\,\eth\alpha^{(0)}=2a_mY_{1m}+2\Phi_{,1}-2m\langle r,x^A\rangle/r^2. \tag{4.6}$$

Comparing this with (4.5), we have $\eth\,\bar{\eth}c^{(0)}=2(m-m\langle x^A\rangle)$, where $m\langle x^A\rangle=m\langle\infty,x^A\rangle$. To find $P^{(0)}$, first note that (3.15) implies

$$\eth\,\bar{\eth}\int\alpha^{(0)}=2ra_mY_{1m}+2a+2\Phi-2\int\frac{m\langle r,x^A\rangle}{r^2},$$

which gives

$$\eth\,\bar{\eth}\int\alpha^{(0)}\triangleq 2ra_mY_{1m}+k^{(0)}-\frac{2(m-m\langle x^A\rangle)}{r}+\frac{2d_mY_{1m}}{r^2}+\mathscr{P}\frac{2\phi_{lm}Y_{lm}}{r^{l+1}}, \tag{4.7}$$

where $k^{(0)}=2a-2\int^\infty m\langle r,x^A\rangle/r^2$. Before finding $P^{(0)}$, we must remove the $\eth\,\bar{\eth}$ from the left side of (4.7). We first need a formula for $a=-\Phi_0$. Setting $m\langle r\rangle=(1/4\pi)\times\oint d\Omega\, m\langle r,x^A\rangle$, we have

$$\begin{aligned}\Phi_0&=\frac{1}{4\pi}\oint d\Omega\,\Phi_0=\frac{1}{4\pi}\int^\infty\oint d\Omega\,\Phi_{,1}\\&=-\frac{1}{4\pi}\int^\infty\oint\frac{d\Omega\,(r^2\Phi_{,1})_{,1}}{r}\\&=-\frac{1}{4\pi}\int^\infty\oint d\Omega\, r\nabla^2\Phi=-\int^\infty\oint d\Omega\,\rho r\\&=-\int^\infty\frac{1}{r}(m\langle r\rangle)_{,1}=-\int^\infty\frac{m\langle r\rangle}{r^2}.\end{aligned} \tag{4.8}$$

Thus $k^{(0)}=2\int^\infty(m\langle r\rangle-m\langle r,x^A\rangle)/r^2$, which has no $l=1$ part, so that we may set $k^{(0)}=\eth\,\bar{\eth}K^{(0)}$. Then, combining (3.14) and (4.7), we obtain

$$P^{(0)}\triangleq\eth^2\left(-\frac{K^{(0)}}{2}+\frac{c^{(0)}}{2r}+\frac{\phi_{lm}Y_{lm}}{l(l+1)r^{l+1}}\right). \tag{4.9}$$

Formulas for the remaining $n=0$ variables follow quite readily from inserting these results into (3.8)–(3.12):

$$\beta^{(0)}\triangleq a/2-k^{(0)}/4\equiv H^{(0)}, \tag{4.10}$$

$$Z^{(0)}\triangleq\alpha^{(0)}+2H^{(0)}/r-m\langle x^A\rangle/r^2, \tag{4.11}$$

$$W^{(0)}\triangleq 2r(\Phi^*-H^{(0)}), \tag{4.12}$$

where $H^{(0)}$, defined in (4.10), is independent of $r$.

Thus all the physically relevant $n=0$ quantities have analytic asymptotic expansions in $(1/r)$. The important features are that $\int^\infty m\langle r,x^A\rangle/r^2$ determines the asymptotic geometry of the $r=$ const surfaces, that $m\langle x^A\rangle$ determines the asymptotic shear of the null cone, and that the Newtonian quadrupole moments determine the asymptotic Weyl data. (Formulas for the Weyl data are given in Sec. V.)

It is also true that the order $n=1$ quantities have analytic $1/r$ dependence in the region exterior to the matter. This follows from an inspection of the source terms in (3.16). In the region exterior to the matter, (3.16) reduces to

$$\begin{aligned}\nabla^2(r^2\mathscr{P}\alpha^{(1)})_{,1}&\triangleq-4\mathscr{P}(r^3\Phi_{,0})_{,1}/r^3\\&\triangleq 4\mathscr{P}(l-2)\phi_{lm,0}Y_{lm}/r^{l+2},\end{aligned} \tag{4.13}$$

so that the exterior gravitational source is analytic in $1/r$. The solution of (4.13) consists of two pieces, $(r^2\mathscr{P}\alpha^{(1)})_{,1}=A^{(1)}+B^{(1)}$, where, in the exterior, $\nabla^2A^{(1)}=0$ so that $A^{(1)}=O(1/r^3)$ is analytic and $B^{(1)}$ is the solution of (4.13) given by

$$B^{(1)}\triangleq 2[(2-l)/l\,]\mathscr{P}\phi_{lm,0}Y_{lm}/r^l=O(r^{-3}). \tag{4.14}$$

Thus, in the exterior, $\mathscr{P}\alpha^{(1)}$ is analytic and has the form $\mathscr{P}\alpha^{(1)}=c^{(1)}/r^2+O(1/r^4)$. It is now easy to check that all other $n=1$ quantities are analytic in the exterior. The details will not be needed.

For $n=2$, the analysis of asymptotic behavior is considerably more difficult because the source term $S^{(2)}$, in (4.1),

contains quadratic combinations of $n = 0$ quantities. Since, in the exterior, all order $n = 0$ and $n = 1$ quantities are analytic in $1/r$, so is $S^{(2)}$. However, that is not sufficient to guarantee an analytic solution for $(r^2\mathscr{P}\alpha^{(2)})_{,1}$. It is adequate to investigate $S^{(2)}$ to the leading order in $1/r$ to establish that no substantial problems arise. We expand $S^{(2)} \triangleq \Sigma S^{(2,k)}/r^k$. The possible leading terms are $k = 4$ and $k = 5$. The Newtonian monopole and dipole quantities $a_m$, $a$, $m$, and $d_m$ do not enter into either $S^{(2,4)}$ or $S^{(2,5)}$, even though the quadratic coupling would allow them to contribute to $l \geqslant 2$ source terms. Also, an inspection of (4.2) shows that $J_\alpha{}^{(2)}$ cannot contribute to $S^{(2,4)}$ because of its analyticity in $1/r$. This simplifies the calculation of $S^{(2,4)}$. By inserting the asymptotic expansions (4.4) and (4.5) and (4.9)–(4.12) into Eqs. (3.23) and (3.24) for $J_\beta{}^{(2)}$ and $J_Z{}^{(2)}$, we in fact find that $S^{(2,4)} = 0$.

We next expand $S^{(2,5)} = S^{(2,5)}{}_{lm} Y_{lm}{}^{(2,5)}$, with $l \geqslant 2$. The inhomogeneous solution of $\nabla^2 B^{(2,5)} = S^{(2,5)}/r^5$ has a possible $\ln r$ term:

$$B^{(2,5)} = -\frac{S^{(2,5)}_{2m} Y_{2m} \ln r}{5r^3} - \sum_{l=3}^{\infty} \frac{S^{(2,5)}_{lm} Y_{lm}}{(l^2 + l - 6)r^3}. \tag{4.15}$$

In order to investigate the existence of the $\ln r$ term, we reexpress $S^{(2,5)}$ in terms of Newtonian data. It is easy to check that $J_\beta{}^{(2)}$ makes no contribution so that $S^{(2,5)} = S_Z{}^{(2,5)} + S_\alpha{}^{(2,5)}$, in terms of the contributions from $J_Z{}^{(2)}$ and $J_\alpha{}^{(2)}$. By inserting the asymptotic expressions for the order $n = 0$ quantities into (3.24), we obtain

$$\begin{aligned} S_Z^{(2,5)} = \mathscr{P}\eth^{-1}[&4(\eth\,\overline{\eth}K^{(0)})\eth\phi_{2m} Y_{2m} \\ &- (\overline{\eth}^2\eth K^{(0)})\eth^2\phi_{2m} Y_{2m} + (\overline{\eth}^2 K^{(0)})\eth^3\phi_{2m} Y_{2m} \\ &- 2(\eth\,\overline{\eth}^2 K^{(0)})\eth^2\phi_{2m} Y_{2m} + (\eth^2 K^{(0)})\eth\,\overline{\eth}^2\phi_{2m} Y_{2m} \\ &- 8a\eth\phi_{2m} Y_{2m} + 4(\eth^2 c^{(0)})\overline{\eth} d_m Y_{1m} \\ &- \tfrac{4}{3}(\overline{\eth} a_{m'} Y_{1m'})\eth^2\phi_{3m} Y_{3m}]. \end{aligned} \tag{4.16}$$

Referring to (3.25), $S_\alpha{}^{(2,5)}$ consists of two pieces. The hypersurface part is determined in the same manner just used to find $S_Z{}^{(2,5)}$. To express the time-derivative contribution, we set $(r^2\mathscr{P}\alpha^{(1)})_{,1} \triangleq \mathscr{P}A^{(1)}{}_{lm} Y_{lm}/r^{l+1} + \mathscr{P}B^{(1)}{}_{lm} Y_{lm}/r^l$, in accordance with the discussion following (4.13). We then obtain

$$\begin{aligned} S_\alpha^{(2,5)} = &-2(A_{3m} Y_{3m} + B_{4m} Y_{4m})_{,0} \\ &+ 4\mathscr{P}\eth^{-2}[(\eth(3\eth\,\overline{\eth} - 2)K^{(0)})\eth\phi_{2m} Y_{2m} \\ &- 12(\eth^2 K^{(0)})\phi_{2m} Y_{2m} + (\eth\,\overline{\eth} K_{(0)})\eth^2\phi_{2m} Y_{2m} \\ &+ (\eth^3 K^{(0)})\overline{\eth}\phi_{2m} Y_{2m} + 2a\eth^2\phi_{2m} Y_{2m} \\ &+ 8(\eth^2 c^{(0)})d_m Y_{1m} - (\eth^3 c^{(0)})\overline{\eth} d_m Y_{1m} \\ &- 3(\eth(2 + \eth\,\overline{\eth})c^{(0)})\eth d_m Y_{1m} + \tfrac{10}{3} a_{m'} Y_{1m'}\eth^2\phi_{3m} Y_{3m} \\ &+ 5(\eth a_{m'} Y_{1m'})\eth\phi_{3m} Y_{3m} - \tfrac{1}{6}(\overline{\eth} a_{m'} Y_{1m'})\eth^3\phi_{3m} Y_{3m}]. \end{aligned} \tag{4.17}$$

We are interested in $S_{2m}{}^{(2,5)}\ Y_{2m}$. When we combine (4.16) and (4.17) and extract out the $l = 2$ part, only the terms involving products of $K^{(0)}$ and $\phi_{2m} Y_{2m}$ survive. The cancellation of some of the terms in (4.16) and (4.17) is obvious. In other cases, it is helpful to use the identity $\frac{1}{24}\eth^2\overline{\eth}^2\ Y_{2m} = Y_{2m}$. Thus we can apply $\frac{1}{24}\eth^2\overline{\eth}^2$ to the sum of (4.16) and (4.17) without changing the $l = 2$ part. After doing this, any term of the form $\overline{\eth}^3 f$ may be discarded since $\overline{\eth}^3$ annihilates $l = 2$ terms.[18] As an example of how this technique works consider the terms $C$ involving $c^{(0)}$ in $\eth^2\overline{\eth}^2 S^{(2,5)}$:

$$\begin{aligned} C &= \overline{\eth}^2[4(\eth^2 c^{(0)})\eth\overline{\eth} d_m Y_{1m} + 32(\eth^2 c^{(0)}) d_m Y_{1m} \\ &\quad - 12(2\eth c^{(0)} + \eth\,\overline{\eth}\eth c^{(0)})\eth d_m Y_{1m}] \\ &= \overline{\eth}^2[-12(\eth^2 c^{(0)})\eth\,\overline{\eth} d_m Y_{1m} \\ &\quad - 12((2 + \eth\,\overline{\eth})\eth c^{(0)})\eth d_m Y_{1m}] \\ &= -12\overline{\eth}^3[(\eth^2 c^{(0)})\eth d_m Y_{1m}] \\ &\quad - 12\overline{\eth}^2[((2 + \eth\,\overline{\eth} - \overline{\eth}\,\eth)\eth c^{(0)})\eth d_m Y_{1m}] \\ &= -12\overline{\eth}^3[(\eth^2 c^{(0)})\eth d_m Y_{1m}]. \end{aligned}$$

In this way, we find from (4.16) and (4.17)

$$\begin{aligned} S^{(2,5)}_{2m} Y_{2m} = \tfrac{1}{6}\mathscr{P}_2\overline{\eth}^2[&(4\eth^2\overline{\eth}K^{(0)} - 2\eth K^{(0)})\eth\phi_{2m} Y_{2m} \\ &- 24(\eth^2 K^{(0)})\phi_{2m} Y_{2m} \\ &+ \tfrac{5}{2}(\eth\,\overline{\eth} K^{(0)})\eth^2\phi_{2m} Y_{2m}], \end{aligned}$$

where $\mathscr{P}_2$ is the projection operator for $l = 2$ harmonics. Using the technique of dropping terms of the form $\overline{\eth}^3 f$, we can then establish that $S_{2m}{}^{(2,5)} = 0$. Thus the $\ln r$ term in (4.15) vanishes.

As a result, the Poisson equation for $(r^2\mathscr{P}\alpha^{(2)})_{,1}$ admits a unique solution with $O(1/r^3)$ asymptotic behavior. To each source term $S_{lm}{}^{(2,k)}$ there corresponds a $(1/r^{k-2})$-type solution for $k \neq l + 3$ and a possible $(\ln r/r^{k-2})$-type solution for $k = l + 3$. The solution has the form $(r^2 P\alpha^{(2)})_{,1} \triangleq A^{(2)} + B^{(2)}$, where $\nabla^2 A^{(2)} = 0$ and $\nabla^2 B^{(2)} = S^{(2)}$, with $A^{(2)} = O(1/r^3)$ and $B^{(2)} = O(1/r^3)$. As discussed in the next section, the order $(n = 2)$ data are consistent with asymptotic flatness.

## V. DISCUSSION

As shown in the preceeding section, the iteration scheme produces unique null data up through $\alpha^{(2)}$, with asymptotic behavior $\alpha^{(0)} = c^{(0)}/r^2 + O(1/r^4)$, $\alpha^{(1)} = c^{(1)}/r^2 + O(1/r^4)$ and $\alpha^{(2)} = c^{(2)}/r^2 + O(1/r^4)$. To what order $\ln r/r^k$ terms appear in $\alpha^{(2)}$ and to what order the uniqueness of the iteration scheme can be extended are two of the questions emerging from this work that deserve further attention.

In principle the existence of unique data at the $\alpha^{(2)}$ level should lead to unique corrections, of the corresponding order, to the Euler equations of Newtonian hydrodynamics. These are obtained from a $\lambda$-expansion of $T^{\mu\nu}{}_{;\nu} = 0$. As discussed in paper I, the leading-order terms give the Euler equations. The next-order terms would give the corrections introduced by $\rho^{(1)}$ and $v_\alpha{}^{(1)}$, and so forth. However, a preliminary investigation reveals enormous technical difficulties in comparing such corrections with those obtained by spacelike methods. This is due to the complicated transformation relating null coordinates to harmonic coordinates. Even in the flat-space case, the scalar $\rho_h$, representing the density in harmonic coordinates, has the awkward expansion $\rho_h(t) = \rho(u + \lambda r) = \rho^{(0)} + \lambda r\rho^{(1)} + \cdots$. Unfortunately, there are no known gauge invariant consequences arising from post-Newtonian hydrodynamics that different formalisms

can apply as a common algorithm for purposes of comparison.

The corrections to Newtonian hydrodynamics do not require detailed asymptotic knowledge of the null data.[19] For some neighborhood $u_0 \leqslant u < u_1$ of the initial cone, only the initial null data in a region of compact support can influence the fluid. However, the *radiation fields* produced by the fluid at $u = u_0$ are influenced by the null data on the entire initial cone. Furthermore, the only known geometrically meaningful description of radiation fields is based upon asymptotic flatness, which requires asymptotically flat null data.[14,20] This latter requirement is expressed most simply in terms of the Weyl data, the component of the Weyl tensor intrinsic to an outgoing null hypersurface. One consequence of the peeling theorem associated with asymptotic flatness is that this Weyl data has a certain asymptotic falloff. In our present null coordinate system, this condition takes the form

$$\Psi \equiv (1/\lambda^4)\, C_{1AB1} m^A m^B = O(1/r^3). \tag{5.1}$$

In terms of metric quantities,

$$\Psi = -\tfrac{1}{2}(r^2 m^A m^B c_{AB})_{,1} + \lambda^2 \beta_{,1} c_{AB} m^A m^B.$$

A $\lambda$-expansion then gives, in the exterior region, $\Psi \triangleq -\frac{1}{2}\eth^2(r^2\alpha)_{,1} + O(\lambda^4)$, which satisfies (5.1) if $(r^2 \mathscr{P}\alpha^{(n)})_{,1} = O(1/r^3)$. The results of Sec. IV establish that this asymptotic flatness condition is satisfied for $n$ equals 0, 1, and 2. The fact that asymptotic flatness is maintained for $n = 2$ is quite encouraging since, at this order, the source terms include quadratic self-coupling to order ($n = 0$) fields. The Weyl tensor component $\Psi$ is invariantly picked out by the intrinsic geometry of the null hypersurface. Thus, given that $\Psi^{(0)}$ and $\Psi^{(1)}$ satisfy (5.1), the result that $\Psi^{(2)}$ satisfies (5.1) is completely gauge invariant.

In the axisymmetric case, an explicit formula for the Bondi[21] news function in terms of null cone initial data is given in Ref. 15. Calculation of the news function for the $\lambda$-dependent null data obtained in Sec. III gives evidence that, to leading order, these null data properly exclude incoming radiation. It is known that, in general, the null data $\alpha = 0$ contain incoming radiation which acts as one of the sources for the outgoing radiation. The choice $\alpha = \alpha^{(0)}$ (with the $\alpha^{(0)}$ given in Sec. III) changes the $\lambda$-dependence of the news by a factor of $\lambda$, reflecting the exclusion of incoming radiation to leading order. In addition, analytic and numerical studies[16] of special cases indicate that the choice $\alpha = \alpha^{(0)} + \lambda\alpha^{(1)}$ introduces an additional factor of $\lambda$ in the news function.

These preliminary results up to the ($n = 2$) order give encouragement that the approach of this paper is on the right track. However, at some sufficiently high order, it would seem likely that the delicate analytic conditions, which successfully eliminated incoming fields in the simple model discussed in the Introduction, break down on account of nonlinearity. Such nonlinearity generally produces backscattering and radiation tails which obscure any precise attempt to exclude the contribution of incoming fields from null data on an outgoing null hypersurface. In the space-like approach, the evaluation of certain integrals force Futamase and Schutz[8,9] to introduce an asymptotic approximation to Newtonian theory, which allows $\lambda^n \ln \lambda$ terms, rather than a series expansion valid about $\lambda = 0$. Some aspect of this might conceivably arise at higher order in the null approach. For instance, the asymptotic behavior of the source terms might eventually lead to nonunique null data. This, in fact, does occur to leading order in the space-like approach in which solutions are obtained by averaging over all possible Cauchy data.[8,9] Also, ln $r$ terms, which are commonplace in the space-like approach, may eventually arise. However, there is some reason to hope they stem from the particular nature of harmonic coordinates[22] and would not appear in null coordinates to any asymptotic order affecting radiative properties.

A preliminary study indicates that, for quasi-Newtonian radiative systems, the leading term in $\lambda$ of the news function should depend upon the time-derivative of $c^{(2)}$ and the asymptotic value of $P^{(3)}$. All the theoretical issues are now resolved to carry out such a calculation unambiguously. In particular, the pieces are in place to check the Einstein quadrupole formula for radiation. Regrettably the calculations are too complicated to carry out at this time.

Ehlers has recently emphasized[23]: "Since exact models of radiating systems are not in sight, it appears to be urgent to connect the concepts and constructions of asymptotics with approximation methods, even if that requires sacrificing some of its geometrical beauty or modifying its assumptions."

At present, asymptotics provides the only context for a meaningful description of radiation fields in general relativity. To abandon its geometric spirit, without some viable alternative, would reopen the door to the same kind of confusion concerning radiation fields that existed prior to the pioneering work of Bondi.[21] The $\lambda$-expansion offers an approach to approximating quasi-Newtonian systems without sacrificing a rigorous description of null infinity.

One very attractive strategy is to view our approach as providing "exact models of radiating systems" in the full spirit of general relativity. The role of the Newtonian limit, from this viewpoint, would only be to provide initial null data to approximate some quasi-Newtonian astrophysical system. Rather than demanding that the Newtonian evolution of the Newtonian data exactly agree with the $\lambda = 0$ limit of the Einsteinian evolution, only a certain degree of tangency in time would be required. This would free up the null data at higher $n$ orders for which its physical importance might be questionable. Given such quasi-Newtonian null data, Einstein's equation would then be treated exactly. Ultimately, this might even lead to a way to discuss the behavior of a quasi-Newtonian astrophysical system in general relativistic rather than Newtonian terms. What lends this strategy some hope of practicability is the recent success with numerical calculations of gravitational radiation for null data obtained by this means.[16]

## ACKNOWLEDGMENTS

I have had the advantage and pleasure of discussions of the space-like approach with B. F. Schutz. I have drawn from ideas generated in ongoing projects with R. A. Isaacson and J. S. Welling. And I have had the benefits of the criticism and reaction of my fellow members of the Pittsburgh Relativity Group—C. Kozameh, A. I. Janis, O. Moreschi, E. T. Newman, J. Porter, and G. Sparling.

This work was supported in part by NSF Contract PHY800823.

[1]J. Winicour, J. Math. Phys. **24**, 1193 (1983).
[2]V. Fock, *The Theory of Space Time and Gravitation* (Pergamon, New York, 1959).
[3]S. Chandrasekhar, Astrophys. J. **142**, 1488 (1965).
[4]S. Chandrasekhar and E. P. Esposito, Astrophys. J. **160**, 153 (1970).
[5]J. L. Anderson and T. C. Decanio, Gen. Rel. Grav. **6**, 197 (1975).
[6]G. D. Kerlick, Gen. Rel. Grav. **12**, 467 (1980).
[7]G. D. Kerlick, Gen. Rel. Grav. **12**, 521 (1980).
[8]T. Futamase and B. F. Schutz, "Newtonian and post-Newtonian Approximations are Asymptotic to General Relativity," Phys. Rev. (to appear).
[9]T. Futamase, "Gravitational Radiation Reaction in the Newtonian Limit," Phys. Rev. (to appear).
[10]A derivation based upon a particle rather than hydrodynamic approach has also recently been given by T. Damour and N. Deruelle, Phys. Lett. A **87**, 81 (1981).
[11]S. Persides, Astrophys. J. **170**, 479 (1971).
[12]B. F. Schutz has informed me that analogous constraints in the space-like case may also lead to unique Cauchy data (private communication).
[13]In Ref. 8, $\tilde{\rho}$, $\tilde{p}$, and $\tilde{v}^{\alpha}$ determine a "regular, asymptotically Newtonian sequence."
[14]R. Penrose, Phys. Rev. Lett. **10**, 66 (1963).
[15]R. A. Isaacson, J. S. Welling, and J. Winicour, J. Math. Phys. **24**, 1824 (1983).
[16]J. S. Welling, Ph.D. thesis, University of Pittsburgh, 1983.
[17]For $u > u_0$, we must make the substitutions $(\rho v_1^2 + p) \to \rho^{(0)}[2v_1^{(1)} + (v_1^{(0)})^2] + 2\rho^{(1)}v_1^{(0)} + \rho^{(2)} + p^{(0)}$ in (3.23); $\rho v_1 v_A \to \rho^{(0)}(v_1^{(0)}v_A^{(0)} + v_A^{(1)}) + \rho^{(1)}v_A^{(0)}$ in (3.24); and $\rho(v_A q^A)^2 \to \rho^{(0)}(v_A^{(0)}q^A)^2$, in (3.25).
[18]This technique substitutes for a more cumbersome analysis in terms of Clebsch–Gordan coefficients.
[19]Only the necessary asymptotic behavior to guarantee a unique solution of the Poisson equation is required.
[20]R. Geroch, in *Asymptotic Structure of Space-Time*, edited by F. P. Esposito and L. Witten (Plenum, New York, 1977).
[21]H. Bondi, M. G. J. Van der Burg, and A. W. K. Metzner, Proc. R. Soc. London Ser. A **270**, 103 (1962).
[22]R. A. Isaacson and J. Winicour, Phys. Rev. **168**, 1451 (1968).
[23]J. Ehlers, Ann. N. Y. Acad. Sci. **336**, 279 (1980).

# Self-gravitating isotropic and anisotropic fluid distributions

K. D. Krori
*Mathematical Physics Forum, Cotton College, Gauhati-781001, India and Institute of Advanced Study in Science and Technology, Assam Science Society, Gauhati-781001, India*

D. Nandy
*Department of Physics, St. Anthony's College, Shillong-793001, India*

(Received 13 September 1983; accepted for publication 23 December 1983)

Some exact solutions of isotropic and anisotropic fluid distributions in general relativity have been generated in this paper on the basis of the Letelier theory of fluids.

PACS numbers: 04.20.Jb, 97.10.Nf

## I. INTRODUCTION

Isotropy is a common feature in the astrophysical studies of massive objects in general relativity. Recent investigations on more realistic stellar objects, however, indicate that the interior matter distributions may be anisotropic at least in certain density ranges.[1,2]

Recently, Letelier[3] has developed a theory of anisotropic fluids the stress-energy tensor of which is obtained by combining the stress tensors of two irrotational perfect fluids. Each of the two fluid components obeys a stiff equation of state (pressure = energy density) and is described in terms of a massless scalar field. If $\Psi$ and $\Phi$ are the massless scalar fields, they are given by the following field equations:

$$\Psi^{,\mu}_{;\mu} = 0, \tag{1}$$

$$\Phi^{,\mu}_{;\mu} = 0, \tag{2}$$

and their stress-energy tensor is

$$T_{\mu\nu} = \Psi_{,\mu}\Psi_{,\nu} + \Phi_{,\mu}\Phi_{,\nu} - \tfrac{1}{2}(\Psi_{,\alpha}\Psi^{,\alpha} + \Phi_{,\beta}\Phi^{,\beta})g_{\mu\nu}, \tag{3}$$

with $T^{\mu\nu}_{;\nu} = 0$. The Einstein field equations are (with $C = 1$ and $8\pi G = 1$)

$$R_{\mu\nu} - \tfrac{1}{2}g_{\mu\nu}R = -T_{\mu\nu}. \tag{4}$$

The four-velocities of the perfect fluid components are

$$u_\mu = \Psi_{,\mu}/(\Psi_{,\alpha}\Psi^{,\alpha})^{1/2}, \tag{5}$$

$$v_\mu = \Phi_{,\mu}/(\Phi_{,\alpha}\Phi^{,\alpha})^{1/2}, \tag{6}$$

where $\Psi_{,\alpha}\Psi^{,\alpha} \neq 0$ and $\Phi_{,\alpha}\Phi^{,\alpha} \neq 0$.

The pressure and energy of the fluid components are, respectively,

$$P_\Psi = \rho_\Psi = \tfrac{1}{2}\Psi_{,\alpha}\Psi^{,\alpha}, \tag{7}$$

$$P_\Phi = \rho_\Phi = \tfrac{1}{2}\Phi_{,\alpha}\Phi^{,\alpha}. \tag{8}$$

Defining a complex scalar field $\Lambda$ as

$$\Lambda = \Psi + i\Phi, \quad \bar{\Lambda} = \Psi - i\Phi, \tag{9}$$

the following anisotropic fluid variables are obtained in terms of $\Lambda$:

$$T_{\mu\nu} = \rho U_\mu U_\nu + (\sigma - \pi)\chi_\mu\chi_\nu - \pi(g_{\mu\nu} - U_\mu U_\nu), \tag{10}$$

$$\rho = \sigma = \tfrac{1}{2}|\Lambda^{,\mu}\Lambda_{,\mu}|, \tag{11}$$

$$\pi = \tfrac{1}{2}\bar{\Lambda}_{,\mu}\Lambda^{,\mu}, \tag{12}$$

$$U_\mu = \mathrm{Re}(e^{i\alpha}\bar{\Lambda}_{,\mu})/[\mathrm{Re}(e^{i\alpha}\bar{\Lambda}_{,\nu})\mathrm{Re}(e^{i\alpha}\bar{\Lambda}^{,\nu})]^{1/2}, \tag{13}$$

$$\chi_\mu = -\mathrm{Im}(e^{i\alpha}\bar{\Lambda}_{,\mu})/[-\mathrm{Im}(e^{i\alpha}\bar{\Lambda}_{,\nu})\mathrm{Im}(e^{i\alpha}\bar{\Lambda}^{,\nu})]^{1/2}, \tag{14}$$

$$\sqrt{2}e^{i\alpha} = \left[1 + \frac{\mathrm{Re}(\Lambda_{,\mu}\Lambda^{,\mu})}{|\Lambda_{,\beta}\Lambda^{,\beta}|}\right]^{1/2} + i\left[1 - \frac{\mathrm{Re}(\Lambda_{,\mu}\Lambda^{,\mu})}{|\Lambda_{,\beta}\Lambda^{,\beta}|}\right]^{1/2}, \tag{15}$$

where $U^\mu$ is the anisotropic fluid flux velocity, $\chi_\mu$ is a spacelike unit four-vector that points in the direction of anisotropy, $\rho$ is the usual rest energy density of the fluid, $\pi$ is the pressure on a plane perpendicular to the anisotropy direction, and $\sigma$ is the pressure along the anisotropy direction.

In view of what has been stated at the beginning, exact solutions of anisotropic fluid distributions acquire some importance. Letelier and Machado[4] and Letelier[5] have recently given solutions for anisotropic fluids in the Letelier theory.[3] We present in this paper two methods to generate exact solutions of fluid distributions in the same theory. The methods are developed in Secs. II and III. Some illustrative solutions are presented in Sec. IV. The solutions obtained by the method of Sec. II represent isotropic fluids whereas those obtained by the method of Sec. III represent anisotropic fluids.

## II. GENERATION OF NEW SOLUTIONS FROM VACUUM EINSTEIN SOLUTIONS

We consider a metric of the form

$$ds^2 = \epsilon[\exp(2V)(dx^k)^2 + \exp(-2V)(\nu_{\alpha\beta}\,dx^\alpha\,dx^\beta)], \tag{16}$$

where $\epsilon = +1$ or $-1$ depending on the signature of $g_{kk}$ and $k$ is either 0, 1, 2, or 3 (see Ref. 6). Greek indices take the values 0, 1, 2, and 3.

Field equations (4) in terms of (16) may be written as

$$P_{\alpha\beta} + 2V_{,\alpha}V_{,\beta} = -[\Psi_{,\alpha}\Psi_{,\beta} + \Phi_{,\alpha}\Phi_{,\beta}], \tag{17}$$

$$V^{,\delta}_{;\delta} = 0, \tag{18}$$

where $V, \nu_{\alpha\beta}$ are functions of $x^\alpha$, $P_{\alpha\beta}$ is the Ricci tensor formed with respect to $\nu_{\alpha\beta}$, and covariant derivatives are also taken with respect to $\nu_{\alpha\beta}$.

In the axially symmetric cases Eqs. (1) and (2) reduce to the two-dimensional Laplace equations when suitable coordinates are used. We assume

$$\Phi = AV + C \quad \text{and} \quad \Psi = BV + D, \tag{19}$$

where $A$, $B$, $C$, and $D$ are constants. Therefore Eqs. (1), (2), (17), and (18) reduce to

$$P_{\alpha\beta} + 2F^2V_{,\alpha}V_{,\beta} = 0, \tag{20}$$

$$V^{,\delta}_{;\delta} = 0, \tag{21}$$

where

 0022-2488/84/082515-04$02.50  

$$F^2 = \{1 + \tfrac{1}{2}(A^2 + B^2)\}.$$

Again let

$$K = FV. \tag{22}$$

With this the field equations (20) and (21) become

$$P_{\alpha\beta} + 2K_{,\alpha}K_{,\beta} = 0,$$
$$K^{;\delta}_{;\delta} = 0. \tag{23}$$

These are the field equations $R_{\mu\nu} = 0$ of the Einstein theory for the line element

$$ds^2 = \epsilon[\exp(2K)(dx^k)^2 + \exp(-2K)(\nu_{\alpha\beta}\,dx^\alpha\,dx^\beta)]. \tag{24}$$

From Eqs. (9), (19), and (22) one obtains

$$\Lambda = (B/F)K + D + i[(A/F)K + C], \tag{25}$$

$$\overline{\Lambda} = (B/F)K + D - i[(A/F)K + C]. \tag{26}$$

The fluid variables can now be determined from Eqs. (11)–(15).

We therefore arrive at the following theorem: Corresponding to any diagonalizable solution of Einstein vacuum field equations in which the field and metric coefficients are functions of not more than three variables we can generate a solution of Einstein's equations coupled to a fluid described by two perfect-fluid components in the case that each fluid component is irrotational and each obeys the equation of state pressure = energy density.

The linear relationship between $\Psi$ and $\Phi$ assumed in Eq. (19) imposes some restrictions on the fluid components. In the following section we shall develop a more general method in which they are free from such restrictions.

## III. A MORE GENERAL METHOD FOR GENERATION OF SOLUTIONS

We consider a metric of the form

$$ds^2 = \epsilon[\exp(2V)(dx^k)^2 + \exp(-2V)(\nu_{\alpha\beta}\,dx^\alpha\,dx^\beta)], \tag{27}$$

where $V$ is a function of all the four coordinates $(x^1,x^2,x^3,x^0)$, and $k$ is either 0, 1, 2, or 3. Greek indices take the values 0, 1, 2, and 3.

Hence the field equations (4) become

$$P_{\alpha\beta} - \nu_{\alpha\beta}V^{;\delta}_{;\delta} + 2V_{,\alpha}V_{,\beta} + 4\exp[-4V]\nu_{\alpha\beta}V^2_{,k} - \exp[-4V]\nu_{\alpha\beta}V_{,kk} = -\Psi_{,\alpha}\Psi_{,\beta} - \Phi_{,\alpha}\Phi_{,\beta}, \tag{28}$$

$$\exp[4V]V^{;\delta}_{;\delta} + 6V^2_{,k} - 3V_{,kk} = -\Psi^2_{,k} - \Phi^2_{,k}, \tag{29}$$

$$2(V_{,\alpha}V_{,k} - V_{,k\alpha}) = -\Psi_{,k}\Psi_{,\alpha} - \Phi_{,k}\Phi_{,\alpha}, \tag{30}$$

where $P_{\alpha\beta}$ is the Ricci tensor formed by the metric tensor $\nu_{\alpha\beta}$ and covariant derivatives are also taken with respect to $\nu_{\alpha\beta}$.

Now when $V, \nu_{\alpha\beta}, \Psi, \Phi$ are functions of $x^\alpha$ only and not of $x^k$ the field equations to be satisfied are

$$P_{\alpha\beta} + 2V_{,\alpha}V_{,\beta} = -\Psi_{,\alpha}\Psi_{,\beta} - \Phi_{,\alpha}\Phi_{,\beta}, \tag{31}$$

$$V^{;\delta}_{;\delta} = 0, \tag{32}$$

and Eqs. (1) and (2). We assume now that

$$\Phi = MV \tag{33}$$

and

$$\Psi = HV, \tag{34}$$

where $M$ and $H$ are constants. Equations (1), (2), (31), and (32) now reduce to

$$P_{\alpha\beta} + 2J^2V_{,\alpha}V_{,\beta} = 0, \tag{35}$$

$$V^{;\delta}_{;\delta} = 0, \tag{36}$$

where

$$J^2 = \{1 + \tfrac{1}{2}(M^2 + H^2)\}.$$

Again let

$$K = JV. \tag{37}$$

With this the field equations (35) and (36) become

$$P_{\alpha\beta} + 2K_{,\alpha}K_{,\beta} = 0,$$
$$K^{;\delta}_{;\delta} = 0. \tag{38}$$

These are the field equations $R_{\mu\nu} = 0$ of the Einstein theory for the line element

$$ds^2 = \epsilon[\exp(2K)(dx^k)^2 + \exp(-2K)(\nu_{\alpha\beta}\,dx^\alpha\,dx^\beta)]. \tag{39}$$

If the metric (39) is known, then the solution of Eqs. (28)–(30) and that of Eqs. (1) and (2) can be obtained if we solve the following equations:

$$4V^2_{,k} - V_{,kk} = 0, \tag{40}$$

$$6V^2_{,k} - 3V_{,kk} = -\Psi^2_{,k} - \Phi^2_{,k}, \tag{41}$$

$$2(V_{,\alpha}V_{,k} - V_{,k\alpha}) = -\Psi_{,k}\Psi_{,\alpha} - \Phi_{,k}\Phi_{,\alpha}, \tag{42}$$

because when the gravitational equations (4) are satisfied, Eqs. (1) and (2) are a consequence of Bianchi identities.

Integration of Eq. (40) gives

$$V = -\tfrac{1}{4}\log(4x^k + E) + Y(x^\alpha), \tag{43}$$

where $E$ is an arbitrary constant and $Y(x^\alpha)$ is a function of $x^\alpha$ and can be obtained from the metric (39) in terms of Eq. (37).

In view of Eq. (43), Eq. (42) reduces to

$$2V_{,\alpha}V_{,k} = -\Psi_{,k}\Psi_{,\alpha} - \Phi_{,k}\Phi_{,\alpha}. \tag{44}$$

Now from Eqs. (33), (34), and (44) we have

$$2V_{,k} = -H\psi_{,k} - M\Phi_{,k}. \tag{45}$$

Again from Eqs. (40), (41), and (45) one gets

$$\Psi_{,k} = X\Phi_{,k}, \tag{46}$$

where

$$X = \left\{ -\frac{\frac{3}{2}HM}{\frac{3}{2}H^2 - 1} \pm \frac{1}{2}\left[\frac{9H^2M^2}{(\frac{3}{2}H^2 - 1)^2} - 4\left(\frac{\frac{3}{2}M^2 - 1}{\frac{3}{2}H^2 - 1}\right)\right]^{1/2}\right\}. \tag{47}$$

In order that $X$ may be real we must have

$$9H^2M^2 \geqslant 4(\tfrac{3}{2}M^2 - 1)(\tfrac{3}{2}H^2 - 1). \tag{48}$$

Now Eqs. (43), (45), and (46) ultimately give

$$\Phi = [1/2(HX + M)]\log(4x^k + E) + MY(x^\alpha), \tag{49}$$

$$\Psi = [X/2(HX + M)]\log(4x^k + E) + HY(x^\alpha), \tag{50}$$

where $Y(x^\alpha)$ is the same as in Eq. (43). The anisotropic fluid variables may now be obtained from Eqs. (11)–(15) in terms of Eqs. (9), (49), and (50).

The result of this section may be summarized as follows: If a solution of Einstein vacuum field equations is available in terms of any three variables, then a general solution of the Einstein's equations coupled to an anisotropic fluid described by two perfect-fluid components where each fluid component is irrotational and each obeys the equation of state pressure = energy density will be given in terms of all the four variables by Eqs. (43), (49), and (50).

## IV. SOME ILLUSTRATIVE SOLUTIONS

### A. Method of Sec. II

(i) The homogeneous and anisotropic Kasner model is given by

$$ds^2 = dt^2 - t^{2p_1}dx^2 - t^{2p_2}dy^2 - t^{2p_3}dz^2, \tag{51}$$

where the constants satisfy

$$\sum p_i = 1 \quad \text{and} \quad \sum p_i^2 = 1. \tag{52}$$

The corresponding solution of Einstein field equations in the presence of a fluid is given by

$$ds^2 = t^{2p_1[(F-1)/F]}\,dt^2 - t^{2p_1/F}\,dx^2 - t^{2p_2+2p_1[(F-1)/F]}\,dy^2 - t^{2p_3+2p_1[(F-1)/F]}\,dz^2, \tag{53}$$

with

$$\Phi = (A/F)\log t^{p_1} + C \tag{54}$$

and

$$\Psi = (B/F)\log t^{p_1} + D. \tag{55}$$

Now from Eqs. (11) and (12), $\rho$, $\sigma$, and $\pi$ are given by

$$\rho = \sigma = [\,p_1^2(A^2+B^2)/2F^2]\,t^{-2\{p_1[(F-1)/F]+1\}}, \tag{56}$$

$$\pi = [\,p_1^2(A^2+B^2)/2F^2]\,t^{-2\{p_1[(F-1)/F]+1\}}. \tag{57}$$

(ii) The Einstein vacuum solution of Mishra and Radhakrishna[7] is given by

$$ds^2 = \exp(m^2r^2/4 + mt)(dt^2 - dr^2) - \exp(mt)r^2\,d\theta^2 - \exp(-mt)dz^2, \tag{58}$$

where $m$ is a constant. The corresponding solution in the presence of a fluid is given by

$$ds^2 = \exp[m^2r^2/4 + (m/F)t\,](dt^2 - dr^2) - \exp[(m/F)t\,]r^2\,d\theta^2 - \exp[-(m/F)t\,]dz^2, \tag{59}$$

with

$$\Phi = -Amt/2F + C, \tag{60}$$

$$\Psi = -Bmt/2F + D, \tag{61}$$

$$\rho = \sigma = \frac{m^2(A^2+B^2)}{8F^2}\exp\left\{-\left(\frac{m^2r^2}{4} + \frac{m}{F}t\right)\right\}, \tag{62}$$

$$\pi = \frac{m^2(A^2+B^2)}{8F^2}\exp\left\{-\left(\frac{m^2r^2}{4} + \frac{m}{F}t\right)\right\}. \tag{63}$$

Equations (56) and (57) and (62) and (63) show that both the examples (i) and (ii) represent isotropic fluids.

### B. Method of Sec. III

(i) We consider the Kasner model given by the metric (51). For an anisotropic fluid the solution to Einstein field equations are given by

$$ds^2 = -[(4x+E)^{-1/2}t^{2p_1/J}\,dx^2 + (4x+E)^{1/2}t^{-2p_1/J} \times\{t^{2(p_2+p_1)}\,dy^2 + t^{2(p_3+p_1)}\,dz^2 - t^{2p_1}\,dt^2\}], \tag{64}$$

with

$$\Phi = \log\{(4x+E)^{1/2(HX+M)}t^{Mp_1/J}\}, \tag{65}$$

$$\Psi = \log\{(4x+E)^{X/2(HX+M)}t^{Hp_1/J}\}. \tag{66}$$

The rest energy density and the pressure along the anisotropy direction is given by

$$\rho = \sigma = \frac{1}{2}\left[\left\{\frac{4(1-X^2)t^{-2p_1/J}}{(HX+M)^2(4x+E)^{3/2}} + \frac{p_1^2(H^2-M^2)}{J^2(4x+E)^{1/2}}\,t^{2\{p_1[(1-J)/J]-1\}}\right\}^2 + \left\{\frac{2HMp_1^2}{J^2(4x+E)^{1/2}}\,t^{2\{p_1[(1-J)/J]-1\}} - \frac{8Xt^{-2p_1/J}}{(HX+M)^2(4x+E)^{3/2}}\right\}^2\right]^{1/2}, \tag{67}$$

and $\pi$, the pressure on a plane perpendicular to the anisotropy direction, is given by

$$\pi = \frac{p_1^2(H^2+M^2)}{2J^2(4x+E)^{1/2}}\,t^{2\{p_1[(1-J)/J]-1\}} - \frac{2(X^2+1)t^{-2p_1/J}}{(HX+M)^2(4x+E)^{3/2}}. \tag{68}$$

(ii) We consider the metric of Mishra and Radhakrishna[7] given by Eq. (58). The corresponding metric for an anisotropic fluid will be given by

$$ds^2 = -\left[(4z+E)^{-1/2}\exp\left(-\frac{m}{J}t\right)dz^2 + (4z+E)^{1/2}\exp\left(\frac{m}{J}t\right) \times\left\{\exp\left(\frac{m^2r^2}{4}\right)(dr^2 - dt^2) + r^2\,d\theta^2\right\}\right], \tag{69}$$

with

$$\Phi = \frac{1}{2(HX+M)}\log(4z+E) - \frac{Mm}{2J}t, \tag{70}$$

$$\Psi = \frac{X}{2(HX+M)}\log(4z+E) - \frac{Hm}{2J}t, \tag{71}$$

$$\rho = \sigma = \frac{1}{2}\left[\left\{\frac{4(1-X^2)\exp[(m/J)t\,]}{(HX+M)(4z+E)^{3/2}} + \frac{m^2(H^2-M^2)}{4J^2(4z+E)^{1/2}} \times\exp\left[-\left(\frac{m^2r^2}{4} + \frac{m}{J}t\right)\right]\right\}^2 + \left\{\frac{HMm^2}{2J^2(4z+E)^{1/2}}\exp\left[-\left(\frac{m^2r^2}{4} + \frac{m}{J}t\right)\right] - \frac{8X\exp[(m/J)t\,]}{(HX+M)^2(4z+E)^{3/2}}\right\}^2\right]^{1/2}, \tag{72}$$

and

$$\pi = \frac{m^2(H^2+M^2)}{8J^2(4z+E)^{1/2}}\exp\left[-\left(\frac{m^2r^2}{4} + \frac{m}{J}t\right)\right] - \frac{2(X^2+1)\exp[(m/J)t\,]}{(HX+M)^2(4z+E)^{3/2}}. \tag{73}$$

(iii) The static plane-symmetric Taub solution[8] is given by the metric

$$ds = (k_1x + k_2)^{-1/2}(dt^2 - dx^2) - (k_1x + k_2)(dy^2 + dz^2), \quad (74)$$

where $k_1$ and $k_2$ are constants. The corresponding solution for an anisotropic fluid is

$$\begin{aligned} ds^2 = {} & (4t + E)^{-1/2}(k_1x + k_2)^{-1/2J}dt^2 \\ & - (4t + E)^{1/2}(k_1x + k_2)^{1/2J}[(k_1x + k_2)^{-1}dx^2 \\ & + (k_1x + k_2)^{1/2}(dy^2 + dz^2)], \end{aligned} \quad (75)$$

with

$$\Phi = \log\{(4t + E)^{1/2(HX+M)}(k_1x + k_2)^{-M/4J}\}, \quad (76)$$

$$\Psi = \log\{(4t + E)^{X/2(HX+M)}(k_1x + k_2)^{-H/4J}\}, \quad (77)$$

$$\begin{aligned} \rho = \sigma = \frac{1}{2}\Bigg[ & \left\{\frac{k_1^2(M^2 - H^2)}{16J^2(k_1x + k_2)^{1+1/2J}(4t + E)^{1/2}} + \frac{4(X^2 - 1)(k_1x + k_2)^{1/2J}}{(HX + M)^2(4t + E)^{3/2}}\right\}^2 \\ & + \left\{\frac{8X(k_1x + k_2)^{1/2J}}{(HX + M)^2(4t + E)^{3/2}} - \frac{HMk_1^2}{8J^2(k_1x + k_2)^{1+1/2J}(4t + E)^{1/2}}\right\}^2\Bigg]^{1/2}, \end{aligned} \quad (78)$$

$$\pi = \frac{2(X^2 + 1)(k_1x + k_2)^{1/2J}}{(HX + M)^2(4t + E)^{3/2}} - \frac{k_1^2(H^2 + M^2)}{32J^2(k_1x + k_2)^{1+1/2J}(4t + E)^{1/2}}. \quad (79)$$

It is obvious that all the examples of this subsection represent anisotropic fluids.

## ACKNOWLEDGMENTS

The authors sincerely appreciate encouragement from Professor B. K. Barua. They express their profound gratitude to the Government of Assam, Dispur, for all facilities provided at Cotton College Gauhati-781001, India to work out this paper.

[1]M. Ruderman. Ann. Rev. Astron. Astrophys. **10**, 427 (1972).
[2]V. Canuto, *Neutron Stars: General Review* (Solvay Conference on Astrophysics and Gravitation, Brussels, 1973).
[3]P. S. Letelier, Phys. Rev. D **22**, 807 (1980).
[4]P. S. Letelier and R. Machado, J. Math. Phys. **22**, 827 (1981).
[5]P. S. Letelier, Nuovo Cimento B **69**, 145 (1982).
[6]$\epsilon = +1$ for $k = 0$ [see, e.g., Eq. (75)] and $\epsilon = -1$ for $k = 1$, 2, or 3 [see, e.g., Eq. (69)].
[7]M. Misra, and L.Radhakrishna, Proc. Natl. Inst. Sci. India A **28**, 632 (1962).
[8]A. H. Taub, Ann. Math. **53**, 472 (1951).

# Exactly solvable irreversible processes on one-dimensional lattices

N. O. Wolf,[a] J. W. Evans, and D. K. Hoffman
*Ames Laboratory and Department of Chemistry, Iowa State University, Ames, Iowa 50010*

(Received 7 September 1983; accepted for publication 16 March 1984)

We consider the kinetics of a process where the sites of an infinite 1-D lattice are filled irreversibly and, in general, cooperatively by $N$-mers (taking $N$ consecutive sites at a time). We extend the previously available exact solution for nearest neighbor cooperative effects to range $N$ cooperative effects. Connection with the continuous "cooperative car parking problem" is indicated. Both uniform and periodic lattices, and empty and certain partially filled lattice initial conditions are considered. We also treat monomer "filling in stages" for certain highly autoinhibitory cooperative effects of arbitrary range.

PACS numbers: 05.50. + q, 05.70.Ln

## I. INTRODUCTION

There are many physical and chemical processes where events occur irreversibly and, in general, cooperatively at localized sites which form a lattice.[1] For unification, we shall use the terminology of adsorption processes and "$o$"("$a$") will denote an empty (filled) site. Both irreversible monomer filling of single sites as well as irreversible $N$-mer filling, taking $N$ consecutive sites at a time, can be considered. If the filling is random, only a single nonzero rate needs to be specified. More generally, for filling with range $R$ cooperative effects, the adsorption rate depends on the state of all sites within range $R$ of those being filled. Thus for each state of this (finite) influencing environment, a corresponding rate must be given.

For a finite lattice, it is a straightforward matter to write down a finite coupled set of master equations for probabilities of total lattice configurations (states) incorporating these rates. By the usual methods employed in deriving equations for reduced distribution functions, these master equations can be used to obtain a hierarchy of rate equations for the various subconfiguration probabilities. This hierarchy (being well defined, although infinite, in the infinite lattice limit) provides the natural starting point for our present discussions, since we are primarily interested in infinite lattices. Since these hierarchial equations can be written down intuitively for finite range cooperative effects,[1] there is no need here to detail their derivation from the master equations. The structure of the hierarchy for infinite range cooperative effects is less transparent and will be discussed in detail in later work.

Here we discuss processes on infinite 1-D lattices for which exact solution of the hierarchy is possible. Previous exact analyses, predominantly in the context of reactions on 1-D polymer chains, are reviewed below. Here $\sigma$ will denote a general subconfiguration of sites, each specified either filled or empty and $f_\sigma$ the corresponding probability.

*1-D random polyatomic filling:* Random dimer filling was first considered by Flory[2] in 1939 in the context of a polymer condensation reaction. He developed recursion relations for the finite lattice case based on the observation that after the first dimer lands, one is effectively left with two smaller isolated unreacted polymers on which the same problem must be solved. A saturation coverage of $1 - e^{-2}$ was obtained for the infinite lattice limit.[2] More detailed analyses of the kinetics for finite lattices have been obtained using recursion relation or generating function techniques.[3,4] For an infinite lattice, the hierarchy for probabilities of empty $n$-tuples of sites can be solved by inspection by virtue of the shielding property of a single empty site.[1,5,6] It has also been noted for random dimer filling that (a) filling a linear lattice of length $M$ and a ring of length $M + 2$ are isomorphic,[3] and (b) if $f_i(M)$ is the probability that the $i$th site of an $M$ site lattice is empty and the end site probabilities $f_1(M) = f_M(M) \equiv p_M$, then $f_i(M) = p_i\, p_{M-i}$ which relates $f_i$(M) to end site probabilities for smaller lattices.[7] Other analyses of a statistical nature have also been given.[7-9]

The extension to random $N$-mer filling, $N$ consecutive sites at a time, has also been considered via similar techniques.[10,11] It has been shown that the probability of $N-1$ consecutive empty sites at saturation equals $\exp(-2\Sigma_{j=1}^{N-1} 1/j)$[12,5] which is a natural generalization of the above-mentioned Flory result. Note that the $N \to \infty$ limit corresponds to the continuous "car parking problem" of randomly filling a line with nonoverlapping unit intervals.[13]

*1-D cooperative filling:* Monomer filling an infinite lattice with nearest neighbor (NN) cooperative effects was first considered in 1962–3 (see Ref. 14). Here the hierarchy of rate equations can be solved exactly for the probabilities of $n$-tuples of unreacted sites (in terms of incomplete $\Gamma$-functions[15]) by virtue of a shielding property of adjacent pairs of empty sites.[1,5,16] More recent statistical treatments[17] emphasize this feature, the extension of which is the basis for our analysis. Another approach yielding exact solutions exploits a related "principle of independence" of empty neighbors,[18] namely: the probability of any subconfiguration $\sigma oo\sigma'$, including an adjacent pair of empty sites, can be factored as $f_{\sigma oo\sigma'} = f_{\sigma\cdot} f_{\cdot o\sigma'}$. Here "$\cdot$" indicates that we have split the original lattice creating two "artificial" sublattices. Rates for landing on the new end sites *must* be chosen as if the "missing" neighbor was empty and those for other sites correspond to the original infinite lattice values. Solution also exploits the simple time dependence of $f_{\cdot o}$ (see Ref. 18).

[a] Present address: Conoco Inc., P. O. Box 1267, Ponca City, OK 74603.

 0022-2488/84/082519-08$02.50  

Schwartz[19] has implemented an approximate "triplet closure rule" to truncate the hierarchy. This amounts to a 2nd-order spatial Markovian assumption on the distribution, and is an approximation since both sites in an adjacent pair must be specified empty to shield exactly.

Several workers have commented on the difficulty of obtaining, e.g., probabilities for filled, cf. empty, $n$-tuples.[11,20] A particularly useful approximate analysis has been given[21] but an *exact* treatment is also possible.[22,5] A related problem of determining the large separation behavior of spatial correlations has also been solved.[23] The extension of this problem to lattices with several types of periodically or stochastically distributed sites has also been given using the "principle of independence" of empty neighbors.[24]

The only exact 1-D analysis, to date, for more general cooperative effects are for monomer filling with range $R$ cooperative effects incorporating range $R-1$ blocking[5,12,25,26] (range $b$ blocking here means that filling cannot occur at a site if there are one or more already filled sites within $b$ lattice vectors). We note that this process is isomorphic to the adsorption of $R$-mers onto $R$-consecutive sites with corresponding NN cooperative effects (see Fig. 1).

Exact closed form solutions have also been given for some rather special monomer filling processes on higher-dimensional lattices. These include the case of NN cooperative effects where one adsorption rate, $\tau$, operates if the site being filled has at least one empty NN and a second, $\alpha\tau$, operates when all NN are already filled ($\alpha=1$ corresponds to trivial random filling).[1,27] It also seems possible to exactly solve the hierarchy equations for various competing irreversible processes where the cooperativity is such that the individual processes are solvable.[27,28] We also note the formal "density" expansions of solutions (albeit with potential convergence problems) are always available.[29] A variety of processes amenable to exact solution on Bethe lattices are discussed in the following paper.[30]

Here we discuss the most general irreversible cooperative processes, on infinite 1-D lattices, amenable to exact solution via hierarchy truncation. These cases are characterized by the probabilities for $n$-tuples of empty sites satisfying a (minimal) *closed* subhierarchy. The basis of our solution procedure is an empty site shielding property discussed, in a general context, in Ref. 1. For an $N$-mer landing on an 1-D lattice with range $R$ cooperative effects, this tells us that a block of sites specified *empty* of width $N-1+2R$ shields sites on one side from the influence of those on the other (*provided* this is compatible with the initial conditions). This condition is expressed most naturally in terms of the conditional probabilities $q_{\sigma\bar{\sigma}'}(\equiv f_{\sigma+\sigma'}/f_{\sigma'})$ of $\boldsymbol{\sigma}$ *given* $\boldsymbol{\sigma}'$ (for typographic convenience, empty conditioning sites $\bar{o}$ are sometimes denoted $\phi$). The proof via self-consistency with respect to the $q$-hierarchy equations is indicated in the Appendix.

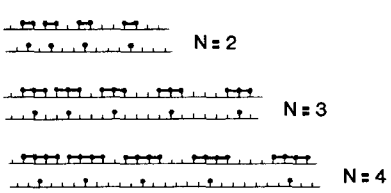



FIG. 1. Equivalence of $N$-mer filling and monomer filling with range $N-1$ blocking. Corresponding (nonblocking) cooperative effects are required.

A special case of shielding, of particular relevance to hierarchy truncation in translationally invariant systems, is that

$$q_m \equiv q_{\underbrace{o\phi\phi\cdots\phi}_{m}} \equiv q_{\underbrace{\phi\phi\cdots\phi o}_{m}} \equiv f_{\mathbf{o}_{m+1}}/f_{\mathbf{o}_m}\,, \tag{1.1}$$

for all $m \geqslant N-1+2R$, are equal (to $q$, say) in the case of $N$-mer adsorption with range $R$ cooperative effects. (Here $\mathbf{o}_m$ denotes an $m$-tuple of empty sites.) Furthermore,

$$\frac{d}{dt}\ln q = -\tau\,, \tag{1.2}$$

where $\tau$ is the adsorption rate with no filled sites in the cooperative range. For $N$-mer filling of a periodic lattice, the analysis of the Appendix can be readily extended to show that again $N-1+2R$ sites shield for range $R$ cooperative effects. For a lattice of periodicity $P$, $q^i_{\underbrace{o\phi\phi\cdots\phi}_{m}}$ (where $i=1,2,\ldots,P$ indicates the type of the $o$-site) are equal (to $q^i$, say) for all $m \geqslant N-1+2R$ and a closed coupled set of equations can be obtained for these after applying shielding. Here the rates $\tau^i\ldots$ are labeled by the adsorption site type and probabilities of subconfigurations are only invariant under translations of (multiples of) $P$ sites.

In Sec. II, we first demonstrate in detail exact solution of the hierarchy for dimer filling with range 2 cooperative effects. We then sketch the analogous solution for $N$-mer filling with range $N$ cooperative effects and indicate why exact solution is *not* possible for general range $N+1$ cooperative effects. Connection with the continuous "cooperative car parking problem" is also discussed. The extension to periodic lattices and processes for which the initial conditions satisfy a modified shielding condition is described in Sec. III. Monomer "filling in stages" for highly autoinhibitory rates is considered in Sec. IV and extensions of this work are discussed in Sec. V.

Before proceeding with these analyses, we make two further observations. First, we note that $N$-mer filling with range $R$ cooperative effects incorporating range $r$ blocking is equivalent to monomer filling with corresponding range $N-1+R$ cooperative effects incorporating range $N-1+r$ blocking (see Fig. 1). Second, we remark that the "principle of independence" of empty neighbors can be extended to monomer filling with range $R$ cooperative effects (or more generally to $N$-mer filling). For the former, we have that $f_{\sigma o_{2R}\sigma'} \equiv f_{\sigma o_R \cdot} f_{\cdot o_R \sigma'}$. Again "$\cdot$" indicates that we have split the lattice and rates for landing within $R$ sites of these new ends *must* be chosen as if the lattice continued past these ends with empty sites. Proof parallels that for $R=1$ (Ref. 18). This approach also provides exact solutions for certain choices of cooperative effects after noting, for example, that $f_{\sigma o_{R+k}\cdot} = f_{\sigma o_R \cdot} f_{\cdot o_k \cdot}$ and $f_{\cdot o_k \cdot}$ has simple time dependence. The approach we choose, however, is more direct avoiding the introduction of "artificial" sublattices.

## II. EXACTLY SOLVABLE 1-D IRREVERSIBLE PROCESSES

We consider here translationally invariant, irreversible $N$-mer filling of an infinite, uniform 1-D lattice (taking $N$

consecutive sites at a time) with range $R$ cooperative effects. For monomer filling ($N = 1$), random filling ($R = 0$) is trivial and, for $R = 1$, exact solution of the hierarchy equations for probabilities of empty $n$-tuples is well documented.[14–18] For $R = 2$, an exact solution is not possible[31] even though empty 4-tuples shield exactly [see (1.1) and (1.2)]. Corresponding exact solutions for dimer filling ($N = 2$) with $R = 0$ (Flory) and $R = 1$ have been given previously.[5,26]

Thus here we begin by giving the first treatment of dimer filling with range two cooperative effects ($R = 2$). The input to the hierarchial rate equations is a set of rates $\tau_{\sigma_1\sigma_2\cdot\sigma_3\sigma_4}$, where $\sigma_i = o,a$ indicate the state of the influencing pairs of sites to the left and right of the two empty sites being filled by the dimer. For convenience (only), we consider here rates with reflection symmetry. This in turn implies the $f$'s are reflection invariant. The rate equation for the probability of a single empty site is thus given by[1]

$$-\frac{d}{dt}f_o = 2\sum_{\sigma_i=o,a}\tau_{\sigma_1\sigma_2\cdot\sigma_3\sigma_4}f_{\sigma_1\sigma_2 oo\sigma_3\sigma_4}\,. \tag{2.1}$$

The factor of 2 arises since either of the constituent atoms of the dimer could land on "$o$."

There is a rather subtle implicit restriction on the sum of the rhs of (2.1) which is essential to the exact solution of this problem. This arises because for dimer filling, subconfigurations $\cdots oao\cdots$ cannot occur, so, e.g., $f_{oaoooo}\equiv 0$ and the term $2\tau_{oa\cdot oo}f_{oaoooo}$ does not contribute to (2.1). Taking account of all such cases, as well as reflection symmetry, (2.1) becomes

$$\begin{aligned}-\frac{1}{2}\frac{d}{dt}f_o = {}&\tau_{oo\cdot oo}f_{oooooo} + 2\tau_{ao\cdot oo}f_{aooooo}\\ &+ 2\tau_{aa\cdot oo}f_{aaoooo} + \tau_{ao\cdot oa}f_{aooooa}\\ &+ 2\tau_{aa\cdot oa}f_{aaoooa} + \tau_{aa\cdot aa}f_{aaooaa}\,,\end{aligned} \tag{2.2}$$

where the six independent rates for this process now appear explicitly. One further essential observation is that, for dimer filling, we must have that

$$f_{aaoooo} = f_{aoooo},\quad f_{aaoooa} = f_{aoooa},\quad f_{aaooaa} = f_{aooa}\,, \tag{2.3}$$

since, given a subconfiguration $oa(ao)$, we know that the site to the right (left) must also be filled. Then using conservation of probability, e.g.,

$$f_{aooo} = f_{ooo} - f_{oooo},\quad f_{aooa} = f_{oo} - 2f_{ooo} + f_{oooo}\,, \tag{2.4}$$

we can express the rhs of (2.2) in terms of probabilities, $f_{\mathbf{o}_n} = f_n$, for $n$-tuples $\mathbf{o}_n$ of empty sites. Thus, after some rearrangement, (2.2) becomes

$$\begin{aligned}-\frac{1}{2}\frac{d}{dt}f_1 = {}&\tau_{aa\cdot aa}f_2 + 2(\tau_{aa\cdot oa} - \tau_{aa\cdot aa})f_3\\ &+ (\tau_{aa\cdot aa} - 4\tau_{aa\cdot oa} + \tau_{ao\cdot oa} + 2\tau_{aa\cdot oo})f_4\\ &+ 2(\tau_{aa\cdot oa} - \tau_{ao\cdot oa} - \tau_{aa\cdot oo} + \tau_{ao\cdot oo})f_5\\ &+ (\tau_{ao\cdot oa} - 2\tau_{ao\cdot oo} + \tau_{oo\cdot oo})f_6\,.\end{aligned} \tag{2.5}$$

One must next consider the rate equations for $f_2, f_3, \ldots$ which have the form

$$\begin{aligned}-\frac{d}{dt}f_2 = -\frac{d}{dt}f_{oo} = {}&\sum_{\sigma_i=o,a}\tau_{\sigma_1\sigma_2\cdot\sigma_3\sigma_4}f_{\sigma_1\sigma_2 oo\sigma_3\sigma_4}\\ &+ 2\sum_{\sigma_i=o,a}\tau_{\sigma_1\sigma_2\cdot oo\sigma_3}f_{\sigma_1\sigma_2 ooo\sigma_3}\,,\end{aligned} \tag{2.6a}$$

$$\begin{aligned}-\frac{d}{dt}f_3 = {}&2\sum_{\sigma_i=o,a}\tau_{\sigma_1\sigma_2\cdot oo\sigma_3}f_{\sigma_1\sigma_2 ooo\sigma_3}\\ &+ 2\sum_{\sigma_i=o,a}\tau_{\sigma_1\sigma_2\cdot oo}f_{\sigma_1\sigma_2 oooo}\,,\end{aligned} \tag{2.6b}$$

$$\begin{aligned}-\frac{d}{dt}f_4 = {}&\sum_{\sigma_i=o,a}\tau_{\sigma_1 o\cdot o\sigma_2}f_{\sigma_1 oooo\sigma_2}\\ &+ 2\sum_{\sigma_i=o,a}\tau_{\sigma_1\sigma_2\cdot oo}f_{\sigma_1\sigma_2 oooo}\\ &+ 2\sum_{\sigma_i=o,a}\tau_{\sigma_1\sigma_2\cdot oo}f_{\sigma_1\sigma_2 ooooo}\,,\end{aligned} \tag{2.6c}$$

where the first (second) term of (2.6a) corresponds to the dimer landing squarely on (partly overlapping) the pair $oo$, etc. For $n \geqslant 5$, the rate equations have the generic form

$$\begin{aligned}-\frac{d}{dt}f_n = {}&(n-5)\tau_{oo\cdot oo}f_n + 2\sum_{\sigma_1=o,a}\tau_{\sigma_1 o\cdot oo}f_{\sigma_1\mathbf{o}_n}\\ &+ 2\sum_{\sigma_1=o,a}\tau_{\sigma_1\sigma_2\cdot oo}f_{\sigma_1\sigma_2\mathbf{o}_n}\\ &+ 2\sum_{\sigma_1=o,a}\tau_{\sigma_1\sigma_2\cdot oo}f_{\sigma_1\sigma_2\mathbf{o}_{n+1}}\,,\end{aligned} \tag{2.7}$$

where the 1st, 2nd, 3rd, and 4th terms correspond to dimers landing in the interior of, one site from the ends, at the ends, and overlapping the ends of the $n$-tuple $\mathbf{o}_n$ (respectively).

Terms appearing on the rhs of these equations can be expressed in terms of $f_{\mathbf{o}_n} = f_n$ using

$$f_{a\mathbf{o}_m} = f_{\mathbf{o}_m a} = f_{aa\mathbf{o}_m} = f_{\mathbf{o}_m aa} = f_m - f_{m+1}\,, \tag{2.8}$$

$$f_{a\mathbf{o}_m a} = f_{aa\mathbf{o}_m a} = f_{a\mathbf{o}_m aa} = f_{aa\mathbf{o}_m aa} = f_m - 2f_{m+1} + f_{m+2}\,.$$

For example, for $n \geqslant 5$, one has

$$\begin{aligned}-\frac{d}{dt}f_n = {}&(n-5)\tau_{oo\cdot oo}f_n + 2(\tau_{ao\cdot oo} + \tau_{aa\cdot oo})f_n\\ &+ 2\tau_{oo\cdot oo}f_{n+1} + 2(\tau_{oo\cdot oo} - \tau_{aa\cdot oo})f_{n+2}\\ &+ 2(\tau_{oo\cdot oo} - \tau_{ao\cdot oo})f_{n+3}\,.\end{aligned} \tag{2.9}$$

Thus we have shown, for this process, that one can obtain a closed subhierarchy for the $\{f_n\}$. This feature is essential in obtaining exact solutions for the $f_n$ which follow immediately from the observation that the identity

$$f_n(t)/f_5(t) = [\exp(-\tau_{oo\cdot oo}t)q_5(0)]^{n-5} \tag{2.10}$$

is compatible with (2.9) [where here, by assumption, $q_5(0) = f_6(0)/f_5(0) = f_{n+1}(0)/f_n(0)$ for $n \geqslant 5$]. Clearly use of (2.10) leads to exact hierarchy truncation yielding a closed coupled set of equations for $f_1, f_2, \ldots, f_5$ which can be straightforwardly integrated. We can then, using (2.10), also calculate any of the quantities in (2.8).

Equation (2.10) is, of course, an explicit example of the shielding property (1.1) and (1.2) discussed in the Introduc-

tion. To make contact with the general proof of shielding described in the Appendix, note that (2.9) can be rewritten to give the following rate equations for the conditional probabilities $q_m = f_{m+1}/f_m$:

$$-\frac{d}{dt}\ln q_m = -\frac{d}{dt}\ln f_{m+1} + \frac{d}{dt}\ln f_m$$
$$= \tau_{oo\cdot oo} + 2\tau_{oo\cdot oo}(q_{m+1} - q_m)$$
$$+ 2(\tau_{oo\cdot oo} - \tau_{aa\cdot oo})(q_{m+2}q_{m+1} - q_{m+1}q_m)$$
$$+ 2(\tau_{oo\cdot oo} - \tau_{ao\cdot oo})(q_{m+3}q_{m+2}q_{m+1}$$
$$- q_{m+2}q_{m+1}q_m), \quad \text{for } m \geqslant 5, \qquad (2.11)$$

which, by inspection, has the desired solution $q_m = q_5(0) \times \exp(-\tau_{oo\cdot oo}t)$ for $m \geqslant 5$.

We now turn our attention to $N$-mer filling with range $N$ cooperative effects for general $N$. All the basic ingredients for the solution of this problem appear in the above special (dimer filling) case so we only sketch the treatment. In Fig. 2 we consider the filling of the empty $N$-tuple $\mathbf{o}_N$, enumerate all possible configurations of the influencing environment, and assign an appropriate notation for the corresponding rates.

Using such identities as

$$f_{\mathbf{a}_k\mathbf{o}_m} = f_{a\mathbf{o}_m} = f_m - f_{m+1}, \quad \text{for } 1 \leqslant k \leqslant N, \qquad (2.12)$$

$$f_{\mathbf{a}_k\mathbf{o}_m\mathbf{a}_j} = f_{a\mathbf{o}_m a} = f_m - 2f_{m+1} + f_{m+2}, \quad \text{for } 1 \leqslant k, j \leqslant N$$

(where $\mathbf{a}_k$ denotes a filled $k$-tuple), one can readily express $(d/dt)f_n$ in terms of other $f_m$, for any $n$, thus again obtaining a closed subhierarchy for the $\{f_n\}$. These equations are necessarily complicated because of the general nature of the cooperative effects. Here we just mention that they have the generic form

$$-\frac{d}{dt}f_n = (n - 3N + 1)\tau f_n + \sum_{k=0}^{2N-1} \gamma_k f_{n+k}, \quad \text{for } n \geqslant 3N - 1, \qquad (2.13)$$

where the coefficients $\gamma_k$ in the "boundary terms" depend on the rates $\tau(i-)$, $\tau(j+)$, $\tau$, but *not* on "$n$." Exact hierarchy truncation and solution for the $f_n$'s follows from the observation that the identity

$$f_n(t)/f_{3N-1}(t) = [e^{-\tau t}q_{3N-1}(0)]^{(n-3N+1)}, \quad \text{for } n \geqslant 3N - 1, \qquad (2.14)$$

is compatible with (2.13) being, of course, an example of the shielding property (1.12) [and where, by assumption, $q_{3N-1}(0) = f_{3N}(0)/f_{3N-1}(0) = f_{n+1}(0)/f_n(0)$, for $n \geqslant 3N - 1$]. Using (2.14), a closed coupled set of equations are obtained for $f_1, f_2, \ldots, f_{3N-1}$, which can be straightforwardly integrated.

The key to the exact solution of the above problem was that the minimal closed hierarchy involved only $f_m$ which then allowed use of the shielding property (1.1) and (1.2) to achieve exact hierarchy truncation. For $N$-mer filling, with cooperative range greater than $N$, this is not the case. For example, with range $N+1$, $(d/dt)f_o$ has a contribution proportional to $f_{o\mathbf{a}_N\mathbf{o}_N\mathbf{a}_N o}$ which *cannot* be written in terms of $f_m$. The simplest problem *not* amenable to exact solution is monomer filling with range two cooperative effects.[31] Finally, we note that if the $N$-mer cooperative effects incorporate a range $b$ blocking, then we can exactly truncate the hierarchy for range $N + 2b$ cooperative effects. The treatment of this case is analogous to that given above.

Let us now consider the $N \to \infty$ continuous limit of the above problem. Various workers have commented that, for random $N$-mer filling, this limit yields the "car parking problem."[5,25] Of course this is also true for any finite, $N$-independent range, cooperative effects. Here, however, we consider the case where

$$\tau(i-) \to \tau_-(i/N), \quad \tau(j+) \to \tau_+(j/N),$$
$$\tau(i-|j+) \to \tau(i/N\,|\,j/N), \quad \text{as } N \to \infty, \qquad (2.15)$$

for some $\tau_\mp(\ )$ defined on [0,1] and $\tau(\ |\ )$ defined on $[0,1]\times[0,1]$ [it is natural to choose $\tau(x|1) = \tau_-(x)$, $\tau(1|y) = \tau_+(y)$ and $\tau_-(1) = \tau_+(1) = \tau$]. The $N \to \infty$ limit corre-

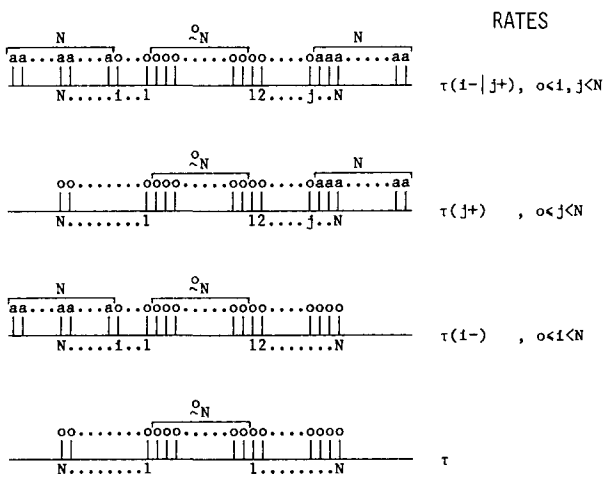



FIG. 2. $N$-mer filling $\mathbf{o}_N$ with range $N$ cooperative effects; enumeration of influencing environments and corresponding rates.

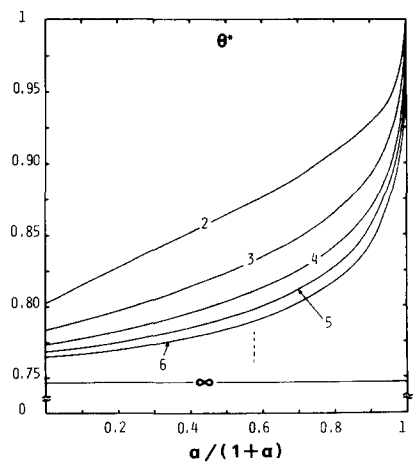



FIG. 3. Saturation coverages $\theta^*(N;\alpha)$ for $N$-mer filling of an initially empty, infinite lattice with NN cooperative effects where each filled NN changes the rate by a factor of $\alpha$. $N = 2,3,\ldots$ are shown.

TABLE I. $N$-mer filling an initially empty, infinite lattice in stages with 0, 1, then 2 filled NN; stages having final $\theta$-values of $\theta_1^* = \theta^*(N;0) = N(N+1)^{-1}\theta^*(N+1;1)$, $\theta_2^* = \theta_1^* + N\exp(-2\Sigma_{j=1}^{N-1} 1/j)$, $\theta_3^* = \theta^*(N;0+)$, respectively [see Fig. 3 for $\theta^*(N;\alpha)$].

| $n$ | 1 | 2 | 3 | 4 | 5 | 6 | ⋯ | ∞ |
|---|---|---|---|---|---|---|---|---|
| $\theta_1^*$ | 0.432 33... | 0.549 10... | 0.602 92... | 0.633 82... | 0.653 85... | 0.667 90... | ⋯ | 0.7476... |
| $\theta_2^*$ | 0.567 66... | 0.648 68... | 0.679 60... | 0.695 83... | 0.705 82... | 0.712 58... | ⋯ | 0.7476... |
| $\theta_3^*$ | 1.000 0 | 0.802 22... | 0.782 92... | 0.773 66... | 0.768 24... | 0.764 68... | ⋯ | 0.7476... |

sponds to the "cooperative car parking" where nonoverlapping unit intervals fill the line with range one cooperative effects determined by $\tau$, $\tau_-(\ )$, $\tau_+(\ )$, and $\tau(\ |\ )$. A more direct analysis of this problem, however, would write down appropriate continuous integral equations and use a continuous analog of the shielding condition (2.2) and (2.3) (see Ref. 5). This will be done in later work.

For the most general exactly solvable $N$-mer filling problem described above, there is great flexibility in the choice of rates. Thus, in presenting some numerical results here, we restrict our attention to the case of NN cooperative effects where each filled NN changes the rate (from the zero NN case) by a factor of $\alpha$. All these results are obtained by simply integrating the closed coupled sets of equations from the infinite $f_n$-hierarchy after exact truncation and choosing initial conditions corresponding to a completely empty lattice. In Fig. 3, we have shown the saturation coverage as a function of $\alpha/1+\alpha$ for various $N$. Note the uniform constant lower bound of 0.7476... (see Ref. 13) from the $N \to \infty$ case. The $\alpha \to \infty$ limit with saturation coverage of unity can be thought of as a single contiguous island of $N$-mers growing about a single nucleating $N$-mer. In Table I, we consider the low $\alpha$ regime in detail where filling occurs in three stages with 0,1, and 2 filled NN, consecutively. The $\alpha = 1$ case corresponds to random $N$-mer filling. In Table II, we catalog saturation values (some of which have been determined previously) for various isomorphic random filling processes.

## III. EXTENSIONS TO PERIODIC LATTICES AND TO INITIAL CONDITIONS WITH CORRELATIONS

Irreversible processes on certain 1-D periodic lattices have been considered previously for NN cooperative effects via the "principle of independence" of empty neighbors.[24] Here, however, we indicate a more direct approach using our form of the shielding condition. Suppose that the lattice has periodicity $P$ and the different types of sites are labeled by $i = 1,2,\ldots, P$. Rates $\tau^i\cdots$ must now be specified for filling each type of site $i$ (for all configurations of the influencing environment). Probabilities for various subconfigurations now have only restricted translational invariance and so also carry a "position" label $i$.

Consider the case of $N$-mer filling with range $N$ cooperative effects. Since, as mentioned in the Introduction, any $(3N-1)$-tuple of empty sites shields, if $q_m^i \equiv q^i_{o\underbrace{\phi\phi\cdots\phi}_{m}}$ ($i$ denotes the type of the "$o$" site), then

$$q_m^i = q_{3N-1}^i \equiv q^i, \quad \text{say,} \quad \text{for } m \geqslant 3N-1, \tag{3.1}$$

and a closed set of equations can be obtained for these (see example below). Furthermore it is easy to show that an infinite closed subhierarchy of equations can again be obtained for the probabilities of $n$-tuples of empty sites. Exact truncation follows immediately upon application of the shielding condition.

As a simple example, consider monomer filling with NN cooperative effects (assumed reflection invariant, for convenience) on a lattice with periodicity 2. Then $\tau^i_{o\cdot o}$, $\tau^i_{a\cdot o}$, and $\tau^i_{a\cdot a}$ denote the rates for filling type $i(=1,2)$ sites with 0, 1, and 2 filled NN, respectively. The rate equation for the probability $f_o^i$ that a type $i$ site is empty, becomes

$$-\frac{d}{dt}f_o^i = \tau^i_{o\cdot o} f^{\,i}_{ooo} + 2\tau^i_{a\cdot o} f^{\,i}_{aoo} + \tau^i_{a\cdot a} f^{\,i}_{aoa}, \quad i = 1,2, \tag{3.2}$$

where the superscript again denotes the type of site. Probabilities for empty pairs are independent of position, i.e., $f^1_{oo} = f^2_{oo} \equiv f_{oo}$ for reflection invariant rates and satisfy the equation

$$-\frac{d}{dt}f_{oo} = \sum_{i=1}^{2} (\tau^i_{o\cdot o} f^{\,i}_{ooo} + \tau^i_{a\cdot o} f^{\,i}_{aoo}). \tag{3.3}$$

Using shielding, the following closed pair of equations are obtained for $q^i = f^i_{ooo}/f_{oo}$, $i = 1,2$:

TABLE II. Saturation coverages $\theta^*(N|r')$ for random $N$-mer filling with range $r'$ blocking of an initially empty, infinite lattice. From isomorphism arguments $\theta^*(N|r') = N(N+r')^{-1}\{\theta^*(N+r'|0) = \theta^*(N+r';1)\}$. Values for $r' = 0$ are given in Ref. 25.

| $r$ | 0 | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|---|
| monomer, range $r$ blocking | 1.0000 | 0.432 33... | 0.274 55... | 0.200 97... | 0.158 45... | 0.130 77... | 0.111 31... |
| dimer, range $r-1$ blocking | | 0.864 66... | 0.549 10... | 0.401 94... | 0.316 91... | 0.261 54... | 0.222 63... |
| trimer, range $r-2$ blocking | | | 0.823 65... | 0.602 91... | 0.475 36... | 0.392 31... | 0.333 95... |
| 4-mer, range $r-3$ blocking | | | | 0.803 89... | 0.633 82... | 0.523 08... | 0.445 26... |
| 5-mer, range $r-4$ blocking | | | | | 0.792 27... | 0.653 85... | 0.556 58... |
| 6-mer, range $r-5$ blocking | | | | | | 0.784 63... | 0.667 90... |
| 7-mer, range $r-6$ blocking | | | | | | | 0.779 21... |

$$-\frac{d}{dt}\ln q^1 = -\frac{d}{dt}\ln f^1_{ooo} + \frac{d}{dt}\ln f_{oo}$$
$$= \tau^1_{o\cdot o} q^2_{o\phi\phi\phi} + \tau^1_{a\cdot o} q^2_{a\phi\phi\phi} + \tau^2_{o\cdot o}$$
$$- \tau^2_{o\cdot o} q^1_{o\phi\phi} - \tau^2_{a\cdot o} q^1_{a\phi\phi}$$
$$+ \tau^1_{o\cdot a}(q^2_{\phi\phi\phi a} - q^2_{\phi\phi a}) + \tau^1_{o\cdot o}(q^2_{\phi\phi\phi o} - q^2_{\phi\phi o})$$
$$= \tau^2_{o\cdot o} + (\tau^1_{o.o} - \tau^1_{a\cdot o})q^2 - (\tau^2_{o\cdot o} - \tau^2_{a\cdot o})q^1$$
$$+ \tau^1_{a\cdot o} - \tau^2_{a\cdot o}\,, \tag{3.4}$$

and the $q^2$ equation follows from interchanging 1 and 2. After conversion of the rhs's of (3.2) and (3.3) to probabilities for empty $n$-tuples and use of the identities $f^1_{ooo} = q^2 f_{oo}$, $f^2_{ooo} = q^1 f_{oo}$, these and the $q^i$ equations form a closed set.

To this point, we have assumed that the initial conditons are compatible with the "standard" shielding condition which, for $N$-mer filling with range $R$ cooperative effects, states that empty $(N-1+2R)$-tuples shield. This is certainly true for the important case where the lattice is initially completely empty and, in fact, is true for any $m$th-order spatially Markovian distribution where $m \leqslant N-1+2R$. If, however, the (more esoteric) initial conditions are only compatible with the shielding of empty $m$-tuples only for $m \geqslant m^*$ and $m^* > N-1+2R$, then it is straightforward to show that empty $m^*$-tuples shield for all later times. Furthermore, this modified shielding condition implies

$$\frac{d}{dt}\ln q_{m^*} = -\tau, \tag{3.5}$$

for translationally invariant processes, which can be used to exactly truncate the $f_m$ subhierarchy for $N$-mer filling with range $N$ cooperative effects. Of course, a larger number of equations are now retained after truncation. For example, consider monomer filling with NN cooperative effects and rates $\tau_{o\cdot o}, \tau_{a\cdot o} \equiv \tau_{o\cdot a}$ (by assumption), $\tau_{a\cdot a}$ for 0,1,2 filled NN, respectively. Then we must solve the finite coupled set of equations

$$-\frac{d}{dt}f_1 = -\frac{d}{dt}f_o = \tau_{o\cdot o} f_{ooo} + 2\tau_{a\cdot o} f_{aoo} + \tau_{a\cdot a} f_{aoa}$$
$$= \tau_{a\cdot a} f_1 + 2(\tau_{a\cdot o} - \tau_{o\cdot o})f_2$$
$$+ (\tau_{o\cdot o} - 2\tau_{a\cdot o} + \tau_{a\cdot a})f_3, \tag{3.6a}$$
$$-\frac{d}{dt}f_n = (n-2)f_n + 2[\tau_{o\cdot o} f_{n+1} + \tau_{a\cdot o} f_{ao_n}]$$
$$= (n-2)f_n + 2[\tau_{a\cdot o} f_n + (\tau_{o\cdot o} - \tau_{a\cdot o})f_{n+1}],$$
$$\text{for } 2 \leqslant n \leqslant m^*, \tag{3.6b}$$

and

$$f_{m^*+1}(t)/f_{m^*}(t) = q_{m^*}(0)e^{-\tau_{o\cdot o}t}. \tag{3.6c}$$

[Equation (3.6b) is, in fact, valid for all $n \geqslant 2$ and can be used to trivially prove the modified shielding condition for this case.]

## IV. FILLING IN STAGES (HIGHLY AUTOINHIBITORY RATES)

In this section, we treat only the case of monomer filling. From previous discussion, we know that exact solution is possible here for NN, but *not* longer-range general cooperative effects. However if we restrict our attention to certain limiting regimes for longer range cooperative effects, again exact results can be obtained. Here we consider the special case of range $R$ (arbitrary) cooperative effects where the rates are greatly inhibited by filled sites (within the cooperative range) and, the closer the (already) filled site to the adsorption site, the greater the inhibition. Specifically, we consider the extreme case where filling occurs in stages, consecutively, with no filled sites in the cooperative range $R$, with just one, then two, filled $R$-th NN, with one filled $(R-1)$-th NN (and the $R$-th NN on the other side filled), with two filled $(R-1)$-th NN, with one filled $(R-2)$-th, etc.

For $R = 1$, this three stage filling has been described in (Refs. 1, 16, and 32) and, for $R = 2$, this five stage filling has been described in Ref. 31. For general $R$, the following detailed characterization of the stages, labeled (1), (2),...,$(2R+1)$, is possible ($\theta = 1 - f_o$ denotes the coverage).

(1) $0 \leqslant \theta \leqslant \theta^*(1)$, $\underset{\downarrow}{o_R o o_R}$, i.e., filling with $R$ th NN blocking (here $\downarrow$ indicates the adsorption site).

$$\frac{d}{d\theta}f_m = -m, \quad m = 1,2,\ldots, R+1,$$
$$f_m \to 0 \quad \text{as } \theta \to \theta^*(1)-, \quad m \geqslant 2R+1,$$
$$\theta^*(1) = \int_0^1 dq \exp\left[-2\sum_{l=1}^{R} l^{-1}(1-q^l)\right]$$

(see Refs. 5 and 25).

(2) $\theta^*(1) \leqslant \theta \leqslant \theta^*(2)$, $a\underset{\downarrow}{o_{R-1}} o o_R(a)$,

$$\frac{d}{d\theta}f_m = -m, \quad m = 1,2,\ldots, R,$$
$$\frac{d}{d\theta}f_{R+m} = -R+m-1, \quad m = 1,2,\ldots, R,$$
$$f_{2R} \to 0 \quad \text{as } \theta \to \theta^*(2)-,$$
$$\theta^*(2) = \theta^*(1) + f_{2R}|_{\theta^*(1)}.$$

(3) $\theta^*(2) \leqslant \theta \leqslant \theta^*(3)$, $a\underset{\downarrow}{o_{R-1}} o o_{R-1} a$,

$$\frac{d}{d\theta}f_m = -m, \quad m = 1,2,\ldots, R,$$
$$\frac{d}{d\theta}f_{R+m} = -R+m, \quad m = 1,2,\ldots, R,$$
$$f_{2R-1} \to 0 \quad \text{as } \theta \to \theta^*(3)-,$$
$$\theta^*(3) = \theta^*(2) + f_{2R-1}|_{\theta^*(2)}$$
$$= \theta^*(1) + f_{2R-1}|_{\theta^*(1)} - f_{2R}|_{\theta^*(1)}.$$

(4) $\theta^*(3) \leqslant \theta \leqslant \theta^*(4)$, $a\underset{\downarrow}{o_{R-2}} o o_{R-1}(a)$,

$$\frac{d}{d\theta}f_m = -m, \quad m = 1,2,\ldots, R-1,$$
$$\frac{d}{d\theta}f_{R+m} = -R+m+1, \quad m = 0,1,\ldots, R-2,$$
$$\vdots$$
$$(2R+1)\ \theta^*(2R) = 1 - \theta^*(2R-1) \leqslant \theta$$
$$\leqslant \theta^*(2R+1) = 1, \quad \underset{\downarrow}{aoa}\,.$$

TABLE III. Monomer filling an initially empty, infinite lattice in stages (with highly autoinhibitory range $R$ cooperative effects); $\theta^*(m)$ is the coverage at the end of the $m$th stage.

| | $\theta^*(2R)$ | $\theta^*(2R-1)$ | $\theta^*(2R-2)$ | $\theta^*(2R-3)$ | $\theta^*(2R-4)$ | $\theta^*(2R-5)$ | $\theta^*(2R-6)$ | $\theta^*(2R-7)$ | $\theta^*(2R-8)$ | $\theta^*(2R-9)$ |
|---|---|---|---|---|---|---|---|---|---|---|
| $R=1$ | 0.567 67 | 0.432 33 | | | | | | | | |
| $R=2$ | 0.598 89 | 0.401 11 | 0.324 34 | 0.274 55 | | | | | | |
| $R=3$ | 0.572 49 | 0.427 51 | 0.311 52 | 0.260 98 | 0.226 53 | 0.200 97 | | | | |
| $R=4$ | 0.587 62 | 0.412 38 | 0.332 41 | 0.255 20 | 0.218 96 | 0.193 42 | 0.173 96 | 0.158 45 | | |
| $R=5$ | 0.594 68 | 0.405 34 | 0.322 72 | 0.271 94 | 0.216 41 | 0.188 93 | 0.169 07 | 0.153 65 | 0.141 16 | 0.130 77 |

In summary, we note that the $f_m$ are piecewise linear after the first stage and can be calculated directly and trivially knowing $\theta^*(1)$ and their values at $\theta^*(1)$. The latter are obtained from exact solution of the monomer filling problem with range $R$ blocking [which is isomorphic to random $(R+1)$-mer filling] using the methods of Sec. II. In Table III, we have given $\theta^*(1),\ldots,\theta^*(2R)$ values for $R = 1$–5.

## V. DISCUSSION

Here we have shown that exact solution via hierarchy truncation is possible for $N$-mer filling, not only with NN cooperative effects (as realized previously), but also with general range $N$ cooperative effects. Exact solution follows from the shielding property of sufficiently large blocks of empty sites together with the observation that a "minimal" closed hierarchy can be obtained for the $f_m$. For processes on higher-dimensional lattices, a shielding property still exists for suitable thick walls of sites specified empty which separate the lattice into two disconnected regions.[1] However this condition is not "strong" enough to allow exact truncation. On the other hand, for Bethe lattices, we show in the following paper[30] that various random and cooperative processes are also amenable to exact solution.

We also note that, for the processes solved exactly here, we have only considered the minimal closed hierarchy for probabilities of empty $n$-tuples. As indicated in the introduction, this is not sufficient to calculate exact values for probabilities of, e.g., filled $n$-tuples or two-point correlations. However these can be obtained by including further rate equations for appropriate "effectively disconnected empty configurations"[1] and again using shielding to exactly truncate these equations (see Refs. 5, 22, and 23 for examples of these manipulations). We emphasize that the more general statement of shielding [rather than just (1.1) and (1.2)] is required here.

Finally we remark that certain other irreversible processes in 1-D lattice are amenable to exact solution, e.g., some involving competing events[28] and some on lattices with a stochastic distribution of different types of sites. It is also possible to exactly solve the infinite hierarchies of rate equations for certain processes on semi-infinite lattices[5] and on lattices with defective clusters of sites (here there is no translational invariance). These will be discussed in later work.


## ACKNOWLEDGMENTS

Ames Laboratory is operated for the U.S. Department of Energy under Contract No. W-7405-ENG-82. This work was supported by the Office of Basic Energy Sciences.


## APPENDIX: THE SHIELDING PROPERTY

Using the $q$-hierarchy in the form

$$\frac{d}{dt}\ln q_{\sigma\bar{\sigma}} = S(\sigma+\sigma') - S(\sigma'), \tag{A1}$$

where $S(\sigma'') = [(d/dt)f_{\sigma''}]/f_{\sigma''}$, we prove the shielding property for monomer filling ($N=1$) with range $R$ cooperative effects via self-consistency. Consider $q$'s of the form $q_{\overline{\sigma \mathbf{o}_{2R}\sigma'}\sigma''}$, where $\sigma$, $\sigma'$, $\sigma''$ involve only empty sites and, as the notation implies, $\sigma$ is to the left and $\sigma'$, $\sigma''$ are to the right of $\mathbf{o}_{2R}$. These $q$'s together with the corresponding reflected quantities satisfy a closed set of hierarchy equations (after using conservation of probability to convert any filled to empty sites). If $\tau_{\sigma^j}$ denotes the adsorption rate at site $j$ given the influencing environment in state $\sigma^j$ and if $\sigma^j_*$ denotes the part of $\sigma^j$ not overlapping the other conditioning sites, and $\mathbf{o}_{2R} = \mathbf{o}_{R-}\mathbf{o}_{R+}$ denotes a decomposition into left and right empty $R$-tuples, then[1]

$$\begin{aligned}
-\frac{d}{dt}\ln q_{\overline{\sigma \mathbf{o}_{2R}\sigma'}\sigma''} &= \sum_{j\in\sigma\mathbf{o}_{R-}}\sum_{\sigma^j_*}\tau_{\sigma^j}\left(q_{\sigma^j_*\,\overline{\sigma\mathbf{o}_{2R}\sigma'\sigma''}} - q_{\sigma^j_*\,\overline{\sigma\mathbf{o}_{2R}\sigma'}}\right)\\
&+ \sum_{j\in\mathbf{o}_{R+}\sigma'\sigma''}\sum_{\sigma^j_*}\tau_{\sigma^j} q_{\overline{\sigma\mathbf{o}_{2R}\sigma'\sigma''}\sigma^j_*}\\
&- \sum_{j\in\mathbf{o}_{R+}\sigma'}\sum_{\sigma^j_*}\tau_{\sigma^j} q_{\overline{\sigma\mathbf{o}_{2R}\sigma'}\sigma^j_*}\,.
\end{aligned} \tag{A2}$$

The analogous reflected quantities satisfy a similar set of equations. If the initial conditions satisfy the condition that an empty $2R$-tuple shields, then it is easy to see that the above equations are consistent with this condition for all time. For example, in (2), the first term cancels and the last two are independent of $\sigma$ consistent with $q_{\overline{\sigma\mathbf{o}_{2R}\sigma'}\sigma''}$ being independent of $\sigma$. Equations (1.1) and (1.2) follow as a special case. An analogous treatment of the special case $R=2$ appears in Ref. 31. If an empty $2R$-tuple does not shield initially but an empty $m^*$-tuple does ($m^* > 2R$), then the proof is easily modified to show that an empty $m^*$-tuple shields for all times.

We can relax the condition that $\sigma$, $\sigma'$, and $\sigma''$ be empty which gives more complicated $q$-equations[1] but the proof goes through unchanged in spirit. Similarly, extension to the $N$-mer case is more complicated in detail but the same in essence.

[1]J. W. Evans, D. R. Burgess, and D. K. Hoffman, J. Chem. Phys. **79**, 5011 (1983); specific examples of the hierarchial rate equations also appear in many of the following references.
[2]P. J. Flory, J. Am. Chem. Soc. **61**, 1518 (1939).
[3]E. R. Cohen and H. Reiss, J. Chem. Phys. **38**, 680 (1963).
[4]R. B. McQuistan and D. Lichtman, J. Math. Phys. **38**, 680 (1963); T. H. K. Barron and E. A. Boucher, Trans. Faraday Soc. **65**, 3301 (1969).
[5]N. O. Wolf, Ph.D. thesis, Iowa State University, 1979.
[6]K. J. Vette, T. W. Orent, D. K. Hoffman, and R. S. Hansen, J. Chem. Phys. **60**, 4854 (1974).
[7]E. S. Page, J. R. Stat. Soc. B **21**, 364 (1959).
[8]F. Downton, J. R. Stat. Soc. **23**, 207 (1961).
[9]W. H. Olson, J. Appl. Probab. **15**, 835 (1978).
[10]M. Gordon and I. H. Hillier, J. Chem. Phys. **38**, 1376 (1963); J. K. MacKenzie, *ibid.* **37**, 723 (1962); E. A. Boucher, Chem. Phys. Lett. **17**, 221 (1972).
[11]E. A. Boucher, J. Chem. Phys. **59**, 3848 (1972).
[12]E. A. Boucher, Faraday Trans. II **69**, 1839 (1973).
[13]A. Renyi, Sel. Translat. Math. Stat. Probab. **4**, 205 (1963); C. Domb, Proc. Philos. Soc. **43**, 329 (1947); P. E. Ney, Ann. Math. Stat. **33**, 702 (1962).
[14]J. B. Keller, J. Chem. Phys. **38**, 325 (1963) and references therein.
[15]D. A. McQuarrie, J. P. McTague, and H. Reiss, Biopolymers **3**, 657 (1965).
[16]N. O. Wolf, D. R. Burgess, and D. K. Hoffman, Surf. Sci. **100**, 453 (1980).
[17]R. L. Dobrushin, Probl. Peredachi Inf. **7**, 57 (1971); L. G. Mityushin, *ibid.* **9**, 81 (1973).
[18]See, for example, P. C. Hemmer and J. J. Gonzalez, J. Polymer Sci. **15**, 321 (1977).
[19]G. Schwartz, Ber. Bunsenges. Phys. Chem. **75**, 40 (1971).
[20]T. H. K. Barron, R. J. Bawden, and E. A. Boucher, J. Chem. Soc. **70**, 651 (1974).
[21]E. Klesper, W. Gronski, and V. Barth, Makromol. Chem. **150**, 223 (1971) and **160**, 167 (1972).
[22]N. A. Plate, A. D. Litmanovich, O. V. Noah, A. L. Toam, and N. B. Vasileyev, J. Polym. Sci. **12**, 2165 (1974).
[23]J. W. Evans, D. R. Burgess, and D. K. Hoffman, J. Math. Phys. (in press).
[24]J. J. Gonzalez and P. C. Hemmer, Polym. Lett. Ed. **14**, 645 (1976); J. Chem. Phys. **67**, 2496 and 2509 (1977).
[25]J. J. Gonzalez, P. C. Hemmer, and J. S. Høye, Chem. Phys. **3**, 228 (1974).
[26]I. R. Epstein, Biopolymers **18**, 765 (1979).
[27]J. W. Evans and D. K. Hoffman, J. Stat. Phys. (in press, 1984).
[28]J. W. Evans, D. K. Hoffman, and D. R. Burgess, J. Chem. Phys. **80**, 936 (1984).
[29]D. K. Hoffman, J. Chem. Phys. **69**, 765 (1978); J. W. Evans, Physica **123A**, 297 (1984).
[30]J. W. Evans, J. Math. Phys. **25**, 2527 (1984).
[31]J. W. Evans and D. R. Burgess, J. Chem. Phys. **79**, 5023 (1983).
[32]M. Higuchi and R. Senju, Polym. J. **3**, 370 (1972).

# Exactly solvable irreversible processes on Bethe lattices

J. W. Evans
*Ames Laboratory and Department of Chemistry, Iowa State University, Ames, Iowa 50010*

(Received 7 September 1983; accepted for publication 16 March 1984)

We consider the kinetics of processes where the sites of a Bethe lattice are filled irreversibly and, in general, cooperatively by monomers, dimers, or polyatomics. For nearest neighbor and sometimes more general cooperative effects (including random filling as a special case), we show that the infinite hierarchy of rate equations for probabilities of empty subconfigurations can be exacty truncated and solved using a shielding property of empty sites. We indicate, in certain cases, a connection between these Bethe lattice solutions and certain approximate truncation solutions for corresponding processes on "physical" 2-D and 3-D lattices with the same coordination number.

PACS numbers: 05.50. + q, 05.70.Ln

## I. INTRODUCTION

We consider here processes where the sites of a Bethe lattice are irreversibly filled, either randomly or cooperatively, by monomers, dimers, or polyatomics. In the language of graph theory, a Bethe lattice is an infinite, regular (Cayley) tree, i.e., a connected graph without circuits.[1] We must also specify the coordination number $c$ ($c = 2$ corresponds to a linear lattice). Thus for $c > 2$, Bethe lattices incorporate the connectivity features of a 1-D lattice with some of the combinatorial properties of higher-dimensional lattices. The role of Bethe lattices in the theory of disordered systems, e.g., in Ising model or percolation theory calculations, is well known as providing an artificial mathematical model for which some theoretical techniques give exact results.[2] We will show here that the same is true for certain *dynamic* irreversible processes.

For finite range cooperative effects or random processes, one can immediately write down an infinite hierarchy of rate equations for the probabilities of various subconfigurations[3] (which is trivial only for random monomer filling). Using conservation of probability, one can instead deal with the subhierarchy for the probabilities of subconfigurations where all specified sites are empty.[3] We show here that it is possible to solve this hierarchy (in closed form) by exact truncation for a variety of processes.

The basic tool used to obtain solutions is the following shielding property of empty sites[3]: a connected cluster of sites specified empty separates the lattice into several disconnected parts; suppose the cluster is sufficiently large that an event occuring on the lattice is not simultaneously affected by the state of sites in a particular pair of these disconnected parts; then sites in either of these two parts are completely shielded from the effect of those in the other. Just as in the preceding paper,[4] those processes amenable to exact solution via truncation are characterized by the existence of a minimal *closed* subhierarchy for the probabilities of connected clusters of empty sites.

In Sec. II, we consider in detail the exact solution for random dimer filling of a Bethe lattice with arbitrary coordination number $c$ and, in Sec. III, monomer filling with nearest neighbor (NN) cooperative effects for $c = 3$ (the solution for general $c$ is also discussed here). Explicit examples of the shielding condition are given and its role in exact hierarchy truncation is elucidated. The connection between the solutions for these processes and certain approximate solutions of corresponding processes on "physical" 2-D and 3-D lattices with the same coordination number is also described. In Sec. IV, we discuss more general exactly solvable processes and also describe those not amenable to exact solution. Finally, in Sec. V, we give some concluding remarks, in particular, indicating the utility of exact Bethe lattice solutions as providing a basis for resummation of formal density expansions for solutions of corresponding processes on physical 2-D and 3-D lattices.

The following notation is used here. Empty (filled) sites are denoted by "$o$" ("$a$"). Let $f_\sigma$ denote the probability of a subconfiguration $\sigma$ of sites specified filled or empty, and $q_{\sigma\bar{\sigma}'} = f_{\sigma+\sigma'}/f_{\sigma'}$, the conditional probability of $\sigma$ *given* $\sigma'$ (for typographic convenience, empty conditioning sites $\bar{o}$ are denoted $\phi$ here). All processes considered have translational invariance (as well as other natural symmetries) so $f$'s and $q$'s are position independent.

## II. RANDOM DIMER FILLING

Here we consider the random dimer filling of NN pairs of sites on a Bethe lattice with coordination number $c$ (arbitrary, but fixed). The rate of decrease of the probability of some subconfiguration of sites specified empty consists simply of a sum of terms corresponding to all possible ways a dimer can land completely within or partly overlapping this configuration (cf. Ref. 5). To provide some examples we need the following notation. Let $o_n$ denote a connected string of $n$ empty sites, e.g., $o$, $oo$, $ooo$, ..., and $\overset{o}{\mathbf{o}_i o \mathbf{o}_j}$ denote a string of $i + 1 + j$ empty sites where, in addition, one of the remaining $c - 2$ NN of a non-end site of the string is also specified empty. Then if $\tau$ denotes the single filling rate, one has

$$-\tau^{-1}\frac{d}{dt}f_o = cf_{oo}, \tag{2.1a}$$

$$-\tau^{-1}\frac{d}{dt}f_{oo} = f_{oo} + 2(c-1)f_{ooo}, \tag{2.1b}$$

$$-\tau^{-1}\frac{d}{dt}f_{o_n} = (n-1)f_{o_n} + 2(c-1)f_{o_{n+1}} + (c-2)\sum_{\substack{i+j=n-1 \\ i,j>1}} f_{\overset{o}{\mathbf{o}_i o \mathbf{o}_j}}, \quad n \geqslant 3, \tag{2.1c}$$

$$\vdots$$

 0022-2488/84/082527-06$02.50  

(using translation, etc., invariance of the $f$'s).

The key to exact solution of this problem is the sheilding property of empty sites (as stated in the introduction) which is embedded in (2.1) in a rather subtle way. Since each site of a Bethe lattice is an articulation point (i.e., separates the lattice into disconnected parts), for random dimer filling, one has that a single site specified empty shields. Thus, for example,

$$q_m \equiv q_{o\phi_m} \equiv f_{o_{m+1}}/f_{o_m} = q_1 \equiv q_{o\phi}, \quad m \geqslant 1. \tag{2.2}$$

The shielding property can most easily be proved by writing down the $q$-hierarchy and observing self-consistency (assuming compatible initial conditions). Because of notational difficulties, we only give the following examples of the general $q$-equations $(d/dt)\ln q_{\sigma\bar{\sigma}'} = S(\sigma+\sigma') - S(\sigma')$, where $S(\sigma'') = (d/dt)\ln f_{\sigma''}$:

$$-\frac{1}{\tau}\frac{d}{dt}\ln q_{o\phi} = 1 + (c-2)q_{o\phi} + 2(c-1)(q_{o\phi\phi} - q_{o\phi}), \tag{2.3a}$$

$$\begin{aligned} -\frac{1}{\tau}\frac{d}{dt}\ln q_{o\phi_{2n}} &= 1 + (c-2)q_{o\phi_n\phi\phi_n} \\ &\quad + 2(c-1)(q_{o\phi_{2n+1}} - q_{o\phi_{2n}}) \\ &\quad + 2(c-2)\sum_{i=1}^{n-1}(q_{o\phi_i\phi\phi_{2n-i}} - q_{o\phi_i\phi\phi_{2n-1-i}}), \end{aligned} \tag{2.3b}$$

$n \geqslant 1$

$$\begin{aligned} -\frac{1}{\tau}\frac{d}{dt}\ln q_{o\phi_{2n+1}} &= 1 + 2(c-1)(q_{o\phi_{2n+2}} - q_{o\phi_{2n+1}}) \\ &\quad + 2(c-2)\sum_{i=1}^{n-1}(q_{o\phi_i\phi\phi_{2n-1-i}} - q_{o\phi_i\phi\phi_{2n-2-i}}) \\ &\quad + (c-2)(2q_{o\phi_n\phi\phi_{n+1}} - q_{o\phi_n\phi\phi_n}), \end{aligned} \tag{2.3c}$$

$n \geqslant 1$

$\vdots$

where $q_{o\phi_i\phi\phi_j} = f_{o\phi_i\phi o_j}/f_{o_{i+1+j}}$. The full $q$-hierarchy written out in this fashion is clearly compatible with the stated shielding condition [note that the third term in (2.3a), the third and fourth terms in (2.3b), and the second and third terms in (2.3c), cancel] and furthermore leads to the closed equation

$$-\frac{1}{\tau}\frac{d}{dt}\ln q_{0\phi} = 1 + (c-2)q_{o\phi}, \tag{2.4}$$

which can be integrated together with (2.1a) rewritten as

$$-\frac{1}{\tau}\frac{d}{dt}\ln f_o = cq_{o\phi}. \tag{2.5}$$

For an initially empty lattice, these have the solution for $q = q_{o\phi}$ in terms of the coverage $\theta = 1 - f_o$ of

$$q = \begin{cases} 1 + \frac{1}{2}\ln(1-\theta), & \text{for } c = 2, \quad (2.6a) \\ (c-2)^{-1}[(c-1)(1-\theta)^{(c-2)/c} - 1], & \text{for } c > 2. \quad (2.6b) \end{cases}$$

Setting $q = 0$ gives a saturation coverage $\theta^*$ satisfying

$$1 - \theta^* = \begin{cases} e^{-2}, & \text{for } c = 2, \quad (2.7a) \\ (c-1)^{-c/(c-2)}, & \text{for } c > 2 \quad (2.7b) \end{cases}$$

[e.g., $= \frac{1}{8}$ for $c = 3$, $= \frac{1}{9}$ for $c = 4$, $= 1/(5\sqrt{5})$ for $c = 6$]. We can also readily calculate various other probabilities. For example, using shielding, it follows that the probability of any empty, connected cluster of $n \geqslant 1$ sites is given by $q^{n-1}f_o$. Also by converting filled to empty sites using conservation of probability and using shielding, one can show that the probability for a single empty site surrounded by $c$ filled sites is given by

$$\sum_{n=0}^{c}(-1)^n\binom{c}{n}q^n f_o = (1-q)^c f_o.$$

The expressions (2.6a) and (2.7a) for $c = 2$ are of course exact results for dimer filling on a linear lattice.[5] In particular, (2.7a) was first obtained by Flory in 1939 (see Ref. 6). For $c > 2$, some of these expressions have appeared previously in the context of "first shell truncation approximations" for random dimer filling on higher-dimensional lattices, in particular, $c = 3$ for the 2-D hexagonal and $c = 4$ for the 2-D square lattice.[5] For $c = 6$, these correspond to the "first shell truncation approximation" on the 3-D cubic[7] but *not* the 2-D triangular lattice.[5]

## III. MONOMER FILLING WITH NN COOPERATIVE EFFECTS

We first illustrate in detail this process for a Bethe lattice with coordination number 3. Here the rates for filling a site with 0,1,2,3 occupied NN are denoted by $\tau_{o\cdot_o^o}$, $\tau_{a\cdot_o^o}$, $\tau_{o\cdot_a^a}$, $\tau_{a\cdot_a^a}$, respectively. The rate equations for the probability of a single empty site $f_o$, of an adjacent pair of empty sites $f_{oo}$, of a string of three empty sites $f_{ooo} = f_{o_o^o}$, of a site and all three of its NN empty $f_{oo_o^o} = f_{o_{oo}^o}$,... are

$$\begin{aligned} -\frac{d}{dt}f_o &= \tau_{o\cdot_o^o}f_{oo_o^o} + 3\tau_{a\cdot_o^o}f_{ao_o^o} + 3\tau_{o\cdot_a^a}f_{oo_a^a} + \tau_{a\cdot_a^a}f_{ao_a^a} \\ &= (\tau_{o\cdot_o^o} - 3\tau_{a\cdot_o^o} + 3\tau_{o\cdot_a^a} - \tau_{a\cdot_a^a})f_{oo_o^o} \\ &\quad + 3(\tau_{a\cdot_o^o} - 2\tau_{o\cdot_a^a} + \tau_{a\cdot_a^a})f_{ooo} \\ &\quad + 3(\tau_{o\cdot_a^a} - \tau_{a\cdot_a^a})f_{oo} + \tau_{a\cdot_a^a}f_o, \end{aligned} \tag{3.1a}$$

$$\begin{aligned} -\frac{d}{dt}f_{oo} &= 2\left[\tau_{o\cdot_o^o}f_{oo_o^o} + 2\tau_{o\cdot_o^a}f_{oo_o^a} + \tau_{o\cdot_a^a}f_{oo_a^a}\right] \\ &= 2\left[(\tau_{o\cdot_o^o} - 2\tau_{o\cdot_o^a} + \tau_{o\cdot_a^a})f_{oo_o^o}\right. \\ &\quad \left. + 2(\tau_{o\cdot_o^a} - \tau_{o\cdot_a^a})f_{ooo} + \tau_{o\cdot_a^a}f_{oo}\right], \end{aligned} \tag{3.1b}$$

$$\begin{aligned} -\frac{d}{dt}f_{ooo} &= -\frac{d}{dt}f_{o_o^o} = \tau_{o\cdot_o^o}f_{oo_o^o} + \tau_{a\cdot_o^o}f_{ao_o^o} \\ &\quad + 2\left[\tau_{o\cdot_o^o}f_{o_{oo}^o} + 2\tau_{o\cdot_o^a}f_{o_{oo}^a} + \tau_{o\cdot_a^a}f_{o_{oo}^a}\right] \\ &= \tau_{a\cdot_o^o}f_{ooo} + (\tau_{o\cdot_o^o} - \tau_{a\cdot_o^o})f_{oo_o^o} \\ &\quad + 2\left[(\tau_{o\cdot_o^o} - 2\tau_{o\cdot_o^a} + \tau_{o\cdot_a^a})f_{o_{oo}^o}\right. \\ &\quad \left. + 2(\tau_{o\cdot_o^a} - \tau_{o\cdot_a^a})f_{oo_o^o} + \tau_{o\cdot_a^a}f_{ooo}\right], \end{aligned} \tag{3.1c}$$

$$-\frac{d}{dt}f_{o_{o}^{oo}} = \tau_{o\cdot_o^o} f_{oo_o^o} + 3\left[\tau_{o\cdot_o^o} f_{o_o^{oo}{}_o^o} + 2\tau_{o\cdot_o^a} f_{o_o^{oo}{}_o^a} + \tau_{o\cdot_a^a} f_{o_a^{oo}{}_o^a}\right]$$
$$= \cdots,$$
$$\vdots. \qquad (3.1d)$$

Examples of shielding used here to obtain an exact truncation of this hierarchy are shown in Fig. 1. These include such identities as

$$q_{\phi_{\phi\phi_\phi^o}} = q_{\phi_{\phi\phi_\phi^o}} = q_{\phi\phi_\phi^o} (= q_{o\phi_\phi^\phi}), \qquad (3.2a)$$

$$q_{\phi_{\phi\phi^o}} = q_{\phi_{\phi\phi^o}} = q_{\phi\phi^o} (= q_{o\phi\phi}), \qquad (3.2b)$$

and thus also imply such quantities as

$$q_{\phi_{\phi\phi_o^o}} \equiv q_{\phi_{\phi\phi_\phi^o}} q_{\phi_{\phi\phi_o^\phi}}, \quad q_{\phi_{\phi\phi_o^o}} \equiv q_{\phi_{\phi\phi_\phi^o}} q_{\phi\phi_o^\phi},$$
$$q_{\phi\phi_o^o} = q_{\phi\phi_\phi^o} q_{\phi\phi_o^\phi} \qquad (3.3)$$

are equal. Further identities such as $q_{\phi_{\phi\phi_o^a}} = q_{\phi_{\phi\phi_o^a}}, q_{\phi_{\phi\phi_a^a}} = q_{\phi\phi_a^a}$,..., all compatible with the statement of shielding given in the Introduction, then follow from conservation of probability.

Again these properties are imbedded in the hierarchy equations in a rather subtle way. The simplest demonstration of the results required here comes from writing down the closed $q$-subhierarchy for conditional probabilities of a single empty site *given* various adjacent connected clusters of sites are empty (factorization [cf. (3.3)] and conservation of probability must be used to obtain this reduced form). When suitably arranged, consistency with the shielding condition becomes obvious [cf. (2.3) and Ref. 4] assuming compatible initial conditions. This procedure is straightforward though notationally complex. Thus instead of presenting these general $q$-equations, we consider in detail only the particularly important cases for $q_{o\phi\phi}$ and $q_{o\phi_\phi^\phi}$. These equations are first simply obtained from (3.1b)–(3.1d) and then closed applying shielding and factorization [cf. (3.2) and (3.3)] as shown below:

$$\frac{d}{dt}\ln q_{o\phi\phi} = \frac{(d/dt)f_{ooo}}{f_{ooo}} - \frac{(d/dt)f_{oo}}{f_{oo}}$$
$$= -\tau_{o\cdot_o^o} q_{o\phi_\phi^\phi} - \tau_{a\cdot_o^o} q_{a\phi_\phi^\phi} - 2\tau_{o\cdot_o^o}(q_{\phi_{\phi\phi_o^o}} - q_{\phi\phi_o^o})$$
$$- 4\tau_{o\cdot_o^a}(q_{\phi_{\phi\phi_o^a}} - q_{\phi\phi_o^a}) - 2\tau_{o\cdot_a^a}(q_{\phi_{\phi\phi_a^a}} - q_{\phi\phi_a^a})$$
$$= -\tau_{o\cdot_o^o} q_{o\phi_\phi^\phi} - \tau_{a\cdot_o^o} q_{a\phi_\phi^\phi}$$
$$= -\tau_{a\cdot_o^o} - (\tau_{o\cdot_o^o} - \tau_{a\cdot_o^o}) q_{o\phi_\phi^\phi}, \qquad (3.4a)$$

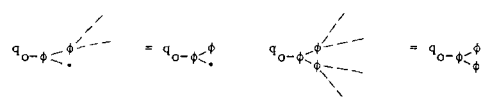


FIG. 1. Examples of shielding for monomer filling with NN cooperative effects on a Bethe lattice of coordination number 3; here - - - denotes other conditioning sites which could be specified either filled or empty.

$$\frac{d}{dt}\ln q_{o\phi_\phi^\phi} = \frac{(d/dt)f_{oo_o^o}}{f_{oo_o^o}} - \frac{(d/dt)f_{o_o^o}}{f_{o_o^o}}$$
$$= -\tau_{o\cdot_o^o} - \tau_{o\cdot_o^o}(3q_{\phi_{\phi\phi_o^o}} - 2q_{\phi_{\phi\phi_o^o}})$$
$$- 2\tau_{a\cdot_o^o}(3q_{\phi_{\phi\phi_o^a}} - 2q_{\phi_{\phi\phi_o^a}})$$
$$- \tau_{o\cdot_a^a}(3q_{\phi_{\phi\phi_a^a}} - 2q_{\phi_{\phi\phi_a^a}})$$
$$+ \tau_{o\cdot_o^o} q_{o\phi_\phi^\phi} + \tau_{a\cdot_o^o} q_{a\phi\cdot_\phi^\phi}$$
$$= -\tau_{o\cdot_o^o} - \tau_{o\cdot_a^a} - 2(\tau_{a\cdot_o^o} - \tau_{o\cdot_a^a}) q_{o\phi\phi}$$
$$- (\tau_{o\cdot_o^o} - 2\tau_{a\cdot_o^o} + \tau_{o\cdot_a^a}) q_{o\phi\phi} q_{o\phi_\phi^\phi}$$
$$+ (\tau_{o\cdot_o^o} - \tau_{a\cdot_o^o}) q_{o\phi_\phi^\phi} + \tau_{a\cdot_o^o}. \qquad (3.4b)$$

Equations (3.1a), (3.1b), and (3.4), together with the identities $f_{ooo} = f_{oo} q_{o\phi\phi}$, $f_{oo_o^o} = f_{o_o^o} q_{o\phi_\phi^\phi}$, yield an exact closed set of equations for $f_o, f_{oo}, q_{o\phi\phi}$, and $q_{o\phi_\phi^\phi}$. From these quantities, together with the shielding condition, we can readily calculate, e.g., the probability for any empty, connected cluster, e.g., $f_{o_{oo_o^o}} = (q_{o\phi\phi})^2 (q_{o\phi_\phi^\phi})^2 f_{oo}$.

We now analyze in detail the filling of an initially empty lattice for various choices of rates. All numerical results presented come from simultaneous integration of the closed set of equations described above [using a variable-order Adams predictor-corrector (Gear's) method]. First we consider the case where each filled NN changes the rate from $\tau_{o\cdot_o^o}$ by a factor of $\alpha$, i.e., $\tau_{o\cdot_o^o}:\tau_{a\cdot_o^o}:\tau_{o\cdot_a^a}:\tau_{a\cdot_a^a} = 1:\alpha:\alpha^2:\alpha^3$. The behavior $f_{oo}$ as a function of the coverage $\theta = 1 - f_o$ is shown in Fig. 2 for various $\alpha$.

*(a) Highly autoinhibitory rates $\alpha \ll 1$*: Here the process effectively takes place in four stages, namely, filling of sites with $m = 0,1,2,3$ occupied NN (consecutively). In the $m$th stage,

$$\frac{d}{d\theta} f_{oo} = -\tfrac{2}{3}(3 - m), \qquad (3.5)$$

since $3 - m$ $oo$ pairs are destroyed by each monomer adsorbing and there are $\frac{2}{3}$ as many sites as there are (adjacent) pairs of sites (cf. Ref. 3). We denote the $\theta$ values of the ends of these stages by $\theta_m^*$. Then $f_{oo_o^o} \to 0$, $f_{oo} \to 0.2500$, and $f_{o_o^o} \to 0.0767$ as $\theta \to \theta_0^* = 0.3750$ [these are saturation values for filling with NN blocking and are determined (numerically) after setting $\alpha = 0$]; $f_{o_o^o} \to 0$ and $f_{oo} \to 0.0732$ as $\theta \to \theta_1^* \approx 0.5076$ [numerical determination of these values actually comes from a modified choice of rates (i), described below, when $\alpha \to 0$]; $f_{oo} \to 0$ as $\theta \to \theta_2^* = \theta_1^* + \frac{3}{2} f_{oo}|_{\theta_1^*} \approx 0.6174$ and $\theta_3^* = 1$.

*(b) Random filling $\alpha = 1$:* Here, trivially, $f_{oo} = f_o^2$, $f_{o_o^o} = f_o^3$, etc.

*(c) Highly autocatalytic rates $\alpha \gg 1$*: Here the process involves competition between filling of a small fraction of sites with empty NN (island birth) and the rapid formation of

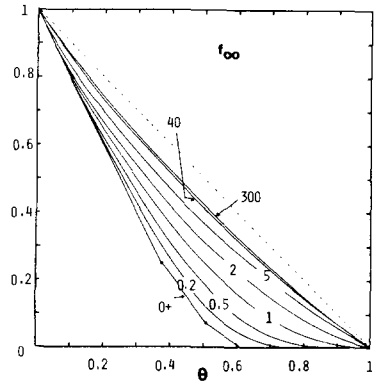



FIG. 2. $f_\infty$ for monomer filling of a Bethe lattice of coordination number $c = 3$ with NN cooperative effects where each filled NN changes the adsorption rate by a factor of $\alpha$ (shown); $\alpha = 300$ and $\infty$ behavior differ negligibly.

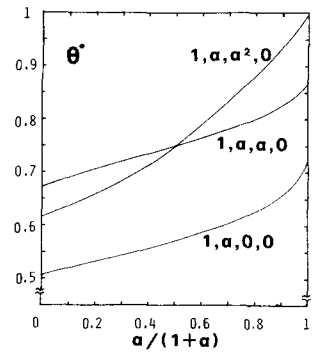



FIG. 3. Saturation coverage $\theta^*$ for monomer filling of a Bethe lattice with $c = 3$ for choices of rates $\tau_{o\cdot o}$, $\tau_{a\cdot o}, \tau_{o\cdot a}, \tau_{a\cdot a}$ as (i) $1,\alpha,0,0$, (ii) $1,\alpha,\alpha,0$, and (ii)′ $1,\alpha,\alpha^2,0$ (or any multiple of these).

contiguous clusters of filled sites around these (island growth). We now consider the $\alpha\to\infty$ limit which we show *cannot* be thought of as a single island growing. For if this were the case, except for the nucleating atom, every later atom adsorbed would destroy two *oo* pairs so $(d/d\theta)f_{oo} = -\frac{2}{3}\times 2 = -\frac{4}{3}$. This is clearly not correct for all $\theta$ since $f_{oo}\to 0$ only when $\theta\to 1$. However, we do anticipate that $[\lim_{\alpha\to\infty}(d/d\theta)f_{oo}]|_{\theta=0} = -\frac{4}{3}$. Deviation of $(d/d\theta)f_{oo}$ from $-\frac{4}{3}$ for $\theta \neq 0$ (see Fig. 2) must be due to infinitely large islands merging. In constrast, the $\alpha\to\infty$ behavior of $f_{oo}$ for corresponding processes on 1-D and "physical" 2-D or 3-D lattices can be determined by regarding these as occurring by a single island growing.[3]

Next we consider the modified choices for the rates $\tau_{o\cdot o}$: $\tau_{a\cdot o}:\tau_{o\cdot a}:\tau_{a\cdot a}$, of (i) $1:\alpha:0:0$, (ii) $1:\alpha:\alpha:0$, and (ii)′ $1:\alpha:\alpha^2:0$. Here we focus our attention on the saturation coverage $\theta^*(\alpha) < 1$ for finite $\alpha$ (see Fig. 3). Of course $\lim_{\alpha\to 0}\theta^*(\alpha)$ equals $\theta_1^*$ for (i) and $\theta_2^*$ for (ii)′ where $\theta_i^*$, defined above, are associated with filling in stages. For (ii) and (ii)′, when $\alpha = 1$, filling occurs randomly except that it is blocked when all NN are occupied. It has been shown previously[3] that the saturation coverage for this sort of "almost random" filling on an infinite, uniform lattice with coordination number $c$ is given by $c/(c+1)$, i.e., $\theta^*(1) = \frac{3}{4}$ for (ii) and (ii)′. The $\alpha\to\infty$ limit for (i) and (ii) provides an unequivocal demonstration of the anomalous feature of irreversible island growth on Bethe lattices described above, namely, that the highly autocatalytic/clustering limit *cannot* be though of as a single island growing. From Fig. 3 we see that $\lim_{\alpha\to\infty}\theta^*(\alpha) < 1$ for (i) and (ii). [Our numerical verification could be supplemented by a detailed asymptotic analysis of the appropriate solutions of (3.1) and (3.2) cf. Ref. 3.] The deficit from unity is associated with imperfect filling at the boundaries between coalesced infinite contiguous islands. The fact that these boundaries must be associated with a finite fraction of the lattice sites is not that surprising when one realizes that, for a single island on this Bethe lattice containing all atoms $\leqslant N$ lattice vectors from the nucleating atom, $\frac{1}{2}(1-2^{1-N}/3)\sim\frac{1}{2}$ the atoms are boundary atoms.

One final point of significance pertaining to this process is its connection with approximate solution for *monomer filling, with NN cooperative effects, at the sites of a 2-D hexagonal lattice* (also having coordination number 3). In previous work we have developed schemes to approximately truncate the infinite hierarchies of rate equations for such processes.[3] These deal with conditional probabilities involving only empty sites (and a single conditioned $o$-site) and attempt to exploit a shielding propensity of empty (cf. filled) sites. Two such schemes both start with the empty $q$-hierarchy and either (i) factorize $q$'s involving several $o$-sites in terms of those with a single $o$-site then truncate $\phi$ sites further than $n$ lattice vectors from the $o$-site ($FT$-$n$), or (ii) truncate $\phi$-sites further than $n$ lattice vectors from any $o$-site then factorize ($T$-$nF$). Reasonable results are expected for $n\geqslant 2$ which avoids explicit truncation of the $f_o$ equation.

It is easy to show that both the $FT$-2 and $T$-2$F$ equations for the above mentioned monomer filling of a hexagonal lattice are identical to (3.1a), (3.1b), and (3.4) for $f_o = q_o, f_{oo}$ (or $q_{o\phi}$), $q_{o\phi\phi}, q_{o\phi\phi}$ (and, consequently, we are guaranteed "well-behaved" solutions of these). We thus anticipate that the above Bethe lattice solutions bear some similarity to the corresponding exact quantities for the hexagonal lattice. It is however also clear that the $FT$-2 and $T$-2$F$ solutions have "lost knowledge" of the closed loops in the hexagonal lattice. Now consider the case where, again, each filled NN changes the rate from $\tau_{o\cdot o}$ by a factor of $\alpha$. As $\alpha\to 0$, filling on the hexagonal lattice also occurs in four stages and (3.5) holds (cf. Ref. 3). No doubt, the ends of these stages will vary from the Bethe lattice values given above. One can think of the $\alpha\to\infty$ limit as a single contiguous island growing and since, at nonzero coverages, an infinitesimal fraction of the occupied sites are boundary sites it follows that $q_{o\phi}, q_{o\phi\phi},\ldots,\to 1$. Consequently $f_{oo}, f_{ooo},\ldots,\sim f_o$ in contrast to the Bethe lattice case.

We can also consider a choice of rates analogous to (i) and (ii) described above for the Bethe lattice. The qualitative behavior of the saturation coverage $\theta^*(\alpha)$ for these cases is *deceptively* similar to the Bethe lattice behavior. In particular $\lim_{\alpha\to\infty}\theta^*(\alpha) < 1$, but this corresponds to the coverage of a single infinitely large noncontiguous island [which contains isolated empty sites for (i) and (ii) and also isolated empty pairs for (i)].

Finally, in this section, we consider the extension of the above results to monomer filling with NN cooperative effects on a Bethe lattice with arbitrary coordination number $c$. A detailed treatment would require the development of an

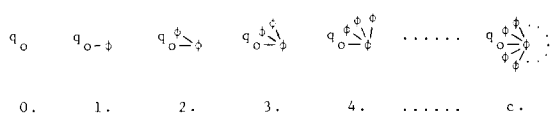


FIG. 4. Monomer filling with NN cooperative effects on a Bethe lattice with coordination number $c$; the $c+1$ empty $q$'s satisfying a closed set of equations (after applying shielding) are shown.

elaborate notational scheme to describe subconfigurations on such a lattice. However since the procedure leading to exact solution parallels the $c=3$ case described above, we just sketch the steps. The input to the hierarchial equations here is a set of rates $\tau_i$, $i=0,1,...,c$, for filling a site with exactly $i$ already filled NN. As for $c=3$, one starts by straightforwardly writing down rate equations for $f_o$, $f_{oo}$, $f_{ooo}$,... [cf. (2.1)]. From the $f_{oo}f_{ooo}$,... equations, one readily obtains a set of $q$-equations for $q_{o\phi\phi}=f_{ooo}/f_{oo}$,... . Applying shielding to these leads immediately to a *closed* set of equations for the $q$'s shown in Fig. 4 [cf. (3.4)], where we have also included $q_o=f_o$ and $q_{o\phi}$ (rather than $f_{oo}$) for completeness. The probability of, e.g., any empty cluster can be calculated from these.

Again we consider the natural choice of rates where each filled NN changes the rate (from the all empty NN case) by a factor of $\alpha$, i.e., $\tau_i/\tau_o=\alpha^i$. If $\alpha \ll 1$, the process effectively occurs in $c+1$ stages, namely, filling of sites with $m=0,1,2,...,c$ occupied NN (consecutively). In the $m$th stage,

$$\frac{d}{d\theta}f_{oo}=-\frac{2}{c}(c-m), \tag{3.6}$$

since $c-m$ $oo$ pairs are destroyed by each monomer adsorbing and there are $2/c$ times as many sites as there are (adjacent) pairs of sites. The $\alpha\to\infty$ limit is again quite complex but one can show that $[\lim_{\alpha\to\infty}(d/d\theta)f_{oo}]|_{\theta=0}=-2(c-1)/c$.

## IV. EXTENSIONS

Let us return first to the consideration of monomer filling processes. In Sec. III, we show how for NN cooperative effects, one can obtain a minimal closed hierarchy for the probabilities of connected clusters of empty sites. It is also possible to show that if the cooperative effects have a range of $2r+1$ lattice vectors incorporating range $r$ blocking, then again a minimal closed hierarchy can be obtained for the probabilities of empty connected clusters (cf. Ref. 4). Exact solution follows after applying the shielding condition. However for range $R>2r+1$ cooperative effects (incorporating range $r$ blocking), the minimal closed hierarchy for empty configurations contains not only connected configurations but also disconnected ones.[3] Just as for 1-D lattices,[4] shielding does not lead to exact truncation here and the simplest such process not amenable to exact solution is mon omer filling with range 2 cooperative effects.[8]

Next consider dimer filling processes. The random filling case has been solved in Sec. II. It is clear that for NN cooperative effects, again the minimal closed hierarchy involves only probabilities for empty connected configurations and thus exact truncation solution via shielding is possible.

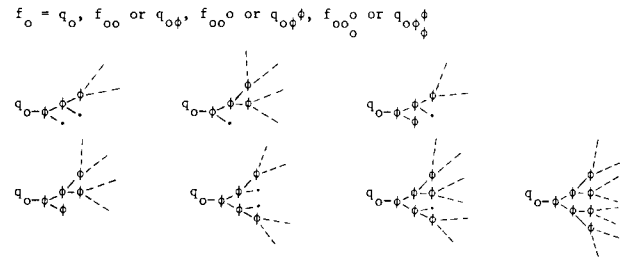


FIG. 5. Dimer filling with NN cooperative effects of a Bethe lattice with coordination number 3; $q$'s satisfying a closed set of equations (after applying shielding); here - - - indicate that the $q$'s are independent of these conditioning sites.

In Fig. 5, we show the finite set of quantities for which closed equations are obtained after applying shielding together with examples of the shielding condition for a Bethe lattice of coordination number $c=3$. For dimer filling with range 2 cooperative effects on a linear lattice, exact truncation of the hierarchy is possible.[4] However on a Bethe lattice with $c\geqslant 3$, this is not the case as it is easily seen that the minimal closed hierarchy does not involve only probabilities of connected empty clusters.

It is possible to consider processes involving filling of larger polyatomic molecules, e.g., trimers, linear or Y-shaped quadmers,... . Again exact hierarchy truncation is possible for NN, but not longer-range, cooperative effects (unless they incorporate a suitable range of blocking). It is quite straightforward to extent the analysis to treat adsorption of mixtures of polyatomics with the same atomic constituents (cf. Ref. 9).

We note that the preceding discussions of exact solutions for certain processes have concentrated on obtaining probabilities of empty, connected clusters of sites. We cannot obtain probabilities of, e.g., filled, connected clusters or two-point correlations from these. However by adding rate equations for suitable disconnected empty configurations and using shielding to truncate these, the latter quantities can also be calcualted exactly. Reference 10 illustrates the analogous procedures for the 1-D monomer filling process with NN cooperative effects. Finally we remark that one could also continue to examine competitive irreversible adsorption of different types of momomers "$a$," "$b$,"..., and, e.g., in the case of NN cooperative effects, also obtain exact solutions. The analogous 1-D calculations can be found in Ref. 11.

## V. DISCUSSION

Here we have shown how to obtain exact solutions via hierarchy truncation for a variety of nontrivial irreversible, random and cooperative, monomer, dimer, and polyatomic filling processes on Bethe lattices. These, being exact solutions, have intrinsic interest in enhancing our understanding of the kinetic behavior of irreversible processes on lattices. In some cases we have commented on the agreement between Bethe lattices solutions and certain natural low-order approximate truncation solutions for corresponding processes on 2-D and 3-D "physical" lattices with the same coordination number. Roughly speaking, this will be the case if the shortest closed loops on the physical lattice are sufficiently

large compared to the size of the polyatomic adsorbing, the cooperative range, and thus the (low-order) truncation range. Then these low-order truncations "cannot tell" that the loops close. This agreement was observed, for example, with random dimer filling on lattices with closed loops of length $> 3$ (so excluding the triangular lattice) for one lattice vector truncations; and for monomer filling with NN cooperative effects on a hexagonal, but not on a square or triangular lattice, for two lattice vector truncations.

For cases where agreement exists, this suggests some similarity between the Bethe lattice and exact "physical" lattice processes at least with regard to the probabilities of corresponding small configurations (not containing closed loops). We can always obtain formal "density" expansions for the $f$'s for physical lattice processes no matter how complex the cooperative effects.[12,13] The above observation suggests the following simple resummation procedure for expansion of $f$'s for small configurations. First we write the physical lattice probability $f^{\text{phys}}$ as the sum of the Bethe lattice one, $f^{\text{Bethe}}$, and a residual power series density expansion. Then coefficients in the residual expansion can be obtained from a knowledge of $f^{\text{Bethe}}$ and the density expansion of $f_{\text{phys}}$. This procedure has proved quite successful for some random dimer filling processes.[13,7]


## ACKNOWLEDGMENTS

Ames Laboratory is operated for the U.S. Department of Energy under Contract No. W-7405-ENG-82. This work was supported by the Office of Basic Energy Sciences.



[1] *Graph Theory and Theoretical Physics*, edited by F. Harary (Academic, London, 1967); R. J. Wilson, *Introduction to Graph Theory* (Oliver & Boyd, Edinburgh, 1972).
[2] J. M. Ziman, *Models of Disorder* (Cambridge U.P., Cambridge, 1979).
[3] J. W. Evans, D. R. Burgess, and D. K. Hoffman, J. Chem. Phys. **79**, 5011 (1983).
[4] N. O. Wolf, J. W. Evans, and D. K. Hoffman, J.Math. Phys. **25**, 2519 (1984).
[5] K. J. Vette, T. W.Orent, D. K. Hoffman, and R. S. Hansen, J. Chem. Phys. **60**, 4854 (1974).
[6] P. J. Flory, J. Chem. Soc. **61**, 1518 (1939).
[7] J. W. Evans and R. S. Nord, J. Stat. Phys. (submitted).
[8] J. W. Evans and D. R. Burgess, J. Chem. Phys. **79**, 5023 (1983).
[9] I. R. Epstein, Biopolymers **18**, 765 (1979).
[10] N. A. Plate, A. D. Litmanovich, O. V. Noah, A. L. Toam, and N. B. Vasileyev, J. Polymer Sci. **12**, 2165 (1974).
[11] J. W. Evans, D. K. Hoffman, and D. R. Burgess, J. Chem. Phys. **80**, 936 (1984).
[12] D. K. Hoffman, J. Chem. Phys. **69**, 3438 (1978).
[13] J. W. Evans, Physica **123A**, 297 (1984).

# Dirac equation for a linear potential

M. L. Glasser and N. Shawagfeh[a)]
*Clarkson University, Department of Mathematics and Computer Science, Potsdam, New York 13676*

(Received 12 October 1983; accepted for publication 6 January 1984)

The Dirac equation is solved for a linear potential. The complete asymptotic behavior of the solution in the nonrelativistic regime is developed by means of the comparison differential equation technique.

PACS numbers: 11.10.Qr, 02.90. + p

## I. INTRODUCTION

The linear potential, which has long been an important quantum mechanical model, has been the focus of renewed interest due to its role in theories of quark confinement. The only treatment of the Dirac equation with a linear potential appears to be that of Sauter[1] in 1931. Sauter's treatment employs the heavy machinery of gamma matrices to derive the four-component spinor wave function in terms of confluent hypergeometric functions. By means of an integral representation and the saddle point method he was able to obtain the leading asymptotic behavior in powers of the reciprocal field strength. Sauter's work is unsatisfactory in that his saddle point procedure leads only to order estimates for the subsequent terms, and that is able to give only a qualitative account of the nonrelativistic limit.

The purpose of this paper is to present a somewhat simplified solution to the Dirac equation and a complete quantitative discussion of the nonrelativistic limit. This requires deriving new uniform asymptotic expansions for the confluent hypergeometric function. Explicit error bounds for this expansion are given in the Appendix.

## II. SCHRÖDINGER EQUATION

For the potential $V(x) = \gamma x$ the Schrödinger equation is

$$\{-(\hbar^2/2m)\partial_x^2 + \gamma x\}\psi_0 = \epsilon\psi_0, \tag{1}$$

where we ignore the (separable) behavior in the orthogonal dimensions. The change of variable $\zeta = (2m/\hbar^2\gamma^2)^{1/2}(\gamma x - \epsilon)$ reduces this to Airy's equation with the familiar general solution

$$\psi_0 = S_1\,\mathrm{Ai}\{(2m/\hbar^2\gamma^2)^{1/3}(\gamma x - \epsilon)\} + S_2\,\mathrm{Ai}_1\{(2m/\hbar^2\gamma^2)^{1/3}(\gamma x - \epsilon)\}, \tag{2}$$

where $\mathrm{Ai}_{\pm 1}(z) = \mathrm{Ai}(ze^{\pm 2\pi i/3})$ and $S_{1,2}$ are arbitrary constants. Our aim is to show that the large components of the corresponding solution to Dirac's equation have this form in the nonrelativistic limit, and to find the leading-order correction.

## III. DIRAC EQUATION

After separating the irrelevant $y$ and $z$ motion and noting that there are no spin-dependent forces we are led to the two-component equation

$$\{\partial_x - (i/\hbar c)(E - \gamma x)\sigma_x - (mc/\hbar)\sigma_y\}\psi = 0, \tag{3}$$

where $\sigma_{x,y}$ are standard Pauli matrices and the derivative is understood to be multiplied by the unit matrix. The two-component spinor

$$\psi = \begin{pmatrix}\psi_1\\ \psi_2\end{pmatrix} \tag{4}$$

describes both spin states, wherein $\psi_1$ is the large component. We introduce the notation

$$u = (\gamma x - E)/(\hbar c\gamma)^{1/2}, \quad a = mc^2/(\hbar c\gamma)^{1/2}, \tag{5}$$

whereby (3) becomes

$$\{\partial_u - iu\sigma_x - a\sigma_y\}\psi = 0, \tag{6}$$

with the ansatz

$$\psi_1 = e^{iu^2/2}[F(u) + iG(u)], \quad \psi_2 = e^{iu^2/2}[F(u) - iG(u)], \tag{7}$$

we are led to the coupled system

$$F' + aG = 0, \quad G' + 2iuG + aF = 0, \tag{8}$$

from which $G$ can be immediately eliminated, giving

$$F'' + 2iuF' - a^2F = 0. \tag{9}$$

Two linearly independent solutions of (9) are

$$F_- = {}_1F_1(\tfrac{1}{4}ia^2, \tfrac{1}{2}, -iu^2), \quad F_+ = u\,{}_1F_1(\tfrac{1}{2} + \tfrac{1}{4}ia^2, \tfrac{3}{2}, -iu^2), \tag{10}$$

so the general solution to (3) is

$$\psi = e^{iu^2/2}\begin{bmatrix} AF_- + BF_+ - (i/a)(AF'_- + BF'_+)\\ AF_- + BF_+ + (i/a)(AF'_- + BF'_+)\end{bmatrix}, \tag{11}$$

where $A$ and $B$ are arbitrary constants.

## IV. ASYMPTOTIC ANALYSIS

In the nonrelativistic limit $c \to \infty$, we have $a \to \infty$. We also write $E = \epsilon + mc^2$, where $\epsilon$ has a finite limit, so $|u| \to \infty$ as well. The change of variable

$$z = -iu^2, \quad w = e^{-z/2}z^{m+1/2}F \tag{12}$$

reduces (9) to Whittaker's equation

$$\frac{d^2w}{dz^2} = \left(\frac{1}{4} - \frac{1}{4z} + \frac{ia^2}{4z} + \frac{m^2 - \frac{1}{4}}{z^2}\right)w, \tag{13}$$

for which the solutions corresponding to (10) are the functions $M_{1/4 - ia^2/4, \pm 1/4}$. Since we use an asymptotic expression for $w$ for large $a$ which is uniformly valid with respect to unbounded $z$ we define

[a)] Present address: P. O. Box 442, Mafraq, Jordan.

 0022-2488/84/082533-05$02.50  

$$t = zi/2N, \quad N = a^2/2, \tag{14}$$

whereby (13) becomes

$$\frac{d^2w}{dt^2} = N^2\left(f_0(t) + \frac{1}{N}f_1(t) + \frac{1}{N^2}f_2(t)\right)w, \tag{15}$$

$$f_0(t) = \frac{1-t}{2}, \quad f_1(t) = \frac{i}{2t}, \quad f_2(t) = \frac{m^2 - \frac{1}{4}}{t^2}.$$

All these functions are analytic for $t$ in the interval $(0,\infty)$ which includes the zero of $f_0(t)$. Our approach is to adopt the comparison differential equation method, which is described in detail by Olver.[2]

Making the Liouville transformation

$$W = \left(\frac{dt}{d\xi}\right)^{-1/2} w, \tag{16a}$$

$$\xi = t^2[(1-t)/t\,], \tag{16b}$$

(15) becomes

$$\frac{d^2W}{d\xi^2} = \{N^2\xi + N\phi_1(\xi) + \phi_2(\xi)\}W, \tag{17}$$

where

$$\phi_1(\xi) = -\tfrac{1}{2}i[\xi/(1-t)],$$

$$\phi_2(\xi) = \frac{5}{16\xi^2} + \left(\frac{m^2-\frac{1}{4}}{t^2} - \frac{5}{16t^2(1-t)^2} + \frac{1}{2t^2(1-t)}\right)\dot{t}^2. \tag{18}$$

From (16b) we find

$$\tfrac{2}{3}\xi^{3/2} = -\sqrt{t(1-t)} + \cos^{-1}\sqrt{t} \quad (0<t\leqslant 1), \tag{19}$$

$$\tfrac{2}{3}(-\xi)^{3/2} = \sqrt{t(1-t)} - \cosh^{-1}\sqrt{t} \quad (1<t<\infty),$$

whereby $t = 0,\infty$ correspond to $\xi = (3\pi/4)^{2/3}, -\infty$, respectively. Therefore

$$\phi_1(\xi) \sim (i/2)(\tfrac{3}{2})^{2/3}t^{-1/3}, \quad t\to\infty, \tag{20}$$

$$\phi_2(\xi) \sim 5/16\xi^2.$$

Following Olver[2] we make the further transformation

$$\bar{\xi}\left(\frac{d\bar{\xi}}{d\xi}\right)^2 = N^2\xi + N\phi_1(\xi) + \frac{\phi_1^2(\xi)}{4\xi}, \tag{21}$$

which gives

$$\bar{\xi} = N^{2/3}\xi + N^{-1/3}\Phi(\xi) + O(N^{-4/3}),$$

$$\Phi = \frac{1}{2}\xi^{-1/2}\int_0^\xi t^{-1/2}\phi_1(t)dt. \tag{22}$$

Then the solutions of (17) with $\phi_2$ replaced by $\phi_1^2(\xi)/4\xi$ are $(d\xi/d\bar{\xi})^{1/2}\,\mathrm{Ai}_j(\bar{\xi})$, $j = 0, \pm 1$, where $\mathrm{Ai}_0(z) = \mathrm{Ai}(z)$. The choice of two linearly independent solutions depends on the domain of the argument: the following sectors are associated with these three Airy functions:

$$\begin{aligned} &S_0: && |\mathrm{Arg}\, z| \leqslant 2\pi/3, \\ &S_{-1}: && -\pi \leqslant \mathrm{Arg}\, z \leqslant -\pi/3, \\ &S_{+1}: && \pi/3 \leqslant \mathrm{Arg}\, z \leqslant \pi, \end{aligned} \tag{23}$$

so that, for example, when $z \in S_0 \cup S_{-1}$, a satisfactory pair of solutions is $A_0(z)$ and $A_{-1}(z)$. By expanding for large $N$ we obtain the approximations to the solutions of (17)

$$W \sim \{1 + O(N^{-1})\}\,\mathrm{Ai}_j\{N^{2/3}\xi + N^{-1/3}\Phi(\xi)\} + O(N^{-4/3})\mathrm{Ai}_j'\{N^{2/3}\xi + N^{-1/3}\Phi\}. \tag{24}$$

Hence, the solution can be written in series form:

$$W_j = \mathrm{Ai}_j(\hat{\xi}) \sum_{s=0}^{\infty} N^{-s}A_s(\xi) + N^{-4/3}\mathrm{Ai}_j'(\hat{\xi}) \times \sum_{s=0}^{\infty} N^{-s}B_s(\xi), \tag{25}$$

$$\hat{\xi} = N^{2/3}\xi + N^{-1/3}\Phi(\xi),$$

where

$$\Phi(\xi) = \begin{cases} \tfrac{1}{2}i\xi^{-1/2}\cos^{-1}\sqrt{t}, & \xi\geqslant 0, \quad 0<t\leqslant 1, \\ \tfrac{1}{2}(-\xi)^{-1/2}\ln(\sqrt{t} + \sqrt{t-1}), & -\infty<\xi\leqslant 0, \\ \qquad 1<t<\infty. \end{cases} \tag{26}$$

Since $\hat{\xi} \in S_0 \cup S_1$, the appropriate solutions to the comparison equation are $W_0$ and $W_1$.

The coefficients in (25) can be evaluated by substituting (25) into (17) and equating like powers of $N$. In particular

$$A_0 = 1,$$

$$B_0 = \begin{cases} \frac{1}{2}\xi^{-1/2}\int_0^\xi [\phi_2(t) - 2\Phi(t)\Phi'(t) - t\Phi'^2(t)]t^{-1/2}\,dt, & \xi>0, \\ \frac{1}{2}(-\xi)^{-1/2}\int_\xi^0 [\phi_2(t) - 2\Phi(t)\Phi'(t) - t\Phi'^2(t)](-t)^{-1/2}\,dt, & \xi<0. \end{cases} \tag{27}$$

Also we see that $\Phi(\xi)$ is analytic and the analyticity of $\Phi_2(\xi)$ assures the analyticity of $A_s$, $B_s$. Error bounds on these solutions are given in the Appendix. Thus, truncating these expansions we have

$$W_{2n+1,0}(\xi) = \mathrm{Ai}(\hat{\xi}) \sum_{s=0}^{n} N^{-s}A_s(\xi) + N^{-4/3} \times \mathrm{Ai}'(\hat{\xi}) \sum_{s=0}^{n-1} N^{-s}B_s(\xi) + \epsilon_{2n+1,0}(\xi,N), \tag{28}$$

$$W_{2n+1,1}(\xi) = \mathrm{Ai}_1(\hat{\xi}) \sum_{s=0}^{n} N^{-s}A_s(\xi) + N^{-4/3} \times \mathrm{Ai}_1' \sum_{s=0}^{n-1} N^{-s}B_s(\xi) + \epsilon_{2n+1,1}(\xi,N).$$

Next we note that in terms of Whittaker functions (10) becomes

$$F_\pm = e^{-iNt}(-2iNt)^{-1/4}M_{k,\pm 1/4}(-2iNt), \tag{29}$$

where $k = \frac{1}{4} - \frac{1}{2}iN$. We write

$$M_{k,m}(-2iNt) = \left(\frac{dt}{d\xi}\right)^{1/2}[a_n W_{2n+1,0}(\xi) + b_n W_{2n+1,1}(\xi)], \tag{30}$$

and note that $a_n$, $b_n$ can be identified by comparing both sides of (30) in the limit $t\to\infty$ ($\xi\to-\infty$). This is facilitated by the known asymptotic formula

$$M_{k,m}(-2iNt)$$
$$\sim \frac{\Gamma(2m+1)}{\Gamma(m+k+\frac{1}{2})} e^{(k-m-1/2)\pi i}(-2iNt)^k e^{iNt}$$
$$+ \frac{\Gamma(2m+1)}{\Gamma(m-k+\frac{1}{2})}(-2iNt)^{-k} e^{-iNt}. \tag{31}$$

Using (19) and (22) we are led to the values

$$a_n = (1/\Omega_n)[e^{-\pi i/3}\Gamma_2^{\pm} + N^{-1}\beta_n e^{\pi i/3}\Gamma_1^{\pm}],$$
$$b_n = (1/\Omega_n)[(1 - iN^{-1}\beta_n\Gamma_1^{\pm}) - (1 + i\beta_n)\Gamma_2^{\pm}], \tag{32}$$

where

$$\Gamma_1^{(\mathrm{sgn}\, m)} = -\frac{\Gamma(2m+1)}{\Gamma(m+k+\frac{1}{2})} e^{(k-m)\pi i}(-2iN)^{k_2 iN - 1/2} e^{iN/2},$$
$$\Gamma_2^{(\mathrm{sgn}\, m)} = \frac{\Gamma(2m+1)}{\Gamma(m-k+\frac{1}{2})}(-2iN)^{-k_2 - iN + 1/2} e^{-iN/2}, \tag{33}$$
$$\Omega_n = 2\pi^{1/2} N^{1/6} e^{i\pi/4}[(1 - iN^{-1}\beta_n)e^{-i\pi/3} + N^{-1}\beta_n(1 + iN^{-1}\beta_n)e^{\pi i/3}],$$
$$\beta_n = \lim_{\xi \to -\infty} \sum_{s=0}^{n-1} \frac{|\xi|^{1/2} B_s(\xi)}{N^s}.$$

Hence we have the desired result

$$F_{\pm} = \Omega_n^{-1} e^{-iNt}(-2iN)^{-1/4}\left(\frac{\xi}{1-t}\right)^{1/4}\Bigg[(e^{-\pi i/3}\Gamma_2^{\pm} + N^{-1}\beta_n e^{\pi i/3}\Gamma_1^{\pm})\Big(\mathrm{Ai}(\hat{\xi}) \sum_{s=0}^{n} N^{-s} A_s(\xi)$$
$$+ N^{-4/3}\,\mathrm{Ai}'(\hat{\xi}) \sum_{s=0}^{n-1} N^{-s} B_s(\xi) + \epsilon_{2n+1,0}\Big) + [(1 - iN^{-1}\beta_n)\Gamma_1^{\pm} - (1 + iN^{-1}\beta_n)\Gamma_2^{\pm}]\Big(\mathrm{Ai}_1(\xi) \sum_{s=0}^{n} N^{-s} A_s(\xi)$$
$$+ N^{-4/3}\,\mathrm{Ai}_1'(\hat{\xi}) \sum_{s=0}^{n-1} N^{-s} B_s(\xi) + \epsilon_{2n+1,1}\Big)\Bigg]. \tag{34}$$

## V. DISCUSSION

The first approximation to (11) corresponds to $n = 0$. In this case (33) gives $\beta_0 = 0, \beta_{-1} = 0, \Omega_0 = 2\pi^{1/2} N^{1/6} e^{\pi i/4}$, so

$$F_{\pm}^{(0)}(t) = \frac{e^{-iNt} e^{-i\pi/4}(-2iN)^{-1/4}}{2\pi^{1/2} N^{1/6}}\left(\frac{\xi}{1-t}\right)^{1/4}$$
$$\times [\Gamma_2^{\pm}\,\mathrm{Ai}(\hat{\xi}) + (\Gamma_1^{\pm} - \Gamma_2^{\pm})\mathrm{Ai}_1(\hat{\xi})], \tag{35}$$

so

$$F^{(0)} \equiv AF_-^{(0)} + BF_+^{(0)} = e^{-iNt}[\xi/(1-t)]^{1/4}$$
$$\times [K_1\,\mathrm{Ai}(\hat{\xi}) + K_2\,\mathrm{Ai}_1(\hat{\xi})],$$

$$F^{(0)\prime} = \frac{2}{a} t^{1/2} e^{-iNt}\left(\frac{\xi}{1-t}\right)^{1/4}\Bigg[\Bigg\{-iN - \frac{1}{4(1-t)}$$
$$- \frac{1}{4\xi}\left(\frac{1-t}{t\xi}\right)^{1/4}\Bigg\}\{K_1\,\mathrm{Ai}(\hat{\xi}) + K_2\,\mathrm{Ai}_1(\hat{\xi})\}$$
$$+ N^{2/3}(1 + N^{-1}\Phi)\{K_1\,\mathrm{Ai}'(\hat{\xi}) + K_2\,\mathrm{Ai}_1'(\hat{\xi})\}\Bigg]. \tag{36}$$

By definition, $u = (\gamma x - \epsilon - mc^2)/(\hbar\gamma c)^{1/2} \equiv q - a$, where $q = (\gamma x - \epsilon)/(\hbar\gamma c)^{1/2}$. Hence,

$$t = 1 - 2q/a + q^2/a^2, \quad \xi = 2q/a + O[(q/a)^2], \tag{37}$$
$$\Phi = 1 + q/12a + O[(q/a)^2],$$
$$\hat{\xi} \sim (2a)^{1/3} q.$$

Therefore as $c \to \infty$,

$$e^{iNt}\left(F^{(0)}(u) + \frac{i}{a}F^{(0)\prime}(u)\right) \to 2[K_1\,\mathrm{Ai}\{(2a)^{1/3}q\} + K_2\,\mathrm{Ai}_1\{(2a)^{1/3}q\}],$$
$$e^{iNt}\left(F^{(0)}(u) - \frac{i}{a}F^{(0)\prime}(u)\right) \to 0, \tag{38}$$

and

$$\psi \to \begin{bmatrix} S_1\,\mathrm{Ai}[(2m/\hbar^2\gamma^2)^{1/3}(\gamma x - \epsilon)] + S_2\,\mathrm{Ai}_1[(2m/\hbar^2\gamma^2)^{1/3}(\gamma x - \epsilon)] \\ 0 \end{bmatrix}, \tag{39}$$

in exact agreement with (2). In the same way we obtain the first-order (in $N^{-1}$) relativistic correction

$$AF_+ + BF_- = F^{(0)}(u) - \frac{1}{2N} e^{-iNt}\left(\frac{\xi}{1-t}\right)^{1/4}\Phi'(\xi)[K_1\,\mathrm{Ai}(\hat{\xi}) + K_2\,\mathrm{Ai}_1(\hat{\xi})] + O(1/N^2). \tag{40}$$

In terms of Sauter's solution, we have the correspondence

$$F_- \leftrightarrow f_1, \quad a \leftrightarrow k,$$
$$F_+ \leftrightarrow g_1^*, \quad u \leftrightarrow \xi, \tag{41}$$

so in our terms Sauter's asymptotic ($k \gg 1$) results are: for $|u| < a$,

$$F_- \sim \left(1 - \frac{u^2}{a^2}\right)^{-1/4} \exp\left\{\frac{a|u|}{2}\sqrt{a^2 - u^2} + \frac{1}{2}(a^2 + i)\sin^{-1}\left(\frac{|u|}{a}\right)\right\},$$

$$F_+ \sim \left(1-\frac{u^2}{a^2}\right)^{-1/4}(\operatorname{sgn} u)\exp\left\{\frac{1}{2}a|u|\sqrt{a^2-u^2}\right.$$
$$\left.+\frac{1}{2}(a^2+i)\sin^{-1}\left(\frac{|u|}{a}\right)\right\}; \tag{42}$$

and for $|u|>a$,

$$F_-$$
$$\sim \frac{a}{|u|}w^{-1/2}\left(i\frac{\exp\{-\frac{1}{2}iu^2w+(ia^2/4)\ln(1+w)/(1-w)\}}{\sqrt{1+w}}\right.$$
$$\left.+\frac{\exp\{\frac{1}{2}iu^2w-(ia^2/4)\ln(1+w)/(1-w)\}}{\sqrt{1-w}}\right),$$

$$F_+$$
$$\sim \frac{a}{|u|}w^{-1/2}\left(\frac{\exp\{-\frac{1}{2}iu^2w+(ia^2/4)\ln(1+w)/(1-w)\}}{\sqrt{1-w}}\right.$$
$$\left.-\frac{i\exp\{\frac{1}{2}iu^2w-(ia^2/4)\ln(1+w)/(1-w)\}}{\sqrt{1+w}}\right),$$

where $w=(1-a^2/u^2)^{1/2}$. These expressions are not valid for $|u|\cong a$. Although (42) may have some computational advantage over (34) (since it involves only elementary functions), we consider the latter to be far more satisfactory not only because it is a complete asymptotic series with explicit error limits and reproduces the exact nonrelativistic limit, but also because it is uniformly valid for all values of $x$ including the classical turning point where (42) breaks down.


## ACKNOWLEDGMENT

The work of the first author was supported by the National Science Foundation.


## APPENDIX

To construct the error bounds $\epsilon_{2n+1,j}$ defined in (28) we need a differential equation with known solutions that approximate (17) up to the term $N\phi(\xi)W$. This can be found by applying the Liouville transformation

$$\hat{\xi}=N^{2/3}\xi+N^{-1/3}\Phi, \quad \hat{W}=\left(\frac{d\hat{\xi}}{d\xi}\right)^{1/2}W \tag{A1}$$

to the Airy equation for $\hat{W}(\hat{\xi})$, which leads to the differential equation

$$\frac{d^2W}{d\xi^2}=\{N^2\xi+N\phi_1(\xi)+G(\xi)\}W,$$
$$G(\xi)=\xi\Phi'^2+2\Phi\Phi'+\frac{\Phi\Phi'^2}{N} \tag{A2}$$
$$+\frac{3\Phi''^2-2\Phi'\Phi''-2N\Phi'''}{4(N+\Phi')^2}$$

with solutions $(N+\Phi')^{-1/2}\mathrm{Ai}_j(\hat{\xi})$, $j=0,1$. Now, by substituting (28) into (17) we find the differential equation

$$\frac{\partial^2}{\partial\xi^2}\epsilon_{2n+1,j}(\xi,N)-\{N^2\xi+N\phi_1+G(\xi)\}\epsilon_{2n+1,j}$$
$$=\{\phi_2-G\}\epsilon_{2n+1,j}+\{F_n\mathrm{Ai}_j(\hat{\xi})+H_n\,\mathrm{Ai}_j'(\hat{\xi})\}N^{-n},$$
$$F_n=(B_n+2\xi B_n')-N^{-1}\{\Phi\Phi'^2A_n+2\Phi\Phi'B_{n-1}' \tag{A3}$$
$$+(\Phi\Phi''+\Phi'^2)B_{n-1}\},$$
$$\mathrm{H}_n=2N^{-4/3}\{A_{n+1}'-N^{-1}\Phi\Phi'^2B_{n-1}\}.$$

We convert (A3) into the integral equation

$$\epsilon_{2n+1,j}=2\pi N^{-2/3}e^{-\pi i/6}\int_L K(\xi,t)[\{\phi_2(t)$$
$$-G(t)\}\epsilon_{2n+1,j}(t,N)+\{H_n(t)\mathrm{Ai}_j'(\hat{t})$$
$$+F_n(t)\mathrm{Ai}_j(\hat{t})\}N^{-n}]dt \tag{A4}$$
$$K(\xi,t)=[1+N^{-1}\Phi'(\xi)]^{-1/2}[1+N^{-1}\Phi'(t)]^{-1/2}$$
$$\times[\mathrm{Ai}(\hat{\xi})\mathrm{Ai}_1(\hat{t})-\mathrm{Ai}(\hat{t})\mathrm{Ai}_1(\hat{\xi})],$$

where the contour $L$ is the interval $(0,\xi)$ for $j=0$ and $(-\infty,\xi)$ for $j=1$, by the method of variation of parameters. We define the weight function

$$E_j(z)=|e^{(2/3)z^{2/3}}|, \quad j=0,\pm1, \tag{A5}$$

where the branch is chosen so that $E_j(z)\geqslant 1$ in the sector $S_j$ and $E_j(z)\leqslant 1$ in the complement of this sector. Next, the modulus and phase functions are defined by

$$E_{j+1}(z)|\mathrm{Ai}_{j+1}(z)|=M_j(z)\sin\theta_j(z), \quad j=-1,0,$$
$$E_{j-1}(z)|\mathrm{Ai}_{j-1}(z)|=M_j(z)\cos\theta_j(z), \quad j=1. \tag{A6}$$

The function $\bar{M}_j$, $\bar{\theta}_j$ are defined in the same way with Ai' replacing Ai. Finally, let

$$\lambda=\sup|1+N^{-1}\Phi'(\xi)|, \quad \xi\epsilon(-\infty,(3\pi/4)^{2/3}). \tag{A7}$$

Then

$$|K(\xi,t)|\leqslant\lambda\left(\frac{M_{-1}(\hat{\xi})M_{-1}(\hat{t})\sin\theta_{-1}(\hat{\xi})\cos\theta_{-1}(\hat{t})}{E_0(\hat{\xi})E_1(\hat{t})}\right.$$
$$\left.-\frac{M_{-1}(\hat{t})M_{-1}(\hat{\xi})\sin\theta_{-1}(\hat{t})\cos\theta_{-1}(\hat{\xi})}{E_0(\hat{t})E_1(\hat{\xi})}\right). \tag{A8}$$

But for $\hat{\xi}$, $\hat{t}\in S_0\cup S_1$, from (26) we have $E_1(\hat{t})=1/E_0(\hat{t})$, and since when $t\leqslant\xi$, $E_0(\hat{\xi})\leqslant E_0(\hat{t})$, (A8) reduces to

$$K(\xi,t)\leqslant\lambda M_{-1}(\hat{\xi})M_{-1}(\hat{t})E_0(\hat{t})E_0^{-1}(\hat{\xi}) \quad (t\leqslant\xi). \tag{A9}$$

Similarly,

$$\left|\frac{\partial}{\partial\xi}K(\xi,t)\right|\leqslant N^{2/3}N_{-1}(\hat{\xi})M_{-1}(\hat{t})E_0(\hat{t})E_0^{-1}(\hat{\xi}),$$
$$t\leqslant\xi,$$
$$|K(\xi,t)|\leqslant\lambda M_{-1}(\hat{\xi})M_{-1}(\hat{t})E_1(\hat{t})E_1^{-1}(\hat{\xi}), \tag{A10}$$
$$\xi\leqslant t,$$
$$\left|\frac{\partial}{\partial\xi}K(\xi,t)\right|\leqslant N^{2/3}N_{-1}(\hat{\xi})M_{-1}(\hat{t})E_1(\hat{t})E_1^{-1}(\hat{\xi}),$$
$$\xi\leqslant t.$$

Now, in order to bound $\epsilon_{2n+1,j}$ using (A4), we need an extension of Olver's theorem[3]: Consider the integral equation

$$h(\xi)=\int_\alpha^\xi K(\xi,t)\{\phi(t)J(t)+\rho(t)J'(t)$$
$$+\psi_0(t)h(t)+\psi_1(t)h'(t)\}dt, \tag{A11}$$

such that (i) $\phi(t)$, $\rho(t)$, $\psi_0(t)$, and $\psi_1(t)$ have at most a finite number of discontinuities in the range of $\xi$ and $J(t)$ is analytic; (ii) $K(\xi,t)$ and its first two partial $\xi$ derivatives are continuous in both variables; (iii) $K(\xi,\xi)=0$; (iv)

$$|K(\xi,t)|\leqslant P_0(\xi)Q(t),$$

$$\left|\frac{\partial}{\partial\xi}K(\xi,t)\right| \leqslant P_1(\xi)Q(t),$$
$$\left|\frac{\partial^2}{\partial\xi^2}K(\xi,t)\right| \leqslant P_2(\xi)Q(t),$$

where $P_j(\xi)>0$, $Q(t)$ are continuous real functions; (v) the suprema

$$x=\sup\{Q(\xi)|J(\xi)|\},\quad x_j=\sup\{P_j(\xi)Q(\xi)\},$$
$$\bar{x}=\sup\{Q(t)|J'(t)|\},$$

are finite; and (vi)

$$\Psi_j(\xi)=\int_\alpha^\xi |\psi_j(t)|dt,\quad j=0,1$$

and

$$\Phi(\xi)=\int_\alpha^\xi |\phi(t)|dt$$

converge. Then

$$|h(\xi)|/p_0(\xi),\ |h'(\xi)|/p_1(\xi)$$
$$\leqslant x\Phi(\xi)+\bar{x}R(\xi)\exp[x_0\Psi_0(\xi)+x_1\Psi_1(\xi)]. \tag{A12}$$

The proof parallels Olver's[4] and will not be given.

Compare now (A4) with (A11) setting

$$\psi_0=\phi_2-G,\quad \psi_1=0,\quad \sigma=H_n(t)N^{-n},$$
$$\rho=F_n(t)N^{-n},\quad J(t)=\mathrm{Ai}_j'(t)$$
$$\bar{x}=\begin{cases}\sup\{\pi E_{j+1}(z)M_j(z)|\mathrm{Ai}_{j+1}'(z)|\}, & j=-1,0,\\ \sup\{\pi E_{j-1}(z)M_j(z)|\mathrm{Ai}_{j-1}'(z)|\}, & j=1.\end{cases}$$

Then we have the following bounds:

$$|\epsilon_{2n+1,j}(\xi,N)|$$
$$\leqslant 2\lambda N^{-n-2/3}M_{-1}(\hat{\xi})E_j^{-1}(\xi)\{\nu S(\xi)+\bar{x}R(\xi)\}\exp\{(\lambda\bar{\nu}/\pi)\Psi_0(\xi)\},$$

where

$$S(\xi)=\int_0^\xi |H_n(t)||dt|,$$
$$R(\xi)=\int_0^\xi |F_n(t)||dt|,$$
$$\Psi_0(\xi)=\int_0^\xi |\phi_2(t)-G(t)||dt|,$$
$$\bar{\nu}=\sup\{\pi|z|^{1/2}M_j^2(z)\},$$
$$\nu=\begin{cases}\sup\{\pi E_{j-1}(z)M_j(z)|z^{1/2}\times\mathrm{Ai}_{j-1}(z)|\}, & j=1,\\ \sup\{\pi E_{j+1}(z)M_j(z)|z^{1/2}\times\mathrm{Ai}_{j+1}(z)|\}, & j=-1,0.\end{cases}$$


[1] F. Sauter, Z. Phys. **69**, 742 (1931).
[2] F. W. J. Olver, *Asymptotics and Special Functions* (Academic, New York, 1974).
[3] Reference 2, p. 218.
[4] N. Shawagfeh, Ph.D. thesis, Clarkson College of Technology, 1982 (unpublished).

# On the validity and practical applicability of derivative analyticity relations

P. Kolář [a)]
*International Centre for Theoretical Physics, Trieste, Italy*

J. Fischer
*Institute of Physics, Czechoslovak Academy of Sciences, Prague, Czechoslovakia*

(Received 25 October 1983; accepted for publication 13 January 1984)

We examine derivative analyticity relations (DAR), which were originally proposed by Bronzan as an alternative to dispersion relations and in which the dispersion integral is replaced by a tangent series of derivatives. We characterize the class of functions satisfying DAR, and show that outside this class the dispersion integral represents a Borel-like sum of tangent series. We point out difficulties connected with the application of DAR.

PACS numbers: 11.20.Fm, 02.30. + g

## I. INTRODUCTION

Ten years ago it was proposed[1] to replace the dispersion relation for the elastic scattering amplitude $F(s,t)$ by a "derivative analyticity relation" (DAR) which, contrary to the dispersion relation, would allow the real part $\mathrm{Re}\,F(s,t)$ at an energy $\sqrt{s}$ to be calculated by differentiating the imaginary part at the same energy, without the need to know the imaginary part at all points of the cuts. This idea aroused interest for practical reasons, promising to simplify theoretical calculations and also to reduce the amount of measurements required for the complete determination of the scattering amplitude. It aroused, on the other hand, critical comments regarding the mathematical background.

If poles and subtraction constants are omitted, the derivative analyticity relation can be written in the form[1]

$$\frac{\mathrm{Re}\,F(E)}{E^{\alpha}} = \tan\left(\frac{\pi}{2}\left(\alpha - 1 + \frac{d}{d\ln E}\right)\right)\frac{\mathrm{Im}\,F(E)}{E^{\alpha}}, \tag{1.1}$$

where $\alpha$ is a real number, $E$ is the laboratory energy of the incoming particle, and $F(E)$, the crossing-symmetric forward-scattering amplitude, is considered as a function of the real variable $E$. In the case of nonforward scattering ($t \neq 0$, $t$ fixed), $E$ in (1.1) is to be replaced by $\nu = (s-u)/4m$, where $m$ is the mass of the target particle. The right-hand side of (1.1) is an infinite sum of odd derivatives with respect to the logarithm of energy; for $\alpha = 1$ it has the form

$$\tan\left(\frac{\pi}{2}\frac{d}{dx}\right)f(x), \tag{1.2}$$

where $x$ and $f(x)$ stand for $\ln E$ and $\mathrm{Im}\,F(E)/E$, respectively.

We discuss in the present paper both the questions of validity and those of practical applicability of the derivative analyticity relations. First of all, Relation (1.1) has sense only if the infinite sum on its right-hand side is convergent; we therefore derive, in Sec. II (Theorem 1), a necessary and sufficient condition of the convergence of (1.2). In the same section, in Theorems 2, 3, and 4 we give conditions under which an infinitely differentiable function $f$ defined on the real axis $R^1$ can be extended from some interval $I \subset R^1$ to the corresponding entire function. According to Theorem 2, such an entire function exists if (1.2) converges on $I$. Further implications of the validity of (1.1) are also derived. In Sec. III, we apply the results to high-energy particle scattering. We conclude that if (1.1) is valid on some interval $I$ of the real axis then both $\mathrm{Re}\,F(E)$ and $\mathrm{Im}\,F(E)$ and, in this, also $F(E)$ must be extensible to entire functions. As Theorems 2, 3, and 4 have the form of necessary conditions, there are, on the other hand, entire functions for which the series (1.2) is divergent. Some examples to illustrate this are also given in Sec. III. A criterion[2] of the convergence of (1.2) in terms of the convergence radius of the Taylor expansion of $f(x)$ is found incorrect and counterexamples are given.

In Sec. IV, we derive the integral dispersion relation from the derivative one, (1.1). This allows us to investigate in detail the essential difference in the validity domains of a dispersion relation and the DAR. The dispersion integral is a generalized sum of the tangent series (1.2); this generalization slightly differs from the Borel summation.

Problems of approximating the true amplitude in terms of functions satisfying DAR are analyzed in Sec. V. We introduce two different fits to data, $\mathrm{Im}\,F^D$ and $\mathrm{Im}\,F^B$, to be used in the dispersion-relation approach and in the DAR approach, respectively, and derive bounds on the difference $\mathrm{Re}\,F^D - \mathrm{Re}\,F^B$. It turns out that in some situations these bounds may grow in an uncontrollable way due to low-energy contributions. (This fact was already pointed out by G. Höhler.[3]) Finally, we give special examples of functions showing that neither of the two approaches is secure from instabilities, though of different kinds. Section VI contains concluding remarks.

## II. ON THE CONVERGENCE OF THE TANGENT SERIES

The first problem to solve is to find a necessary and sufficient condition of the convergence of expression (1.2), which for $x = \ln E$ and $f(x) = \mathrm{Im}\,F(E)/E$ coincides with the right-hand side of (1.1) for $\alpha = 1$. (The case $\alpha \neq 1$ will be discussed in Sec. IV.)

The class of functions for which the tangent series (1.2) is convergent for every $x \in I$ will be referred to as class $\mathscr{A}(I)$ in the following. The symbol $\mathscr{A}$ will be used for $\mathscr{A}(R^1)$. We prove the following:

**Theorem 1:** Let $f: R^1 \to R^1$. The series (1.2) converges at a point $x \in R^1$ if and only if the series

$$\sum_{n=0}^{\infty} f^{(2n+1)}(x) \tag{2.1}$$

is convergent.

[a)] On leave of absence from Institute of Physics, Czechoslovak Academy of Sciences, Prague, Czechoslovakia.

 0022-2488/84/082538-07$02.50  

To prove it, we note that

$$\tan\left(\frac{\pi}{2}\frac{d}{dx}\right)f(x) = \sum_{n=1}^{\infty} a_n f^{(2n-1)}(x), \tag{2.2}$$

where

$$a_n = \frac{2\pi^{2n-1}(2^{2n}-1)}{(2n)!}|B_{2n}| = \frac{4}{\pi}(1-2^{-2n})\zeta(2n). \tag{2.3}$$

Here $B_n$ are the Bernoulli numbers and $\zeta(z)$ is the Riemann zeta function, which satisfies the inequalities[4]

$$1<\zeta(2n)<\tfrac{7}{4} \quad \text{for} \quad n\geqslant 1.$$

Hence $\{a_n\}$ is a bounded sequence.

Furthermore, it is simple to show that $\{a_n\}$ is decreasing for sufficiently large $n$. To see this we note that

$$\zeta(x) = \sum_{n=1}^{\infty} n^{-x}$$

and, for sufficiently large $n$, $a_n$ can be given in the form

$$a_n = (4/\pi)(1+3^{-2n}+\mathcal{O}(5^{-2n})).$$

Hence $a_n$ decreases to $4/\pi$ for $n\to\infty$. The statement of the theorem follows now from Abel's convergence test.[5]

Theorem 1 allows one to investigate the convergence of the tangent series (1.2) by considering the properties of $f(x)$ and its derivatives at the same point. It is, however, particularly useful if it is applied to a whole interval $I$ of the real axis $R^1$. To see this we use the following:

**Theorem 2:** Let the function $f: I\to R^1$ have all derivatives at every point of $I$ [i.e., $f$ belongs to the class $C^\infty(I)$]. If the series (2.1) converges for every $x\in I\subset R^1$, then an entire function exists which assumes the values of $f(x)$ on $I$.

The theorem follows from Ref. 6, where its generalization is proved. (The main steps of the proof are given in the Appendix.)

Of course, the same theorem is also valid if (2.1) is replaced by the sum of all even derivatives,

$$\sum_{n=0}^{\infty} f^{(2n)}(x). \tag{2.4}$$

We shall examine the connection between the properties of the two series. First we prove

**Theorem 3:** Let $f(x)$ be a function regular on a region $D$ of the complex plane $\mathscr{C}$. If the two sums (2.1) and (2.4) converge for some $x = x_0\in D$, then $f$ can be extended to an entire function $\tilde{f}$ ($\tilde{f}(x) = f(x)$ on $D$) such that $\Sigma_{n=0}^{\infty}\tilde{f}^{(2n+1)}(x)$ and $\Sigma_{n=0}^{\infty}\tilde{f}^{(2n)}$ converge uniformly on every bounded subset of $\mathscr{C}$.

The assumptions of the theorem imply the validity of the Taylor expansion of $f(x)$,

$$f(x) = \sum_{n=0}^{\infty}\frac{1}{n!}f^{(n)}(x_0)(x-x_0)^n \tag{2.5}$$

on some neighborhood of $x_0$. Since (2.5) converges for every $x\in\mathscr{C}$, the function $f(x)$ is extensible to an entire function $\tilde{f}(x)$. In the following, we shall prove the convergence of $\Sigma_{n=0}^{\infty}\tilde{f}^{(2n+1)}(x)$ only, because the convergence of the other series is proved analogously. Note that the statement of the theorem is also valid for $\Sigma_{n=0}^{\infty}\tilde{f}^{(n)}(x)$.

Denoting $S_{n,m} = \Sigma_{k=0}^{n} f^{(m+2k+1)}(x_0)$ we find

$$\sum_{k=N}^{M}\tilde{f}^{(2k+1)}(x) = \sum_{k=N}^{M}\sum_{n=0}^{\infty}\frac{1}{n!}f^{(n+2k+1)}(x_0)(x-x_0)^n$$

$$= \sum_{n=0}^{\infty}\frac{1}{n!}(S_{M,n}-S_{N-1,n})(x-x_0)^n.$$

The sum $S_{n,m}$ contains only even or only odd derivatives for $m$ odd or even, respectively. Take any $x\in\mathscr{C}$, $|x-x_0|\leqslant R$ and any positive $\epsilon$. Combining the Cauchy convergence tests for $\Sigma f^{(2k+1)}(x_0)$ and for $\Sigma f^{(2k)}(x_0)$ we easily find that there exists an $N_0$ such that

$$|S_{M,n}-S_{N-1,n}|<\epsilon e^{-R}$$

holds for every $M>N_0$, $N>N_0$ and every $n\geqslant 0$. Hence, we obtain that relation

$$\left|\sum_{k=N}^{M}\tilde{f}^{(2k+1)}(x)\right| < \epsilon e^{-R}e^{|x-x_0|} < \epsilon$$

is valid for any $x$, $|x-x_0|\leqslant R$, $M>N_0$ and $N>N_0$. Then, the theorem is proved by Cauchy's test.

The relation between the two sums (2.1) and (2.4) is now illustrated by the following theorem:

**Theorem 4:** If the series (2.1) converges on some interval $I\subset R^1$, then (2.4) is also convergent on $I$ and both sums can be extended to entire functions in the complex plane.

To prove this, it is sufficient to show the existence of some point $x_0\in I$ such that $\Sigma_{n=0}^{\infty}f^{(2n)}(x_0)$ converges. The theorem will then follow from Theorem 3.

According to the assumption of Theorem 4, $\Sigma_{n=0}^{\infty}f^{(2n+1)}(x)$ converges for every $x\in I$. Since $f^{(n)}\in C^\infty$, there exists an interval $(x_0,x_1)\subset I$ and $M>0$ such that $\Sigma_{n=0}^{N}f^{(2n+1)}(x)\leqslant M$ for every $x\in(x_0,x_1)$ and for every $N>0$. This is a consequence of the fact that (2.1) defines a finite function of the first class according to the Baire classification.[6] Using the identity

$$f^{(2n+1)}(x_1)-f^{(2n+1)}(x_0) = f^{(2n+2)}(x_0)(x_1-x_0) + \int_{x_0}^{x_1} f^{(2n+3)}(t)(x_1-t)\,dt,$$

we conclude that

$$\left|\sum_{n=N}^{M}f^{(2n+2)}(x_0)\right| \leqslant \frac{1}{|x_1-x_0|}\left|\sum_{n=N}^{M}f^{(2n+1)}(x_1) - \sum_{n=N}^{M}f^{(2n+1)}(x_0)\right| + M|x_1-x_0|. \tag{2.6}$$

Let $\epsilon>0$; first we choose $x_1$ to make $M|x_1-x_0|<\epsilon/2$. Then, applying the Cauchy tests on (2.1) for $x=x_0$ and $x=x_1$ we find that the first term on the right-hand side of (2.6) is bounded by $\epsilon/2$ for $N>N_0$ and $M>N_0$. Then we deduce the convergence of $\Sigma_{n=0}^{\infty}f^{(2n+2)}(x_0)$ from the Cauchy test.

The series (2.1) and (2.4) can be extended to the corresponding entire functions because the series of the entire functions are uniformly convergent on any bounded subset of the complex plane (see Theorem 3).

It is easy to see that the role of (2.1) and (2.4) can be interchanged in the theorem.

If Theorems 1, 2, and 4 are applied to an interval $I\subset R^1$, they can be combined to conclude that if (1.2) or $\tan[(\pi/2)(d/dx)]f'(x)$ converges for a function $f(x)$ on $I$ then $f(x)$ as well as the series (2.1), (2.4), (1.2), and $\tan[(\pi/2)(d/dx)]f'(x)$ are all extensible to the corresponding entire functions. The extension of $f(x)$ assumes the values of $f(x)$ on $I$ but may differ from it outside $I$.

## III. APPLICATION TO HIGH-ENERGY SCATTERING

Applying these results to the problem of high-energy scattering, we choose $x = \ln E$ and $f(x) = \operatorname{Im} F(E)/E$; the normalization of $F(E)$ is chosen such that the optical theorem has the form $\operatorname{Im} F(E) = k_{\text{lab}}\sigma(E)$, where $\sigma(E)$ is the symmetric total cross section.

The series (1.2) is often used to represent the scattering amplitude in some high-energy interval $I \subset R^1$. We see from the results obtained that this is possible only if $\operatorname{Im} F(E)$ is extensible from $I$ to a function entire in $\ln E$. This possibility cannot be *a priori* excluded but the practical applicability is limited due to the presence of branch points on $R^1$. [One can rather think of approximating $\operatorname{Im} F(E)$ with an entire function on $I$; we briefly comment on this in Sec. V.]

Furthermore, if we identify, as the relation (1.1) suggests, (1.2) with $\operatorname{Re} F(E)/E$, we see that $\operatorname{Re} F(E)/E$ should also be extensible from $I$ to a function which is entire in $\ln(E/E_0)$. Thus, the whole amplitude $F(E) = \operatorname{Re} F(E) + i \operatorname{Im} F(E)$ must also be extensible to an entire function. This extension, however, must coincide with the physical amplitude. Consequently, if (1.2) is convergent on some interval $I \subset R^1$, then the physical amplitude $F(E)$ should be an entire function in the logarithm.

We conclude that, to meet all the conditions required for the convergence of the tangent series, one must choose the interval $I$ extraordinarily short not to hit any of the branch points on $R^1$. This is hardly interesting from the practical point of view. Note also that the physical amplitude $F(E)$ which possesses cuts on $R^1$ is, in general, different from the entire extension, both in the complex plane and on the real axis outside $I$.

Theorems 1 and 2 do not exclude the possibility of $f(x)$ being extensible from $I$ to an entire function even when (1.2) and, by this, (2.1), is divergent. As an example, we can consider the function $f(x) = e^x$, for which $\sum f^{(2k+1)}(x)$ is divergent for all real $x$. Another such function, which in addition is bounded on $R^1$, is, e.g., $f(x) = \cos x$. The series (2.1) is divergent for it at all real $x$ unless $x = k\pi$.

Another consequence of our theorems is that if, on an interval $I$, $f(x)$ does not assume the values of an entire function, there must be at least one point on $I$ at which (2.1) is divergent. In the same interval, there must also be a point at which the series (2.4) diverges. We have already mentioned that the presence of branch points on the real axis makes the extension of $\operatorname{Im} F(E)$ to an entire function difficult. It is therefore of interest to investigate the situation in which $f(x)$ is holomorphic in a bounded domain only. The problem of the convergence of the tangent series under this assumption was considered by Heidrich and Kazes.[2] They conclude that if the Taylor series of a function $f(x) \in C^\infty$,

$$f(x) = \sum_{n=0}^{\infty} f^{(n)}(x_0)\frac{(x-x_0)^n}{n!},$$

has finite radius of convergence, $0 \leqslant R < \infty$, then the series (1.2) diverges at every $x \in R^1$. However, the statement is incorrect. One can find counterexamples, for instance,

$$f(z) = (1 + a\sin^2 z)^{-1} \tag{3.1}$$

with $a > 0$. The convergence radius $R$,

$$R = \ln\left[(1+\sqrt{1+a})/\sqrt{a}\right],$$

is finite and tends to infinity with $a \to 0$. The series (2.2) converges at all $z = k\pi$ with $k$ being an integer.

It is worth noting the difference between the set $A \subset I$ of points at which (1.2) converges and the set $B = I - A$ of points at which (1.2) diverges. If an interval $J = (a,b)$ is a subset of $A$, then its endpoints $a$ and $b$ also belong to $A$. To prove this, we assume the contrary. Let $a \in B$. One can extend $f(x) \in C^\infty$ from $J$ to an entire function $\tilde{f}(x)$. According to Theorem 3, $\Sigma_n f^{(2n+1)}(z)$ uniformly converges on any bounded region in the complex plane $\mathscr{C}$. Hence, we get

$$\sum_n \tilde{f}^{(2n+1)}(a) = \lim_{\substack{x\to a \\ x\in J}} \sum_n \tilde{f}^{(2n+1)}(x) = \sum_n \lim_{\substack{x\to a \\ x\in J}} \tilde{f}^{(2n+1)}(x)$$

$$= \sum_n \lim_{\substack{x\to a \\ x\in J}} f^{(2n+1)}(x) = \sum_n f^{(2n+1)}(a).$$

Thus, the point $a$ belongs to $A$, and consequently $B$ cannot have isolated points.

Contrary to this, $A$ can contain isolated points. One example to illustrate this is given by formula (3.1). We did not, however, succeed in finding a function for which the set $A$ would have a limit point.

Let us summarize the discussion into the following formal statement:

*Corollary*: If $f(x)$ is regular on some bounded region $D$ containing the interval $I \subset R^1$ and is not extensible to an entire function, then there exists a subset $B$ of $I$ ($B \subset I$) everywhere dense in $I$ such that (1.2) diverges on $B$. Besides this, $I$ may contain points of convergence of (1.2).

## IV. LINK TO DISPERSION RELATIONS

We turn now to some further properties of the function $f(x)$. First we give a simple bound on $f(x)$.

**Theorem 5:** If $f \in \mathscr{A}(I)$, then the entire function $\tilde{f}$ which extends $f$ to the complex plane obeys the bound

$$|\tilde{f}(z)| \leqslant \epsilon e^{|z|} + C(\epsilon); \quad z \in \mathscr{C} \tag{4.1}$$

for every $\epsilon > 0$. Here, $C(\epsilon)$ is a constant that depends on $\epsilon$.

The proof easily follows from the Taylor expansion of $\tilde{f}(z)$ and from the fact that $\tilde{f}^{(n)}(z)$ tends to zero with increasing $n$ for every fixed $z \in \mathscr{C}$.

The relation (4.1) shows that functions belonging to $\mathscr{A}(I)$ are, at most, of the order 1 and of the type 1. These limiting values are saturated within $\mathscr{A}(I)$. For instance, the function $f(x) = \sin x/x$ satisfies (4.1) and is of the order 1 and of the type 1.

The bound (4.1) is valid in the whole complex plane. The same bound also holds for any derivative of $f$. We can also find constraints on the asymptotic behavior of $f(x)$ along the real axis from

**Theorem 6:** Let $f(z)$ be an entire function. If $\Sigma_{n=0}^{\infty} f^{(2n+1)}(x)$ converges for some $x \in R^1$, then

$$\sum_{n=0}^{\infty} f^{(2n+1)}(x) = \frac{1}{2}\int_0^\infty e^{-t}(f(x+t) - f(x-t))\,dt. \tag{4.2}$$

Indeed, since $f(z)$ is an entire function we have

$$\frac{1}{2}(f(x+t) - f(x-t)) = \sum_{n=0}^{\infty} f^{(2n+1)}(x)\frac{t^{2n+1}}{(2n+1)!}.$$

According to the Hardy theorem (see, e.g., Ref. 7) we can multiply this power series by $e^{-t}$ and integrate it term by term. Then Eq. (4.2) is easily established.

Similarly, if $f(z)$ is an entire function, then

$$\sum_{n=0}^{\infty} f^{(2n)}(x) = \frac{1}{2}\int_0^{\infty} e^{-t}(f(x+t)+f(x-t))\,dt \quad (4.3)$$

provided that the sum on the left-hand side of (4.3) is convergent.

A simple application of Theorem 6 is the following:

*Corollary*: Let $I$ be an interval of $R^1$. If $\Sigma_{n=0}^{\infty} f^{(2n+1)}(x)$ converges for every $x\in I$, then the relation

$$\sum_{n=0}^{\infty} f^{(n)}(x) = \int_0^{\infty} e^{-t}\tilde{f}(x+t)\,dt \quad (4.4)$$

holds for every $x\in I$. Here $\tilde{f}$ is the entire function which coincides with $f(x)$ inside $I$. Similarly,

$$\sum_{n=0}^{\infty} f^{(2n+1)}(x) - \sum_{n=0}^{\infty} f^{(2n)}(x) = -\int_0^{\infty} e^{-t}\tilde{f}(x-t)\,dt; \quad x\in I.$$

Of course, if $I = R^1$, then $f=\tilde{f}$, as follows from Theorem 2. If $f$ is replaced by $\tilde{f}$ the equation remains valid for every $x\in\mathscr{C}$.

It is straightforward to generalize (4.2), (4.3), and (4.4) to relations in which $f(z)$ is replaced by some derivative of general order. Thus, for instance, besides (4.4) we also have

$$\sum_{n=k}^{\infty} f^{(n)}(x) = \int_0^{\infty} e^{-t}\tilde{f}^{(k)}(x+t)\,dt$$

under the same assumptions. This gives the following relation for the partial sum:

$$\sum_{n=0}^{k-1} f^{(n)}(x) = \int_0^{\infty} e^{-t}(\tilde{f}(x+t)-\tilde{f}^{(k)}(x+t))\,dt.$$

We turn back now to the tangent series (1.2).

**Theorem 7:** Let $f(z)$ be an entire function. If the series (1.2) converges at a point $x\in R^1$, then

$$\tan\left(\frac{\pi}{2}\frac{d}{dx}\right)f(x) = \int_0^{\infty} a(t)e^{-t}(f(x+t)-f(x-t))\,dt, \quad (4.5)$$

where $a(t) = (2/\pi)(1-e^{-2t})^{-1}$. Note that the integral in (4.5) exists as a consequence of Theorems 1 and 6 and of the boundedness of the integrand in any neighborhood of the point $t=0$. The proof is analogous to that of Theorem 6. We emphasize that the convergence of the integrals (4.2), (4.3), and (4.5) does not imply the convergence of the corresponding sums [see, for instance, the function $f(x)=\sin x$].

The interest of Theorem 7 arises from the fact that we can use it to establish a connection between the series (1.2) and a dispersion relation. The integral in (4.5) can be given in the form

$$\int_0^{\infty} a(t)e^{-t}(f(x+t)-f(x-t))\,dt = \frac{1}{2\pi}\int_{-\infty}^{+\infty}\frac{dt}{\sinh t}(f(x+t)-f(x-t)) = \frac{1}{\pi}\int_{-\infty}^{+\infty}\frac{dt}{\sinh(t-x)}f(t)$$

$$= \frac{2e^x}{\pi}\int_0^{\infty}\frac{f(\ln E')}{E'^2-e^{2x}}\,dE',$$

which resembles the standard dispersion integral. Finally, setting $x=\ln E$ we obtain

$$\tan\left(\frac{\pi}{2}\frac{d}{d\ln E}\right)f(\ln E) = \frac{2E}{\pi}\int_0^{\infty}\frac{f(\ln E')}{E'^2-E^2}\,dE', \quad (4.6)$$

while the dispersion relation for the crossing-even amplitude $F(E)$ reads

$$\mathrm{Re}\,F(E) = \mathrm{Re}\,F_p(E) + \frac{2E^2}{\pi}\int_{E_0}^{\infty}\frac{\mathrm{Im}\,F(E')}{E'(E'^2-E^2)}\,dE', \quad (4.7)$$

where $F_p(E)$ denotes the sum of pole terms and the subtraction constant $F(0)$.

The right-hand side of (4.2) represents the Borel sum of the corresponding series. The dispersion integral (4.6) can be understood as a generalized sum of the tangent series (1.2) but this generalization differs from the Borel summation [because of the extra factors $a_n$ (2.3)].

Putting $f(x) = \mathrm{Im}\,F(E)/E$, we find the following formal differences between (4.6) and (4.7):

(i) the physical (or unphysical) threshold is equal to zero in (4.6),

(ii) the tangent series (1.2) can represent only the difference $\mathrm{Re}\,F(E) - \mathrm{Re}\,F_p(E)$.

A deeper difference is, however, associated with the validity domains of the relations (4.6) and (4.7). While the former applies to functions which are entire in $\ln E$, the latter has a larger domain of validity. For instance, the Hölder condition is sufficient for the validity of (4.7). Both relations assume the same constraint on the asymptotic behavior of $F(E)$. Concerning the point (i), let us remark that one can formally avoid this difference by considering dispersion relations with the lower integration limit equal to zero. This can be achieved, for instance, by replacing the energy variable $E$ by the laboratory momentum $k$ of the incoming particle.[8]

In the remainder of this section we shall briefly discuss the case $\alpha\neq 1$ [see Eq. (1.1)]. To establish the convergence of the series

$$\tan\left(\frac{\pi}{2}\left(\alpha-1+\frac{d}{dx}\right)\right)f(x) \quad (4.8)$$

we use the following identity:

$$\tan\left(\frac{\pi}{2}\left(\alpha-1+\frac{d}{dx}\right)\right)f(x) = e^{-(\alpha-1)x}\tan\left(\frac{\pi}{2}\frac{d}{dx}\right)e^{(\alpha-1)x}f(x). \quad (4.9)$$

The results of Secs. II and III can now be applied to the function $e^{(\alpha-1)x}f(x)$. We obtain, in particular, that if the series (4.8) converges on some interval $I$ then $f(x)$ can be extended to an entire function for any $\alpha$.

The mapping $f(x)\to e^{(\alpha-1)x}f(x)$ changes the inequality (4.1), which is valid in the whole complex plane, into

$$|\tilde{f}(x)| \leqslant e^{|x|-(\alpha-1)\mathrm{Re}\,x} + C(\epsilon)e^{-(\alpha-1)\mathrm{Re}\,x}. \quad (4.10)$$

Comparing (4.10) with (4.1) we see that the bound on $f(x)$ is changed unless $x$ lies on the imaginary axis.

The general derivative analyticity relation (1.1) is now obtained by setting

$$f(x) = \operatorname{Im} F(E)/E^{\alpha}. \tag{4.11}$$

Then (4.9) becomes

$$E^{\alpha} \tan\left(\frac{\pi}{2}\left(\alpha - 1 + \frac{d}{d\ln E}\right)\right) \frac{\operatorname{Im} F(E)}{E^{\alpha}} = E \tan\left(\frac{\pi}{2}\frac{d}{d\ln E}\right) \frac{\operatorname{Im} F(E)}{E} \tag{4.12}$$

for every real $\alpha$. Thus, there is no constraint on the value of $\alpha$. On the other hand, a derivation of (4.6) by expanding $f(x)$ given by (4.11) in a Taylor series and integrating term by term would require $\alpha$ to lie between 0 and 2 (see Ref. 2); a typical integral is $\int_{-\infty}^{+\infty} e^{(\alpha-1)t}(t^n/\sinh t)dt$. In our opinion, the choice of an $\alpha$ different from 1 has no practical advantage.

## V. COMMENTS ON THE PRACTICAL APPLICABILITY OF DAR

As the class of functions to which the derivative analyticity relations apply is narrow, the true amplitude can only be approximated in terms of these functions. Such approximations have been commonly used, in particular, to fit high-energy data.[9]

We shall give a qualitative discussion of the differences between the DAR approach and the ordinary dispersion relations. Let $\operatorname{Im} F^D(E)$ be an all-data fit to the imaginary part of $F(E)$, to be used in the dispersion relation, and let $\operatorname{Re} F^D(E)$ be the corresponding real part calculated from the dispersion relation. Let further $\operatorname{Im} F^B(E)$ be a high-energy fit in terms of entire functions and $\operatorname{Re} F^B(E)$ the corresponding real part calculated by using (1.1). As $\operatorname{Im} F^B(E)$ has been obtained by fitting all experimental data above a certain energy $E_1$, the difference

$$\Delta(E) = \operatorname{Im} F^B(E) - \operatorname{Im} F^D(E)$$

will be small at all energies between $E_1$ and the highest experimental point $E_2 > E_1$. Outside this interval, however, $\operatorname{Im} F^B(E)$ can be considerably different from $\operatorname{Im} F^D(E)$. This difference, in turn, will affect the predicted values of the real part.

Introducing

$$\delta(E) = \operatorname{Re} F^B(E) - \operatorname{Re} F^D(E) \tag{5.1}$$

we represent $\delta(E)$ in the following form:

$$\delta(E) = \delta_1(E) + \delta_2(E) + \delta_3(E), \tag{5.2}$$

where

$$\delta_1(E) \equiv \frac{2E^2}{\pi} \int_0^{E_0} \frac{\operatorname{Im} F^B(E')}{E'(E'^2 - E^2)} dE', \tag{5.3}$$

$$\delta_2(E) \equiv \frac{2E^2}{\pi} \int_{E_0}^{E_1} \frac{\Delta(E')}{E'(E'^2 - E^2)} dE', \tag{5.4}$$

$$\delta_3(E) \equiv \frac{2E^2}{\pi} \int_{E_1}^{\infty} \frac{\Delta(E')}{E'(E'^2 - E^2)} dE', \tag{5.5}$$

$E_0$ being the lowest branch point of the true scattering amplitude $F(E)$.

Let us discuss now the relative importance of the three terms on the right-hand side of (5.2). If $\Delta(E) \neq 0$ for $E$ above $E_1$, $\delta_3(E)$ may play an essential role. Yet, since the conventional fits to $\operatorname{Im} F^D(E)$ have the form of entire functions in $\ln E$, we can consider $\operatorname{Im} F^B(E)$ to be equal to $\operatorname{Im} F^D(E)$ for all $E > E_1$ and, by this, dispense with $\delta_3(E)$ in (5.2).

As we see from (5.3) $\delta_1(E)$ is infinite unless $\operatorname{Im} F^B(0) = 0$.[2] Fortunately, the last relation is easily proved. Inserting (4.11) with $\alpha = 1$ to (4.1) we find that for every positive $\epsilon$ and $E$ sufficiently small the inequality

$$|\operatorname{Im} F^B(E)| \leqslant \epsilon + C(\epsilon)E$$

holds. Hence, we get

$$\lim_{E \to 0} \operatorname{Im} F^B(E) = 0.$$

The conventional parametrizations satisfy the following bound for $0 < E < E_0$:

$$|\operatorname{Im} F^B(E)| \leqslant CE^{\gamma} \ln^n E, \quad 0 < \gamma \leqslant 1, \ n = 0,1,2,\ldots.$$

Then we find

$$|\delta_1(E)| \leqslant 2C \frac{n!}{\pi} \frac{1}{\gamma^{n+1}} + \mathcal{O}\left(\frac{1}{E^2}\right). \tag{5.6}$$

Hence, small values of $\gamma$ allow $\delta_1(E)$ to grow faster than large values of $\gamma$. From this point of view, it is desirable to use such parametrizations of $F^B(E)$ which nearly saturate the Froissart bound. There are reasons to believe that one can meet this requirement by choosing $E_1$ large enough, such that only terms with large values of the exponent $\gamma$ may be used in the parametrization. For instance, this has been carried out by Höhler *et al.*[9] for pion–proton scattering with $E_1 \simeq 10$ GeV.

The term $\delta_2(E)$ can be estimated in the following way. Assuming that $|\Delta(E)/E| \leqslant D$ on the interval $\langle E_0, E_1 \rangle$, where $D$ is a constant, we obtain from (5.4)

$$|\delta_2(E)| \leqslant (2D/\pi)(E_1 - E_0) + \mathcal{O}(1/E). \tag{5.7}$$

While the bound on $\delta_2(E)$ exhibits a moderate growth with increasing $E_1$, that on $\delta_1(E)$ may grow in an uncontrollable way for $E_1$ small enough, due to small values of the exponent $\gamma$ coming into play.

This situation can be illustrated by an example given by Bujak and Dumbrajs,[10] who added the term

$$cE^{\alpha} \tag{5.8}$$

to the parametrization of $\operatorname{Im} F(E)$, choosing $|c|$ so small that no sensible change in the fit to $\operatorname{Im} F(E)$ was produced. The exponent $\alpha$ was taken from the interval

$$0 < \alpha \leqslant 1, \tag{5.9}$$

to preserve the convergence of the series (2.2) and also not to violate the Froissart bound. Because of this term, the real part acquires the term

$$-cE^{\alpha} \cot(\alpha\pi/2), \tag{5.10}$$

which becomes arbitrarily large at $\alpha$ sufficiently near to zero. In this way, a negligible change (5.8) of $\operatorname{Im} F(E)$ may cause an uncontrollable change (5.2) of $\operatorname{Re} F(E)$, if $\alpha$ is given with an experimental error.

Of course, this is no argument against DAR themselves. It only says that DAR should not be applied to a function which, being determined from error-affected data,

lies near to the border of $\mathscr{A}(I)$. Here, the "distance from the border" is measured by the distance of $\alpha$ from the value of 0 beyond which, according to (5.9), DAR do not apply.

Similar problems can also arise in the ordinary dispersion relation. To give an example, let us consider a function defined by

$$\varphi_\alpha(E) = cE(E-a)^\alpha(b-E)^\alpha; \quad b>a>E_0, \ \alpha>0$$

if $E\in(a,b)$ and by

$$\varphi_\alpha(E) = 0 \tag{5.11}$$

elsewhere. When this function is added to $\operatorname{Im} F^D(E)$, it will produce no sensible change of the fit provided that either $|c|$ or $b-a$ are chosen sufficiently small. [In the latter case, $\varphi_\alpha(E)$ is different from zero only between two neighboring energy values at which measurements are carried out.] At any positive $\alpha$, the function $\varphi_\alpha(E)$ satisfies the Hölder condition and the dispersion integral is well defined. Because of $\varphi_\alpha(E)$, the real part $\operatorname{Re} F^D(E)$ acquires a term of the form

$$J_\alpha(E) = \frac{2E^2}{\pi} c \int_a^b \frac{(E'-a)^\alpha(b-E')^\alpha}{E'^2-E^2}\,dE'. \tag{5.12}$$

We shall consider the limit of $J_\alpha(E)$ for $\alpha$ tending to zero. If $E$ equals $a$ or $b$, the leading term of $J_\alpha(E)$ is

$$\frac{ca}{\pi}\frac{1}{\alpha} \tag{5.13}$$

or

$$-\frac{cb}{\pi}\frac{1}{\alpha}, \tag{5.14}$$

respectively. Then, the addition of $\varphi_\alpha(E)$ to $\operatorname{Im} F^D(E)$ introduces an uncontrollable change of the real part. If $E$ differs from $a$ and $b$, the limit is equal to

$$\lim_{\alpha\to 0^+} J_\alpha(E) = \frac{cE}{\pi}\ln\left|\frac{(E+a)(E-b)}{(E-a)(E+b)}\right|. \tag{5.15}$$

The change in $\operatorname{Re} F^D(E)$ is now finite, but it is arbitrarily large near $E=a$ and $E=b$, no matter how small $|c|$ has been chosen.

Of course, the function $\varphi_\alpha(E)$ with $\alpha=0$ is not suitable for a dispersion relation, because it does not satisfy the Hölder condition. But even if $\alpha$ remains positive it can approach zero arbitrarily close without spoiling the quality of the fit to $\operatorname{Im} F^D(E)$. On the other hand, the change in $\operatorname{Re} F^D(E)$ becomes arbitrarily large with $\alpha\to 0$, as follows from (5.13), (5.14), and (5.15), provided that $E$ is sufficiently close to $a$ or $b$.

In this sense, the role of the function $\varphi_\alpha(E)$ in the ordinary dispersion relations is quite analogous to that of the term (5.8) in the derivative analyticity relation. Both (5.8) and (5.11) can easily approach the border of the applicability domain of the respective relation, by choosing $\alpha$ sufficiently close to zero. The main difference is that, while (5.11) is "pathological" at first sight, the term (5.8) belongs to functions which are commonly used to describe experimental data. The main merit of the ordinary dispersion relation remains that they are valid for a much wider class of functions and that the low-energy contribution can be properly taken into account. It is only convention which prevents one from using functions of the type (5.11) in practical applications of dispersion relations.

## VI. CONCLUSION

We have investigated the class of functions for which the tangent series (1.2) is convergent. According to Theorem 1, (1.2) converges if and only if the infinite sum (2.1) converges at the same point. [The sum (2.1) is easier to handle and, moreover, it reveals the connection of the problem of convergence with entire functions.]

Then, we apply this result to an interval $I\subset R^1$, to show that the requirement of the convergence of (1.2) implies a number of necessary conditions to be fulfilled:

(i) the infinite sum (2.1), (2.4), and $\tan[(\pi/2)(d/dx)]f'(x)$ must converge on $I$, and be extensible from $I$ to entire functions;

(ii) both $f(x)$ and (1.2) must be extensible from $I$ to entire functions;

(iii) the entire function which extends $f(x)$ to the complex plane must obey the bound (4.1).

How severe these conditions are in practice can be seen when they are applied to forward scattering; $x=\ln E$, $f(x)=\operatorname{Im} F(E)/E$. If, in addition, the validity of (1.1) is required, both $\operatorname{Im} F(E)$ and $\operatorname{Re} F(E)$ and, by this, the amplitude $F(E)$ itself must be extensible from $I$ to functions entire in $\ln E$.

In Sec. IV, we have analyzed the connection of the convergence of (1.2) with the existence of the corresponding dispersion integral. In Theorems 6 and 7, we derive integral representations of (2.1), (2.4), and (1.2). In this way, a link between DAR and the dispersion relations is established; it follows that the essential difference between them consists in their respective domains of validity. In this sense, the integrals are extensions of the infinite sums to points at which the sums are divergent.

Concerning the practical applicability of DAR, our conclusions do not change the results of the analysis made in Refs. 8 and 9. The applicability is very limited and the predictive power is much lower than that of the dispersion relation. One can point out several reasons for this:

(1) the true amplitude does not satisfy DAR,

(2) the use of DAR requires calculation of higher derivatives of curves which are obtained from experimental data,

(3) some of the functions which are widely used to fit data lie on the border of the class $\mathscr{A}(I)$ and, consequently, may lead to predictions which are unstable against experimental errors. The example of the function $\varphi_\alpha(E)$ defined by (5.11) shows, on the other hand, that such a danger exists also in the case of dispersion relations: but the difference is that functions of the type (5.11) are not used in practice to fit data.

In conclusion, we would like to draw the reader's attention to the interesting possibility of applying the derivative analyticity relations at asymptotic energy. As was shown in Ref. 11, the first term of the infinite sum on the right-hand side of (1.1),

$$\frac{\pi}{2}\left(\alpha-1+\frac{d}{d\ln E}\right)\frac{\operatorname{Im} F(E)}{E^\alpha}, \tag{6.1}$$

correctly gives the leading term of the high-energy behavior of $\operatorname{Re} F(E)/E^\alpha$. This asymptotic relation applies to most of

functions which can be physically interesting. In this way, DAR provide a reasonable approximation at sufficiently high energies and it is not excluded that they will serve as a starting point for an approximation scheme of high-energy scattering.

## ACKNOWLEDGMENTS

The authors acknowledge the stimulating correspondence they had with Professor G. Höhler and are indebted to him for useful remarks on the manuscript. Thanks are also due to Dr. I. Vrkoč for a fruitful collaboration and discussions. One of the authors (P.K.) would like to thank Professor Abdus Salam, the International Atomic Energy Agency, and UNESCO for hospitality at the International Centre for Theoretical Physics, Trieste.

## APPENDIX

For the convenience of the reader we sketch in this Appendix the main points of the proof of Theorem 2. We completely follow Ref. 6, where the details can be found.

Let $I$ be an open interval $I \subset R^1$ and $C^\infty(I)$ denotes the class of functions all derivatives of which exist in $I$. The proof is divided into a series of lemmas. First we present two lemmas on the uniform boundedness of derivatives of $f(x)$. We denote

$$g_n(x) = \sum_{k=0}^{n} f^{(2k+1)}(x), \quad g(x) = \sum_{k=0}^{\infty} f^{(2k+1)}(x).$$

*Lemma 1*: Let $I$ be an interval and $F$ a perfect subset of $I$. Let $g_n$ be continuous functions $g_n: I \to R^1$. If $\lim_{n\to\infty} g_n(x)$ exists at every point $x \in I$, then there exists an open interval $(a,b) \subset I$, $(a,b) \cap F \neq 0$ and a number $M$ so that

$$|g(x)| \leqslant M, \quad |g_n(x)| \leqslant M, \quad x \in (a,b) \cap F,$$

where $g(x) = \lim_{n\to\infty} g_n(x)$.

*Lemma 2*: Let $f \in C^\infty(a,b)$ and let $|f^{(2n+1)}(x)| \leqslant M$ for $x \in (a,b)$, $n = 1,\ldots$. Then $|f^{(2n)}(x)| \leqslant M_1$ for $x \in (a,b)$, $n = 1,\ldots$, where $M_1 = 4M/(b-a) + M(b-a)/4$.

From Lemmas 1 and 2 we conclude that $|f^{(2n+1)}(x)| \leqslant 2M$ and $|f^{(2n)}(x)| \leqslant 2M_1$ on some interval $(a_1,b_1)$.

We define the function

$$f(x,x_0) = \sum_{n=0}^{\infty} f^{(n)}(x_0) \frac{(x-x_0)^n}{n!}.$$

Then we can prove

*Lemma 3*: Let a function $f$ fulfill the condition

$$\sum_{n=0}^{\infty} |f^{(2n+1)}(x)| \quad \text{converges for every } x \in I, \tag{A1}$$

and let an interval $\langle a_0,b_0\rangle \subset I$ be given. There exists an interval $(a_1,b_1)$, $a_0 \leqslant a_1 < b_1 \leqslant b_0$ so that $f(x,x_0)$ is defined for $x \in R^1$, $x_0 \in (a_1,b_1)$. Moreover, $f(x,x_0)$ is an entire function in $x$ if $x_0 \in (a_1,b_1)$ and $f(x) = f(x,x_0)$ for $x,x_0 \in (a_1,b_1)$.

It is straightforward to see that one can ensure the validity of (A1) changing $f(x)$ to $f(\delta x)$, $\delta < 1$. Of course, it does not change the statement of Theorem 2.

*Lemma 4*: Function $f(x,x_1)$ is defined for $x_1 \in \bar{I}_1$ and $f(x) = f(x,x_1)$ is valid for $x,x_1 \in \bar{I}_1$.

Hence $f(x,x_1)$ does not depend on $x_1$ if $x_1 \in \bar{I}_1$ and we can denote $f(x,x_1)$ as $f(x,I_1)$.

If $I_1 \neq I$, there exists a countable set $\mathcal{F}$ of intervals $I_k$ which are maximal in the following sense:

$$f(x) = f(x,I_k), \quad x \in I_k.$$

Obviously $\bar{I}_k \cap \bar{I}_s = \emptyset$ provided $k \neq s$ and the set $\cup_k I_k$ $(I_k \in \mathcal{F})$ is dense in $I$.

*Lemma 5*: Let $f$ be an entire function and $h \in (0,1\rangle$. If

$$|f^{(2n+1)}(0)| \leqslant 1, \quad |f^{(2n+1)}(h)| \leqslant 1,$$

$n = 0,1,\ldots$, and (A1) hold, then

$$|f^{(2n+1)}(x)| \leqslant M_0 \quad \text{for } x \in \langle 0,h\rangle, \quad n = 0,1,\ldots,$$

where $M_0 = 4e(1+e)/3$.

If we assume now, that the set $\mathcal{F}$ contains two intervals $I_1 = (u_1,v_1)$ and $I_2 = (u_2,v_2)$ and introduce the set

$$F = \langle v_1,u_2\rangle - \{\cup_k I_k;\ I_k \in \mathcal{F}\}, \quad v_1 < u_2$$

it can be shown that $F$ is perfect and nowhere dense. Using Lemma 1 to the functions $g_n$ and to the $F$ we can obtain a contradiction. Thus, the following statement is valid.

*Lemma 6*: $\mathcal{F}$ is a one-element set.

Since the set $\cup_k I_k$ is dense in $I$, we conclude that $I_1 = I$ and Theorem 2 is proved.

[1] J. B. Bronzan, Argonne Nat. Lab. Report No. ANL/HEP 7327, 1973; J. B. Bronzan, G. L. Kane, and U. P. Sukhatme, Phys. Lett. B **49**, 272 (1974).
[2] J. Heidrich and E. Kazes, Lett. Nuovo Cimento **12**, 365 (1975).
[3] G. Höhler (private communication).
[4] E. C. Titchmarsh, *The Theory of the Riemann Zeta-Function* (Oxford U.P., London, 1951), p. 177.
[5] E. T. Whittaker and G. N. Watson, *A Course of Modern Analysis* (Cambridge University, Cambridge, England, 1950), 4th ed., p. 17.
[6] I. Vrkoč, "Holmorphic extension of a function whose odd derivatives are summable," Czech. Math. J. (in print).
[7] E. C. Titchmarsh, *The Theory of Functions* (Oxford U.P., London, 1939), 2nd ed., p. 47.
[8] G. K. Eichmann and J. Dronkers, Phys. Lett. B **52**, 428 (1974).
[9] G. Höhler, in *Landolt-Börnstein*, New Series, Group I, Vol. 9b, Part 2, edited by H. Schopper (Springer, Berlin, 1983), p. 41.
[10] A. Bujak and O. Dumbrajs, J. Phys. G **2**, L129 (1976).
[11] J. Fischer and P. Kolář, Phys. Lett. B **64**, 45 (1976); Phys. Rev. D **17**, 2168 (1978).

# On superfield formulation of Euclidean supersymmetry

J. Lukierski[a)]
*International School for Advanced Studies (SISSA), Trieste, Italy*

A. Nowicki
*Institute of Teachers Training and Educational Research, Wrocław, Poland*

(Received 3 October 1983; accepted for publication 30 December 1983)

The superfield formulation of $D = 4$ super-Euclidean QFT is given. The correspondence with Poincaré supersymmetry is explained. The description of supersymmetric Euclidean Yang–Mills theory is presented and the off-shell extension of the topogical charge by fermionic terms is obtained.

PACS numbers: 11.30.Pb, 11.15. − q

## I. INTRODUCTION

The relation between Euclidean and Minkowski QFT is not trivial for fermions because the fundamental O(3,1) spinors are real four-component (Majorana) or complex two-component (Weyl), but the corresponding fundamental O(4) spinors are complex four-component (Dirac) or quaternionic two-component (see, e.g., Ref. 1). Consequently, $D = 4$ Euclidean superalgebra has as the operator basis *four complex* or *two quaternionic* supercharges, and, in order to relate them with $D = 4$ Poincaré supercharges (two complex or four real), one has to explain the cutting by half of the number of fermionic generators.

In this paper we shall discuss first the superspace relations of $D = 4$ Euclidean supersymmetry, and further consider their correspondence with the superspace realization of Poincaré supersymmetry. We shall show that only the Euclidean superfields satisfying Grassmann analyticity conditions (see also Ref. 2) permit the correspondence with Poincaré superfields introduced by Salam and Strathdee.[3] These Grassmann-analytic superfields can be considered as the representations of *non-self-conjugate* $D = 4$ Euclidean superalgebra. As an application of the presented scheme we consider Euclidean sypersymmetric Yang–Mills theory, which should be formulated in Euclidean space, e.g., for the discussion of its topological properties.

We recall here that the correspondence between Euclidean and Minkowski QFT with bosons and fermions was described first by Osterwalder and Schrader (see, e.g., Refs. 4 and 5); the Euclidean formulation of gravity with spinor couplings is due to Hawking and collaborators (see, e.g., Ref. 6). Further the Osterwalder and Schrader approach was applied by Nicolai[7] to the component formulation of $\Phi^3$ supersymmetric theory (self-interacting Wess–Zumino chiral multiplet); Nicolai also first presented in Ref. 8 remarks on the Euclidean superfield formalism with non-self-conjugate (non-Hermitian) superfields. Our aim here is to show how the choice of non-Hermitian Grassman analytic superfields follows from the structure of $D = 4$ Euclidean superalgebra and its superfield representations. However, particular realization of Euclidean supersymmetry on field multiplet was given by Zumino in Ref. 9 some time ago; the Euclidean superfield formulation as derived from $D = 4$ Euclidean superalgebra has not been studied.

## II. EUCLIDEAN NON-SELF-CONJUGATE $D = 4$ EUCLIDEAN SUPERALGEBRAS

The lowest-dimensional faithful SO(4) spinor representations is described by two independent SU(2) spinors. We choose these two spinors as Grassmann coordinates in $D = 4$ Euclidean superspace, and denote them by $\theta_{\alpha;}$ and $\theta_{;\alpha}$ ($\alpha = 1,2$). They transform under $SO(4) = SU(2) \times SU(2)$ as follows:

$$\theta_{\alpha;} = [\exp(\alpha_i^{+} e^i)]_{\alpha}{}^{\beta;} \theta_{\beta;},$$
$$\theta_{i\alpha} = [\exp(\alpha_i^{-} e^i)]_{;\alpha}{}^{\beta} \theta_{i\beta}, \tag{1}$$

where $e^i{}_{\alpha}{}^{\beta;} = e^i{}_{;\alpha}{}^{\beta} = i\sigma_i^{\tau}$. Three real parameters $\alpha_i^{\pm}$ ($i = 1,2,3$) are expressed in terms of O(4) rotations $\alpha_{\mu\nu}$ ($\mu,\nu = 0,1,2,3$) as follows:

$$\alpha_i^{\pm} = \tfrac{1}{2}\epsilon_{ijk}\alpha_{jk} \pm \alpha_{i0}. \tag{2a}$$

The parametrization (2a) corresponds to the well-known decomposition of the O(4) algebra

$$M_{\mu\nu} \to M_i{}^{\pm} = \tfrac{1}{2}\epsilon_{ijk} M_{jk} \pm M_{i0}. \tag{2b}$$

We see that under the time reversal operation $T$ changing the sign of $M_{i0}$, we get $M_i{}^{\pm} \to M_i^{\mp}$.

In order to construct from the SO(4) spinors the vector-like objects, we need also the complex-conjugated spinors $\theta_{;\dot\alpha} = (\theta_{;\alpha})^*$, $\theta_{\dot\alpha;} = (\theta_{\alpha;})^*$ with the transformation laws

$$\theta_{\alpha;} = [\exp(\alpha_i^{+} e^i)]_{\dot\alpha}{}^{\beta;} \theta_{\beta;},$$
$$\theta_{;\dot\alpha} = [\exp(\alpha_i^{-} e^i)]_{;\dot\alpha}{}^{\beta} \theta_{;\beta}, \tag{3}$$

where $e^i{}_{\dot\alpha}{}^{\beta;} = (e^i{}_{\alpha}{}^{\beta;})^* = (e^i{}_{;\dot\alpha}{}^{\beta})^* = e^i{}_{;\dot\alpha}{}^{\beta} = -i\sigma_i$. The real O(4)-vector $X_\mu$ transforms as the mixed $(\frac{1}{2},\frac{1}{2})$ spinor $X^{\dot\alpha;\beta} = (X^{\alpha;\dot\beta})^*$ due to the relation

$$X^{\dot\alpha;\beta} = \tfrac{1}{2} e_\mu^{\dot\alpha;\beta} X_\mu, \tag{4}$$

where $e_\mu^{\dot\alpha;\beta} = (e^{\alpha;\dot\beta})^* = (-i\sigma_i, I_2)$.

The lowest-dimensional self-conjugate (real) $D = 4$ Euclidean superspace looks as follows:

$$S = (x_\mu, \theta_{\alpha;}, \theta_{;\alpha}, \theta_{\dot\alpha;}, \theta_{;\dot\alpha}). \tag{5}$$

In order to obtain the closed Euclidean supergroup law, one can consider only non-self-conjugate Grassmann-analytic chiral superspaces

$$S^{+} = (x_\mu, \theta_{\alpha;}, \theta_{;\dot\alpha}), \quad S^{-} = (x_\mu, \theta_{;\alpha}, \theta_{;\dot\alpha}), \tag{6}$$

[a)] On leave of absence from the Institute for Theoretical Physics, University of Wrocław, ul. Cybulskiego 36, 50-205 Wrocław, Poland.

 0022-2488/84/082545-05$02.50  

where $S = S^+ \cap S^-$. For example, the Euclidean supergroup transformation law in $S^+$ takes the form:

$$\begin{aligned} x'_\mu &= x_\mu + a_\mu - \alpha_{\mu\nu}x_\nu + \tfrac{1}{2}(\epsilon_{\dot\alpha;}e_\mu^{\dot\alpha;\beta}\theta_{\beta;} - \theta_{\dot\alpha;}e_\mu^{\dot\alpha;\beta}\theta_{\beta;}), \\ \theta'_{\dot\alpha;} &= \theta_{\dot\alpha;} + \epsilon_{\dot\alpha;} + \alpha_i^{+}e^{i}{}_{\dot\alpha}{}^{\dot\beta;}\theta_{\dot\beta;}, \\ \theta'_{;\alpha} &= \theta_{;\alpha} + \epsilon_{;\alpha} + \alpha_i^{-}e^{i}{}_{;\alpha}{}^{\beta}\theta_{\beta;}. \end{aligned} \tag{7}$$

The generators of four complex supertranslations

$$\begin{aligned} \epsilon_{\dot\alpha;}:\quad Q^{\dot\alpha;} &= \frac{\partial}{\partial\theta_{\dot\alpha;}} + \frac{1}{2}e_\mu^{\dot\alpha;\beta}\theta_{\beta;}\frac{\partial}{\partial x_\mu}, \\ \epsilon_{;\alpha}:\quad Q^{;\alpha} &= \frac{\partial}{\partial\theta_{;\alpha}} + \frac{1}{2}\theta_{\dot\beta;}e^{\dot\beta;\alpha}\frac{\partial}{\partial x_\mu} \end{aligned} \tag{8}$$

represents the following *non-self-conjugate* $D = 4$ Euclidean superalgebra $\mathscr{E}^+$:

$$\{Q^{\dot\alpha;},Q^{\dot\beta;}\} = \{Q^{;\alpha},Q^{;\beta}\} = 0, \tag{9a}$$

$$\{Q^{\dot\alpha;},Q^{;\beta}\} = e_\mu^{\dot\alpha;\beta}P_\mu, \tag{9b}$$

$$[Q^{\dot\alpha;},P_\mu] = [Q^{;\alpha},P_\mu] = 0, \tag{9c}$$

$$[Q^{\dot\alpha;},M_i^{+}] = e_i{}^{\dot\alpha}{}_{\dot\beta;}Q^{\dot\beta;}, \quad [Q^{\dot\alpha;},M_i^{-}] = 0, \tag{9d}$$

$$[Q^{;\alpha},M_i^{-}] = e_i{}^{;\alpha}{}_{\beta}Q^{;\beta}, \quad [Q^{;\alpha},M_i^{+}] = 0, \tag{9e}$$

where $P_\mu = \partial/\partial x_\mu$ and $M_i^{\pm}$ describe SO(4) generators [see (4)]. The superalgebra (9) is realized in Grassmann analytic superspace $S^+ = (x_\mu,\theta_{\dot\alpha;},\theta_{;\alpha})$. The supertranslations of $x_\mu$ are not real [see (7)]; one can assume that the space-time coordinates are complex, i.e., $x_\mu \to z_\mu = x_\mu + iy_\mu$. One gets $S^+ \to \tilde{S}^+ = (z_\mu,\theta_{\dot\alpha;},\theta_{;\alpha})$ and $P_\mu = \partial/\partial z_\mu = \partial/\partial x_\mu + i\,\partial/\partial y_\mu$. Because we are not restricted by the hermiticity of $P_\mu$, one can extend also $SU(2)\times SU(2) \to SL(2,c)\times SL(2,c) \simeq O(4;c)$, which is obtained in (7) by the complexification of the parameters $\alpha_i^{+},\alpha_i^{-}$.

The relation (9b) with $Q_{\dot\alpha;}$ and $Q^{;\beta}$ transforming under two commuting SL(2,c) groups describes as special cases both $d = 4$ Euclidean and super-Poincaré algebra. Indeed, putting $z_i = x_i$, $z_\mu = ix_0$, one gets

$$\{Q^{\dot\alpha;},Q^{;\beta}\} = \sigma_\mu{}^{\alpha;\beta}\mathscr{P}^\mu, \tag{10}$$

where $\mathscr{P}_\mu = (1/i)\partial/\partial x_\mu$ and $\sigma_\mu = (\sigma_i,1)$. In order to obtain $\mathscr{P}_\mu$ as Minkowski 4-vectors, we should identify the parameters of two independent SL(2,c) groups (i.e., $Q^{\dot\alpha;} \equiv Q^{\dot\alpha}$ and $Q^{;\beta} \equiv Q^{\beta}$) and postulate the Majorana condition $[Q^{\dot\alpha} = (Q^{\alpha})^*]$. In such a way relation (10) becomes the fundamental relation defining $N = 1$ super-Poincaré algebra.

Finally we would like to mention that the $D = 4$ Euclidean formalism also after supersymmetric extension is described in a compact way by quaternions. Indeed, because $SU(2) \simeq Sp(1)$, on the list of fundamental spinor representations of orthogonal groups the $D = 4$ Euclidean spinors are described by two-dimensional quaternionic module $H \oplus H$ (see, e.g., Ref. 1). Because the squares of quaternion-valued Grassmann variables are not zero, the superfield expansion in quaternionic superspace is more complicated than in real and complex case. For completeness the short discussion of $D = 4$ Euclidean superspace in quaternionic language is given in the Appendix.

## III. EUCLIDEAN SELF-CONJUGATE $D = 4$ SUPERALGEBRA

It is easy to see that by complex conjugation of (7) we obtain the transformations in *conjugate* superspaces $S^- = (x_\mu,\theta_{\alpha;},\theta_{;\dot\beta})$, generated by conjugated superalgebra $\mathscr{E}^-$. In order to obtain the self-conjugate (Hermitian) realizations of Euclidean supersymmetry, one should consider the joint realizations of $\mathscr{E} = \mathscr{E}^+ \cap \mathscr{E}^-$ on self-conjugate (Hermitian) superspaces or $S = S^+ \cap S^- = (x_\mu,\theta_{\alpha;},\theta_{\dot\alpha;},\theta_{;\alpha},\theta_{;\dot\alpha})$.[10] An example of such a realization is given in Ref. 8. Introducing anticommuting Dirac spinors

$$\epsilon_A = \begin{pmatrix} \epsilon_{\alpha;} \\ \epsilon_{;\dot\alpha} \end{pmatrix}, \quad \theta_A = \begin{pmatrix} \theta_{\alpha;} \\ \theta_{;\dot\alpha} \end{pmatrix} \quad (A = 1,\ldots,4),$$

the transformation laws of the Hermitian superspace coordinate $S = (x_\mu,\theta_A,\theta_A^*)$ look as follows:

$$\begin{aligned} x'_\mu &= x_\mu + q_\mu - \tfrac{1}{2}\alpha_{\mu\nu}x_\nu + \tfrac{1}{2}(\bar\epsilon_A\Gamma_{\mu AB}\theta_B - \bar\theta_A\Gamma_{\mu AB}\epsilon_B), \\ \theta'_A &= \theta_A + \epsilon_A + \alpha_{\mu\nu}\Sigma_{\mu\nu}^{AB}\theta_B, \\ \bar\theta'_A &= \bar\theta_A + \bar\epsilon_A + \alpha_{\mu\nu}\bar\Sigma_{\mu\nu}^{AB}\bar\theta_B, \end{aligned} \tag{11}$$

where $\Gamma_{\mu AB}$ and $\Sigma_{\mu\nu}^{AB}$ are obtained from (1)–(2) by using Pauli matrix representations of quaternionic imaginary units. The generators of supertranslations (we denote here by bar complex conjugation)

$$\epsilon_A:\quad \bar Q^A = \frac{\partial}{\partial\theta_A} + \frac{1}{2}\bar\Gamma_\mu{}^{AB}\bar\theta_B\frac{\partial}{\partial x_\mu}, \tag{12a}$$

$$\bar\epsilon_A:\quad Q^A = \frac{\partial}{\partial\bar\theta_A} + \frac{1}{2}\Gamma_\mu{}^{AB}\theta_B\frac{\partial}{\partial x_\mu} \tag{12b}$$

form the *complex self-conjugate* $D = 4$ Euclidean superalgebra

$$\begin{aligned} &\{Q^A,Q^B\} = \{\bar Q^A,\bar Q^B\} = 0, \\ &\{\bar Q^A,Q^B\} = \Gamma_\mu{}^{AB}P_\mu, \\ &[P_\mu,Q^A] = [P_\mu,\bar Q^A] = 0, \\ &[M_{\mu\nu},Q^A] = \Sigma_{\mu\nu}^{AB}Q_B\,[M_{\mu\nu},\bar Q^B] = \bar\Sigma_{\mu\nu}^{AB}\bar Q_B \end{aligned} \tag{13}$$

supplemented by known ISO(4) Lie algebra.

In order to introduce a conjugation which maps superspaces $H^{\pm}$ or $S^{\pm}$ into each other, one should supplement the Hermitian conjugation with time reversal $\hat T$, which implies

$$\hat T q\hat T^{-1} = -\bar q, \quad \hat T\theta_1\hat T^{-1} = \theta_2, \quad \hat T\theta_2\hat T^{-1} = \theta_1. \tag{14}$$

A product of Hermitian conjugation and time reversal was introduced to Euclidean QFT in Ref. 4 and we shall call it, following Ref. 8, the Osterwalder–Schrader (OS) conjugation. In Euclidean supersymmetry OS conjugation replaces Hermitian conjugation; in particular OS-self-conjugate Euclidean superfields become Hermitian *after* prolongation to Minkowski space and imposing the Majorana conditions. For example, complex Grassmann analytic $\mathscr{E}^+$-superfield

$$\begin{aligned} V(x,\theta_{\dot\alpha;},\theta_{;\alpha}) = {} & c(x) + \psi_{\dot\alpha;}(x)\theta^{\dot\alpha;} + \psi_{;\alpha}(x)\theta^{;\alpha} \\ & + \theta_{\dot\alpha;}\alpha^{\dot\alpha;}F(x) + \theta_{;\alpha}\theta^{;\alpha}G(x) \\ & + \theta_{\dot\alpha;}\theta^{\dot\alpha;}\theta_{;\alpha}\lambda^{;\alpha}(x) \\ & + \theta_{;\alpha}\theta^{;\alpha}\theta_{\dot\alpha;}\lambda^{\dot\alpha;}(x) + \theta_{\dot\alpha;}e_\mu^{\dot\alpha;\beta}\theta_{;\beta}V_\mu(x) \\ & + \theta_{\dot\alpha;}\theta^{\dot\alpha;}\theta_{;\alpha}\theta^{;\alpha}D(x) \end{aligned} \tag{15}$$

after taking into consideration the time reversal for superspace coordinates as well as field components is restricted by OS self-conjugation as follows:

$$C = C^*, \quad F = G^*, \quad V_\mu = V_\mu^*, \quad D = D^*. \tag{16}$$

We see therefore that we obtain only doubling of spin $\frac{1}{2}$ (in general, half-integer spin) field components, which we remove in the procedure of continuation to Minkowski space by imposing the Majorana conditions

$$\psi_{\alpha;} = \psi_{;\alpha} = \psi_\alpha, \quad \lambda_{\alpha;} = \lambda_{;\alpha} = \lambda_\alpha.$$

The complex scalar superfield does not depend on the complex Grassmann coordinates $\theta_{\alpha;}, \theta_{;\dot\alpha}$, i.e., it is Grassmann-analytic in $\theta_{;\alpha}$ and Grassmann-anti-analytic in $\theta_{\alpha;}$. This superfield can be further restricted by Euclidean chiral conditions to the Euclidean chiral fields $\Phi_1(x,\theta_{\alpha;})$ or $\Phi_2(x,\theta_{\dot\beta;})$, which are OS-conjugated to each other and represent faithfully only respective O(3) subgroups of the full group O(4) = O(3)×O(3) of Euclidean rotations. In order to write down the chirality conditions in $\mathscr{E}^+$ it is convenient to use the chiral realizations of the superalgebra (9b). For example,

$$Q^{\dot\alpha;} = \frac{\partial}{\partial\theta_{\dot\alpha;}} + e_\mu^{\dot\alpha;\beta}\theta_{;\beta}\frac{\partial}{\partial x_\mu}, \quad Q^{;\alpha} = \frac{\partial}{\partial\theta_{;\alpha}}, \tag{17}$$

the covariant derivatives have the form

$$D_{\dot\alpha;} = \frac{\partial}{\partial\theta_{\dot\alpha;}}, \quad D^{;\alpha} = \frac{\partial}{\partial\theta_{;\alpha}} - \theta_{\dot\beta;}e_\mu^{\dot\beta;\alpha}\frac{\partial}{\partial x_\mu} \tag{18}$$

and the chirality condition $D^{\dot\alpha;}\Phi(x_\mu,\theta_{\dot\alpha;},\theta_{;\beta}) = 0$ implies $\Phi = \Phi_1(x_\mu,\theta_{;\beta})$.

## IV. AN EXAMPLE: $D = 4$ EUCLIDEAN SUPERSYMMETRIC YM THEORY

As an application we shall present the superfield formulation of Euclidean $D = 4$ Yang–Mills theory. The Euclidean gauge superfield is described by the $\mathscr{E}^+$-superfield (15) with the reality conditions (16) and the components taking values in the matrix algebra of the adjoint representation of internal symmetry $G$. The gauge-covariant chiral derivative have the standard form (see, e.g., Ref. 11)

$$\nabla^{;\alpha} = e^{-V}D^{;\alpha}e^V, \quad \nabla^{\dot\alpha;} = D^{\dot\alpha;} \tag{19}$$

and the gauge-covariant chiral field strength superfield is defined as follows

$$F^{\dot\alpha\dot\beta;\beta} = 2i[\nabla^{\dot\alpha;}\{\nabla^{\dot\beta;},\nabla^{;\beta}\}] = 2i\epsilon^{\alpha\dot\beta;}W^{;\beta}. \tag{20}$$

If we observe that the following coupling to chiral fields $\Phi_1,\Phi_2$, is OS-self-conjugate,

$$\mathscr{L}_{\text{int}} = \int d^2\theta_{\dot\alpha;}\, d^2\theta_{;\alpha}\, \Phi_1 e^{-V}\Phi_2. \tag{21}$$

We see that the supergauge freedom depends on the *two* chiral superfield parameters $\Lambda_1(x,\theta_{\dot\alpha;})$ and $\Lambda_2(x,\theta_{;\alpha})$, and Euclidean Wess–Zummino gauge can be defined by putting in (15) $C = \Psi_{\dot\alpha;} = \Psi_{;\alpha} = F = G = 0$.

In such a gauge

$$W^{;\alpha}(x,\theta_{;\alpha}) = -\lambda^{;\alpha} - 2D\theta^{;\alpha} + \tfrac{1}{4}\eta_{k,\rho\sigma}e_k^{;\alpha\beta} \times F_{\rho\sigma}\theta_{;\beta} - \tfrac{1}{2}e_\mu^{\dot\beta;\alpha}(\partial_\mu\lambda_{\dot\beta;} + [\lambda_{\dot\beta;},V_\mu])\theta_{;\gamma}\theta^{;\gamma}, \tag{22}$$

where $F_{\mu\nu} = \partial_\mu V_\nu - \partial_\nu V_\mu - [V_\mu,V_\nu]$ and $\eta_{k,\rho\sigma} = -\eta_{k,\sigma\rho}$ $(k = 1,2,3)$ denotes t'Hooft Euclidean self-dual tensor, satisfying the relation

$$\tfrac{1}{2}\epsilon_{\mu\nu\rho\tau}\eta_{k,\rho\tau} = \eta_{k,\mu\nu}. \tag{23}$$

The superfield (22) describes only the self-dual part $F_{\mu\nu}^{(+)} = \frac{1}{2}(F_{\mu\nu} + \epsilon_{\mu\nu\rho\tau}F_{\rho\tau})$ of YM field strength, and the supersymmetric extension of the self-dual part of YM action looks as follows:

$$S^{(+)} = \int d^4x\, d\theta_{;1}\, d\theta_{;2}\, \mathrm{Tr}(W_{;\alpha}W^{;\alpha}) = \int d^4x\, \mathrm{Tr}(\tfrac{1}{4}F_{\mu\nu}^{(+)}F_{\mu\nu}^{(+)} + \lambda_{;\alpha}e_\mu^{\dot\beta;\alpha}D_\mu\lambda_{;\dot\beta} + 4D^2), \tag{24}$$

where $D_\mu$ is the covariant derivative for adjoint representation. The anti-self-dual sypersymmetrically extended action can be obtained by introducing in analogous way the anti-self-dual superfield strength $W^{\dot\beta;}(x,\theta_{\dot\beta;})$. One gets

$$S^{(-)} = \int d^4x\, \mathrm{Tr}\,(\tfrac{1}{4}F_{\mu\nu}^{(-)}F_{\mu\nu}^{(-)} + \lambda_{\dot\alpha;}e_\mu^{\dot\alpha;\beta}D_\mu\lambda_{;\beta} + 4D^2). \tag{25}$$

The Euclidean action is given by $S = S^{(+)} + S^{(-)}$ and the supersymmetrized topological charge is

$$\begin{aligned} Q = S^+ - S^- &= \int d^4x(\tfrac{1}{4}\epsilon_{\mu\nu\rho\tau}F_{\mu\nu}F_{\rho\tau} \\ &\quad + \lambda_{;\alpha}e^{\dot\beta;\alpha}{}_\mu D_\mu\lambda_{\dot\beta;} - \lambda_{\dot\alpha;}e_\mu{}^{\dot\alpha;\beta}D_\mu\lambda_{;\beta}) \\ &= \int d^4x\frac{\partial}{\partial x_\mu}[\epsilon_{\mu\nu\rho\tau}A^k{}_\nu(F^k_{\rho\tau} \\ &\quad - \tfrac{1}{3}C^K_{LM}A^L_\rho A^M_\tau) + 2\lambda_{;\alpha}e_\mu^{\dot\beta;\alpha}\lambda_{\dot\beta;}], \end{aligned} \tag{26}$$

where we used the Grassmann property of $\lambda_{;\alpha},\lambda^{\dot\beta;}$ and the antisymmetry of the internal symmetry generators in the adjoint representation. We see that the topological current for YM field is extended by a fermionic part, which is even in Grassmann variables and vanishes on-shell.

It should be mentioned that because one can write $2S^{(+)} = S + Q$, the formula (24) (multiplied by factor 2) differs only by the 4-divergence from the complete Euclidean action for SS YM theory.

If we pass to Minkowski field theory, the duality operation [see (23)] will have the purely imaginary eigenvalues $\pm i$, and $F_{\mu\nu}^{(+)} \to \tilde F_{\mu\nu}^{(+)} = \frac{1}{2}(F_{\mu\nu} + i\epsilon_{\mu\nu\rho\tau}F_{\rho\tau})$. The contribution from the topological charge to (24) becomes purely imaginary, and the Majorana condition is $\lambda^{;\alpha} = \lambda^\alpha = (\lambda^{\alpha;})^* = \lambda^{\alpha;}$. In Minkowski space one should use rather as the SS YM action the sum of (24) and (25), because the Minkowski fields do not vanish in timelike directions and the meaning of $Q$ as the topological charge in Minkowski space is very obscure.

## V. REMARKS

We would like also to add the following.

(a) However the Euclidean superfield (15) contains complex spinors in place of Majorana spinors in corresponding Minkowski superfield, the matching of fermionic and bosonic degrees of freedom is preserved, due to Grassmann analiticity conditions. This conclusion is in consistency with the functional description of Euclidean QFT in Ref. 6, be-

cause complex variable $z$ can be counted twice only if it is integrated in functional integral over $dz$ and $d\bar{z}$.

(b) Our conclusion is that the topological charge in SS YM theory is an even Grassmann number. Such a topological charge has only meaning if we consider the functional averages over Grassmann variables via Berezin integration, i.e., after quantization of fermionic degrees of freedom.

We suspect that the result found in supergravity that the Grassmann contributions to topological charges cancel (see, e.g., Ref. 12) is not valid in the Euclidean version of supergravity. This problem is under consideration.

(c) The Hermitian superalgebra (13) is not modified if we introduce the following transformations of the four complex supercharges $Q^A$:

(1) noncompact chiral transformations

$$Q' = \alpha Q + \beta \Gamma_5 Q, \quad \alpha^2 - \beta^2 = 1; \tag{27a}$$

(2) compact SU(2) transformations

$$Q' = aQ + bC\bar{Q}, \quad |a|^2 + |b|^2 = 1, \tag{27b}$$

where $C\Gamma_\mu^T C = -\Gamma_\mu$ [in the representation (1) one should choose $C = \left(\begin{smallmatrix} e_2 & 0 \\ 0 & e_2 \end{smallmatrix}\right)$ where $e_2 = i\sigma_2$].

If we consider the non-self-conjugate superalgebra $\mathscr{E}^+$ [see (9)], from the transformations (27a–b) only remains the U(1) group described by $|a| = 1$,

$$b = 0 \,(Q'_{\alpha;} = e^{i\alpha} Q_{\alpha;}, Q'_{;\alpha} = e^{i\alpha} Q_{;\alpha}).$$

The freedom described by transformations (27a–b) has been observed firstly in Ref. 9, where the Hermitian realizations were studied.

(d) The description of $D = 4$ Euclidean supersymmetry by non-self-conjugate superalgebras and Grassman-analytic superfields cannot be extended to the Euclidean supersymmetry with cosmological constant, described by the quaternionic graded supergroup Osp(1,2;$H$).[13] The superalgebra Osp(1,2;$H$) with 11 bosonic generators of O(5) $\oplus$ O(2) lie algebra [or O(4,1) $\oplus$ O(2) if we replace Sp(2;$H$)$\rightarrow$Sp(1,1,$H$) in the quaternionic supergroup[14]] and eight fermionic supercharges can be obtained only as the deformation of the Hermitian Euclidean superalgebra.

(e) The Euclidean superfield formulation of SS YM theory permits the supersymmetric generalization of $D = 4$ quaternionic HP($n$) $\sigma$-model,[13,15] which is related to the ADHM parametrization of SU(2) instanton solutions. The compositeness condition of quaternionic as well as complex $D = 4$ Euclidean $\sigma$-superfields will be discussed elsewhere.

## ACKNOWLEDGMENTS

One author (J.L.) would like to thank Professor Paolo Budini and Professor Luciano Fonda for their hospitality at the International School for Advanced Studies (SISSA) as well as for financial support. The author mentioned above also profited from the discussion and creative atmosphere at the International Centre for Theoretical Physics (ICTP), and would like to thank P. van Nieuwenhuizen for valuable comments.

## APPENDIX

Four-dimensional Euclidean framework favors the use of quaternions because:

—four real coordinates $x_\mu$ are described by real quaternions $q = x_\mu e_\mu$ (see e.g., Ref. 16), where $e_\mu = (e_i, 1)$ ($\mu = 1,...,4$; $i = 1,2,3$), and $e_i e_j = -\delta_{ij} + \epsilon_{ijk} e_k$;

—fundamental O(4) spinor representation is described by two quaternions $\theta_1, \theta_2$,[1] carrying the representation of Sp(1)$\times$Sp(1).[17]

The quaternion-valued Euclidean Dirac matrices, satisfying the relation $\{\Gamma_\mu, \Gamma_\nu\} = 2\delta_{\mu\nu}$, can be chosen as follows:

$$\Gamma_\mu = \begin{pmatrix} 0 & \bar{e}_\mu \\ e_\mu & 0 \end{pmatrix}, \quad e_\mu = (e_i, 1), \ \bar{e}_\mu = (-e_i, 1) \tag{A1}$$

The generators of the covering group Sp(1)$\times$Sp(1) of SO(4)

$$\Sigma_{\mu\nu} = -\tfrac{1}{4}[\Gamma_\mu, \Gamma_\nu] = \begin{pmatrix} e_{\mu\nu} & 0 \\ 0 & \tilde{e}_{\mu\nu} \end{pmatrix},$$
$$e_{\mu\nu} = \tfrac{1}{4}(\bar{e}_\nu e_\mu - \bar{e}_\mu e_\nu), \quad \tilde{e}_{\mu\nu} = \tfrac{1}{4}(e_\nu \bar{e}_\mu - e_\mu \bar{e}_\nu) \tag{A2}$$

define the group elements of Sp(1)$\times$Sp(1)

$$U(\alpha) = \exp \alpha_{\mu\nu} \Sigma_{\mu\nu} = \begin{pmatrix} A^+ & 0 \\ 0 & A^- \end{pmatrix}, \quad A^\pm = \exp \alpha_i^\pm e_i. \tag{A3}$$

The SO(4) transformation laws of $x_\mu$ and $\theta_1, \theta_2$ look as follows:

$$X' = UX\bar{U}, \quad \theta' = U\theta, \tag{A4}$$

where $X = x_\mu \Gamma_\mu$ and $\theta = \left(\begin{smallmatrix} \theta_1 \\ \theta_2 \end{smallmatrix}\right)$. The action of Euclidean group of motions ISO(4) is given by the quaternionic formula $q' = A^- q \bar{A}^+ + a$, where the quaternion $a = a_\mu e_\mu$ describe four Euclidean translations.

In order to extend ISO(4) in a supersymmetric way, we assume that $\theta_1, \theta_2$ are *Grassmann-valued* quaternions $\theta_r = \theta_{r\mu} e_\mu$, with $\{\theta_{r\mu}, \theta_{sv}\} = 0$ ($r,s = 1,2$). We introduce the infinitesimal transformations of super Euclidean group as follows:

$$\begin{aligned} q' &= q + a + \alpha_{\mu\nu}(\bar{e}_{\mu\nu} x + x \bar{e}_{\mu\nu}) + \epsilon_2 \bar{\theta}_1 - \theta_2 \bar{\epsilon}_1, \\ \bar{\theta}'_1 &= \bar{\theta}_1 + \bar{\epsilon}_1 + \alpha_{\mu\nu} \bar{\theta}_1 \bar{e}_{\mu\nu}, \\ \theta'_2 &= \theta_2 + \epsilon_2 + \alpha_{\mu\nu} \tilde{e}_{\mu\nu} \theta_2, \end{aligned} \tag{A5}$$

where $\epsilon_r = \epsilon_{r\mu} e_\mu$ and $\theta_{r\mu}, \epsilon_{sv}$ form the basis of 16-dimensional Grassmann algebra. The transformation laws in quaternionic non-self-conjugate superspace $H^+ = (q, \bar{\theta}_1, \theta_2)$ can be described also in complex parametrization, if we introduce the Pauli matrix realization of quaternionic imaginary units $e_i = i\sigma_i$[18] and describe the Grassmann-valued quaternions by means of complex Grassmann variables. If we introduce complex Grassmann coordinates $(\theta_{\alpha;}, \theta_{;\beta})$ and $(\theta_{\dot{\alpha};}, \theta_{;\dot{\beta}})$ where $\alpha, \beta = 1,2$, $\theta_{\dot{\alpha};} = (\theta_{\alpha;})^*$, $\theta_{;\dot{\beta}} = (\theta_{;\beta})^*$ and $\theta_1 = \theta_{1;} + e_2 \theta_{2;}$, $\theta_2 = \theta_{;1} + e_2 \theta_{;2}$, $\bar{\theta}_1 = \theta_{\dot{1};} - e_2 \theta_2$, $\bar{\theta}_2 = \theta_{;\dot{1}} - e_2 \theta_{;2}$,[19] the transformation law (A5) can be written in the form given by Eq. (7).

[1] M. F. Atiyah, R. Bott, and A. Shapiro, Topology **3**, 1 (1964).
[2] A. Galperin, E. Ivanov, and V. Ogievetski, Pis'ma Zh. Eksp. Teor. Fiz. **33**, 176 (1981).
[3] A. Salam and J. Strathdee, Phys. Rev. L **11**, 1521 (1975).
[4] K. Osterwalder and R.Schrader, Commun. Math. Phys. **31**, 83 (1973).
[5] K. Osterwalder and R. Schrader, Helv. Phys. Acta **46**, 277 (1973).
[6] S. W. Hawking Cargese Lecture Notes, Cargese School, 1978, p. 145.
[7] H. Nicolai, Nucl. Phys. B **140**, 294 (1978).
[8] H. Nicolai, Czech. J. Phys. B **29**, 308 (1979).

[9]B. Zumino, Phys. Lett. B **69**, 369 (1977).
[10]If we use complex coordinates $Z_\mu$, one gets $\tilde{S} = \tilde{S}^+ \oplus \tilde{S}^-$.
[11]M. Roček, in *Supergravity*, Proceedings of Nuttfield Workshop, July 1981, edited by S. W. Hawking and M. Rocek.
[12]P. Van Nieuwenhuizen and P. K. Townsend, Phys. Rev. D **19**, 3592 (1979).
[13]J. Lukierski, CERN Preprint TH 2678, published in *Proceedings of Summer Institute Kaiserslautern*, August 1979, edited by W. Rühl (Springer-Verlag, Berlin, 1980), p. 361.
[14]J. Lukierski and A.Nowicki, Forschr. Phys. **30**, 75 (1982).
[15]F. Gürsey and H. C. Tze, Ann. Phys. (N.Y.) **128**, 29 (1980).
[16]F. Gürsey, *IInd John Hopkins Workshop*, edited by G. Domokos and S. Kövesi-Domokos (John Hopkins U.P., Baltimore, 1978), p. 179.
[17]One can represent quaternions by pairs of complex numbers, but such two-dimensional complex space will carry additional quaternionic structure not allowing the multiplication by an overall phase factor.
[18]The quaternion algebra can be realized by the choices $e_i = i\sigma_i$ or $e_i = -i\sigma_i^T$, where $\sigma_1 = \left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)$, $\sigma_2 = \left(\begin{smallmatrix} 0 & i \\ -i & 0 \end{smallmatrix}\right)$, $\sigma_3 = \left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right)$.
[19]We see that $e_3$ plays the role of imaginary unit. Similar formulas one should write for $e_1, e_2$.

# The cohomology of the generalized Lie algebras

Bani Mitra and K. C. Tripathy
*Department of Physics and Astrophysics, University of Delhi, Delhi-110007, India*

(Received 29 June 1983; accepted for publication 30 December 1983)

In this paper, we develop the cohomology theory of the generalized Lie algebras. Using the identification of $H^2_\varphi(\mathbf{g},V)$ with $\mathrm{Ext}_F(\mathbf{g},\mathbf{h})$, we discuss some of the extended Lie algebras useful in physics, for instance the supersymmetry, quantized field, and su(2) ⊕ u(1) algebras. Finally, we show that for a certain class of representations of some generalized Lie algebras, the cohomology is trivial; in so doing, we generalize the result obtained in our earlier investigation.

PACS numbers: 11.30.Pb, 02.20.Sv, 02.40. + m

## I. INTRODUCTION

Supersymmetry, a powerful tool in particle physics,[1] is based on a $Z_2$-graded algebra, also called a superalgebra. Recently a class of generalized Lie algebras (GLA) have been described[2–5] which include the superalgebras as special cases; it is hoped that they may have a bearing on parastatistics[3] or provide new possibilities for mixing internal symmetry algebras with the Poincaré algebra.[6] It is our hope that some of these problems may be answered by cohomology theory, which has been of interest to physicists in a number of contexts, for instance, in the classification of principal bundles (Chern class theory), in the solution of field equations in nonabelian gauge theories (de Rham theory[7,8]), and in the classification of Lie algebra extensions and Lie group algebra contraction schemes.[9–12]

In this paper, we formulate the algebraic cohomology of the generalized Lie algebras. Introducing the concept of extension of a generalized Lie algebra, we relate it to the second-order cohomology group. We then apply the result to the various interesting algebras on hand, viz., the supersymmetry algebra, the Weinberg–Salam symmetry algebra for weak interactions su(2) ⊕ u(1), and the algebra of a quantized field, focusing attention on the uniqueness of the extension of the algebra by other elements, and the representation dependence of the extensions of these algebras. Finally we digress to the computation of the cohomology of the generalized Lie algebras showing that for a major class of representations of some nilpotent or solvable generalized Lie algebras (nilpotent or solvable in the sense of Ref. 12), the cohomology is trivial. This generalizes the result obtained in Ref. 13.

The plan of the paper is as follows. In Sec. II, we develop our basic concepts of the generalized Lie algebras and formulate the algebraic cohomology of such algebras. In Sec. III, we introduce the concept of the extension of a generalized Lie algebra and state the relation of the second-order cohomology group with the set of extension of the algebra by an abelian ideal (generalizing the result for Lie algebras in Ref. 10). We discuss also, the applications of the above, to the physically useful algebras of supersymmetry, su(2) ⊕ u(1) and the algebra of quantized fields. In Sec. IV, we state a theorem on the computation of the $n$th-order cohomology groups of some generalized Lie algebras and point out some salient features that arise.

## II. FORMULATION OF THE COHOMOLOGY OF THE GENERALIZED LIE ALGEBRAS

In the present introductory section, we give a resumé of the concept of the generalized Lie algebra (as formulated by Scheunert in Ref. 2) and mention two important features of the commutation factors that follow immediately. Next we formulate the cohomology of such algebras.

### A. Resumé of $\Gamma$ graded algebras

The following definitions are well known from the theory of $\Gamma$ graded algebras[2] and are given below for the sake of completeness.

*Definition 1:* A vector $V$ is said to be *$\Gamma$ graded* if we are given a family $(V_\gamma)_{\gamma\in\Gamma}$ of subspace of $V$ such that $V$ is the direct sum

$$V = \bigoplus_{\gamma\in\Gamma} V_\gamma . \tag{2.1}$$

An element of $V$ is said to be homogeneous of degree $\gamma\in\Gamma$ if it is an element of $V_\gamma$. A subspace $V'$ of $V$ is said to be $\Gamma$ graded if

$$V' = \bigoplus_{\gamma\in\Gamma} (V'\cap V_\gamma).$$

*Definition 2:* An algebra $S$ is called $\Gamma$ graded if its underlying vector space is $\Gamma$ graded,

$$S = \bigoplus_{\gamma\in\Gamma} S_\gamma \tag{2.2a}$$

and if

$$S_\alpha S_\beta \subset S_{\alpha+\beta} \quad \forall\alpha,\beta\in\Gamma. \tag{2.2b}$$

If $S$ has a unit element $e$, it follows that $e\in S_0$. A subalgebra $S$ is said to be $\Gamma$ graded if it is graded as a subspace of $S$.

*Definition 3:* A commutation factor on $\Gamma$ is a mapping $\epsilon$: $\Gamma\times\Gamma\to K$ ($K$ being a field) so that

$$(1)\quad \epsilon(\alpha,\beta)\epsilon(\beta,\alpha) = 1, \tag{2.3a}$$

$$(2)\quad \epsilon(\alpha,\beta+\gamma) = \epsilon(\alpha,\beta)\epsilon(\alpha,\gamma), \tag{2.3b}$$

$$(3)\quad \epsilon(\alpha+\beta,\gamma) = \epsilon(\alpha,\gamma)\epsilon(\beta,\gamma), \tag{2.3c}$$

$$\forall\alpha,\beta,\gamma\in\Gamma.$$

From (2.3b),

 0022-2488/84/082550-07$02.50  

$$\epsilon(\alpha,0)\epsilon(\alpha,\beta)=\epsilon(\alpha,\beta)\quad\forall\alpha,\beta$$
$$\Rightarrow \epsilon(\alpha,0)=1=\epsilon(0,\alpha)\quad\forall\alpha. \tag{2.4}$$

From (2.3a);

$$\epsilon(\alpha,\alpha)\epsilon(\alpha,\alpha)=1$$
$$\Rightarrow \epsilon(\alpha,\alpha)=\pm 1. \tag{2.5}$$

*Definition 4:* Let $\Gamma$ be an abelian group, and let $\epsilon$ be a commutation factor on $\Gamma$. A $\Gamma$ graded algebra

$$L=\bigoplus_{\gamma\in\Gamma} L_\gamma$$

product mapping is denoted by an angle bracket ($\langle\ \rangle$) and is called a $\Gamma$ graded $\epsilon$ Lie algebra (or, simply, a $\Gamma$ graded Lie algebra) if the following properties are satisfied:

(1) $\langle A,B\rangle=-\epsilon(\alpha,\beta)\langle B,A\rangle$ ($\epsilon$ skew symmetry), (2.6a)

(2) $\epsilon(\gamma,\alpha)\langle A,\langle B,C\rangle\rangle+$ cyclic permutations $=0$ (2.6b)

($\epsilon$ Jacobi identity)

$\forall A\in L_\alpha,\quad B\in L_\beta,\quad C\in L_\gamma,\quad \alpha,\beta,\gamma\in\Gamma.$

From (2.4) and (2.6a), it is evident that for $X\in L_0$ the bracket operation of $X$ with all elements of the algebra including itself reduces to the well-known commutator. Henceforth, if $X\in S_\alpha$, we denote the degree $\alpha$ of $X$ as $|X|$; thus we write $|X|=\alpha$.

### B. Cohomology of $\Gamma$ graded Lie algebras

Let $\mathbf{g}$ be a $\Gamma$ graded $\epsilon$ Lie algebera over the field $\mathbb{F}$ and $V$ be a $\Gamma$ graded vector space. Let $\varphi:\mathbf{g}\to$ End $V$ be a representation of $\mathbf{g}$ in $V$. We define

$$C^n(\mathbf{g},V)=\{f: g^n\to V\cdots; f \text{ is } \mathbb{F} \text{ linear and } f(X_1,\ldots,X_i,\ldots,X_j,\ldots,X_n)=-\epsilon(\alpha_i,\alpha_j)f(X_1,\ldots,X_j,\ldots,X_i,\ldots,X_n)\},$$

where

$$X_i\in\mathbf{g}_{\alpha_i}\quad\text{and}\quad X_j\in\mathbf{g}_{\alpha_j}. \tag{2.7}$$

We further define

$$\partial^n: C^n(\mathbf{g},V)\to C^{n+1}(\mathbf{g},V)$$

by

$$\partial^n f(X_1,\ldots,X_{n+1})=\sum_{i=1}^{n+1}(-1)^{i+1}\epsilon\left(\sum_{k=1}^{i-1}\alpha_k,\alpha_i\right)\times\phi(X_i)f(X_1,\ldots,\hat{X}_i,\ldots,X_{n+1})+\sum_{i<j=1}^{n+1}(-1)^{i+j}\times\epsilon\left(\sum_{k=1}^{i-1}\alpha_k,\alpha_i\right)\epsilon\left(\sum_{l=1}^{j-1}\alpha_l,\alpha_j\right)\epsilon(\alpha_j,\alpha_i)\times f(\langle X_i,X_j\rangle,\ldots,\hat{X}_i,\ldots,\hat{X}_j,\ldots,X_{n+1}), \tag{2.8}$$

where $\hat{X}_i$ implies that $X_i$ is to be omitted from the arguments in $f$. We further define

$$C^0(\mathbf{g},V)=\{f:\mathbb{F}\to V, f \text{ is } \mathbb{F} \text{ linear and } \partial^0 f(X)=\phi(X)f(1)\forall X\in\mathbf{g}\}.$$

It follows that

$$\partial^{n+1}\circ\partial^n=0. \tag{2.9}$$

Let $Z^n_\varphi(\mathbf{g},V)=$ Kernel $\partial^n$, (2.10a)

and

$$B^n_\varphi(\mathbf{g},V)=\text{Image }\partial^{n-1}. \tag{2.10b}$$

Finally

$$H^n_\varphi(\mathbf{g},V)=Z^n_\varphi(\mathbf{g},V)/B^n_\varphi(\mathbf{g},V) \tag{2.10c}$$

is defined as the *nth-order cohomology group* for the representation $\varphi$ of $\mathbf{g}$ in $V$. $Z^n_\varphi(\mathbf{g},V)$ and $B^n_\varphi(\mathbf{g},V)$ are called the $n$th-order cocycles and the $n$th-order coboundaries, respectively.

The groups $H^n_\varphi(\mathbf{g},V)$ have the natural structure of a $\Gamma$ graded vector space. In contrast to the ungraded case, $n>\dim g$ does not imply $H^n_\varphi(\mathbf{g},V)$ is trivial, unless $\epsilon(\alpha,\alpha)=+1\ \forall\alpha\in\Gamma$. This may be seen as follows. In the case $\epsilon(\alpha,\alpha)=+1\ \forall\alpha\in\Gamma$, for $n>\dim\mathbf{g}$, an $f$ antisymmetric in all indices is identically zero, as some indices are necessarily repeated. When $\epsilon(\alpha,\alpha)=-1$ for some $\alpha$, however, the argument of degree $\alpha$ may be repeated indefinitely in $f$, so that $n$ could exceed dim $\mathbf{g}$. Hence, $H^n_\varphi(\mathbf{g},V)$ need not be zero automatically when $\epsilon(\alpha,\alpha)=-1$ for some $\alpha$.

For future purposes, we adopt the shorthand notation $\epsilon(-\alpha_j)$ for $\epsilon(\Sigma_{l=1}^{j-1}\alpha_l,\alpha_j)$. Elements of degree zero are excluded from the sum $\Sigma_{l=1}^{j-1}\alpha_l$, for convenience.

## III. $H^n_\varphi(\mathbf{g},\mathbf{h})$ AS AN EXTENSION BY AN ABELIAN IDEAL

Let us consider $\mathbb{F}$ linear maps $f:\mathbf{g}\to\mathbf{h}$, $\mathbf{g},\mathbf{h}$ being $\Gamma$ graded Lie algebras which satisfy

(i) $f(X)=|X|$, (ii) $f(\langle X,Y\rangle)=\langle f(X),f(Y)\rangle$.

*Definition 5:* Given a sequence of homomorphisms of the graded vector spaces $E$, $F$, and $G$

$$E\xrightarrow{f}F\xrightarrow{g}G, \tag{3.1}$$

the sequence is said to be *exact* at $F$ if Ker $g=$ Im $f$. A *short exact sequence* of the $\Gamma$ graded vector spaces is a sequence of the following form

$$0\to E\xrightarrow{f}F\xrightarrow{g}G\to 0, \tag{3.2}$$

which is exact at $E$, $F$, and $G$. In particular, $f$ is injective and $g$ is surjective and Im $f=$ Ker $g$ is a subspace of $F$.

*Definition 6:* An *extension* $E=(\theta,\sigma)$ of $\mathbf{g}$ by $\mathbf{h}$ is a short exact sequence

$$E: 0\to\mathbf{h}\xrightarrow{\theta}\mathbf{k}\xrightarrow{\sigma}\mathbf{g}\to 0. \tag{3.3}$$

Two extensions $E$ and $E'$ are congruent if the diagram commutes:

$$\begin{array}{lccccc} E: & 0\to\mathbf{h} & \xrightarrow{\theta}\mathbf{k} & \xrightarrow[\sigma]{}\mathbf{g} & \to 0 \\ & i\downarrow & \downarrow\beta & \downarrow i & \\ E': & 0\to\mathbf{h} & \xrightarrow{\theta'}\mathbf{k}' & \xrightarrow{\sigma'}\mathbf{g} & \to 0, \end{array} \tag{3.4}$$

$i$ being the identity map. It may be shown that $\beta$ is an isomorphism[11]; hence the congruence of extensions is a reflexive, symmetric, and transitive property.

One of the most obvious extensions of a Lie algebra $\mathbf{g}$ by an abelian ideal $\mathbf{h}$, is through the semidirect sum $\mathbf{k} = \mathbf{h} \circledS \mathbf{g}$ the conditions on $\mathbf{h}$ and $\mathbf{g}$ being

(1) $\mathbf{h}$ is an ideal in $\mathbf{k}$,

(2) $\mathbf{k}|\mathbf{h} \approx \mathbf{g}$,

(3) If $\theta$: $\mathbf{h} \to \mathbf{k}$ is an injective homomorphism and $\sigma$:$\mathbf{k} \to \mathbf{g}$ is a surjective homomorphism, then Ker $\sigma$ = Im $\theta$.

Hence $\mathbf{h}$, $\mathbf{k}$, and $\mathbf{g}$ form an exact sequence.

For instance, the Poincaré algebra is an extension of the SO(3,1) algebra by the abelian ideal $P_\mu$ (formed by the space-time translation generators), through the semidirect sum, the sequence $0 \to P_\mu \to \mathrm{SO}(3,1) \circledS P_\mu \to \mathrm{SO}(3,1) \to 0$ is exact. The questions, we attempt to answer in the following are:

(1) Are there any extensions allowed for a $\Gamma$ graded Lie algebra other than through the semidirect sum of the algebra with the ideal?

(2) How do we determine them?

In what follows, we assume that the sequence (3.3) splits, i.e., is isomorphic to

$$0 \to \mathbf{h} \xrightarrow{\theta} \mathbf{h} \oplus \mathbf{g} \xrightarrow{\sigma} \mathbf{g} \to 0. \tag{3.5}$$

This implies[11] that

(1) $\sigma$ has a right inverse $u$ with $\sigma \cdot u = 1g$;

(2) $\theta$ has a left inverse $\rho$ with $\rho \cdot \theta = 1u$.

Let $\mathrm{Ext}_{\mathbb{F}}(\mathbf{g},\mathbf{h})$ be the congruence class of extensions of $\mathbf{g}$ by $\mathbf{h}$ over the field $\mathbb{F}$. Then, we have

**Theorem 1:**

$$H^n_\varphi(\mathbf{g},\mathbf{h}) = \mathrm{Ext}_F(\mathbf{g},\mathbf{h}),$$

if

$$\langle \mathrm{Im}(\mathbf{g},\mathbf{h}), \mathrm{Im}(\mathbf{g},\mathbf{h}) \rangle = 0.$$

$(\mathrm{Im}(\mathbf{g},\mathbf{h}))$ is the map $\rho u$: $\mathbf{g} \to \mathbf{h}$. It does not exist for all elements of $\mathbf{g}$, as the map $\rho$ does not exist for all elements of $\mathbf{k}$.

The proof of the theorem follows on the same lines as in Refs. 10, 11, and 14 with the commutator bracket ([ ]) being replaced by the $\epsilon$ Lie bracket. It is given in the Appendix.

*Remark 1:* We note that $\mathrm{Ext}_{\mathbb{F}}(\mathbf{g},\mathbf{h})$ is dependent on the representation $\varphi$ of $\mathbf{g}$ in $\mathbf{h}$ [(A2)]. One may state a corollary to Theorem 1 as follows:

*Corollary 1:* The extension of an algebra $\mathbf{h}$ by a single element set $\mathbf{g}$ of degree zero, so that $\mathbf{h}$ is an ideal, is unique and given by the semidirect sum.

*Proof*: $C^2(\mathbf{g},\mathbf{h})$ contains the single element zero. This is also obvious in the extension; the new element can only commute with itself.

In such a case, the only choice remains in the action of the element in $\mathbf{g}$ on the elements of $\mathbf{h}$.

Returning the Poincaré algebra as an extension of SO(3,1), we note that $H^n_\varphi(\mathbf{g},\mathbf{h}) = 0$ for all semisimple $\mathbf{g}$ and finite-dimensional $\mathbf{h}$.[15] Identifying $\mathbf{g}$ with SO(3,1) and $\mathbf{h}$ with $P_\mu$, $H^2_\varphi(\mathbf{g},\mathbf{h}) = 0\ \forall\ \varphi$ and hence the extension of the Lorentz group SO(3,1) by the translation generators is uniquely defined and given by the semidirect sum.

*Remark 2:* The extension of the $\Gamma$ graded *Lie algebra* by the semidirect sum *does not* presume to throw any light on the extension of the corresponding graded Lie group. For instance, the extension of the su(2) algebra by u(1) may be shown to be unique but the extension of the group SU(2) by U(1) may be done in two possible ways, the extended groups being U(2) and SU(2) $\otimes$ U(1). They are related by a two-to-one homomorphism.

In practice, the extensions of the algebra are most simply calculated by setting the commutation relations of the elements of $\mathbf{g}$ to zero, unless desired otherwise by the conditions of the problem under consideration. For the elements $X_i, X_j$ of $\mathbf{g}$ for which $\langle X_i, X_j \rangle$ is zero, $f(X_i,X_j)$ is written as a linear combination of the elements in $\mathbf{h}$ and $H^2_\varphi(\mathbf{g},\mathbf{h})$ is calculated by the usual procedure.

Further examples of extensions of $\Gamma$ graded Lie algebras are given in what follows.

### A. Graded Poincaré algegra

Having introduced the generators $Q_\alpha$, $\overline{Q}_\beta$ which take a boson to a fermion and vice versa, we now consider the problem of extension of the algebra formed by the set $\mathbf{g} = \{M_{\mu\nu}, Q_\alpha, \overline{Q}_\beta\}$; we are given that $|M_{\mu\nu}| = 0$, $|Q_\alpha| = |\overline{Q}_\beta| = 1$, $\epsilon(1,1) = -1$. The commutation (anticommutation) relations are as given below:

$$[M_{\mu\nu},M_{\rho\sigma}] = g_{\mu\rho} M_{\nu\sigma} + g_{\nu\sigma} M_{\mu\rho} - g_{\mu\sigma} M_{\nu\rho} - g_{\nu\rho} M_{\mu\sigma}, \tag{3.6a}$$

$$[M_{\mu\nu},Q_\alpha] = \tfrac{1}{2} i(\sigma_{\mu\nu} Q)_\alpha, \quad [\overline{Q}_\beta,M_{\mu\nu}] = \tfrac{1}{2} i(\overline{Q}\sigma_{\mu\nu})_\beta, \tag{3.6b}$$

$$\sigma_{\mu\nu} = \tfrac{1}{2} i[\gamma_\mu,\gamma_\nu], \quad [Q_\alpha,\overline{Q}_\beta]_+ = 0.$$

The algebra $\mathbf{g}$ is to be extended by the abelian ideal $\mathbf{h}$ formed by the set $\{P_\mu,(\gamma_\mu)_{\alpha\beta},(\gamma_\mu P^\mu)_{\alpha\beta}\}$. $P_\mu$ are the space-time translation generators, and $(\gamma_\mu)_{\alpha\beta}$ are the Dirac matrices. Considering the split sequence

$$0 \to \{(\gamma_\mu)_{\alpha\beta}, P_\mu(\gamma_\mu P^\mu)_{\alpha\beta}\} \to \{(\gamma_\mu)_{\alpha\beta},P_\mu(\gamma_\mu, P^\mu)_{\alpha\beta}, M_{\mu\nu},Q_\alpha,\overline{Q}_\beta\} \to \{M_{\mu\nu},Q_\alpha,\overline{Q}_\beta\} \to 0 \tag{3.7}$$

we introduce [from (A2)]

$$\begin{aligned}
&\varphi(M_{\mu\nu})P_\rho = g_{\mu\rho} P_\nu - g_{\nu\rho} P_\mu,\\
&\varphi(M_{\mu\nu})\gamma^\rho P_\rho = (P_\nu,\gamma_\mu - P_\mu \gamma_\nu),\\
&\varphi(M_{\mu\nu})\gamma_\rho = 0, \quad \varphi(Q_\alpha) P_\mu = \varphi(\overline{Q}_\beta) P_\mu = 0,\\
&\varphi(Q_\alpha)(\gamma_\mu)_{\beta\gamma} = \varphi(\overline{Q}_\beta)(\gamma_\mu)_{\alpha\gamma}\\
&\qquad = \varphi(Q_\alpha)(\gamma_\mu P^\mu)_{\beta\gamma} = \varphi(\overline{Q}_\beta)(\gamma_\mu P^\mu)_{\alpha\gamma} = 0.
\end{aligned} \tag{3.8}$$

These follow from elementary 4-vector considerations in field theory. We note that in $\mathbf{h}$, $[P_\mu,P_\nu] = 0$ $[\gamma_\mu,\gamma_\nu]_+ = 2g_{\mu\nu} 1_{\mathbf{h}}$.

We now calculate $H^2_\varphi(\mathbf{g},\mathbf{h})$. Using Lorentz covariance, we know that $f(Q_\alpha,\overline{Q}_\beta)$ should be a Lorentz scalar, so that

$$f(Q_\alpha,\overline{Q}_\beta) = a(\gamma_\rho P^\rho)_{\alpha\beta}. \tag{3.9a}$$

$f(M_{\mu\nu},Q_\alpha)$ should be a Lorentz tensor of rank 2 and a Dirac spinor. Neither of $(\gamma_\mu)_{\alpha\beta}$, $P_\mu$ or $(\gamma_\mu P^\mu)_{\alpha\beta}$ qualify, so that

$$f(M_{\mu\nu},Q_\alpha) = f(M_{\mu\nu},\overline{Q}_\beta) = 0. \tag{3.9b}$$

Using (3.8), (3.9a), and (3.9b),

$$\partial f(M_{\mu\nu},Q_\alpha,\bar{Q}_\beta)$$
$$= \varphi(M_{\mu\nu})f(Q_\alpha,\bar{Q}_\beta) + \tfrac{1}{2}if((\sigma_{\mu\nu}Q)_\alpha,\bar{Q}_\beta)$$
$$- \tfrac{1}{2}if((\bar{Q}\sigma_{\mu\nu})_\beta,Q_\alpha)$$
$$= \tfrac{1}{2}i(\sigma_{\mu\nu})_{\alpha\gamma}a(\gamma_\rho P^\rho)_{\gamma\beta} - \tfrac{1}{2}i(\sigma_{\mu\nu})_{\gamma\beta}(\gamma_\rho P^\rho)_{\alpha\gamma}$$
$$- a(P_\mu\gamma_\nu - P_\nu\gamma_\mu)_{\alpha\beta} = 0.$$

So,

$$f \in Z^2(\mathbf{g},\mathbf{h}).$$

Let

$$f'(Q_\alpha,Q_\beta) = f(Q_\alpha,\bar{Q}_\beta) - \partial V(Q_\alpha,\bar{Q}_\beta)$$
$$= a(\gamma_\rho P^\rho)_{\alpha\beta}$$
$$- \varphi(Q_\alpha)V(\bar{Q}_\beta) - \varphi(\bar{Q}_\beta)V(Q_\alpha)$$
$$+ V([Q_\alpha,\bar{Q}_\beta]_+).$$

In $\mathbf{h}$ the only choice for $V(Q_\alpha)$ and $V(\bar{Q}_\beta)$ is zero, as no member in $\mathbf{h}$ has only one spinor index. Besides $V([Q_\alpha,\bar{Q}_\beta]_+) = V(0) = 0$.

Hence

$$f'(Q_\alpha,\bar{Q}_\beta) = a(\gamma_\rho P^\rho)_{\alpha\beta} \tag{3.10}$$

is the only nonzero choice for $f$ allowed by $H^2_\varphi(\mathbf{g},\mathbf{h})$ for the spinor and tensor transformation properties allowed by $\varphi$: $\mathbf{g} \to \mathbf{h}$. The only extensions for this Lie algebra are

$$(1)\ [Q_\alpha,\bar{Q}_\beta]_+ = (\gamma_\rho P^\rho)_{\alpha\beta}, \tag{3.11a}$$

$$(2)\ [Q_\alpha,\bar{Q}_\beta]_+ = 0. \tag{3.11b}$$

(3.11b) is meaningless if $Q_\alpha(\bar{Q}_\beta)$ are given the interpretation of annihilation (creation) operators of spin-$\frac{1}{2}$ particles.

### B. Extension of su(2) by u(1)

We consider the sequence formed by $\mathbf{h} = \{X_1,X_2,X_3\}$, $\forall\alpha,\beta \in \Gamma$, $\mathbf{g} = \{u(1)\}$, $\epsilon(\alpha,\beta) = +1$ so that $\mathbf{k} = \{X_1,X_2,X_3, u(1)\}$. $X_1, X_2, X_3$ have the usual su(2) algebra $[X_i,X_j] = i\epsilon_{ijk}X_k$, $\epsilon_{ijk}$ being antisymmetric in $i,j$, and $k$. From considerations in (A2) we take

$$\varphi(\mathrm{u}(1))X_i = 0, \quad i = 1,2,3 \tag{3.12}$$

as $[\mathrm{u}(1),X_i]$ is required to be zero in $\mathbf{k}$.

From Corollory 1, $H^2_\varphi(\mathbf{g},\mathbf{h})$ is zero; hence, the extension of the su(2) Lie algebra by u(1) is trivial and accomplished by the semidirect sum.

### C. The extension of an algebra and quantization

(a) The abelian Lie algebra $\mathbf{g} = \{p,q\}$ may be extended by the set $\mathbf{h}$ consisting of the ideal formed by the identity $\{e\}$ in two possible ways. One may check that $H^2_\varphi(\mathbf{g},\mathbf{h})$ gives $f(p,q) = re$ for $r \in \mathbb{F}$ assuming $\varphi(p)e = \varphi(q)e = 0$.

The two possible extensions of $\mathbf{g} = \{p,q\}$ by $\mathbf{h} = \{e\}$ are

(1) the abelian algebra semidirect sum $\mathbf{k} = \{p,q,e\}$, $[p,q] = [p,e] = [q,e] = 0$,

(2) the extended algebra $[p,q] = e$, $[p,e] = [q,e] = 0$.

The latter is tantamount to giving a quantization condition if $q(p)$ were given by the interpretations of coordinate (momentum), respectively.

(b) The extension of $\mathbf{g} = \{H\}$ by the abelian ideal formed by $\mathbf{h} = \{p,q\}$ assuming degree $|H| = 0$ yields

$$H^2_\varphi(\mathbf{g},\mathbf{h}) = 0 \quad \text{(by Corollary 1)}.$$

The extension is trivial; however, one is at liberty to choose $\varphi(H)p = q(p)$, $\varphi(H)q = -p(q)$.

We end up with either the harmonic oscillator or the scaled Lie algebra.

(c) Likewise when the ideal $\mathbf{h} = \{p,q,e\}$ extends $\mathbf{g} = \{H\}$, we end up with a similarly trivial extension as $H^2_\varphi(\mathbf{g},\mathbf{h}) = 0$ by Corollary 1. This is independent of the internal algebra of $\mathbf{h}$.

(d) If $|H| = 1$ and $\epsilon(1,1) = -1$, starting with $[H,H]_+ = 0$ in $\mathbf{g} = \{H\}$ and $\mathbf{h} = \{p,q\}$, the choice

$$\varphi(H)p = p(q), \quad \varphi(H)q = q(-p),$$

nevertheless, gives $H^2_\varphi(\mathbf{g},\mathbf{h}) = 0$, so that the extension of $[H,H]_+$ remains unmodified at zero, as for $|H| = 0$.

(e) The extension of $\mathbf{g} = \{H\}$ $[|H| = 1, \epsilon(1,1) = -1]$ by $\mathbf{h} = \{p,q,e\}$, $[p,q] = e$ $[p(q),e] = 0$ is nontrivial if the action $\varphi$ of $H$ on either $p$, $q$, or $e$ is zero. If $\varphi(\mathrm{H})e = 0$, $\varphi(H)p(q)$ nonzero, then we end up with $f(H,H) = ae, a \in \mathbb{F} f \in H^2_\varphi(\mathbf{g},\mathbf{h})$ so that $\langle H,H\rangle = e$. The number $a$ may be used to set the required scale. From the above, one may state another corollary to Theorem 2 as follows:

*Corollary 2:* The extension of an algebra $\mathbf{h}$ by a single element set $\mathbf{g} = \{H\}$ of degree $\alpha$: $\epsilon(\alpha,\alpha) = -1$ is given by the elements $X_i$, for which $\varphi(H)X_i = 0$.

In other words, $f(H,H) = \Sigma_i a_i X_i$, $a_i \in \mathbb{F}$, and $\varphi(H)X_i = 0$. The proof follows trivially.

(f) Finally, we consider the extension of the algebra $\mathbf{g} = \{b,b^+\}$ by $\mathbf{h} = \{a,a^+,e\}$ the elements of both $\mathbf{g}$ and $\mathbf{h}$ being of degree zero. Given $[b,b^+] = 0$ in $\mathbf{g}$, $[a,a^+] = e$, $[a,e] = [a^+,e] = 0$ in $\mathbf{h}$, and taking $\varphi(b)(a,a^+,e) = 0 = \varphi(b^+)(a,a^+,e)$, we have for $f \in H^2_\varphi(\mathbf{g},\mathbf{h})$

$$f(b,b) = f(b^+,b^+) = 0,$$
$$f(b,b^+) = a_1a + a_2a^+ + a_3e, \quad a_i \in \mathbb{F}, \quad i = 1,2,3.$$

Physically, we require $b$, $b^+$ to be independent and on par with $a$, $a^+$ for different bosons, created (annihilated) by $a^+$ $(a)$ or $b^+(b)$. Hence, for quantization, only the reduced algebra $[b,b^+] = a_3e$ is taken, the parameter $a_3$ being used to set a suitable scale.

If the elements of $\mathbf{g} = \{b,b^+\}$ had degree 1, $\epsilon(1,1) = -1$, and $[b,b]_+ = [b^+,b^+]_+ = [b,b^+]_+ = 0$ in $\mathbf{g}$, then the extended algebra would contain nine extra parameters, three each for $f(b,b)$, $f(b^+,b^+)$, and $f(b,b^+)$, $f \in H^2_\varphi(\mathbf{g},\mathbf{h})$. The problem under consideration decides the choice.

Quantization is the process of extension that sets $[b,b^+] = e$ and all other extended brackets $[b,b]_+ = [b^+,b^+]_+ = 0$.

It may be further noted that if the algebra formed by $\mathbf{h} = \{b,b^+,e\}$ as ideal, the elements of $\mathbf{h}$ being of degree 1, and if $\epsilon(1,1) = -1$ were to be extended by $\mathbf{g} = \{a,a^+\} \times (a,a^+)$ being of degree zero, then in this reverse extension problem, we could extend the algebra only by three independent parameters, as $f(a,a) = f(a^+,a^+) = 0$ because of antisymmetry.

*Remark 4:* It is evident, from the above example, that the extension of an algebra with elements of degree $\alpha$,

$\epsilon(\alpha,\alpha) = -1$ may involve many more arbitrary parameters than for elements for which $\epsilon(\alpha,\alpha) = +1$. In the case of quantization many of these may be put equal to zero. There may exist cases, however, where these extra parameters may assume importance.

## IV. COHOMOLOGY OF SOME $\Gamma$ GRADED LIE ALGEBRAS

In this section, we generalize the result in Ref. 13 to $\Gamma$ graded Lie algebras and enumerate some cases where it is applicable.

**Theorem 2:** Let $\mathbf{g}$ be a $\Gamma$ graded $\epsilon$ Lie algebra over the field $\mathbb{F}$, and $V$ be a $\mathbf{g}$ module. Assume $\exists x \in \mathbf{g}\ (|x| = 0)$ and $\phi$: $\mathbf{g} \to \text{End } V$ such that the following is true: $\exists$ a basis $B$ for $\mathbf{g}$ such that $\forall X_1, X_2,\ldots,X_l \in B, Z \in V$,

(i) $$[X,X_i] = -a(X_i)X_i, \quad i = 1,\ldots,l, \tag{4.1a}$$

(ii) the equation

$$[\varphi(X) + a]\ Y = Z \tag{4.1b}$$

may be uniquely solved for $Y$ in terms of $Z$, $\forall Z \in V$, where $a = \Sigma_i\, a(Xi)$. Then $H^n_\varphi(\mathbf{g},V) = 0$.

*Proof*: Let $f \in Z^n(\mathbf{g},V)$. Then $\forall X_i \in S_{\alpha_i}, \forall \alpha_i \in \Gamma$

$$\begin{aligned}
&\partial^n f(X_1,\ldots,X_{n+1}) = 0 \\
&\Rightarrow \sum_{i=1}^{n+1} (-1)^{i+1}\epsilon(-\alpha_i)\,\varphi(X_i) f(X_1,\ldots,\hat{X}_i,\ldots,X_{n+1}) \\
&\quad + \sum_{i<j=1}^{n+1} (-1)^{i+j}\epsilon(-\alpha_i)\epsilon(-\alpha_j)\epsilon(\alpha_j,\alpha_i) \\
&\quad \times f(\langle X_i,X_j\rangle,\ldots,\hat{X}_i,\ldots,\hat{X}_j,\ldots,X_{n+1}) = 0 \qquad (4.2a) \\
&\quad \forall X_i \neq X,\ i = 1,\ldots,n+1,
\end{aligned}$$

and

$$\begin{aligned}
&\partial^n f(X,X_1,\ldots,X_n) = 0 \\
&\Rightarrow \varphi(X) f(X_1,\ldots,X_n) + \sum_{i=1}^{n} a(X_i) f(X_1,\ldots,X_n) \\
&\quad + \sum_{i=1}^{n} (-1)^i \epsilon(-\alpha_i)\,\varphi(X_i) f(X,\ldots,\hat{X}_i,\ldots,X_n) \\
&\quad - \sum_{i<j=1}^{n} (-1)^{i+j}\epsilon(-\alpha_i)\epsilon(-\alpha_j)\epsilon(\alpha_j;\alpha_i) \\
&\quad \times f(X,\langle X_i,X_j\rangle,\ldots,\hat{X}_i,\ldots,\hat{X}_j,\ldots,X_n). \qquad (4.2b)
\end{aligned}$$

[In the second term we used (4.1a) and $X_i$ has been transferred to its original place.]

From (4.2b) and (4.1b), we know that $f(X_1,\ldots,X_n)$ exists such that

$$\begin{aligned}
&(\varphi(X) + a) f(X_1,\ldots,X_n) \\
&\quad = \sum_{i=1}^{n} (-1)^{i+1}\epsilon(-\alpha_i) \\
&\qquad \times \varphi(X_i) f(X,\ldots,\hat{X}_i,\ldots,X_n) \\
&\qquad + \sum_{i<j=1}^{n} (-1)^{i+j}\epsilon(-\alpha_i)\epsilon(-\alpha_j,)\epsilon(\alpha_j,\alpha_i) \\
&\qquad \times f(X,\langle X_i,X_j\rangle,\ldots,\hat{X}_i,\ldots,\hat{X}_j,\ldots,X_n). \qquad (4.3a)
\end{aligned}$$

Using condition (ii) and (4.3a), we may write

$$\begin{aligned}
&f(X_1,\ldots,X_n) = \sum_{i=1}^{n} (-1)^{i+1}\epsilon(-\alpha_i)[\varphi(X) + a]^{-1} \\
&\qquad \times f(X,\ldots,\hat{X}_i,\ldots,X_n) + \sum_{i<j=1}^{n} (-1)^{i+j}\epsilon(-\alpha_i) \\
&\qquad \times \epsilon(-\alpha_j)\epsilon(\alpha_j,\alpha_i)[\varphi(x) + a]^{-1} \\
&\qquad \times f(X,\langle X_i,X_j\rangle,\ldots,\hat{X}_i,\ldots,\hat{X}_j,\ldots). \qquad (4.3b)
\end{aligned}$$

(4.3b) follows from the uniqueness of the solution to (4.1b). From (4.1a), we also have

$$(\varphi(X) + a)\,\varphi(X_i) = \varphi(X_i)(\varphi(X) + a - a(X_i)), \tag{4.4}$$

where

$$a = \sum_k a(X_k).$$

Using (4.4) and (4.3b), it is possible to show that (4.2a) is satisfied. Now let $f' = f - \partial\omega$, so that

$$\begin{aligned}
f'(X,\ldots,X_{n-1}) &= f(X,\ldots,X_{n-1}) - \partial\omega(X,\ldots,X_{n-1}) \\
&= f(X,\ldots,X_{n-1}) - \sum_{i=1}^{n-1} (-1)^{i+1} \\
&\quad \times \epsilon(-\alpha_i)\,\varphi(X_i)\omega(X,\ldots,\hat{X}_i,\ldots,X_{n-1}) \\
&\quad - \sum_{i=1}^{n-1} a(X_i)\omega(X_1,\ldots,X_{n-1}) \\
&\quad + \sum_{i<j=1}^{n-1} (-1)^{i+j}\epsilon(-\alpha_i)\epsilon(-\alpha_j)\epsilon(\alpha_j,\alpha_i) \\
&\quad \times \omega(X,\langle X_i,X_j\rangle,\ldots,\hat{X}_i,\ldots,\hat{X}_j,\ldots,X_{n-1}) \\
&\quad - \varphi(X)\omega(X_1,\ldots,X_{n-1}) \\
f'(X,\ldots,X_n) &= f(X_1,\ldots,X_n) - \partial\omega(X_1,\ldots,X_n) \\
&= f(X_1,\ldots,X_n) + \sum_{i=1}^{n} (-1)^i\epsilon(-\alpha_i) \\
&\quad \times \varphi(X_i)\omega(X,\ldots,\hat{X}_i,\ldots,X_n) \\
&\quad - \sum_{i<j=1}^{n} (-1)^{i+j}\epsilon(-\alpha_i)\epsilon(-\alpha_j)\epsilon(\alpha_j,\alpha_i) \\
&\quad \times \omega(\langle X_i,X_j\rangle,\ldots,\hat{X}_i,\ldots,\hat{X}_j,\ldots,X_n). \qquad (4.5b)
\end{aligned}$$

We now choose

$$\omega(X,\ldots,\hat{X}_i,\ldots,X_n) = 0, \tag{4.6a}$$

and

$$\begin{aligned}
&\left[\varphi(X) + \sum_{i=1}^{n-1} a(X_i)\right]\omega(X_1,\ldots,X_{n-1}) \\
&\quad = f(X,\ldots,X_{n-1}). \qquad (4.6b)
\end{aligned}$$

From (4.1b) we know that $\omega(X_1,\ldots,X_{n-1})$ exists uniquely. Substituting in (4.5b) and using (4.4), (4.6a), (4.6b), and (4.3b), we have

$$f'(X,\ldots,X_{n-1}) = 0,$$

$$f'(X_1,\ldots,X_n) = 0,$$

and hence $f = \partial\omega$. Thus

$$H^n_\varphi(\mathbf{g},V) = 0 \quad \forall n > 0.$$

Putting $a(X_i) = 0\ \forall X_i$ and $\varphi(X) = 1_V$, $\varphi^{-1}(X)$ exists trivially, and the equation $Y = Z$ is obviously solvable uniquely $\forall Z \in V$. We thus have,

*Corollary 3:* Let $\mathbf{g}$ be a $\Gamma$ graded Lie algebra and $V$ be a $\mathbf{g}$

module. If $\exists X \in \mathbf{g}$, $|X| = 0$ so that $[X,X_i] = 0\ \forall X_i \in \mathbf{g}$, then $H^n_\varphi(\mathbf{g},V) = 0\ \forall\ n > 0$, $\varphi$: $\mathbf{g} \to \text{End } V$, $\varphi(X) = 1_V$.

This is a generalization of the result obtained in Ref. 13 to all $\Gamma$ graded Lie algebras with a central element and independent of the *form of the commutation factor* for the other elements. Hence it holds for all kinds of $\Gamma$ graded Lie algebras discussed in Ref. 3, in particular, the $Z_N \otimes Z_N \otimes \cdots \otimes Z_N$ and the $Z_2 \otimes Z_2 \otimes \cdots \otimes Z_2$ graded algebras, with at least one element belonging to the center.

The theorem has a straightforward application to all quantization algebras of the form $[a,a^+] = e$, $e$ being the identity element, and $a^+$ ($a$) being the creation (annihilation) operators, respectively, with $\varphi(e) = 1_V$. The graded attribute of the vector space enables us to deal with the algebras of different particles with different commutation factors in one package. Hence, if one is considering a system of bosons and fermions with different commutation factors, then for representations $\phi(e) = 1_V$, $H^n_\varphi(\mathbf{g},V) = 0\ \forall n > 0$. In this connection, we see also the explicit computation of $H^n_\varphi(\mathbf{g},V)$ for the Dirac algebra as reported earlier.[14]

Also, all representations $\varphi$ of su(2) $\oplus$ u(1), $\varphi(u(1)) = 1_V$, admit of a trivial cohomology. In particular, this is true for the spinor representation of su(2) $\oplus$ u(1), which is the symmetry group for the weak interaction Weinberg–Salam Hamiltonian.

*Remark 5:* The harmonic oscillator algebra and the Dirac algebra share the property that the identity element in the center $Z(\mathbf{g})$ of the GLA, $e \in \langle \mathbf{g},\mathbf{g} \rangle$. If $\epsilon(\alpha,\alpha) = +1$ $\forall\ \alpha \in \Gamma: e \in \langle \mathbf{g}_\alpha,\mathbf{g}_\alpha \rangle$, then it implies that for $\varphi(e) = 1_V$, $V$ is necessarily infinite-dimensional. This is because $\text{tr}\, e = \dim V = \text{tr}[\varphi(\mathbf{g}), \varphi(\mathbf{g})] = 0$ for $\epsilon(\alpha,\alpha) = +1$ $\forall \alpha \in \Gamma$ in $V$, which implies that either dim $V$ is trivial, or, that tr[$\varphi(e)$] is not well defined. If $\mathbf{g} = \mathbf{h} \oplus \text{u}(1)$ $\mathbf{h}$ semisimple, or $\epsilon(\alpha,\alpha) = -1$, $\forall \alpha \in \Gamma$: $e \in \langle \mathbf{g}_\alpha, \mathbf{g}_\alpha \rangle$, then the representation may be finite-dimensional. In the first case, even if $\epsilon(\alpha,\alpha) = 1$ for some $\alpha \in \Gamma\ e \notin \langle \mathbf{g}, \mathbf{g} \rangle$. So, the above argument does not hold. In the second case, for instance, for the Dirac algebra, we have dim $V = \text{Tr}[\varphi(e)] = \text{Tr}\langle \phi(\mathbf{g}), \phi(\mathbf{g}) \rangle$. The right-hand side is not necessarily zero any longer; hence the algebra may be finite-dimensional, as is the Dirac algebra.

*Remark 6:* Theorem 2 cannot be applied to semisimple groups for adjoint action. For example, we take the SO(3) basis $\{T_+,T_-,T_3\}$. Identifying $T_3$ with $X$ so that $[T_3,T_+] = T_+$, $[T_3,T_-] = -T_-$, $[T_3,T_3] = 0$, and taking $\varphi(T_3)$ to be the adjoint action, $[T_3, Y] = T_3$ admits no solution for $Y$.

## V. CONCLUSION

In connection with the extension problem, it appears that cohomology gives the best idea of all possible extensions of the algebras under consideration, through their ideals. The list obtained will be exhaustive, notwithstanding the fact that most of the extensions are of no immediate interest to physicists. The purpose of the relevant section in this paper is to focus attention on the uniqueness of some important extensions which have been developed, supersymmetry being one very important example. We plan to report more on the importance of $H^n_\varphi(\mathbf{g},V)$ for arbitary $n$ in a later paper.


## ACKNOWLEDGMENT

We would like to thank Dr. B. R. Sitaram for some very fruitful discussions and especially for pointing out the relevance of condition (4.1b).


## APPENDIX: PROOF OF THEOREM 1

Let the elements of $\mathbf{h}$ be labelled by $X_i$, $i = 1,...,l$, those of $\mathbf{k}$ by $y_i$, $i = 1,...,m$, and those of $\mathbf{g}$ by $Z_i$, $i = 1,...,n$.

Then as $\mathbf{h}$ is an ideal

$$\langle \text{u}(Z_i),\theta(X_j) \rangle \in \text{Im}\ \theta = \theta(X_s) \quad \text{(say)}, \quad s = 1,...,l.$$

A homomorphism $\varphi$: $\mathbf{g} \to \text{End } \mathbf{h}$ is thus induced so that the diagram

$$\begin{array}{ccc} & \sigma & \\ \mathbf{k} & \to & \mathbf{g} \\ \rho \downarrow & \swarrow \varphi & \\ \text{End } \mathbf{h} & & \end{array} \tag{A1}$$

commutes. The

$$\langle \rho\text{u}(Z_i),X_j \rangle_a = \varphi(Z_i)X_j \cdots. \tag{A2}$$

Now the sequence (3.3) is exact, so that

$$\begin{aligned} \sigma[\text{u}(Z_1)\text{u}(Z_2) - \epsilon(|Z|,|Z_2|)\text{u}(Z_2)\text{u}(Z_1) - \text{u}(\langle Z_1,Z_2 \rangle)] \\ = Z_1Z_2 - \epsilon(|Z_1|,|Z_2|)Z_2Z_1 - (Z_1,Z_2) \\ = 0. \end{aligned}$$

[$\because$ $\sigma$ is a homomorphism].

Hence,

$$\begin{aligned} &\text{u}(Z_1)\text{u}(Z_2) - \epsilon(|Z_1|,|Z_2|)\text{u}(Z_2)\text{u}(Z_1) \\ &\quad - \text{u}(\langle Z_1,Z_2 \rangle) \in \ker \sigma = \text{Im}\ \theta. \end{aligned} \tag{A3}$$

Let

$$\begin{aligned} f(Z_1,Z_2) &= \text{u}(Z_1)\text{u}(Z_2) \\ &\quad - \epsilon(|Z_1|,|Z_2|)\text{u}(Z_2)\text{u}(Z_1) - \text{u}(\langle Z_1,Z_2 \rangle) \end{aligned} \tag{A4}$$

$$\Rightarrow \langle \rho\text{u}(Z_1), \rho\text{u}(Z_2) \rangle = \rho\text{u}(\langle Z_1,Z_2 \rangle) + \rho f(Z_1,Z_2), \cdots. \tag{A5}$$

The bracket of $\langle \text{u}(Z_1),\text{u}(Z_2) \rangle$ in $\mathbf{k}$ has been redefined by (A4); hence from (A), $\rho f(Z_1,Z_2)$ defines an extension of the $\Gamma$ graded algebra $\mathbf{g}$ in $\mathbf{k}$ by $f(Z_1,Z_2)$ which has an image in $\mathbf{h}$.

$f(Z_1,Z_2)$ from (A4) is seen to be (1) linear and (2) $\epsilon$ antisymmetric in $Z_1$ and $Z_2$. Hence $\rho f$ belongs to $C^2(\mathbf{g},\mathbf{h})$. Also

$$\begin{aligned} &\langle \rho\text{u}(Z), \langle \rho\text{u}(Z_2), \rho\text{u}(Z_3) \rangle \rangle \\ &\quad = \langle \rho\text{u}(Z_1), \rho\text{u}(\langle Z_2,Z_3 \rangle) + \rho f(Z_2,Z_3) \rangle \\ &\quad = [\rho\text{u}(\langle Z_1, \langle Z_2,Z_3 \rangle \rangle) + \rho f(Z_1, \langle Z_2,Z_3 \rangle) \\ &\qquad + \phi(Z_1)\rho f(Z_2,Z_3)]. \end{aligned} \tag{A6}$$

(A2) has been used to obtain the last term. Using the Jacobi identity, (A2) and (A6), we have

$$\begin{aligned} &\rho\text{u}(\langle Z_1, \langle Z_2,Z_3 \rangle \rangle) + \rho f(Z_1, \langle Z_2,Z_3 \rangle) + \phi(Z_1)\rho f(Z_2,Z_3) \\ &\quad + \epsilon(|Z_1| + |Z_2|,|Z_3|)\rho\text{u}(\langle Z_3, \langle Z_1,Z_2 \rangle \rangle) \\ &\quad + \epsilon(|Z_1| + |Z_2|,|Z_3|)\rho f(Z_3, \langle Z_1,Z_2 \rangle) \\ &\quad + \epsilon(|Z_1|,|Z_2| + |Z_3|)\rho\text{u}(\langle Z_2, \langle Z_3,Z_1 \rangle \rangle) \\ &\quad + \epsilon(|Z_1| + |Z_2|,|Z_3|)\rho f(Z_2, \langle Z_3,Z_1 \rangle) \end{aligned}$$

$$+\epsilon(|Z_1|+|Z_2|,|Z_3|)\,\phi(Z_3)\,\rho f(Z_1,Z_2)$$
$$+\epsilon(|Z_1|,|Z_2|+|Z_3|)\,\phi(Z_2)\,\rho f(Z_3,Z_1)=0.$$

The u terms drop out because of linearity and the Jacobi identity and we are left with

$$\partial^2 \rho f(Z_1,Z_2,Z_3)=0. \tag{A7}$$

Changing to $\rho u(Z)=\rho u(Z)-\rho V(Z)$ so that $\rho V: \mathbf{g}\to\mathbf{h}$, we have

$$\begin{aligned}\rho f'(Z_1,Z_2)&=\rho u'(Z_1)\rho u'(Z_2)-\epsilon(|Z_1|,|Z_2|)\\&\quad\times\rho u'(Z_2)\rho u(Z_1)-\rho\, u'(\langle Z_1,Z_2\rangle)\\&=\rho f(Z_1,Z_2)-\varphi(Z_1)\rho V(Z_2)\\&\quad+\epsilon(|Z_1|,|Z_2|)\,\varphi(Z_2)\rho V(Z_1)\\&\quad+\rho V(\langle Z_1,Z_2\rangle)+\langle\rho V(Z_1),\rho V(Z_2)\rangle.\end{aligned}$$

If $\mathbf{h}$ is graded abelian, $\langle\rho V(Z_1),\rho V(Z_2)\rangle=0$. If we start with $\langle \mathrm{Im}\,\mathbf{g},\mathrm{Im}\,\mathbf{g}\rangle=0$ in $\mathbf{h}$, we have

$$\rho f'(Z_1,Z_2)=\rho f(Z_1,Z_2)-\partial\rho V(Z_1,Z_2). \tag{A8}$$

Thus, once the choice $f(Z_1,Z_2)$ is made, a different map from within $\mathbf{g}$ does not define a new cohomology class.

So far, we have, starting from a member of the congruence class of extensions, obtained a member of the cohomology class of $H^2_\varphi(\mathbf{g},\mathbf{h})$. To prove the converse, we note that $\rho f(Z_1,Z_2)\in C^2(\mathbf{g},\mathbf{h})$ and is $\epsilon$ antisymmetric in $Z_1$ and $Z_2$. Besides $\sigma f(Z_1,Z_2)=0$. Hence the only choice for $f(Z_1,Z_2)$ is given by (A4). Starting from (A8) and working backwards, it is easy to show that the extensions defined by $f'$ and $f$ are related by a linear transformation with $\mathbf{g}$. As the basis of $\mathbf{g}$ can always be redefined, so can the map $\sigma'$ in (3.4) and hence the extensions $f$ and $f'$ are congruent. This completes the proof.


[1] P. Fayet and S. Ferrara, Phys. Rep. C **32**, 249 (1977).
[2] M. Scheunert, J. Math. Phys. **20**, 712 (1979).
[3] V. Rittenberg and D. Wyler, Nucl. Phys. B **139**, 189 (1978).
[4] V. Rittenberg, "Beyond Supersymmetry"—John Hopkins University Workshop on Current Problems in High Energy Particle Theory, Baltimore, 1978, p. 254.
[5] G. Dixon, J. Math. Phys. **19**, 2103 (1978).
[6] J. Lukierski and V. Rittenberg, Phys. Rev. D **18**, 385 (1978).
[7] S. S. Chern, *Complex Manifolds without Potential Theory* (Springer-Verlag, Berlin, 1979).
[8] M. E. Mayer, in *Fibre bundles and Gauge Theories*, Lecture Notes in Physics 67 (Springer-Verlag, Berlin, 1977).
[9] J. Erven and B. Falkowski, *Low Order Cohomology and Applications*, Lecture Notes in Physics 877 (Springer-Verlag, Berlin, 1979).
[10] R. Hermann, Comm. Math. Phys. **2**, 251 (1966); **3**, 53 (1966); **3**, 75 (1966).
[11] S. MacLane, *Homology, Die Grundlehren der Matematischen Wissenchaften* V. 114 (Springer-Verlag, Berlin, 1963).
[12] B. R. Sitaram and K. C. Tripathy, "Representations of Graded Lie Groups, II" (unpublished).
[13] B. Mitra, B. R. Sitaram, and K. C. Tripathy, "Cohomology of Lie Algebras with a non-trivial center," J. Math. Phys. **25**, 443 (1984).
[14] B. R. Sitaram and K. C. Tripathy, Indian J. Pure Appl. Math. **13**, 672 (1982).
[15] N. Jacobson, *Lie Algebras* (Dover, New York, 1979).

# Some exact solutions of nonlinear chiral field equations

Pranab Krishna Chanda and Dipankar Ray
*Department of Applied Mathematics, Calcutta University, 92, A. P. C. Road, Calcutta 700 009, India*
Utpal Kumar De
*Department of Physics, Jadavpur University, Calcutta 700 032, India*

(Received 19 July 1983; accepted for publication 4 November 1983)

Some new exact solutions of the nonlinear field equations for the chiral invariant model of pion dynamics are presented here. These solutions are a further generalization of some previous works presented by one of the authors (Ray). It is interesting to note that equations in (3.2) obtained by Ray (1978) are conformally invariant. Hence from any solution of these equations one can immediately generate infinitely many other solutions of these equations simply by replacing $(x^1, x^2)$ by $(y, z)$, where $y$ and $z$ are any two mutually conjugate solutions of Laplace's equations. Further, a striking similarity in form of these equations with one of the two generalized Lund–Regge equations makes the study of the solutions of these equations more worthwhile with the view that the study of the solutions of these equations will eventually lead to the study of the solution of a larger class of equations that will include these equations and generalized Lund–Regge equations as special cases.

PACS numbers: 11.30.Rd, 11.10.Lm

## 1. INTRODUCTION

Under tangential parametrization (Charap, 1973)[1] the field equations for the Chiral invariant model of the pion dynamics take the form (Charap, 1976)[2]

$$\Box\phi = \eta^{\mu\nu}\frac{\partial\phi}{\partial x^\mu}\frac{\partial\beta}{\partial x^\mu},$$
$$\Box\psi = \eta^{\mu\nu}\frac{\partial\psi}{\partial x^\mu}\frac{\partial\beta}{\partial x^\nu},$$
$$\Box\chi = \eta^{\mu\nu}\frac{\partial\chi}{\partial x^\mu}\frac{\partial\beta}{\partial x^\nu}, \tag{1}$$

where

$$\eta^{\mu\nu} = 0 \quad \text{for } \mu \neq \nu,$$
$$= 1 \quad \text{for } \mu = \nu \neq 4,$$
$$= -1 \quad \text{for } \mu = \nu = 4,$$
$$\beta = \ln(f_\pi^2 + \phi^2 + \psi^2 + \chi^2),$$
$$f_\pi = \text{const.}$$

Under the assumptions

$$\phi = \phi(x^1, x^2, x^3 - x^4),$$
$$\psi = \psi(x^1, x^2, x^3 - x^4), \tag{2}$$
$$\chi = \chi(x^1, x^2, x^3 - x^4),$$

the set of differential equations in four variables, namely, Eqs. (1), reduces to a set of differential equations in two variables as follows (Ray, 1978)[3]:

$$\phi_{11} + \phi_{22} = \beta_1\phi_1 + \beta_2\phi_2, \tag{3a}$$
$$\psi_{11} + \psi_{22} = \beta_1\psi_1 + \beta_2\psi_2, \tag{3b}$$
$$\chi_{11} + \chi_{22} = \beta_1\chi_1 + \beta_2\chi_2, \tag{3c}$$

where

$$\beta = \ln(f_\pi^2 + \phi^2 + \psi^2 + \chi^2), \tag{3d}$$
$$f_\pi = \text{const}, \tag{3e}$$

$$\phi_1 = \frac{\partial\phi}{\partial\chi^1}, \quad \beta_1 = \frac{\partial\beta}{\partial\chi^1}, \quad \text{and so on.}$$

Apart from the physical significance of Eqs. (3) already described, Eqs. (3) seem to be of considerable mathematical interest as well. This is first because Eqs. (3) are conformally invariant, i.e., the form of Eqs. (3) remain invariant under transformation $(x^1, x^2) \to (y, z)$, where $y$ and $z$ are two mutually conjugate solutions of Laplace's equations in $x^1$ and $x^2$. Hence from any solution of Eqs. (3) one can immediately generate infinitely many other solutions of Eqs. (3) simply by replacing $(x^1, x^2)$ by $(y, z)$, where $y$ and $z$ are two mutually conjugate solutions of Laplace's equations.

In this context attention should also be drawn to another set of coupled equations, namely, the generalized Lund–Regge equations, which are[4,5]

$$\text{(a)} \quad \theta_{11} \pm \theta_{22} - 4g(\theta) + h(\theta)(\lambda_1^2 + \lambda_2^2) = 0,$$
$$\text{(b)} \quad \lambda_{11} \pm \lambda_{22} = 2p(\theta)(\lambda_1\theta_1 + \lambda_2\theta_2), \tag{4}$$

where

$$\theta = \theta(x^1, x^2), \quad \lambda = \lambda(x^1, x^2),$$
$$\theta_1 = \frac{\partial\theta}{\partial x^1} \quad \text{and so on.}$$

Particular examples of the generalized Lund–Regge equations include a good number of physically interesting equations, e.g., equations of two-dimensional Heisenberg ferromagnets,[6,7] Ginzburg–Pitaevski equations for superfluids,[8,9] stationary wave envelope in nonlinear optics,[10] and so on. It is interesting to note that Eqs. (3) are of the same form as one of the two generalized Lund–Regge equations, namely, (4b). Since a good number of studies of the solutions of the generalized equations have been made it is quite possible that a study of the solutions of (3) will eventually lead to the study of the solution of a larger class of equations that will include (3) and generalized Lund–Regge equations as special cases. For all these reasons the study of the solution of (3) seems worthwhile.

 0022-2488/84/082557-06$02.50  

## 2. SOLUTIONS

Equation (3) may be rewritten as

$$(e^{-\beta}\phi_1)_1 + (e^{-\beta}\phi_2)_2 = 0, \tag{5a}$$

$$\psi_{11} + \psi_{22} = \psi_1\beta_1 + \psi_2\beta_2, \tag{5b}$$

$$\chi_{11} + \chi_{22} = \chi_1\beta_1 + \chi_2\beta_2, \tag{5c}$$

where

$$\beta = \ln(f_\pi^2 + \phi^2 + \psi^2 + \chi^2). \tag{5d}$$

*Case I*: Assuming

$$\beta = \beta(\phi), \tag{6}$$

one can set

$$X = \int e^{-\beta}\, d\phi. \tag{7}$$

Then Eq. (5a) with the help of Eq. (7) takes the form

$$X_{11} + X_{22} = 0, \tag{8}$$

which is the Laplace's equation and standard solutions for $X$ are obtainable. Now, one can write

$$\psi^2 + \chi^2 = \alpha^2, \tag{9a}$$

when (5d) reduces to

$$\beta = \ln(f_\pi^2 + \phi^2 + \alpha^2). \tag{9b}$$

Then one gets from (9b), (6), and (7),

$$\alpha = \alpha(X), \tag{9c}$$

and from (9a),

$$\psi = \alpha\cos\theta, \tag{9d}$$

$$\chi = \alpha\sin\theta. \tag{9e}$$

One can take

$$\theta = \theta(X, Y), \tag{10}$$

where $Y$ is the solution of the Laplace's equation conjugate to $X$, i.e., $X_1 = Y_2$ and $X_2 = -Y_1$. Using (6), (7), (9d), (9e), and (10) one may reduce (5b) and (5c) to the following forms:

$$\{\alpha_{XX} - \alpha(\theta_X^2 + \theta_Y^2) - \alpha_X\beta_X\}\cos\theta = \{2\alpha_X\theta_X + \alpha(\theta_{XX} + \theta_{YY}) - \alpha\beta_X\theta_X\}\sin\theta, \tag{11a}$$

$$\{\alpha_{XX} - \alpha(\theta_X^2 + \theta_Y^2) - \alpha_X\beta_X\}\sin\theta = -\{2\alpha_X\theta_X + \alpha(\theta_{XX} + \theta_{YY}) - \alpha\beta_X\theta_X\}\cos\theta. \tag{11b}$$

(11a) and (11b) are equivalent to

$$\alpha_{XX} - \alpha(\theta_X^2 + \theta_Y^2) - \alpha_X\beta_X = 0, \tag{12a}$$

$$2\alpha_X\theta_X + \alpha(\theta_{XX} + \theta_{YY}) - \alpha\beta_X\theta_X = 0. \tag{12b}$$

Equation (12) may be rewritten as

$$\theta_X^2 + \theta_Y^2 = \delta(X), \tag{13a}$$

$$\delta(X) = \frac{\alpha_{XX} - \alpha_X\beta_X}{\alpha} \neq 0, \tag{13b}$$

$$\alpha(\theta_{XX} + \theta_{YY}) + 2\alpha_X\theta_X - \alpha\beta_X\theta_X = 0. \tag{13c}$$

The above equation leads to

$$\theta_Y = \text{const} = B \quad \text{(say)}. \tag{14}$$

To verify it, let us consider three possibilities separately.

(i) If $\theta_X \neq 0$, $\theta_Y = 0$, then (14) is automatically satisfied with $B = 0$.

(ii) If $\theta_X = 0$, $\theta_Y \neq 0$, the left-hand side of (13a) is a function of $Y$ only and the right-hand side is a function of $X$ only. Since both are constants. Hence, $\theta_Y = \text{const}$ and (14) is satisfied.

(iii) If $\theta_X \neq 0$, $\theta_Y \neq 0$, one can proceed as follows: Differentiating (13a) with respect to $Y$,

$$\theta_X\theta_{XY} + \theta_Y\theta_{YY} = 0.$$

Eliminating $\theta_{YY}$ from (13c) with the use of the above, dividing throughout by $[\theta_Y e^{-(2\ln\alpha - \beta)}]$ and integrating, $\alpha^2\theta_X e^{-\beta}/\theta_Y = \gamma(Y)$, where $\gamma(Y)$ is an unspecified function of $Y$.

This readily gives $\theta = \theta(w)$, where

$$w = u + v,$$

$$u = \int \frac{e^\beta}{\alpha^2}\, dX,$$

$$v = \int \frac{dY}{\gamma(Y)}.$$

Using the above relations, we get from (13a)

$$\frac{u_X^2}{\delta(X)} + \frac{v_Y^2}{\delta(X)} = \frac{1}{\theta_w^2}.$$

Differentiating the above expression separately with respect to $u$ and $v$, respectively, and comparing the results,

$$\delta\left(\frac{u_X^2}{\delta}\right)_u + \delta\left(\frac{1}{\delta}\right)_u v_Y^2 = (v_Y^2)_v.$$

Differentiating this successively with respect to $u$ and $v$,

$$[\delta(1/\delta)_u]_u (v_Y^2)_v = 0.$$

Hence,

$$\delta(1/\delta)_u = \text{const},$$

or

$$v_Y = \text{const}.$$

That $\delta(1/\delta)_u = \text{const}$ is not possible will be shown in Appendix A.

Differentiating $[u_X^2/\delta(X) + v_Y^2/\delta(X)] = 1/\theta_w^2$ with respect to $v$ and using $v_Y = \text{const}$, one can get

$$\left(\frac{1}{\theta_w^2}\right)_w = 0,$$

which gives $\theta_w = \text{const}$. Hence, $\theta_Y = \theta_w \cdot w_Y = \theta_w \cdot v_Y = \text{const}$. Thus, in this case of $\theta_X \neq 0$, $\theta_Y \neq 0$ too we see that (14) is satisfied. Using (14) in (13c) and after integrating once, one may get

$$\theta_X = Ae^\beta/\alpha^2, \quad \text{where } A = \text{const}. \tag{15}$$

Generalizing (15) and (14), one gets

$$\theta = \int \frac{Ae^\beta}{\alpha^2}\, dX + BY + C, \tag{16}$$

where $B$ and $C$ are also constants. Using (16) in (13a) and (13b),

$$(\alpha_{XX} - \alpha_X\beta_X)/\alpha = A^2e^{2\beta}/\alpha^4 + B^2,$$

or $(e^{-\beta}\alpha_X)_X = (A^2e^\beta/\alpha^4 + B^2/e^\beta)\alpha$, which may be rewritten with the use of (9b) and (7) as

$$\alpha_{\phi\phi} = \frac{A^2}{\alpha^3} + \frac{B^2\alpha}{(f_\pi^2 + \phi^2 + \alpha^2)^2}. \tag{17}$$

Solution of Eq. (17) gives the dependence of $\alpha$ in terms of $\phi$. It may be pointed out that Eq. (17) is easily solvable with the use of $\alpha_{\phi\phi} = \frac{1}{2}(\alpha_\phi^2)_\alpha$ in the particular case $B = 0$. Then all the unknown quantities like $\psi, \chi, \phi, \beta$ become a function of $X$ only, which in turn will be a function of $(x^1, x^2, x^3 - x^4)$.

We also get from (9b) and (7),

$$\phi_X = f_\pi^2 + \phi^2 + \alpha^2. \tag{18}$$

Thus, we can conclude that Eqs. (18) and (17) give the solution of $\phi$ in terms of $X$. As $e^\beta = \phi_X, \beta$ is also known in terms of $X$. The use of the expression for $\phi$ in Eq. (17) will let us determine $\alpha$ in terms of $X$. Then Eq. (16) gives $\theta$ in terms of $X$ and $Y$, where $Y$ is conjugate Laplace solution to $X$. Lastly, $X$ is obtained as a solution of Laplace's equation $X_{11} + X_{22} = 0$. These completely solve the unknown quantities of Eq. (1).

Another case of interest may be added, when $\delta(X) = 0$. Then $\theta$ becomes a constant. Since now $(\alpha_{XX} - \alpha_X\beta_X)/\alpha = 0$, one finally gets, with the help of Eq. (7), $\alpha_{\phi\phi} = 0$. On integration, we get $\alpha$ in terms of $\phi$. Using this in (18), and on integration, which is easily possible, one may get $\phi$ in terms of $X$, which is a solution of Laplace's equation $X_{11} + X_{22} = 0$. Now, the unknown quantities of Eq. (1) become known.

*Case II*: Without loss of generality one can write, from Eq. (5a)

$$e^{-\beta}\phi_1 = \sigma_2, \tag{19a}$$

$$e^{-\beta}\phi_2 = -\sigma_1. \tag{19b}$$

Now, we introduce an assumption that

$$\psi^2 + \chi^2 = (f_\pi^2 + \phi^2)\xi^2(\sigma), \tag{20a}$$

where $\xi(\sigma) =$ some unspecified function of $\sigma$. Then it turns out that

$$\psi = (f_\pi^2 + \phi^2)^{1/2}\xi(\sigma)\cos\Theta, \tag{20b}$$

$$\chi = (f_\pi^2 + \phi^2)^{1/2}\xi(\sigma)\sin\Theta, \tag{20c}$$

$$\Theta = \Theta(\sigma, \phi), \tag{20d}$$

$$\beta = \ln(f_\pi^2 + \phi^2) + \ln(1 + \xi^2). \tag{20e}$$

Using (20b), (20c), (20d), (20e), and (19), one may reduce (5b) and (5c) to the following forms:

$$\begin{aligned}&[\Phi_{\phi\phi}\xi\Phi^4(1+\xi^2)^2 + \Phi\xi_{\sigma\sigma} - \xi\Theta_\phi^2\Phi^5(1+\xi^2)^2\\ &\quad - \Phi\xi\Theta_\sigma^2 - 2\Phi\xi_\sigma(1+\xi^2)_\sigma/(1+\xi^2)]\cos\Theta\\ &= [2\Phi\xi_\sigma\Theta_\sigma + \xi\Theta_{\phi\phi}\Phi^5(1+\xi^2)^2 + \Phi\xi\Theta_{\sigma\sigma}\\ &\quad + 2\Phi^4\xi\Theta_\phi\Phi_\phi(1+\xi^2)^2\\ &\quad - 2\Phi\xi\Theta_\sigma(1+\xi^2)_\sigma/(1+\xi^2)]\sin\Theta,\end{aligned} \tag{21a}$$

and

$$\begin{aligned}&[\Phi_{\phi\phi}\xi\Phi^4(1+\xi^2)^2 + \Phi\xi_{\sigma\sigma} - \xi\Theta_\phi^2\Phi^5(1+\xi^2)^2\\ &\quad - \Phi\xi\Theta_\sigma^2 - 2\Phi\xi_\sigma(1+\xi^2)_\sigma/(1+\xi^2)]\sin\Theta\\ &= -[2\Phi\xi_\sigma\Theta_\sigma + \xi\Theta_{\phi\phi}\Phi^5(1+\xi^2)^2 + \Phi\xi\Theta_{\sigma\sigma}\\ &\quad + 2\Phi^4\xi\Theta_\phi\Phi_\phi(1+\xi^2)^2\\ &\quad - 2\Phi\xi\Theta_\sigma(1+\xi^2)_\sigma/(1+\xi^2)]\cos\Theta,\end{aligned} \tag{21b}$$

where

$$\Phi = (f_\pi^2 + \phi^2)^{1/2}. \tag{22}$$

(21a) and (21b) are equivalent to

$$\begin{aligned}&\Phi_{\phi\phi}\xi\Phi^4(1+\xi^2)^2 + \Phi\xi_{\sigma\sigma} - \xi\Theta_\phi^2\Phi^5(1+\xi^2)^2\\ &\quad - \Phi\xi\Theta_\sigma^2 - 2\Phi\xi_\sigma(1+\xi^2)_\sigma/(1+\xi^2) = 0\end{aligned} \tag{23a}$$

and

$$\begin{aligned}&2\Phi\xi_\sigma\Theta_\sigma + \xi\Theta_{\phi\phi}\Phi^5(1+\xi^2)^2 + \Phi\xi\Theta_{\sigma\sigma}\\ &\quad + 2\Phi^4\xi\Theta_\phi\Phi_\phi(1+\xi^2)^2\\ &\quad - 2\Phi\xi\Theta_\sigma(1+\xi^2)_\sigma/(1+\xi^2) = 0,\end{aligned} \tag{23b}$$

which may be rewritten with the use of (22) as

$$\begin{aligned}&f_\pi^2 + \frac{\xi_{\sigma\sigma}}{\xi(1+\xi^2)^2} - \Phi^4\Theta_\phi^2 - \frac{\Theta_\sigma^2}{(1+\xi^2)^2}\\ &\quad - \frac{2\xi_\sigma(1+\xi^2)_\sigma}{\xi(1+\xi^2)^3} = 0,\end{aligned} \tag{24a}$$

and

$$\begin{aligned}&\frac{\xi^2(\xi^2\Theta_\sigma)_\sigma}{\xi^4(1+\xi^2)^2} + (f_\pi^2 + \phi^2)[(f_\pi^2 + \phi^2)\Theta_\phi]_\phi\\ &\quad - \frac{2(1+\xi^2)_\sigma}{(1+\xi^2)^3}\Theta_\sigma = 0.\end{aligned} \tag{24b}$$

Defining

$$\int \frac{d\sigma}{\xi^2} = p \tag{25a}$$

and

$$\int \frac{d\phi}{f_\pi^2 + \phi^2} = q, \tag{25b}$$

Eqs. (24) may be reduced to

$$\begin{aligned}\frac{\Theta_p^2}{\xi^4(1+\xi^2)^2} + \Theta_q^2 &= f_\pi^2 + \frac{\xi_{\sigma\sigma}}{\xi(1+\xi^2)^2}\\ &\quad - \frac{2\xi_\sigma(1+\xi^2)_\sigma}{\xi(1+\xi^2)^3}\end{aligned} \tag{26a}$$

and

$$\frac{\Theta_{pp}}{\xi^4(1+\xi^2)^2} + \Theta_{qq} = \frac{2(1+\xi^2)_\sigma}{(1+\xi^2)^3\xi^2}\Theta_p. \tag{26b}$$

Again defining

$$\frac{1}{\xi^4(1+\xi^2)^2} = U(p), \tag{27a}$$

$$f_\pi^2 + \frac{\xi_{\sigma\sigma}}{\xi(1+\xi^2)^2} - \frac{2\xi_\sigma(1+\xi^2)_\sigma}{\xi(1+\xi^2)^3} = V(p), \tag{27b}$$

and

$$\frac{2(1+\xi^2)_\sigma}{(1+\xi^2)^3\xi^2} = W(p), \tag{27c}$$

Eqs. (26) may be reduced to

$$U\Theta_p^2 + \Theta_q^2 = V, \tag{28a}$$

$$U\Theta_{pp} + \Theta_{qq} = W\Theta_p. \tag{28b}$$

From Eq. (28a) one can examine four cases separately. However, the following three cases, i.e., (i) $\Theta_p = \Theta_q = 0$, (ii) $\Theta_p \neq 0$ and $\Theta_q = 0$, and (iii) $\Theta_p = 0$ and $\Theta_q \neq 0$, can be grouped under one head of $\Theta_q = \text{const}$, where the constant may even take the value zero.

To study the fourth case, i.e., $\Theta_p \neq 0$, $\Theta_q \neq 0$, one can proceed as follows.

Differentiating (28a) with respect to $q$,

$$U\Theta_p\Theta_{pq} + \Theta_q\Theta_{qq} = 0.$$

Eliminating $\Theta_{qq}$ from (28b) with the use of the above, dividing throughout by $[U\Theta_q]$, expressing $U$ and $W$ in terms of $\xi$ with the help of (27), and integrating,

$$\frac{1}{(1+\xi^2)^2}\cdot\frac{\Theta_p}{\Theta_q} = \epsilon(q),$$

where $\epsilon(q)$ is an unspecified function of $q$. This readily gives

$$\Theta = \Theta(\lambda),$$

where

$$\lambda = m + n,$$

$$m = \int \frac{dq}{\epsilon},$$

$$n = \int (1+\xi^2)^2\, dp.$$

Using the above relations one gets from (28a)

$$n_p^2\left(\frac{U}{V}\right) + \frac{m_q^2}{V} = \frac{1}{\Theta_\lambda^2}.$$

Differentiating the above expression separately with respect to $n$ and $m$, respectively, and comparing the results, one gets

$$V\left(\frac{U}{V}n_p^2\right)_n + V\left(\frac{1}{V}\right)_n m_q^2 = (m_q^2)_m.$$

Differentiating this successively with respect to $n$ and $m$, one can conclude that

$$[V(1/V)_n]_n (m_q^2)_m = 0.$$

Hence,

$$V(1/V)_n = \text{const},$$

or,

$$m_q = \text{const}.$$

At first let us take, $m_q = \text{const}$. Then, differentiating $[n_p^2(U/V) + m_q^2/V] = 1/\Theta_\lambda^2$ with respect to $m$ and using $m_q = \text{const}$, one gets

$$(1/\Theta_\lambda^2)_\lambda = 0,$$

which gives $\Theta_\lambda = \text{const}$. Hence, $\Theta_q = \Theta_\lambda \cdot \lambda_q = \Theta_\lambda \cdot m_q = \text{const}$. Thus, in the case $\Theta_p \neq 0$, $\Theta_q \neq 0$, we immediately find that again the $\Theta_q = \text{const}$ case is satisfied.

So, let us proceed to find the solution in the general case when

$$\Theta_q = \text{const}. \tag{29}$$

Using $\Theta_q = \text{const}$ in (28b) and after integration once, one gets

$$\Theta_p = E(1+\xi^2)^2, \tag{30}$$

where $E$ is a const. Generalizing (29) and (30), one gets

$$\Theta = E\int (1+\xi^2)\, dp + Fq + G, \tag{31}$$

where $F$ and $G$ are also constants. The use of (25) reduces (31) to

$$\Theta = E\int \frac{(1+\xi^2)^2}{\xi^2}\, d\sigma + \frac{F}{f_\pi}\tan^{-1}\left(\frac{\phi}{f_\pi}\right) + G. \tag{32}$$

Using (31) in (28a), expressing $U$ and $V$ with the help of (27) and rearranging, one can get

$$\xi_{\sigma\sigma} - [4\xi/(1+\xi^2)](\xi_\sigma^2) = E^2(1+\xi^2)^4\xi^{-3} + \xi(1+\xi^2)^2(F^2 - f_\pi^2),$$

which with the use of $\xi_{\sigma\sigma} = \frac{1}{2}(\xi_\sigma^2)_\xi$ can easily be integrated to give

$$\sigma = \frac{1}{2}\int \frac{dZ}{[HZ(1+Z)^4 - E^2(1+Z)^4 - (F^2 - f_\pi^2)(1+Z)^3 Z]^{1/2}} + \text{const}, \tag{33}$$

where $H = \text{const}$ and $Z = \xi^2$. With the use of (20e), Eqs. (19a) and (19b) may be reduced to

$$\phi_1/(f_\pi^2 + \phi^2) = (1+\xi^2)\sigma_2$$

and

$$\phi_2/(f_\pi^2 + \phi^2) = -(1+\xi^2)\sigma_1.$$

One may define

$$\int (1+\xi^2)\, d\sigma = r. \tag{34}$$

Then using (25b) and (34), the above two relations may be reduced to $q_1 = r_2$ and $q_2 = -r_1$, which indicate that $q$ and $r$ are mutually conjugate solutions of Laplace's equation.

Further, (25b) may be integrated to give

$$\phi = f_\pi \tan(qf_\pi + Kf_\pi), \tag{35}$$

where $K$ is a constant and $q$ is a solution of Laplace's equation. Thus, one may conclude that (35) gives the solution of $\phi$ in terms of $q$, where $q$ is obtained as a solution of Laplace's equation $q_{11} + q_{22} = 0$. Expressing $\sigma$ in terms of $\xi$ from (33), then $r$ in terms of $\xi$ from (34), and remembering that $r$ is obtained as a solution of Laplace's equation $r_{11} + r_{22} = 0$, $\xi$ becomes known. With $\phi$ and $\xi$ known, one gets immediately $\beta$ and $\Theta$ from (20e) and (32), respectively. With all these being known, solutions of $\psi$ and $\chi$ are readily available through Eqs. (20b) and (20c), respectively. These completely solve the unknown quantities of Eq. (1). That the other possibility of the situation $\Theta_p \neq 0$, $\Theta_q \neq 0$, i.e., $V(1/V)_n = \text{const}$ does not give rise to any new case will be shown in Appendix B.

## 3. CONCLUSION

In our case the solutions of our basic equations given by (3), can be classified under two broad cases, depending on the assumptions being made.

*Case I*: Here it was only assumed that $\beta$ is a function of the single variable $\phi$. The entire problem can be solved exactly if the solution of a single-variable second-order differential equation could be performed [namely, Eq. (17)]. However, in two special cases, the complete solution could be

presented in very simple forms.

*Case II*: Here it was assumed that $\psi^2+\chi^2=(f_\pi^2+\phi^2)\xi^2(\sigma)$, where $\sigma$ is a new variable. Then the problem can be solved completely with $\sigma$ being expressed in terms of $\xi$ in the form of an integral equation. However, one can easily find the solution of $\sigma$ in terms of the usual variables like $x^1, x^2$ if the exact relation of $\xi$ in terms of $\sigma$ could be established. This is possible in principle.

It may be noted that, due to the absence of $x^3$ and $x^4$ in the reduced form (3) of (1), all the constants involved in the solutions are functions of $(x^3-x^4)$.

## APPENDIX A

With

$$\delta(1/\delta)_u=\text{const}=L \quad \text{(say)}, \tag{A1}$$

it is evident from

$$\delta\left(\frac{u_X^2}{\delta}\right)_u+\delta\left(\frac{1}{\delta}\right)_u v_Y^2=(v_Y^2)_v$$

that

$$\delta\left(\frac{u_X^2}{\delta}\right)_u=\text{const}=M \quad \text{(say)} \tag{A2}$$

and

$$(v_Y^2)_v=Lv_Y^2+M. \tag{A3}$$

From (A1), we get

$$\delta=Ne^{-Lu}, \tag{A4}$$

where $N=\text{const}\neq 0$.

In the following it will be shown that the above equations are not satisfied simultaneously and hence $\delta(1/\delta)_u=\text{const}$ is not permissible.

### Case (i): $L\neq 0$

Expanding (A2) and then using (A1), one gets

$$(u_X^2)_u=M-L(u_X^2).$$

Integrating the above expression and with the help of $u=\int(e^\beta/\alpha^2)\,dX$, one finally gets

$$e^{2\beta}/\alpha^4=(1/P)(Q-e^{-Lu}), \tag{A5}$$

where $P=\text{const}\neq 0$, $Q=\text{const}$. Using (7) and $u=\int(e^\beta/\alpha^2)\,dX$, we have

$$\phi_u=\phi_X X_u=\alpha^2. \tag{A6}$$

Using (9b), we have $e^\beta=f_\pi^2+\phi^2+\alpha^2$. Eliminating $e^\beta$ from the above with the use of (A5), one gets

$$\alpha^2=\frac{\phi^2+f_\pi^2}{[(Q-e^{-Lu})/P]^{1/2}-1}. \tag{A7}$$

Eliminating $\alpha^2$ from (A7) with the use of (A6) and then on integration, one gets

$$\phi=f_\pi\tan\left\{f_\pi\int\frac{du}{[(Q-e^{-Lu})/P]^{1/2}-1}-Rf_\pi\right\}, \tag{A8}$$

where $R=\text{const}$. Using (A8) in (A7)

$$\frac{1}{\alpha}=\frac{1}{f_\pi}\left\{\left[\frac{(Q-e^{-Lu})}{P}\right]^{1/2}-1\right\}^{1/2}\times\cos\left\{f_\pi\int\frac{du}{[(Q-e^{-Lu})/P]^{1/2}-1}-Rf_\pi\right\}. \tag{A9}$$

Further, considering (A4) and (13b), one gets

$$\frac{\alpha_{XX}-\alpha_X\beta_X}{\alpha}=Ne^{-Lu},$$

which may be written as

$$\frac{e^\beta(e^{-\beta}\alpha_X)_X}{\alpha}=Ne^{-Lu}.$$

With the change of variable from $X$ to $u$ with the help of $u=\int(e^\beta/\alpha^2)\,dX$, the above is reduced to

$$\frac{e^{2\beta}}{\alpha^4}\left(\frac{1}{\alpha}\right)_{uu}+\frac{1}{\alpha}Ne^{-Lu}=0,$$

which may be rewritten with the use of (A5) as

$$(Q-e^{-Lu})(1/\alpha)_{uu}+(1/\alpha)PNe^{-Lu}=0. \tag{A10}$$

The integration in (A9) may be done easily. Then the elimination of $1/\alpha$ from (A10) with the help of (A9) indicates as if the tangent of an angle is expressible in terms of a rational function in amalgamation with roots. This is not permissible in a physical situation.

### Case (ii): $L=0$

From (A3) and (A4), respectively, one can get

$$(v_Y^2)_v=M, \tag{A11}$$

$$\delta=N. \tag{A12}$$

Expanding (A2) and then using (A12), one gets $(u_X^2)_u=M$. On integration and then with the use of $u=\int(e^\beta/\alpha^2)\,dX$ the above reduces to

$$e^{2\beta}/\alpha^4=Mu+S, \tag{A13}$$

where $S=\text{const}$. As in (A6),

$$\phi_u=\phi_X X_u=\alpha^2. \tag{A14}$$

Using (9b), we have $e^\beta=f_\pi^2+\phi^2+\alpha^2$. Eliminating $e^\beta$ from the above with the use of (A13), one gets

$$\alpha^2=(\phi^2+f_\pi^2)/[(Mu+S)^{1/2}-1]. \tag{A15}$$

Eliminating $\alpha^2$ from (A15) with the use of (A14) and then on integration, one gets

$$\phi=f_\pi\tan\left\{f_\pi\int\frac{du}{(Mu+S)^{1/2}-1}-Jf_\pi\right\}, \tag{A16}$$

where $J=\text{const}$. Using (A15) and (A16),

$$\frac{1}{\alpha}=\frac{1}{f_\pi}[(Mu+S)^{1/2}-1]^{1/2}\times\cos\left\{f_\pi\int\frac{du}{(Mu+S)^{1/2}-1}-Jf_\pi\right\}. \tag{A17}$$

Further, considering (A12) and (13b),

$$(\alpha_{XX}-\alpha_X\beta_X)/\alpha=N,$$

which may be reduced as in case (i) to

$$\frac{e^{2\beta}}{\alpha^4}\left(\frac{1}{\alpha}\right)_{uu} + \frac{N}{\alpha} = 0.$$

This may further be rewritten with the use of (A13) as

$$(Mu + S)(1/\alpha)_{uu} + N/\alpha = 0. \tag{A18}$$

The integration of (A17) may be done easily. The elimination of $1/\alpha$ from (A18) with the help of (A17) indicates if the tangent of an angle is expressible in terms of a rational function in amalgamation with roots. As in the previous case, this is not permissible.

## APPENDIX B

With

$$V(1/V)_n = \text{const} = \Gamma \text{(say)}, \tag{B1}$$

it becomes evident from

$$V\left(\frac{U}{V}n_p^2\right)_n + V\left(\frac{1}{V}\right)_n m_q^2 = (m_q^2)_m$$

that

$$V\left(\frac{U}{V}n_p^2\right)_n = \text{const} = \Sigma \quad \text{(say)}. \tag{B2}$$

Expanding (B2) and then using (B1), (27a), (25a), and

$$n = \int (1+\xi^2)^2\, dp,$$

one gets

$$\xi_\sigma = \tfrac{1}{4}[(\Gamma - \Sigma)\xi^5 + (3\Gamma - \Sigma)\xi^3 + 3\Gamma\xi + \Gamma\xi^{-1}]. \tag{B3}$$

When $\Gamma$ and $\Sigma$ are simultaneously zero, (B3) leads to $\xi = \text{const}$. When $\Gamma$ and $\Sigma$ are not simultaneously zero one can separate the situation into two cases depending on whether $\Gamma \neq 0$ or $\Gamma = 0$.

*Case (i)*: $\Gamma \neq 0$

Dividing (B2) by (B1), integrating, using (27), and $n = \int (1+\xi^2)^2\, dp$ one gets

$$\Gamma\cdot d\cdot\xi^3(1+\xi^2)\xi_{\sigma\sigma} - 4\Gamma\cdot d\cdot\xi^4(\xi_\sigma)^2 + (\Sigma + \Gamma\cdot d\cdot f_\lambda^2)\xi^4(1+\xi^2)^3 - \Gamma(1+\xi^2)^5 = 0. \tag{B4}$$

Using (B3) in (B4) one arrives at an equation in $\xi$ which is satisfied for discrete values of $\xi$ only.

*Case (ii)*: $\Gamma = 0$

From (B1) one can now write

$$V = \text{const} = k \text{ (say)} \neq 0. \tag{B5}$$

Using (27b) one can reduce (B5) to

$$(1+\xi^2)\xi_{\sigma\sigma} - 4\xi(\xi_\sigma)^2 + \xi(1+\xi^2)^3(f_\lambda^2 - k) = 0. \tag{B6}$$

Using (B3), with $\Gamma = 0$, in (B6) one arrives at an equation in $\xi$ which is satisfied for discrete values of $\xi$ only.

Thus, for any value of $\Gamma$ and $\Sigma$ one gets $\xi = \text{const}$. It is easy to check from (20e) that $\xi = \text{const}$ represents a special situation of case I [$\beta = \beta(\phi)$] and hence need not be considered separately.


[1] J. M. Charap, J. Phys. A: Math, Nucl. Gen. **6**, 987 (1973).
[2] J. M. Charap, J. Phys. A: Math. Gen. **9**, 1331 (1976).
[3] D. Ray, J. Phys. A: Math. Gen. **11**, (5), 995 (1978).
[4] J. Corones, J. Math. Phys. **19**, 2431 (1978).
[5] D. Ray, J. Math. Phys. **23**, 2155 (1982).
[6] S. Trimper, Phys. Lett. A **70**, 114 (1979).
[7] D. Ray, Phys. Lett. A **76**, 274 (1980).
[8] V. L. Ginzburg and L. V. Pitaevski, Zh. Eksp. Theor. Fiz. **34**, 1240 (1958). [Sov. Phys. JETP **7**, 858 (1958)].
[9] L. V. Kinknade and Yu. G. Mamaladze, Phys. Lett. A **72**, 351 (1979).
[10] P. K. Chanda and D. Ray, J. Math. Phys. **24**, 1923 (1983).

# Tight bounds for the critical screening parameters of the generalized exponential cosine screened Coulomb potential

Ranabir Dutt
*Department of Physics, Visva-Bharati University, Santiniketan-731235, West Bengal, India*
Y. P. Varshni
*Department of Physics, University of Ottawa, Ottawa, Canada K1N 6N5*

(Received 20 December 1983; accepted for publication 6 April 1984)

Analytic expressions have been obtained for the lower and upper bounds to the critical screening parameter $\lambda_c$ associated with the ground state of the generalized exponential cosine screened Coulomb potential, $V(r) = -e^{-\lambda r}\cos(\epsilon\lambda r)/r$ employing the methods of Hulthén and Laurikainen and of Hemmer. It is found that the predicted numerical results for the range $0 \leqslant \epsilon \leqslant 1$ for the lower bound for $\lambda_c$ are within 0.08% of the exact results, and for the upper bound, within 0.7% of the exact results. It emerges further that the critical screening parameters for the excited $s$-states can be determined as well in an approximate way. For the static screened Coulomb potential, we obtain $\lambda_c^{ns} = 1.283\,664n^{-2} - 0.092\,793n^{-3}$, where $n$ is the principal quantum number. The predicted $\lambda_c$ for various quantum states ($n = 1$ to 9) are in excellent agreement with the values obtained numerically by Rogers *et al.* [Phys. Rev. A **1**, 1577 (1970)].

PACS numbers: 31.15. + q

## I. INTRODUCTION

The generalized exponential cosine screened Coulomb (GECSC) potential, mathematically expressed as

$$V(r) = -Ze^2e^{-\lambda r}\cos(\epsilon\lambda r)/r, \tag{1.1}$$

where $\lambda$ is the screening parameter, represents a wide class of potentials often encountered in plasma physics,[1,2] nuclear,[3,4] and solid state physics.[5–10] Typical examples of this class of potentials are the well-known Yukawa or the static screened Coulomb potential (SSCP),

$$V^{\mathrm{SSCP}}(r) = -Ze^2e^{-\lambda r}/r, \tag{1.2}$$

and the exponential cosine screened Coulomb (ECSC) potential,

$$V^{\mathrm{ECSC}}(r) = -Ze^2e^{-\lambda r}\cos(\lambda r)/r, \tag{1.3}$$

which have received considerable attention in recent years. A number of workers have studied the bound state properties of these potentials using a variety of approximate analytic methods.[11–20] However, all these methods have the common feature that the numerical accuracy of the predicted bound state energies decreases rapidly as the screening parameter approaches its critical value $\lambda_c$. Particular emphasis has been devoted to the calculation of the critical screening parameter $\lambda_c$,[21–26] the value of $\lambda$ for which the state in question is just bound. A parameter of special interest is the critical screening parameter for the ground state which is the largest value of the screening parameter $\lambda$ beyond which no state will be bound by the potential.

The parameter $\lambda_c$ for the ground state has been determined quite accurately for the SSCP by several analytic techniques.[21–24] Hulthén and Laurikainen[27] first obtained the accurate lower bound to $\lambda_c$ using the Rayleigh–Ritz variational method. On the other hand, a method for obtaining the upper bound was suggested by Dyson and Lenard[28] based on the requirement of non-negativity of the Yukawa Hamiltonian. Although this bound was subsequently improved by Piepenbrink,[29] the best result was obtained by Hemmer[30] who invoked an elegant iterative procedure as applied to an integral equation. From the third-order iterative calculation, he obtains

$$\lambda_c^{1s} \simeq 2\left[1 + 6\log 2 - \tfrac{3}{2}\log 3\right]^{1/3}, \tag{1.4}$$

where the equality sign corresponding to the upper bound gives result accurate to within 1%. However, no such precise bounds to $\lambda_c$ have been derived either for the ECSC potential or any other variant of the GECSC potential in (1.1). The reported analytic results for $\lambda_c$ in the ground state of the ECSC potential fall in the range of about 3% accuracy in comparison to the nearly exact results obtained by Singh and Varshni[26] using numerical methods. Very recently, Lai[20] has shown that even the Padé approximant method fails to give accurate result for $\lambda_c$ for the ECSC potential.

The purpose of the present paper is to demonstrate that a very accurate lower as well as upper bound to $\lambda_c$ can be obtained for the screened Coulomb potential even when it contains in general a nonvanishing oscillatory (cosine) term as in (1.1). For lower and upper bound calculations, we essentially extend the techniques discussed in Refs. 27 and 30, respectively. The details of these calculations are presented in Secs. II and III, where we also discuss the numerical accuracy of our predicted lower and upper bounds to $\lambda_c^{1s}$ for various values of $\epsilon$ in comparison to the exact result obtained numerically.[26] The accuracy of the present work is found to be very good.

Finally, we discuss the possibility of determining the critical screening parameters, $\lambda_c^{ns}$ for higher excited $s$-states. It emerges that this can be achieved in a consistent manner only for the SSCP potential (i.e., for $\epsilon = 0$). The expression thus obtained predicts excellent results for excited states, $n = 1$ to 9.

## II. VARIATIONAL CALCULATION FOR THE LOWER BOUND

We begin with the radial Schrödinger equation for the $s$-wave bound states of the GECSC potential (in atomic units)

 0022-2488/84/082563-04$02.50  

$$\frac{d^2\chi}{dr^2} + 2\left(E + \frac{e^{-\lambda r}}{r}\cos(\epsilon\lambda r)\right)\chi(r) = 0, \tag{2.1}$$

where the radial function $\chi(r) = rR(r)$ satisfies the boundary condition

$$\chi(0) = \chi(\infty) = 0. \tag{2.2}$$

Changing the variable

$$x = \lambda r, \tag{2.3}$$

Eq. (2.1) becomes

$$\frac{d^2\chi}{dx^2} + \frac{2}{\lambda}\left(E + \frac{e^{-x}}{x}\cos(\epsilon x)\right)\chi(x) = 0. \tag{2.4}$$

For $\lambda = \lambda_c$, the energy eigenvalue $E$ vanishes identically. The non-negativity of the energy then leads to the condition[29]

$$2/\lambda_c = (J/N)_{\text{minimum extremal}}, \tag{2.5}$$

where

$$J = \int_0^\infty \chi'^2\, dx, \tag{2.6}$$

and

$$N = \int_0^\infty \frac{e^{-x}}{x}\cos(\epsilon x)\chi^2(x)dx. \tag{2.7}$$

To compute the least extremum of the ratio $J/N$, we follow a variational principle using the trial function[27]

$$\chi(x) = (1 - e^{-x})\sum_{\nu=0}^{n} h_\nu e^{-\nu x}. \tag{2.8}$$

Minimization with this trial function yields a linear homogeneous system of equations for the parameters $h_\nu$. The lower bound for the screening parameter is the largest value of $\lambda_c$ for which the system determinant of these equations vanishes. Keeping only two or three variational parameters $h_\nu$ we obtain the results given in Table I. They are compared with the "exact" numerical values obtained by Singh and Varshni[26] in Table I for various values of $\epsilon$ between 0 and 1 covering the variety of screened potentials, SSCP to ECSC. For the sake of comparison, we use their parameter $\theta = \tan^{-1}\epsilon$ instead of $\epsilon$ in the table. It will be noticed that the errors in $\lambda_c^{1s}$ from the two-parameter calculation are reduced by a factor between 3 and 7 when three parameters are taken into consideration. Except in the neighborhood of $\theta = 45°$, the discrepancy from the three-parameter calculation is generally less than 0.007%. Even for $\theta = 45°$ (ECSC potential) the present result is much more accurate than previous analytic results.

## III. DETERMINATION OF THE UPPER BOUND

For this purpose, we extend the work of Hemmer[30] which requires us to convert Eq. (2.1) into a Fredholm integral equation with a symmetric kernel.

Changing the variable

$$x = \beta r, \tag{3.1}$$

with

$$\beta = \lambda\sqrt{1+\epsilon^2}, \tag{3.2}$$

and combining the cosine part with the exponential factor in the potential, we get from Eq. (2.1)

$$\beta\frac{d^2\chi}{dx^2} + \left(2E + \frac{1}{x}(e^{-ax} + e^{-bx})\right)\chi(x) = 0, \tag{3.3}$$

in which

$$a = b^* = (1 + i\epsilon)/\sqrt{1+\epsilon^2}. \tag{3.4}$$

For determining the critical screening parameters, we require to set $E = 0$ in Eq. (3.3). The eigenfunctions are then not square integrable since zero energy belongs to the continuous spectrum. However, bounded solutions of Eq. (3.3) with $E = 0$ exist only for a discrete set $\beta_1, \beta_2, \ldots$ corresponding to the critical screening parameters, $\lambda_c^{ns}$ associated with various $s$-states. To determine these discrete values of $\beta$, we recast the Eq. (3.3) into an integral equation incorporating the boundary conditions (2.2),

$$\beta\chi(x) = x\int_x^\infty \frac{\chi(y)}{y}(e^{-ay} + e^{-by})dy + \int_0^x \chi(y)(e^{-ay} + e^{-by})dy. \tag{3.5}$$

Defining

$$\chi(x) = \left(\frac{2x}{e^{-ax} + e^{-bx}}\right)^{1/2}\psi(x), \tag{3.6}$$

TABLE I. Calculated values of lower and upper bounds to $\lambda_c^{1s}$ for the GECSC potential compared with the exact results of Singh and Varshni[26] for various values of $\epsilon$.

| $\theta = \tan^{-1}\epsilon$ (degrees) | $\lambda_c^{1s}$ Numerical results of Singh and Varshni | $\lambda_c^{1s}$ Lower bound (two parameter) | $\lambda_c^{1s}$ Lower bound (three-parameter) | $\lambda_c^{1s}$ Upper bound [Eq. (3.13)] |
|---|---|---|---|---|
| 0 | 1.190 61 | 1.190 52 | 1.190 58 | 1.198 38 |
| 5 | 1.183 15 | 1.138 05 | 1.183 12 | 1.190 82 |
| 10 | 1.161 07 | 1.160 92 | 1.161 02 | 1.168 44 |
| 15 | 1.125 22 | 1.124 99 | 1.125 15 | 1.132 11 |
| 20 | 1.076 98 | 1.076 71 | 1.076 90 | 1.083 27 |
| 25 | 1.018 17 | 1.017 93 | 1.018 11 | 1.025 79 |
| 30 | 0.950 91 | 0.950 82 | 0.950 89 | 0.955 84 |
| 35 | 0.877 46 | 0.877 42 | 0.877 44 | 0.881 70 |
| 40 | 0.800 01 | 0.799 23 | 0.799 84 | 0.803 44 |
| 45 | 0.720 52 | 0.716 74 | 0.719 98 | 0.722 57 |

Eq. (3.5) is transformed to a homogeneous Fredholm integral equation

$$\nu\psi(x) = \int_0^\infty K(x,y)\psi(y)dy, \tag{3.7}$$

where the kernel

$$K(x,y) = \begin{cases} \frac{1}{2}\sqrt{(x/y)}(e^{-ax}+e^{-bx})(e^{-ay}+e^{-by}), & y>x, \\ \frac{1}{2}\sqrt{(y/x)}(e^{-ax}+e^{-bx})(e^{-ay}+e^{-by}), & y<x \end{cases} \tag{3.8}$$

is real, symmetric, and continuous, and

$$\nu = \beta/2. \tag{3.9}$$

Using the iterated kernels $K_{j+1}(x,y)$ defined as[31,32]

$$K_{j+1}(x,y) = \int_0^\infty K_j(x,t)K(t,y)dt, \tag{3.10}$$

one can evaluate the traces

$$T_j = \int_0^\infty K_j(x,x)dx, \tag{3.11}$$

in closed form for several low $j$'s.

Since the kernel is continuous, all the iterated kernels are continuous. It then follows[31]

$$T_j = \sum_{k=1}^{\infty} \nu_k^j \quad (j = 2,3,4,\ldots), \tag{3.12}$$

where $\nu_k$'s are reciprocals of the eigenvalues of the basic kernel. According to Mercer's theorem,[31,32] the relation (3.12) is fine even for the first-order trace, i.e., $j = 1$ provided the kernel is positive (or negative) continuous over the entire space. It is interesting to note that our kernel in (3.8) is positive continuous over the entire range of the variables only for the SSCP, i.e., for $\epsilon = 0$ which means $a = b = 1$. However, for $\epsilon \neq 0$, the signature of the iterated kernel $K_j(x,x)$ may be either positive or negative due to presence of the oscillatory piece.

Using (3.2) and (3.9) in (3.12), one then obtains that the largest value of $\lambda_c$ is bounded as

$$\lambda_c^{1s} \leqslant (2/\sqrt{1+\epsilon^2})\,T_j^{1/j}. \tag{3.13}$$

Clearly, $\lambda_c^{1s}$ can be computed in a sequential manner by obtaining analytic expressions for several orders of $T_j$. The calculations are however tedious and lengthy and hence we present here only the first three traces:

$$T_1 = 1/\sqrt{1+\epsilon^2}, \tag{3.14a}$$

$$T_2 = [1/(1+\epsilon^2)][-1 + 2(1-\epsilon^2)\log 2 - \tfrac{1}{2}(1-\epsilon^2)\log(1+\epsilon^2) + 2\epsilon\tan^{-1}\epsilon], \tag{3.14b}$$

$$T_3 = \frac{1}{4\sqrt{1+\epsilon^2}}\left[\frac{4}{(1+\epsilon^2)} - \frac{9}{2}\left(\frac{1-3\epsilon^2}{1+\epsilon^2}\right)\log 3 + 12\left(\frac{1-\epsilon^2}{1+\epsilon^2}\right)\log 2 - \frac{3}{2}(3-\epsilon^2)\log\left(\frac{9+\epsilon^2}{1+\epsilon^2}\right) + 12\left(\frac{1-\epsilon^2}{1+\epsilon^2}\right)\log\left(\frac{2}{\sqrt{1+\epsilon^2}}\right) - \frac{9}{4}\left\{\left(\frac{1-3\epsilon^2}{1+\epsilon^2}\right)\log\left(\frac{9+\epsilon^2}{1+\epsilon^2}\right) + 2\epsilon\left(\frac{3-\epsilon^2}{1+\epsilon^2}\right)\tan^{-1}\left(\frac{\epsilon}{3}\right)\right\} - \frac{3}{2}\epsilon\left(\frac{1+5\epsilon^2}{1+\epsilon^2}\right)\tan^{-1}\epsilon\right]. \tag{3.14c}$$

TABLE II. Critical screening parameters for the $ns$ states of the static screened Coulomb potential.

| $n$ | $\lambda_c^{ns}$ Ours [Eq. (4.2)] | $\lambda_c^{ns}$ Numerical (Rogers *et al.*[a]) |
|---|---|---|
| 1 | 1.190 87 | 1.190 61 |
| 2 | 0.309 32 | 0.310 27 |
| 3 | 0.139 19 | 0.139 45 |
| 4 | 0.078 78 | 0.078 82 |
| 5 | 0.050 61 | 0.050 58 |
| 6 | 0.035 23 | 0.035 18 |
| 7 | 0.025 93 | 0.025 88 |
| 8 | 0.019 88 | 0.019 83 |
| 9 | 0.015 72 | 0.015 67 |

[a] Ref. 33.

One may easily check that our expressions (3.14) for $\epsilon = 0$ reproduce the corresponding traces obtained by Hemmer[30] for the Yukawa potential.

The calculated values of the upper bound for $\lambda_c^{1s}$ from (3.14c) are compared to the exact values in Table I. It is found that the percent difference with the exact values is always less than 0.7% for $0 \leqslant \epsilon \leqslant 1$.

## IV. CRITICAL SCREENING PARAMETERS FOR THE EXCITED LEVELS

The critical screening parameters for the excited $s$-states can be computed in an approximate way if one assumes on empirical form

$$\lambda_c^{ns} = \alpha_1 n^{-2} + \alpha_2 n^{-3}, \tag{4.1}$$

with unknown parameters $\alpha_1$ and $\alpha_2$. Mercer's theorem (3.12) for $j = 1$ and 2 furnishes two equations for these two unknown parameters. Solving for these we find for the static screened Coulomb potential ($\epsilon = 0$)

$$\lambda_c^{ns} = 1.283\,664n^{-2} - 0.092\,793n^{-3}. \tag{4.2}$$

The screening parameters obtained from (4.2) for $n < 10$ compare very well with the numerical results of Rogers *et al.*[33] (see Table II). No other analytical calculation has so far reported such accurate results for the critical screening parameters for the excited $s$-states of the Yukawa potential.

## ACKNOWLEDGMENT

This work was supported in part by a research grant from the Natural Sciences and Engineering Research Council of Canada to one of the authors (Y.P.V).

[1] G. M. Harris, Phys. Rev. **125**, 1131 (1962).
[2] H. Margenau and M. Lewis, Rev. Mod. Phys. **31**, 569 (1959).
[3] H. Yukawa, Proc. Phys. Math. Soc. Jpn. **17**, 48 (1935).
[4] A. E. S. Green, Phys. Rev. **75**, 1926 (1949).
[5] J. B. Krieger, Phys. Rev. **178**, 1337 (1969).
[6] V. L. Bonch-Bruevich and V. B. Glasko, Dokl. Akad. Nauk SSSR **124**, 1015 (1959) [Sov. Phys. Dokl. **4**, 147 (1959)].
[7] N. Takimoto, J. Phys. Soc. Jpn. **14**, 1142 (1959).
[8] E. C. McIrvine, J. Phys. Soc. Jpn. **15**, 928 (1960).
[9] G. L. Hall, J. Phys. Chem. Solids **23**, 1147 (1962).
[10] V. Bonch-Bruevich and S. Tyablikov, *The Green's Function Method in Statistical Mechanics* (Interscience, New York, 1962).

[11]G. J. Iafrate and L. B. Mendelsohn, Phys. Rev. **182**, 244 (1969); Phys. Rev. A **2**, 561 (1970).
[12]C. R. Smith, Phys. Rev. **134**, A1235 (1964).
[13]C. S. Lam and Y. P. Varshni, Phys. Rev. A **4**, 1875 (1971); **6**, 1391 (1972); **19**, 413 (1979).
[14]B. I. Bunlap and L. Armstrong, Jr., Phys. Rev. A **6**, 1370 (1972).
[15]J. McEnnan, L. Kissel, and R. H. Pratt, Phys. Rev. A **13**, 532 (1976).
[16]M. Grant and C. S. Lai, Phys. Rev. A **20**, 718 (1979); C. S. Lai and W. C. Lin, Phys. Lett. A **78**, 335 (1980).
[17]R. Dutt, Phys. Lett. A **73**, 310 (1979).
[18]P. P. Ray and A. Ray, Phys. Lett. A **78**, 443 (1980); A. Ray and P. P. Ray, Phys. Lett. A **83**, 383 (1981).
[19]R. Dutt, A. Ray, and P. P. Ray, Phys. Lett. A **83**, 65 (1981).
[20]C. S. Lai, Phys. Rev. A **26**, 2245 (1982).
[21]G. J. Iafrate, Phys. Rev. A **8**, 1138 (1973).
[22]N. Bessis, G. Bessis, G. Corbel, and B. Dakhel, J. Chem. Phys. **63**, 3477 (1975).
[23]R. N. Kesarwani and Y. P. Varshni, J. Math. Phys. **19**, 19 (1978), and references given therein.
[24]V. L. Sedov, Sov. Phys. Dokl. **14**, 1140 (1970).
[25]R. Dutt, Phys. Lett. A **77**, 229 (1980).
[26]D. Singh and Y. P. Varshni, Phys. Rev. A **28**, 2606 (1983).
[27]L. Hulthén and K. V. Laurikainen, Rev. Mod. Phys. **23**, 1 (1951).
[28]F. J. Dyson and A. Lenard, J. Math. Phys. **8**, 423 (1967).
[29]J. Piepenbrink, J. Math. Phys. **13**, 1825 (1972).
[30]P. C. Hemmer, J. Math. Phys. **14**, 1140 (1973).
[31]V. I. Smirnov, *A Course of Higher Mathematics* (Pergamon, Oxford, 1964), Vol. IV.
[32]F. G. Tricomi, *Integral Equations* (Interscience, New York, 1957).
[33]F. J. Rogers, H. C. Graboske, Jr., and D. J. Harwood, Phys. Rev. A **1**, 1577 (1970). Also D. Singh and Y. P. Varshni, Phys. Rev. A **29**, 2895 (1984).

# Kinematics of the CS method for the treatment of molecular collisions

G. E. Hahne[a)]
*Aircraft Aerodynamics Branch, NASA Ames Research Center, Moffett Field, California 94035*

(Received 22 November 1983; accepted for publication 23 March 1984)

This paper presents the kinematical part of a proposal for founding the CS method (for one definition, see V. Khare and D. J. Kouri [J. Chem. Phys. **72**, 2017 (1980)]) for the quantum treatment of molecular collisions on a certain testable physical approximation scheme. That is, let the molecules be rigid diatoms $A_1A_2$ and $B_1B_2$, with internal angular momentum $\mathbf{j}_a$ and $\mathbf{j}_b$, respectively; with relative position, momentum, and angular momentum $\mathbf{r}$, $\mathbf{p}$, and $\mathbf{l} = \mathbf{r}\times\mathbf{p}$; and with total angular momentum $\mathbf{J} = \mathbf{j}_a + \mathbf{j}_b + \mathbf{l}$. Then the motivating conjecture is that, in addition to commuting exactly with $\mathbf{J}$, and as a result of dynamical properties not discussed here, the fully off-the-energy-shell $T(E)$ operator approximately commutes with either $\mathbf{r}$, $\hat{r}$, or $\mathbf{J}\cdot\hat{r} = \omega$ = "the $r$-helicity." The principal results obtained in the paper are these: First, the definitions of, and transformations between, certain complete sets of system basis states in which the $r$-helicity $\omega$ or the $p$-helicity $\lambda = \mathbf{J}\cdot\hat{p}$ is diagonal are established by methods similar to those of Jacob and Wick [Ann. Phys. (N.Y.) **7**, 404 (1959)]. Second, it is argued that in several papers in the literature of the CS method an explicitly or presumptively incorrect kinematical law was applied to derive the matrix elements of $T(E)$ operators for atom–molecule collisions in a basis in which $\omega$ was said to be diagonal from the fully on-the-energy-shell matrix elements of a given $T(E)$. It is clear from the contexts that the quantities tested for conservation could not have been the $r$-helicity in a quantum-mechanical sense. Thus, there is no foundation to the corresponding assertions in these papers that, even if the CS method works fairly well, $r$-helicity conservation is usually badly violated. Third, the Wigner–Mackey theory of induced representations of continuous groups and Schur's lemma are applied to determine the limitations (analogous to the Wigner–Eckart theorem for single operators invariant under rotations) that commuting with $\mathbf{J}$ and with either of the sets of operators $\mathbf{r}$, $\hat{r}$, or $\omega$, imposes on the matrix of a $T(E)$ in a basis in which $r$ is diagonal. Fourth, and finally, it is shown that the on-the-energy-shell matrix of a $T(E)$ that commutes with $\mathbf{J}$ and $\mathbf{r}$ has the property that its matrix elements are zero unless the angular momentum transfer $\mathbf{j}_a' + \mathbf{j}_b' - \mathbf{j}_a - \mathbf{j}_b$ (prime indicates post-, no prime indicates precollision) is perpendicular in a quantum sense to $\mathbf{p}' - \mathbf{p}$, a result that corresponds to one obtained for an analogous classical atom–molecule collision by V. Khare *et al.* [J. Chem. Phys. **74**, 2275 (1981)].

PACS numbers: 34.50.Ez, 03.65.Nk, 02.20. + b

## I. INTRODUCTION

The mathematics presented in the following chapters is addressed to the physical problem of the quantum treatment of nonreactive collisions of two simple molecules, or of an atom and a molecule. The primary objective is the derivation of kinematical structure theorems (analogous to the Wigner–Eckart theorem[1] for scalar operators) for the transition operator ($T$-operator) of the two-molecule system where, besides the usual exact conservation laws, certain additional conservation laws are presumed satisfied in a collision. The justification for studying the particular augmented sets of conservation laws treated below derives from their association with the so-called CS method[2–4] for the approximate treatment of molecular collisions.

The remainder of this paper is organized as follows: This introductory section continues with a specification of the elementary kinematics of a system of two rigid diatomic molecules, and with a brief sketch of the ideas, and some relevant aspects, of the CS method. Section II presents a treatment of certain complete sets of observables and states of the two-molecule system, including sets of states in which the so-called "$r$-helicity" or the "$p$-helicity" is diagonal. Section III contains the derivation of the previously mentioned structure theorems, which result if in addition to the usual conservation laws, either (A) the vector $\mathbf{r}$, (B) the unit vector $\hat{r}$, or (C) the $r$-helicity $\omega$ is assumed conserved by, say, the transition operator $T(E)$ [the symbols $\mathbf{r}$, $\omega$, and $T(E)$ are defined below in Fig. 1, Eq. (2), and Eq. (42)]. Section IV contains some remarks on the momentum space realizations of operators that satisfy (A), (B), or (C), and a quantum derivation of a useful structural property of the on-the-energy-shell matrix of a $T(E)$ of type (A). Section V summarizes the salient points of the paper.

Let us abstract from the physical problem only those aspects that are essential for the purposes here, and consider an idealized physical system of two distinguishable, rigid, electrically neutral diatoms $A_1A_2$ and $B_1B_2$. The overall

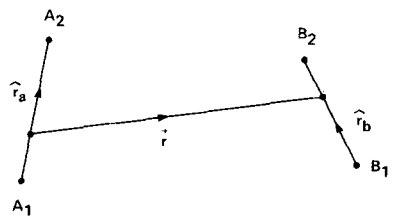



FIG. 1. Collision kinematics.

[a)] National Research Council Associate, 1981–1982.

 0022-2488/84/082567-09$02.50  

center-of-mass (CM) motion of the system will be ignored. There are then seven degrees of freedom, as portrayed in Fig. 1: $\hat{r}_a$ and $\hat{r}_b$ are unit vectors pointing from $A_1$ to $A_2$ and from $B_1$ to $B_2$, respectively, and $\mathbf{r} = r\hat{r}$ is the vector from the internal CM of $A_1 A_2$ to the internal CM of $B_1 B_2$. Let $\mathbf{j}_a$ and $\mathbf{j}_b$ be the internal "dumbbell" angular momentum vector of $A_1 A_2$ and $B_1 B_2$, respectively (in the atom–diatom case we set $\mathbf{j}_a = \mathbf{0}$), and let $\mathbf{p}$ be the momentum conjugate to $\mathbf{r}$. We define

$$\begin{aligned} \mathbf{l} &= \mathbf{r}\times\mathbf{p}, \\ \mathbf{j}_{ab} &= \mathbf{j}_a + \mathbf{j}_b, \\ \mathbf{J} &= \mathbf{j}_{ab} + \mathbf{l}, \end{aligned} \tag{1}$$

where $\mathbf{l}$ and $\mathbf{J}$ will be called the system's orbital and total angular momentum, respectively. We also define certain types of helicity for the system:

$$\begin{aligned} &\omega = \mathbf{J}\cdot\hat{r} = \mathbf{j}_{ab}\cdot\hat{r}, \quad \lambda = \mathbf{J}\cdot\hat{p} = \mathbf{j}_{ab}\cdot\hat{p}. \\ &\omega_a = -\mathbf{j}_a\cdot\hat{r}, \quad \omega_b = \mathbf{j}_b\cdot\hat{r}, \\ &\lambda_a = -\mathbf{j}_a\cdot\hat{p}, \quad \lambda_b = \mathbf{j}_b\cdot\hat{p}, \end{aligned} \tag{2}$$

where $\omega$ and $\lambda$ will be termed "the $r$-helicity" and "the $p$-helicity" for the system, respectively. The operators $\omega_a$, $\omega_b$ and $\lambda_a, \lambda_b$ are individual particle $r$-helicities and $p$-helicities; these quantities are sometimes used,[5–8] but we shall make no use of them here.

The dynamics of the system will be assumed controlled by a Hamiltonian $H$ of the form (note: $\hbar = 1$)

$$\begin{aligned} &H = H_0 + V, \quad H_0 = K + H_a + H_b, \\ &K = \left(\frac{p^2}{2\mu}\right) = -\frac{1}{2\mu}\left(\frac{1}{r}\frac{\partial^2}{\partial r^2}r - \frac{\mathbf{l}\cdot\mathbf{l}}{r^2}\right), \\ &\mu = (m_{A_1} + m_{A_2})(m_{B_1} + m_{B_2})/(m_{A_1} + m_{A_2} + m_{B_1} + m_{B_2}), \\ &H_a = h_a(\mathbf{j}_a\cdot\mathbf{j}_a), \quad H_b = h_b(\mathbf{j}_b\cdot\mathbf{j}_b), \\ &V = V[\hat{r}_a\cdot\hat{r}, \hat{r}_b\cdot\hat{r}, (\hat{r}_a\times\hat{r})\cdot(\hat{r}_b\times\hat{r}), r]. \end{aligned} \tag{3}$$

In Eq. (3), $\mu$ is the reduced mass of the system; the potential function $V$, of the Born–Oppenheimer type, is presumed to be known at the outset.

The objective of the theoretical treatment of this system is to obtain sufficiently accurate wave functions and $S$-matrices for collisions of $A_1 A_2$ and $B_1 B_2$ over a physically interesting range of values of total energy $E$ and total angular momentum. The attempt to solve the Schrödinger equation $H\Psi = E\Psi$ typically involves expanding $\Psi$ in simultaneous eigenstates of $j_a^2$, $j_b^2$, $j_{ab}^2$, $l^2$, $J^2$, and $J_3$, and leads to an infinite set of so-called coupled-channel or close-coupling (CC) ordinary differential equations (ODE) in the continuous variable $r$ (Ref. 6). For the great majority of pairs of chemical species and energies of interest, it does not appear to be physically justified merely to truncate these sets to a tractable size. Thus, the proliferation of rotational states may easily lead to the apparent necessity of coupling $10^3$ channels or more[9]; these sets cannot be treated in full by any existing computer and solution algorithm, nor are they likely to be treated in full any time soon.

Kouri,[2] Gianturco,[3] and Dickinson[4] each reviewed the state of development of this field in 1979. The directions taken by many investigations of this problem have been (1) the search for approximate conservation laws for the system that would allow breaking up the large sets of CC equations into smaller, approximately independent subsets; and (2) the search for scaling laws, or approximate simple regularities, among the submatrices of the wave-function matrix or $S$-matrix for the system.

The "coupled states" or "centrifugal sudden" (CS) method was proposed independently by McGuire and Kouri[10] and by Pack[11] in 1974. A considerable body of literature published since those first papers has dealt with advancing the theory of, or reporting numerical applications of, the CS method. If the results and costs of accurate CC calculations are used as a standard, the CS method has compiled a respectable record of yielding good results with markedly lower costs (see Refs. 2–4 for details and references).

The CS method, in its original form,[10,11] achieved a partial decoupling of the full CC equations by presuming an approximate conservation law, that of the $r$-helicity $\omega$ in an atom–diatom collision. Secrest[12] later proposed a theoretical scheme in which the kinetic energy operator $K$ is modified at the outset by replacing its constituent operator $l^2$ with a constant operator $\bar{l}(\bar{l}+1)I$ ($I$ is the identity operator) so that $K$, and thus H, conserve the unit vector $\hat{r}$. The initial assumption of constant $\bar{l}$ is typically dropped at a later stage of the calculation in favor of one or another so-called labeling scheme. A formal set of rules for obtaining approximate wave functions and $S$-matrices has been developed (see Refs. 2 and 13 for details).

The possible physical relevance of a third level of approximation, that of $\mathbf{r}$ conservation in a useful class of molecular collisions, may be inferred from a series of papers by Khare *et al.*,[14–17] who observed that in a classical atom–molecule collision that is purely impulsive (in that an entire interaction takes place at a single value of $\mathbf{r}$), the internal angular momentum transfer from atom to molecule is perpendicular to the linear momentum transfer (see Sec. IV for a proof of a quantum-mechanical version of this result). Khare *et al.* then transformed some $S$-matrices obtained from accurate CC calculations to show the near-perpendicularity of these quantities in a number of special cases of collisions[15–17] (however, these investigators proposed a different, and somewhat formal, explanation of this phenomenon on the quantum level[14]).

The above remarks are intended to give some idea why the particular augmented sets of conservation laws treated in Secs. III and IV may be involved with the presently unknown[18] mathematical and physical mechanisms that make the CS method work as well as it does. The reader is cautioned that the literature of the CS method contains a number of incorrect statements concerning the approximate validity of $r$-helicity conservation in atom–molecule collisions; namely, it has been asserted that, with few exceptions, the $r$-helicity is not conserved quantum-mechanically, even if a variant of the CS method gives good results. We cite the following representative cases in this regard: Ref. 19, Chap. II B; Ref. 20, Chap. V B; Refs 21 and 22, passim; and the summary statement in Ref. 2, p. 346. In each of these cases, careful scrutiny of the bases for these negative outcomes shows that the authors have inspected certain on-the-energy-shell matrix elements of the $T(E)$ operator in which they

employ a kind of $p$-helicity rather than $r$-helicity. Equations (24) or (26) of Sec. II, or an equivalent form, must be used to transform between $r$- and $p$-helicity states, and an on-the-energy-shell $T(E)$ matrix contains insufficient information to infer the detailed structure of this operator in a basis in which $\omega$ is diagonal; conversely, as remarked in Sec. IV, it is questionable whether there exist any useful general regularities of on-the-energy-shell matrix elements of a $T(E)$ that exactly commutes with $\omega$, or even commutes with the unit vector $\hat{r}$.

The apparent source of the incorrect assertions concerning $r$-helicity conservation in Refs. 2 and 19–22 is a highly simplified and, at least in the present context, inadequate conception of the relationship between $r$- and $p$-helicity. For example, the lines following Eq. II 22 of Ref. 22 are an explicit and generally incorrect statement: "Thus, the only difference between the $P$ and $R$ helicities is a change in the sign of the initial state label. This sign is a matter of convention and cannot affect any observable quantity." Equation (25), which is the asymptotic form for large $rp$ and fixed $J$ of a transformation amplitude between $r$-helicity and $p$-helicity states, yields a result that is plausibly related to this statement of Ref. 22 which apparently derives from a classical picture of the asymptotic relation between these helicities. But even if the factors $\exp(\pm irp)$ are retained. Equation (25) cannot generally be used to carry out the desired transformation, since physically, the matrix elements $\langle \cdots \omega' \mathbf{r}' | T(E) | \cdots \omega \mathbf{r} \rangle$ (see Secs. II and IV for the notation) are likely to be very small in magnitude, except for a range of moderate to small values of both $r'$ and $r$. For a classical straight-line trajectory with impact parameter $b$ we have $bp \approx l$, so that for fixed $b$, as $p$ becomes large, $l$ becomes large as well, and there is then no guarantee that the approximation Eq. (25) will be uniformly valid for the entire region of $r'$, $r$-space for which the magnitude of $T(E)$ is not small. Thus, except possibly for very small $J$ values, one cannot expect that Eq. (25) will be a useful approximation of Eq. (24) in transforming $T(E)$ from an $\omega JMr$ basis to a basis (say) $\langle \cdots \lambda' J' M' p' | T(E) | \cdots \lambda JMp \rangle$.[23] Accordingly, the presumption of Ref. 22 for transforming beween $p$- and $r$-helicity states is generally invalid, and the assertions of Refs. 2 and 19–22 concerning the violation of $r$-helicity conservation in certain cases of molecular collisions are without foundation. A more thorough treatment of these latter points will be given in a later publication, in connection with a discussion of the physics of $\omega$, $\hat{r}$, and $\mathbf{r}$ conservation.

In summary, the present paper is part of an investigation that followed from the observation that not only is the case not proven against approximate conservation of $r$-helicity or $\hat{r}$ in connection with successful uses of the CS method, but proper tests for the presence of these seemingly formal properties would likely confirm their approximate and useful validity in a range of interesting physical cases.

## II. CONSTRUCTION OF SYSTEM EIGENKETS OF $r$-HELICITY OR OF $p$-HELICITY

We shall consider the following complete sets of commuting observables and the associated sets of basis kets for the system of Fig. 1:

(i-$r$) $\hat{r}_a, \hat{r}_b, \mathbf{r}$: $|\hat{r}_a\, \hat{r}_b \mathbf{r}\rangle$,

(ii-$r$) $j_a^2, j_b^2, j_{ab}^2, (j_{ab})_3, \mathbf{r}$: $|j_a\, j_b\, j_{ab}\, m_{ab} \mathbf{r}\rangle$,

(iii-$r$) $j_a^2, j_b^2, j_{ab}^2, l^2, J^2, J_3, r$: $|j_a\, j_b\, j_{ab}\, lJMr\rangle$,

(iv-$r$) $j_a^2, j_b^2, j_{ab}^2, \omega, \mathbf{r}$: $|j_a\, j_b\, j_{ab}\, \omega \mathbf{r}\rangle$,

(v-$r$) $j_a^2, j_b^2, j_{ab}^2, \omega, J^2, J_3, r$: $|j_a\, j_b\, j_{ab}\, \omega JMr\rangle$.

Corresponding to each of the above bases, moreover, is a complementary (dual) basis in which the unsubscripted symbol $r$ is replaced by $p$ and $\omega$ by $\lambda$ (however, $l$ remains the same). The latter bases will be denoted by (i-$p$),...,(v-$p$), respectively. We shall normalize the (i-$r$) and (i-$p$) states such that

$$\langle \hat{r}_a' \hat{r}_b' \mathbf{r}' | \hat{r}_a \hat{r}_b \mathbf{r} \rangle = \delta^2(\hat{r}_a' - \hat{r}_a)\delta^2(\hat{r}_b' - \hat{r}_b)\delta^3(\mathbf{r}' - \mathbf{r}), \tag{4}$$

$$\langle \hat{r}_a' \hat{r}_b' \mathbf{r} | \hat{r}_a \hat{r}_b \mathbf{p} \rangle = \delta^2(\hat{r}_a' - \hat{r}_a)\delta^2(\hat{r}_b' - \hat{r}_b)(2\pi)^{-3/2} \times \exp(i\mathbf{r} \cdot \mathbf{p}). \tag{5}$$

In Eqs. (4) and (5), $\delta^2$ is the Dirac delta function on the unit two-sphere $S^2$.

In what follows, Edmonds'[24] notation $\langle j_1 m_1 j_2 m_2 | j_1 j_2 j m \rangle$ and

$$\begin{pmatrix} j_1 & j_2 & j_3 \\ m_1 & m_2 & m_3 \end{pmatrix}$$

will be used for Clebsch–Gordan coefficients and $3j$ coefficients, respectively, and Brink and Satchler's[25] conventions for spherical harmonics $Y_{lm}(\hat{r})$ and rotation matrices $\mathscr{D}^{j}_{m'm}(\alpha,\beta,\gamma)$ will be followed, with the proviso that all rotations and translations are to be understood as point transformations rather than as coordinate transformations.

The orthonormal bases (ii-$r$) and (iii-$r$) may be defined by the following well-known unitary transformations:

$$|j_a\, j_b\, j_{ab}\, m_{ab} \mathbf{r}\rangle = \sum_{m_a m_b} \langle j_a m_a\, j_b m_b | j_a\, j_b\, j_{ab}\, m_{ab} \rangle \times \int d^2\hat{r}_a \int d^2\hat{r}_b\; Y_{j_a m_a}(\hat{r}_a) \times Y_{j_b m_b}(\hat{r}_b) | \hat{r}_a \hat{r}_b \mathbf{r} \rangle, \tag{6}$$

and

$$|j_a\, j_b\, j_{ab}\, lJMr\rangle = \sum_{m_{ab} m} \langle j_{ab} m_{ab}\, lm | j_{ab}\, lJM \rangle \times \int d^2\hat{r}\; Y_{lm}(\hat{r}) | j_a\, j_b\, j_{ab} m_{ab} \mathbf{r} \rangle. \tag{7}$$

In order to construct the bases (iv-$r$) and (v-$r$), it is useful to associate with each $\hat{r} \in S^2$ a $3 \times 3$ rotation matrix

$$R(\hat{r}) \overset{\text{def}}{=} R[\alpha(\hat{r}), \beta(\hat{r}), \gamma(\hat{r})] \overset{\text{def}}{=} R_z[\alpha(\hat{r})] R_y[\beta(\hat{r})] R_z[\gamma(\hat{r})],$$

such that the $3 \times 3$ matrix $R(\hat{r})$ maps $\hat{e}_z$ into $\hat{r}$, that is, such that $R_{i3}(\hat{r}) = (\hat{r})_i$. This association cannot be made continuous for all $\hat{r} \in S^2$; we need $\alpha(\hat{r}) = \phi$ and $\beta(\hat{r}) = \theta$, where $\theta, \phi$ are the spherical polar coordinates of $\hat{r}$, and we choose $\gamma(\hat{r}) = 0$ [the choice $\gamma(\hat{r}) = -\phi$ is sometimes made[5,26]]. Ac-

cordingly, for each $\hat{r} \neq \pm \hat{e}_z$ it is meaningful to define unique rotation matrices as follows:

$$\mathscr{D}^{j}_{m'm}(\hat{r}) \overset{\text{def}}{=} \mathscr{D}^{j}_{m'm}(\phi,\theta,0). \tag{8}$$

Let us now define certain unitary operators $U$ and $W$ in the space of system states, by their (i-$r$) and (i-$p$) matrix elements, respectively, as follows:

$$\langle \hat{r}'_a \hat{r}'_b \mathbf{r}' | U | \hat{r}_a \hat{r}_b \mathbf{r} \rangle = \delta^2[\hat{r}'_a - R(\hat{r})\hat{r}_a]\delta^2[\hat{r}'_b - R(\hat{r})\hat{r}_b]\delta^3(\mathbf{r}' - \mathbf{r}), \tag{9}$$

and

$$\langle \hat{r}'_a \hat{r}'_b \mathbf{p}' | W | \hat{r}_a \hat{r}_b \mathbf{p} \rangle = \delta^2[\hat{r}'_a - R(\hat{p})\hat{r}_a]\delta^2[\hat{r}'_b - R(\hat{p})\hat{r}_b]\delta^3(\mathbf{p}' - \mathbf{p}). \tag{10}$$

Given a basis such that $\mathbf{r}$ or, respectively, $\mathbf{p}$, is diagonal, $U$ or $W$ is block diagonal and may be represented in each subspace by a product of exponentials[26]:

$$\langle \cdots \mathbf{r}' | U | \cdots \mathbf{r} \rangle = \exp[-i\phi\hat{e}_z \cdot \mathbf{j}_{ab}] \exp[-i\theta\hat{e}_y \cdot \mathbf{j}_{ab}] \delta^3(\mathbf{r}' - \mathbf{r}), \tag{11}$$

where $\mathbf{j}_{ab}$ is understood to be a set of matrices. A similar result holds for $W$ in terms of the spherical polar coordinates of $\hat{p}$, in a basis in which $\mathbf{p}$ is diagonal. By means of calculations similar to those of Goldberger and Watson,[26] it can now be shown that if $\zeta$ is any of the vector operators $\hat{r}_a$, $\hat{r}_b$, $\mathbf{j}_a$, $\mathbf{j}_b$, or $\mathbf{j}_{ab}$, then for any r-compatible basis we have

$$\langle \cdots \mathbf{r}' | U \zeta_i U^{\dagger} | \cdots \mathbf{r} \rangle = \sum_j \langle \cdots | \zeta_j | \cdots \rangle R_{ji}(\hat{r}) \delta^3(\mathbf{r}' - \mathbf{r}). \tag{12}$$

Thus, in particular,

$$U(j_{ab})_3 U^{\dagger} = \omega, \tag{13}$$

and, similarly,

$$W(j_{ab})_3 W^{\dagger} = \lambda. \tag{14}$$

The results of Eqs. (13) and (14) show that the kets of the basis (iv-$r$) or (iv-$p$) may be defined in terms of (ii-$r$) and (ii-$p$) kets, respectively, as follows:

$$|j_a j_b j_{ab}\ \omega\mathbf{r}\rangle_{(\mathrm{iv})} = U |j_a j_b j_{ab}\ \omega\mathbf{r}\rangle_{(\mathrm{ii})}, \tag{15}$$

$$|j_a j_b j_{ab} \lambda\mathbf{p}\rangle_{(\mathrm{iv})} = W |j_a j_b j_{ab} \lambda\mathbf{p}\rangle_{(\mathrm{ii})}, \tag{16}$$

where the kets have been subscripted for clarity.

It follows from Eqs. (9) and (15), and from the property [Ref. 25, p. 28 and p. 53 (see footnote)]

$$Y_{kq}[R(\alpha,\beta,\gamma)^{-1} \cdot (\theta,\phi)] = \sum_{q'} Y_{kq'}(\theta,\phi) \mathscr{D}^{k}_{q'q}(\alpha,\beta,\gamma), \tag{17}$$

of the spherical harmonics that

$$|j_a j_b j_{ab}\ \omega\mathbf{r}\rangle = \sum_{m_{ab}} |j_a j_b j_{ab} m_{ab} \mathbf{r}\rangle \mathscr{D}^{j_{ab}}_{m_{ab}\omega}(\hat{r}); \tag{18}$$

a corresponding result holds in the dual case.

We define for each fixed $j$ and $J$, the real, orthogonal matrix $M_{jJ}$ (with indices $l$ and $\omega$)

$$M_{jJ}(l,\omega) = ([l]/[J])^{1/2} \langle j\omega l\, 0 | jlJ\omega \rangle, \tag{19}$$

where $(2x+1)$ is abbreviated by $[x]$. For any fixed $\omega = 0, \pm 1, \pm 2, \ldots$, the set of complex-valued functions $([J]/4\pi)^{1/2} \mathscr{D}^{J}_{M\omega}(\hat{r})^*$ with domain $S^2$ for all $J \geqslant |\omega|$ and all $|M| \leqslant J$, forms a complete orthonormal set of functions on $S^2$.

Further manipulations of Eqs. (18) and (7) can now be made to yield

$$|j_a j_b j_{ab}\ lJMr\rangle = \sum_{\omega} \int d^2\hat{r} |j_a j_b j_{ab}\ \omega\mathbf{r}\rangle \left(\frac{[J]}{4\pi}\right)^{1/2} \times \mathscr{D}^{J}_{M\omega}(\hat{r})^* M_{j_{ab}J}(l,\omega). \tag{20}$$

With the use of Eq. (20) it can be shown that the operator $\omega$ is diagonal, and has the correct eigenvalues for a set of basis states $|j_a j_b j_{ab}\ \omega JMr\rangle$ of type (v-$r$), if the states are defined by

$$|j_a j_b j_{ab}\ \omega JMr\rangle = \sum_l |j_a j_b j_{ab}\ lJMr\rangle M_{j_{ab}J}(l,\omega). \tag{21}$$

Correspondingly, in the dual case we define (v-$p$) states by

$$|j_a j_b j_{ab} \lambda JMp\rangle = \sum_l |j_a j_b j_{ab}\ lJMp\rangle M_{j_{ab}J}(l,\lambda). \tag{22}$$

Together, the orthogonality of the $M_{jJ}$ matrix and Eq. (20) entail that

$$|j_a j_b j_{ab}\ \omega\mathbf{r}\rangle = \sum_{JM} |j_a j_b j_{ab}\ \omega JMr\rangle \left(\frac{[J]}{4\pi}\right)^{1/2} \mathscr{D}^{J}_{M\omega}(\hat{r}). \tag{23}$$

We also observe that, together, Eqs. (5)–(7), (21), and (22) imply that

$$\langle j'_a j'_b j'_{ab}\ \omega J'M'r | j_a j_b j_{ab} \lambda JMp\rangle = \delta_{j'_a j_a}\delta_{j'_b j_b}\delta_{j'_{ab} j_{ab}}\delta_{J'J}\delta_{M'M} \times \sum_l M_{j_{ab}J}(l,\omega) M_{j_{ab}J}(l,\lambda) i^l \sqrt{2/\pi} j_l(rp), \tag{24}$$

where $j_l$ is a spherical Bessel function.[27] Equation (24) reduces for $\mathbf{j}_a = \mathbf{0}$ to a formula of Moses.[28] For large $rp$, $j_l(rp) \sim (rp)^{-1} \sin(rp - l\pi/2)$, and as a result the sum over $l$ on the right-hand side (rhs) of Eq. (24) becomes

$$(rp)^{-1}[\delta_{\omega\lambda} \exp(irp) - (-1)^{J + j_{ab}} \delta_{\omega,-\lambda} \exp(-irp)], \tag{25}$$

which entails the physically reasonable result that for large $rp$ and fixed $J$, $\omega$ and $\lambda$ are, respectively, parallel and antiparallel to one another for the outgoing and ingoing parts of the wave corresponding to the state $|j_a j_b j_{ab}\ \lambda JMp\rangle$.[29] Moreover, Eq. (18) and its dual, Eq. (5), and Eq. (6) and its dual, imply that

$$\langle j'_a j'_b j'_{ab}\ \omega\mathbf{r} | j_a j_b j_{ab}\ \lambda\mathbf{p}\rangle = \delta_{j'_a j_a}\delta_{j'_b j_b}\delta_{j'_{ab} j_{ab}} (2\pi)^{-3/2} \exp(i\mathbf{r} \cdot \mathbf{p}) \times \sum_{m_{ab}} \mathscr{D}^{j_{ab}}_{m_{ab}\omega}(\hat{r})^* \mathscr{D}^{j_{ab}}_{m_{ab}\lambda}(\hat{p}). \tag{26}$$

The summation on the rhs of Eq. (26) simplifies if $\hat{r} = +\hat{p}$ to $\delta_{\omega\lambda}$ and if $\hat{r} = -\hat{p}$ to $(-1)^{j_{ab}}\delta_{\omega,-\lambda}$; these results are in accord with an intuitive picture of the significance of $r$- and $p$-helicities.

## III. STRUCTURE OF SYSTEM OPERATORS THAT COMMUTE WITH J AND WITH r, $\hat{r}$, OR $\omega$

Let us consider a certain operator $T$ for the system of two rigid rotors; $T$ may be an approximation to the $T(E)$ operator for the system [see Eq. (42), below]. We presume that $T$ conserves each component of the total angular mo-

mentum of the system, that is, that $[\mathbf{J},T]=\mathbf{0}$. Let us suppose, moreover, that as a result of some unspecified physical considerations, an operator $T$ has been obtained that also conserves either (A) $\mathbf{r}$, (B) $\hat{r}$, or (C) $\omega$. We shall now undertake to derive the limitations that these commutation relations impose on the matrix structure of $T$ in the bases (iv-$r$) and (v-$r$) described in Sec. II.

In view of the well-known commutation relations among the components of $\mathbf{J}$, $\mathbf{r}$, $\hat{r}$, and $\omega$, each of the sets (A) $\{J_1,J_2,J_3,r_1,r_2,r_3\}$, (B) $\{J_1,J_2,J_3,(\hat{r})_1,(\hat{r})_2,(\hat{r})_3\}$, and (C) $\{J_1,J_2,J_3,\omega\}$ is a basis for a Lie algebra. In case (A), the algebra comprises the set of generators for the overall system rotations and for the translations in $\mathbf{p}$-space, and there are two fundamental Casimir operators,[30] $\mathbf{J}\cdot\mathbf{r}=r\omega$ and $\mathbf{r}\cdot\mathbf{r}=r^2$. Note that the operator $\omega$ generates rotations of the system around the axis $\hat{r}$.

The Lie algebra properties of the sets (A), (B), and (C) suggest the use of group-theoretical means to analyze the problem of describing the possible structures of the matrix of the operator $T$. It is convenient to treat case (A) first, and then to specialize the Lie algebra and generalize the operator $T$, to cases (B) and (C).

Our approach will be based on the theory of induced representations of the group associated with the six-dimensional Lie algebra of case (A) [the so-called Euclidean group, or inhomogeneous rotation group IR(3)] of rigid motions of Euclidean three-space $E^3$ into itself. The earliest application of induced representations to noncompact continuous groups was performed by Wigner,[31] who studied the Poincaré group. Mackey[32–34] and others[35] greatly generalized the theory, and it is Mackey's methods, reduced to an explicit coordinate form, that inform the steps to follow; related material may be found in Refs. 30 and 36–42.

We introduce coordinates into IR(3) as follows: Let $\mathbf{q}$ be a point in momentum space; a general operation $(\phi,\theta,\psi,\mathbf{p})\in\mathrm{IR}(3)$ will consist of a translation followed by rotations, such that $\mathbf{q}$ is mapped into $\mathbf{q}_1$, as follows:

$$\mathbf{q}_1=R_z(\phi)R_y(\theta)R_z(\psi)(\mathbf{q}+\mathbf{p}). \tag{27}$$

The specification of Eq. (27) serves to determine the multiplication law of group elements in terms of the chosen coordinates.

An infinitesimal left group translation consists of multiplying $(\phi,\theta,\psi,\mathbf{p})$ on the left by the group element corresponding to the product (in any order) of an infinitesimal translation by $\delta\mathbf{p}$ and infinitesimal rotations by $R_x(\delta\Omega_1)$, $R_y(\delta\Omega_2)$, and $R_z(\delta\Omega_3)$; we call the result $(\phi+\Delta\phi,\theta+\Delta\theta,\psi+\Delta\psi,\mathbf{p}+\Delta\mathbf{p})$, where Eq. (27) implies that

$$\begin{aligned}
\Delta\phi&=\delta\Omega_3-\cot\theta\,(\delta\Omega_1\cos\phi+\delta\Omega_2\sin\phi),\\
\Delta\theta&=-\delta\Omega_1\sin\phi+\delta\Omega_2\cos\phi,\\
\Delta\psi&=\csc\theta\,(\delta\Omega_1\cos\phi+\delta\Omega_2\sin\phi),\\
\Delta p_j&=\sum_k R_{kj}(\phi,\theta,\psi)\delta p_k.
\end{aligned}$$

The change $\delta F$ in a smooth, complex-valued function $F(\phi,\theta,\psi,\mathbf{p})$ on the group IR(3) under the infinitesimal left translation is defined to be

$$\begin{aligned}
\delta F&=F(\phi-\Delta\phi,\theta-\Delta\theta,\psi-\Delta\psi,\mathbf{p}-\Delta\mathbf{p})-F(\phi,\theta,\psi,\mathbf{p})\\
&=-i\sum_j\left[\delta\Omega_j J_j F-\delta p_j r_j F\right],
\end{aligned}$$

where the last line neglects second-order terms and where

$$\begin{aligned}
J_1&=i\left(\cot\theta\cos\phi\frac{\partial}{\partial\phi}+\sin\phi\frac{\partial}{\partial\theta}-\csc\theta\cos\phi\frac{\partial}{\partial\psi}\right),\\
J_2&=i\left(\cot\theta\sin\phi\frac{\partial}{\partial\phi}-\cos\phi\frac{\partial}{\partial\theta}-\csc\theta\sin\phi\frac{\partial}{\partial\psi}\right),\\
J_3&=-i\frac{\partial}{\partial\phi},\\
r_j&=i\sum_k R_{jk}(\phi,\theta,\psi)\frac{\partial}{\partial p_k}.
\end{aligned} \tag{28}$$

The operators given by Eq. (28) are Hermitian, and the derived group representation is unitary, with respect to smooth, square-integrable functions on IR(3) with a measure of $\sin\theta\,d\phi\,d\theta\,d\psi\,d^3p$.

Now we may apply Mackey's construction to the present case. Note that the subset of elements $(0,0,\psi,\mathbf{p})$ with $\phi=0$ and $\theta=0$, which we call $L$, forms a (nonabelian) subgroup of IR(3). We consider the special, one-dimensional, unitary (with respect to $d\psi\,d^3p$), irreducible representations $\Gamma^{r,\omega}$ of this subgroup, defined by

$$(0,0,\psi,\mathbf{p})\xrightarrow{\Gamma^{r,\omega}}\exp(-i\omega\psi+irp_3), \tag{29}$$

where $\omega=0,\pm1,\pm2,\ldots$, and $r>0$. Correspondingly, let us consider the linear space of functions $F(\phi,\theta,\psi,\mathbf{p})$ on IR(3) of the special type

$$F(\phi,\theta,\psi,\mathbf{p})=f(\phi,\theta)\exp(i\omega\psi-irp_3), \tag{30}$$

where $f(\phi,\theta)$ is complex-valued, square integrable, and smooth on $S^2$. A calculation shows that the linear space of functions of the type of Eq. (30) remains invariant under the action of the six operators of Eq. (28). Moreover, one may verify that the action of the differential operators of Eq. (28) on the functions of the type of Eq. (30) may be construed as an action on the functions[43] $f(\phi,\theta)$ themselves by the operators obtained by the substitutions $\partial/\partial\psi\to i\omega$ and $\partial/\partial p_k\to-ir\delta_{k3}$ on the rhs of Eq. (28); we call the resulting operators $J_j(r,\omega)$ and $r_j(r,\omega)$, where

$$\begin{aligned}
J_1(r,\omega)&=i\left(\cot\theta\cos\phi\frac{\partial}{\partial\phi}+\sin\phi\frac{\partial}{\partial\theta}\right)+\omega\csc\theta\cos\phi,\\
J_2(r,\omega)&=i\left(\cot\theta\sin\phi\frac{\partial}{\partial\phi}-\cos\phi\frac{\partial}{\partial\theta}\right)+\omega\csc\theta\sin\phi,\\
J_3(r,\omega)&=-i\frac{\partial}{\partial\phi},\\
\mathbf{r}(r,\omega)&=(r\sin\theta\cos\phi,r\sin\theta\sin\phi,r\cos\theta).
\end{aligned} \tag{31}$$

These six operators are Hermitian, and the associated group representation, which we call $U^{r,\omega}$, is unitary with respect to square-integrable, complex-valued functions in the coset space $\mathrm{IR}(3)/L\approx S^2$ with the measure $\sin\theta\,d\theta\,d\phi$. The representation $U^{r,\omega}$ is called the representation of the group IR(3) *induced* from the representation given by Eq. (29) of the subgroup $L$. Note that in Eq. (31) the Casimir operators satisfy $\mathbf{J}\cdot\mathbf{r}=r\omega$, and $\mathbf{r}\cdot\mathbf{r}=r^2$. We shall not treat the physically uninteresting case of $r=0$. With this exception, according to Mackey[32–34] the set of unitary representations $U^{r,\omega}$ of IR(3) are all irreducible, and are inequivalent for distinct

pairs $(r',\omega')$ and $(r,\omega)$; moreover, except where $\mathbf{r}$ is represented by $\mathbf{0}$, each equivalence class of UIR's of IR(3) is realized by one of the $U^{r,\omega}$ representations.

To understand the structure of the unitary, irreducible representation (UIR) $U^{r,\omega}$ or IR(3), we can determine how it "subduces" UIR's of the subgroup under restriction to (1) the subgroup of pure rotations [i.e., of elements of IR(3) with $\mathbf{p}$ coordinates $= \mathbf{0}$], and (2) the subgroup of pure translations [i.e., of elements of IR(3) with $\phi = \theta = \psi = 0$]. Let the former subgroup be called $K$ and the latter $Q$, and, following Coleman,[36,37] let $U^{r,\omega}\downarrow K$ be the representation of $K$ obtained by restricting $U^{r,\omega}$ to $K$. It follows from a theorem of Mackey[44] that $U^{r,\omega}\downarrow K$ is equivalent to the representation of $K$ induced from the representation $R_z(\psi) \to \exp(-i\omega\psi)$ of this SO(2)-type subgroup of $K$. An application of the Frobenius reciprocity theorem[32,36,37] then reveals that $U^{r,\omega}\downarrow K$, when decomposed into irreducible constituents, is equivalent to the following infinite direct sum of representations of $K \approx$ SO(3):

$$U^{r,\omega}\downarrow K \cong \mathscr{D}^{|\omega|} \oplus \mathscr{D}^{|\omega|+1} \oplus \mathscr{D}^{|\omega|+2} \oplus \cdots . \tag{32}$$

Note that no $J < |\omega|$ appears in Eq. (32), and each $J \geqslant |\omega|$ appears exactly once. Similarly, Mackey's theorem implies that[44]

$$U^{r,\omega}\downarrow Q = \int_{\text{all } \hat{r}\in S^2}^{\oplus} \underline{\exp[ir\hat{r}\cdot\mathbf{p}]}\, d^2\hat{r}, \tag{33}$$

where $\mathbf{p} \in Q$ and where the rhs of Eq. (33) is a direct integral over all those indicated (underlined) one-dimensional UIR's of $Q$, such that the character-determining vector $\mathbf{r}$ has a fixed magnitude $r$ and takes each possible directional value $\hat{r}$ exactly once.

We may now infer the structure of the operator $T$ in case (A) by an application of Schur's lemma. It is convenient to work in the bases (iv-$r$) or (v-$r$), such that the Casimir operators $\omega$ and $r$ are diagonal. For fixed $j_a, j_b, j_{ab}$, and for given eigenvalues of $r$ and $\omega$, the set of kets

$$\{\,|\,j_a\, j_b\, j_{ab}\omega JMr\rangle\,|\,\forall J \geqslant |\omega|,\quad \forall M \leqslant J\,\}$$

or, equivalently,

$$\{\,|\,j_a\, j_b\, j_{ab}\omega\hat{r}r\rangle\,|\,\forall\hat{r}\in S^2\}$$

provides a linear subspace of states that supports exactly one equivalent form of the UIR $U^{r,\omega}$ of IR(3). Denote the six operators $\mathbf{J}$ and $\mathbf{r}$ by $\xi_\alpha$, $\alpha = 1,2,\ldots,6$. In a (v-$r$) basis the $\xi_\alpha$ must have matrix elements that satisfy

$$\begin{aligned}\langle j'_a\, j'_b\, j'_{ab}\omega' J'M'r'|\xi_\alpha|j_a\, j_b\, j_{ab}\omega JMr\rangle &= \delta_{j'_a j_a}\delta_{j'_b j_b}\delta_{j'_{ab} j_{ab}}\delta_{\omega'\omega}(1/r'r)\delta(r'-r)\\ &\times\langle J'M'\|\xi_\alpha(r,\omega)\|JM\rangle,\end{aligned} \tag{34}$$

here $\langle J'M'\|\xi_\alpha(r,\omega)\|JM\rangle$ are the matrices in a $J, M$ basis of the generators of the UIR $U^{r,\omega}$ of IR(3); if desired, these matrices may be calculated from the $\hat{r}$-basis matrix elements given in Eq. (31). Now let $T$ be a system operator such that for each $\alpha = 1,2,\ldots,6$ we have $[\xi_\alpha, T] = 0$ so that

$$\begin{aligned}0 = \sum_{J''M''} \{&\langle J'M'\|\xi_\alpha(r',\omega')\|J''M''\rangle\\ &\times\langle j'_a\, j'_b\, j'_{ab}\omega' J''M''r'|T|j_a\, j_b\, j_{ab}\omega JMr\rangle\\ &-\langle j'_a\, j'_b\, j'_{ab}\omega' J'M'r'|T|j_a\, j_b\, j_{ab}\omega J''M''r\rangle\\ &\times\langle J''M''\|\xi_\alpha(r,\omega)\|JM\rangle\}.\end{aligned} \tag{35}$$

Hence, for each set of values of $j'_a, j'_b, j'_{ab}, j_a, j_b, j_{ab}$, the corresponding rectangular subblock of $T$-matrix elements intertwines the two UIR's $U^{r',\omega'}$ and $U^{r,\omega}$ of IR(3); by Schur's lemma,[45] then, these matrix elements of $T$ must all be zero if $r' \neq r$ or $\omega' \neq \omega$ and if $r' = r$ and $\omega' = \omega$, then the associated square subblock of $T$ must be a *constant* multiple of the identity matrix. Thus, we have inferred that there must exist reduced matrix elements $\langle j'_a\, j'_b\, j'_{ab}\|T^{r,\omega}\|j_a\, j_b\, j_{ab}\rangle$ such that

$$\begin{aligned}\langle j'_a\, j'_b\, j'_{ab}\omega' J'M'r'|T|j_a\, j_b\, j_{ab}\omega JMr\rangle &= \delta_{\omega'\omega}(1/r'r)\delta(r'-r)\delta_{J'J}\delta_{M'M}\\ &\times\langle j'_a\, j'_b\, j'_{ab}\|T^{r,\omega}\|j_a\, j_b\, j_{ab}\rangle.\end{aligned} \tag{36}$$

We emphasize that the reduced matrix elements on the rhs of Eq. (36) must be independent of both $J$ and $M$.[46] Similarly, in an $\hat{r}$ basis, Schur's lemma implies that

$$\begin{aligned}\langle j'_a\, j'_b\, j'_{ab}\omega'\hat{r}'r'|T|j_a\, j_b\, j_{ab}\omega\hat{r}r\rangle &= \delta_{\omega'\omega}\delta^3(\mathbf{r}'-\mathbf{r})\langle j'_a\, j'_b\, j'_{ab}\|T^{r,\omega}\|j_a\, j_b\, j_{ab}\rangle,\end{aligned} \tag{37}$$

where the reduced matrix must be independent of $\hat{r}$. It follows from Eq. (23) that the reduced matrices on the rhs of Eqs. (36) and (37) must be the same sets of numbers. The results of Eqs. (36) and (37) are the desired structure theorems for case (A).

The full set of inequivalent UIR's of the Lie algebra for case (B) may be inferred from those of case (A) by specializing to $r = 1$, with $\omega = 0, \pm 1, \pm 2, \ldots$, as before. A similar argument based on Schur's lemma now implies the existence of reduced matrices $\langle j'_a\, j'_b\, j'_{ab}r'\|T^{\omega}\|j_a\, j_b\, j_{ab}r\rangle$ such that

$$\begin{aligned}\langle j'_a\, j'_b\, j'_{ab}\omega' J'M'r'|T|j_a\, j_b\, j_{ab}\omega JMr\rangle &= \delta_{\omega'\omega}\delta_{J'J}\delta_{M'M}\langle j'_a\, j'_b\, j'_{ab}r'\|T^{\omega}\|j_a\, j_b\, j_{ab}r\rangle,\end{aligned} \tag{38}$$

or, similarly,

$$\begin{aligned}\langle j'_a\, j'_b\, j'_{ab}\omega'\hat{r}'r'|T|j_a\, j_b\, j_{ab}\omega\hat{r}r\rangle &= \delta_{\omega'\omega}{}^2(\hat{r}'-\hat{r})\langle j'_a\, j'_b\, j'_{ab}r'\|T^{\omega}\|j_a\, j_b\, j_{ab}r\rangle.\end{aligned} \tag{39}$$

The reduced matrix elements on the rhs of Eq. (38) must be independent of $J$ and $M$, and those on the rhs of Eq. (39) must be independent of $\hat{r}$; again, by Eq. (23), these two reduced matrices must be the same sets of numbers.

In case (C) we infer that

$$\begin{aligned}\langle j'_a\, j'_b\, j'_{ab}\omega' J'M'r'|T|j_a\, j_b\, j_{ab}\omega JMr\rangle &= \delta_{\omega'\omega}\delta_{J'J}\delta_{M'M}\langle j'_a j'_b j'_{ab}r'\|T^{\omega J}\|j_a\, j_b\, j_{ab}r\rangle,\end{aligned} \tag{40}$$

for some reduced matrix $T^{\omega J}$, which may now be $J$-dependent. Such an operator is nonlocal in $\hat{r}$-space.

Finally, it is clear that the dual case of a system operator that commutes with $\mathbf{J}$ and with (A') $\mathbf{p}$, (B') $\hat{p}$, or (C') $\lambda$, may be treated in a manner that differs formally from the above derivation only by suitable changes of symbols and occasional algebraic signs. The corresponding results may be of some relevance to collision theory, since a quantum approximation of type (B') appears to have a semiclassical realization in the form of a simple variant of the eikonal approximation (Ref. 26, p. 330; Ref. 47, p. 339): The wave functions [but not the $T(E)$ matrix] for a two-body collision are derived within the approximate physical constraint that the relative motion of the colliding particles takes place on a directed straight line, that is, $\hat{p}$ is conserved.[48]

## IV. MOMENTUM SPACE REALIZATIONS OF AN r-CONSERVING OPERATOR

The complete causal Green's function $G^{(+)}(E)$ associated with the dynamical system Eq. (3) is given by[49]

$$G^{(+)}(E) = \lim_{\eta \to 0+} (E - H + i\eta)^{-1}. \tag{41}$$

We define a wave operator $\Omega^{(+)}(E)$ and a transition operator $T(E)$ by[49,50]

$$\Omega^{(+)}(E) = I + G^{(+)}(E)V,$$
$$T(E) = V\Omega^{(+)}(E). \tag{42}$$

If one is willing to calculate both the regular and irregular solutions to the CC ODE's, the (v-$r$) matrix, say, of $G^{(+)}(E)$ may be obtained,[51] and hence the (v-$r$) matrix of $T(E)$ may be obtained, so that the properties (A), (B), or (C), of Sec. III may be tested for "by inspection" [note that $V$ as given in Eq. (3) commutes with $\mathbf{r}$]. However, the quantities normally obtained in treatments of the collisions of two molecules are the (regular) wave functions and the scattering amplitudes. Except for multiplying factors of no interest here, these entities may be construed[52,53] in the former case as mixed-basis matrices of $\Omega^{(+)}(E)$ of the type $\langle j'_a j'_b j'_{ab}\, l'J'M'r|\Omega^{(+)}(E)|j_a j_b j_{ab}\, lJMp\rangle$, and in the latter case as two-sided momentum space matrices[49] of $T(E)$, such as $\langle j'_a j'_b j'_{ab} m'_{ab}\mathbf{p}'|T(E)|j_a j_b j_{ab} m_{ab}\mathbf{p}\rangle$. In both cases, the momentum states, which are eigenstates of $H_0$, are on-the-energy-shell; in other words, $j_a$, $j_b$, and $p$ for the $\Omega^{(+)}(E)$ matrix, and $j_a, j_b$, and $p$ as well as $j'_a, j'_b$, and $p'$ for the $T(E)$ matrix, are constrained so that each state's total energy ( = eigenvalue of $H_0$) is $E$.

A mathematical problem of physical consequence for the CS method is that of characterizing such momentum space realizations of an operator $\tilde{T}(E)$ [presumed approximately equal to the exact $T(E)$] that satisfies the criteria (A), (B), or (C) of Sec. III, so that the presence of these approximate conservation laws may be conveniently tested for in "exact" solutions of the CC equations. I have been unable to obtain useful general results for an operator of type (B) or (C); some special results, the derivation and analysis of which depend on physical arguments, have been obtained for such operators and will be discussed in a separate paper, along with the structure of the wave function matrix in all three cases. In what follows, we consider only the case of a $\tilde{T}(E)$ operator of type (A).[54]

Let us define[55]

$$\boldsymbol{\Delta} = \Delta\hat{\Delta} = \mathbf{p}' - \mathbf{p}, \quad \mathbf{q} = q\hat{q} = (\mathbf{p}' + \mathbf{p})/2, \quad \boldsymbol{\xi} = \mathbf{j}'_{ab} - \mathbf{j}_{ab}. \tag{43}$$

Then, for any given quantum numbers $j'_a, j'_b, j'_{ab}, m'_{ab}, j_a, j_b, j_{ab}$, and $m_{ab}$, the matrix $\langle j'_a j'_b j'_{ab} m'_{ab}\mathbf{p}'|\tilde{T}(E)|j_a j_b j_{ab} m_{ab}\mathbf{p}\rangle$ of an operator of type (A), taken as a function of $\boldsymbol{\Delta}$ and $\mathbf{q}$, must be independent of $\mathbf{q}$, since the operators $\mathbf{r}$ generate translations in momentum space. This fact, combined with the rotational invariance of $\tilde{T}(E)$, implies (see Ref. 56) the existence of reduced matrices $\langle j'_a j'_b j'_{ab}\|\tilde{T}^{\xi}(E,\Delta)\|j_a j_b j_{ab}\rangle$ such that

$$\langle j'_a j'_b j'_{ab} m'_{ab}\mathbf{p}'|\tilde{T}(E)|j_a j_b j_{ab} m_{ab}\mathbf{p}\rangle$$
$$= \sum_{\xi\eta} (-1)^{m'_{ab}+\eta}[\xi]\begin{pmatrix} j'_{ab} & \xi & j_{ab} \\ -m'_{ab} & -\eta & m_{ab}\end{pmatrix} Y^*_{\xi\eta}(\hat{\Delta})$$
$$\times \langle j'_a j'_b j'_{ab}\|\tilde{T}^{\xi}(E,\Delta)\|j_a j_b j_{ab}\rangle. \tag{44}$$

The usual circumstance that only on-the-energy-shell amplitudes of the exact $T(E)$ are available would make it difficult to establish the $\mathbf{q}$-independence of the matrix of a $T(E)$ even if it were nearly of type (A), for given $\boldsymbol{\Delta}$, $p'$, and $p$, the only degree of freedom remaining to $\mathbf{q}$ is rotation around $\hat{\Delta}$ as an axis. This observation suggests that the amplitude for the internal angular momentum transfer $\boldsymbol{\xi}$ to be along the linear momentum transfer $\boldsymbol{\Delta}$ may have special properties. This amplitude may be obtained by analogy with Eq. (18); using Eqs. (44) and (17), we obtain

$$\sum_{\eta m'_{ab} m_{ab}} \mathscr{D}^{\xi}_{\eta\nu}(\hat{\Delta})(-1)^{m'_{ab}}\begin{pmatrix} j'_{ab} & \xi & j_{ab} \\ -m'_{ab} & -\eta & m_{ab}\end{pmatrix}$$
$$\times \langle j'_a j'_b j'_{ab} m'_{ab}\mathbf{p}'|\tilde{T}(E)|j_a j_b j_{ab} m_{ab}\mathbf{p}\rangle$$
$$= Y_{\xi\nu}(\hat{e}_z)\langle j'_a j'_b j'_{ab}\|\tilde{T}^{\xi}(E,\Delta)\|j_a j_b j_{ab}\rangle, \tag{45}$$

where $\nu$ is the value of $\boldsymbol{\xi}\cdot\hat{\Delta}$. But $Y_{\xi\nu}(\hat{e}_z) \propto \delta_{\nu 0}$, and thus, for a $\tilde{T}(E)$ of type (A), the angular momentum transfer $\boldsymbol{\xi}$ is (in a quantum-mechanical sense) perpendicular to the momentum transfer $\boldsymbol{\Delta}$. This result appears to be generally false for operators of type (B). Hence, we have obtained a necessary and partly distinguishing condition for the "type (A)-ness" of a candidate operator $T(E)$ in terms of its on-the-energy-shell matrix elements.

The circumstance that $\nu = 0$ for classical atom–molecule collisions that take place with a single impulsive (that is, concentrated at a single value of $\mathbf{r}$) exchange of momentum and angular momentum was previously noted by Khare *et al.*[14]; similar results for more complex classical collisions have been given by Hoffman *et al.*[57] Subsequent calculations by Khare *et al.*,[15–17] with a selected set of $S$-matrices obtained from accurate CC calculations, showed that in a number of collisions treated, the $\nu = 0$ amplitude strongly dominated the others; however, what appears to be a completely different explanation was advanced by these investigators as the cause of this phenomenon on the quantum level.[14]

## V. SUMMARY

The fundamental motivation for the above investigation is the conjecture that the successes of the CS method in describing molecular-collision phenomena result from underlying physical mechanisms that, in one or another case, entail the approximate conservation of, in decreasing order of severity, either (A) $\mathbf{r}$, (B) $\hat{r}$, or (C) the $r$-helicity $\omega$, by the $T(E)$ operator in a collision.

It was noted in Sec. I that previous studies purporting to show the invalidity of (C) in atom–molecule collisions were based on a generally incorrect law of transformation between $r$- and $p$-helicity amplitudes; hence, the validity of (C), and of (A) and (B), for any particular $T(E)$ is still an open question.

In Sec. II, $r$- and $p$-helicity basis states for a two-rigid-rotor system were defined and unitary transformations established between these and conventional bases.

In Sec. III, structure theorems, analogous to the Wigner–Eckart theorem for scalar operators, were derived for case (A) in Eqs. (36) and (37), for case (B) in Eqs. (38) and (39), and for case (C) in Eq. (40).

In Sec. IV it was pointed out that certain momentum space realizations of a $T(E)$ operator of type (A) depend on the initial and final momentum $\mathbf{p}$ and $\mathbf{p}'$, respectively, only in the combination $\mathbf{p}' - \mathbf{p}$. It was shown that this property and rotational invariance entail at least one testable conservation law for a kind of $p$-helicity, that is, the projection of $\mathbf{j}_{ab}$ along the momentum transfer is conserved in a collision. Finally, it was noted in Sec. IV that it is questionable whether such clear-cut and generally valid criteria exist to test the on-the-energy-shell matrix of a $T(E)$ operator for the presence of properties (B) or (C).

## ACKNOWLEDGMENTS

I would like to thank Professor Robert Hermann and Professor Hans Samelson for helpful mathematical advice. I wish to thank the referee for pointing out Ref. 17.

I would also like to thank Dr. Charles Chackerian, Jr., and the National Research Council for support of this research.

[1]L. C. Biedenharn and J. D. Louck, *Angular Momentum in Quantum Physics* (Addison–Wesley, Reading, MA, 1981), Chap. 3.15.
[2]D. J. Kouri, in *Atom–Molecule Collision Theory*, edited by R. B. Bernstein (Plenum, New York, 1979), Chap. 9.
[3]F. A. Gianturco, *The Transfer of Molecular Energies by Collision* (Springer-Verlag, Berlin, 1979), Chap. 4.
[4]A. S. Dickinson, Comput. Phys. Commun. **17**, 51 (1979).
[5]M. Jacob and G. C. Wick, Ann. Phys. (N.Y.) **7**, 404 (1959).
[6]H. Klar, Z. Phys. **228**, 59 (1969).
[7]H. Klar, Preprint, Turin, 1970; Nuovo Cimento A **4**, 529 (1971); J. Phys. B **6**, 2139 (1973).
[8]J. Andres, U. Buck, H. Meyer, and J. M. Launay, J. Chem. Phys. **76**, 1417 (1982).
[9]G. Zarur and H. Rabitz, J. Chem. Phys. **60**, 2057 (1974).
[10]P. McGuire and D. J. Kouri, J. Chem. Phys. **60**, 2488 (1974).
[11]R. T. Pack, J. Chem. Phys. **60**, 633 (1974).
[12]D. Secrest, J. Chem. Phys. **62**, 710 (1975). Secrest's "spherical rotor approximation" is a form of the CS method.
[13]V. Khare and D. J. Kouri, J. Chem. Phys. **72**, 2017 (1980).
[14]V. Khare, D. J. Kouri, and D. K. Hoffman, J. Chem. Phys. **74**, 2275 (1981).
[15]V. Khare, D. J. Kouri, and D. K. Hoffman, J. Chem. Phys. **74**, 2656 (1981).
[16]V. Khare, D. J. Kouri, and D. K. Hoffman, J. Chem. Phys. **76**, 4493 (1982).
[17]V. Khare, D. E. Fitz, D. J. Kouri, D. Evans, and D. K. Hoffman, in *Potential Energy Surfaces and Dynamics Calculations*, edited by D. G. Truhlar (Plenum, New York, 1981), p. 717.
[18]J. V. Lill, G. A. Parker, and J. C. Light, Chem. Phys. Lett. **89**, 483 (1982).
[19]S. M. Tarr, H. Rabitz, D. E. Fitz, and R. A. Marcus, J. Chem. Phys. **66**, 2854 (1977).
[20]M. H. Alexander, J. Chem. Phys. **67**, 2703 (1977).
[21]Y. Shimoni and D. J. Kouri, J. Chem. Phys. **66**, 2841 (1977).
[22]D. J. Kouri and Y. Shimoni, J. Chem. Phys. **67**, 86 (1977).
[23]It is of interest to note that Tarr *et al.* (cf. the final paragraph of Chap. II B of Ref. 19) did observe a tendency to diagonality in $|\lambda|$ of some on-the-energy-shell $T$-matrices for very small orbital angular momentum values.
[24]A. R. Edmonds, *Angular Momentum in Quantum Mechanics* (Princeton U.P., Princeton, NJ, 1960), 2nd ed.
[25]D. M. Brink and G. R. Satchler, *Angular Momentum* (Oxford U.P., London, 1968), 2nd ed.
[26]M. L. Goldberger and K. M. Watson, *Collision Theory* (Wiley, New York, 1964), Appendix E.
[27]M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions* (Dover, New York, 1965), p. 437.
[28]H. E. Moses, Ann. Phys. (N.Y.) **60**, 275 (1970), Theorem 25. Also D. J. Kouri, T. G. Heil, and Y. Shimoni, J. Chem. Phys. **65**, 226 (1977), Eq. (25'), have defined *ad hoc* what they call "body frame Bessel functions," which closely resemble the rhs of Eq. (24), in the atom–molecule case.
[29]See Ref. 8 for the corresponding result for the transformation amplitudes between states for which the individual particle helicities are diagonal.
[30]R. Hermann, *Lie Groups for Physicists* (Benjamin, New York, 1966); Casimir operators are defined on p. 73.
[31]E. P. Wigner, Ann. Math. **40**, 149 (1939).
[32]G. W. Mackey, *Induced Representations of Groups and Quantum Mechanics* (Benjamin, New York, 1968).
[33]G. W. Mackey, *Unitary Group Representations in Physics, Probability, and Number Theory* (Benjamin, Reading, MA, 1978), Chap. 10.
[34]G. W. Mackey, in M. F. Atiyah, R. Bott, S. Helgason, D. Kazhdan, B. Kostant, G. Lusztig, I. G. MacDonald, G. W. Mackey, W. Schmid, and D. J. Simms, *Representation Theory of Lie Groups* (Cambridge U. P., Cambridge, 1979), p. 20.
[35]Overviews of the subject of infinite-dimensional group representations and harmonic analysis, with selected literature references, were given by G. W. Mackey, Bull. Am. Math. Soc. **69**, 628 (1963); Bull. Am. Math. Soc. (2) **3**, 543 (1980); and by K. I. Gross, Am. Math. Mon. **85**, 525 (1978).
[36]A. J. Coleman, in *Group Theory and its Applications*, edited by E. M. Loebl (Academic, New York, 1968), Vol. I.
[37]A. J. Coleman, *Induced Representations with Applications to $S_n$ and GL (n)*, Queen's Papers in Pure and Applied Math. No. 4 (Queen's U., Kingston, Ontario, 1966).
[38]W. Miller, Jr., *Lie Theory and Special Functions* (Academic, New York, 1968), Chap. 6.
[39]J. D. Talman, *Special Functions* (Benjamin, New York, 1968), Chap. 12.
[40]U. H. Niederer and L. O'Raifeartaigh, Fortschr. Phys. **22**, 111 (1974).
[41]A. S. Wightman, Rev. Mod. Phys. **34**, 845 (1962).
[42]A. O. Barut and R. Raczka, *Theory of Group Representations and Applications* (Polish Sci. Publ., Warsaw, 1977), Chaps. 16–18.
[43]Strictly speaking the objects dealt with here are sections of a vector bundle over $S^2$, where the vector space here is $\mathbb{C}$, the complex number field. See Refs. 30 or 32 for definitions.
[44]Ref. 32, p. 128, Theorem E', or Ref. 36, Theorem 7: If in Mackey's notation, $H_1 = L$ and $H_2 = K$, then there is only one double coset, that is, $\mathrm{IR}(3) = H_2 e H_1$, with $e =$ the group identity. If we choose $\Gamma^{r,\omega}$ to be the representation of $H_1$ and $e$ to be the double coset representative, Mackey's theorem says that the representation of $H_2$ obtained (1) by restricting $\Gamma^{r,\omega}$ to $H_1 \cap H_2 =$ the subgroup with coordinates $(0,0,\psi,\mathbf{0})$, and (2) by inducing from this representation to a representation of $H_2$, is equivalent to $U^{r,\omega}{\downarrow}H_2$. For the case of the translation subgroup, let $H_1 = L$ and $H_2 = Q$; then each $\hat{r} \in S^2$ with polar coordinates $(\phi,\theta)$ determines a group element $(\phi,\theta,0,\mathbf{0}) \in \mathrm{IR}(3)$ that uniquely labels a double coset $H_2(\phi,\theta,0,\mathbf{0})H_1$ in IR(3), such that the union of these double cosets exhausts IR(3). It can now be shown that Eq. (33) follows from Mackey's theorem.
[45]See Ref. 36, p. 71, and Ref. 32, p. 11.
[46]The $J$-independence of these reduced matrices was noted in the case of the $V$-operator of Eq. (3) by Klar (Ref. 7).
[47]L. I. Schiff, *Quantum Mechanics* (McGraw–Hill, New York, 1968), 3rd ed.
[48]In a series of papers, (a) M. Tamir and M. Shapiro, Chem. Phys. Lett. **31**, 166 (1975), (b) Chem. Phys. Lett. **39**, 79 (1976), and (c) M. Shapiro and M. Tamir, Chem. Phys. **13**, 215 (1976), a theoretical scheme was proposed and applied that its authors called the "$p$-helicity decoupling" (PHD) approximation for atom–molecule scattering; that is, it seems that a dynamical approximation scheme of type (C') was engendered. Subsequently, D. J. Kouri, Chem. Phys. Lett. **31**, 599 (1975) and A. E. DePristo and M. H. Alexander, Chem. Phys. **19**, 181 (1977) pointed out a close resemblance between Tamir and Shapiro's theory and results and the theory and results derived from the CS method, and concluded the PHD approximation was a variant of an approximation of the CS, that is $r$-helicity decoupling type. On physical grounds, however, it is reasonable to expect that $p$-helicity or $r$-helicity conservation would prevail under very different circumstances, and the kinematical relations derived here in Secs. III and IV imply that a dynamics based on one or the other approximation would take substantially different forms, with only a partial mathematical analogy between the two approaches being present due to the duality of position and momentum space. The resolution of this discrepancy appears to be the following: In Tamir and Shapiro's paper (a), they proceed from formal developments of Klar (Refs. 6 and 7), specialized to the atom–molecule case. Klar used a basis of free-particle momentum states in which the individual particle $p$-helicities, as defined here in Eq. (2), were diagonal, and expressed the two-molecule Hamiltonian [of the type of Eq. (3)] in terms of a basis in which the individual particle $r$-helicities were diagonal. Klar, however, used indistinguishable notation—subscripted $\lambda$'s—for both

types of helicity, and at no point in the texts of Refs. 6 and 7 is the physical distinction between these two kinds of entities discussed. I infer that Tamir and Shapiro were not aware of this distinction, and advanced a theory that is physically of $r$-helicity conserving type [as is clear from scrutiny of their Hamiltonian in Eq. (10) of paper (a)], while asserting incorrectly that the quantities being conserved were physical $p$-helicities.

[49]R. G. Newton, *Scattering Theory of Waves and Particles* (McGraw–Hill, New York, 1966), Chap. 7.

[50]R. D. Levine, *Quantum Mechanics of Molecular Rate Processes* (Oxford U. P., London, 1969), Chap. 2.5.0.

[51]J.-T. Hwang and H. Rabitz, J. Chem. Phys. **70**, 4609 (1979), Appendix.

[52]L. W. Hunter and R. F. Snider, J. Chem. Phys. **61**, 5250 (1974).

[53]K. M. Watson and J. Nuttall, *Topics in Several Particle Dynamics* (Holden–Day, San Francisco, 1967), Eq. (1.5).

[54]Transition operators with type (A) structure have been considered by R. T. Ling, J. C. Y. Chen, J. L. Preacher, and K. M. Watson, Phys. Rev. A **6**, 650 (1972), Eq. (1.11), and in Liouville space by A. Ben-Reuven, Adv. Chem. Phys. **33**, 235 (1975), Sec. IX.D.

[55]The vector $\hat{\Delta}$ has been termed the "kinematic apse" by Khare *et al.* (Refs. 14–17).

[56]A. P. Yutsis, I. B. Levinson, and V. V. Vanagas, *Mathematical Apparatus of the Theory of Angular Momentum* (Israel Program for Scientific Translations, Jerusalem, 1962), Chaps. II and III.

[57]D. K. Hoffman, J. W. Evans, and D. J. Kouri, J. Chem. Phys. **80**, 144 (1984).

# Locally attractive normal modes for chemical process

Ariel Fernández and Oktay Sinanoğlu
*Department of Chemistry, Yale University, New Haven, Connecticut 06511*

(Received 16 December 1982; accepted for publication 20 April 1984)

It is shown that if along a normal mode of perturbation of a steady state which is far from equilibrium, the mixed second variation of the entropy is a constant of motion, this mode is locally attractive and the system becomes structurally unstable (in the sense of Adronov and Pontriaguin). In general, the existence of such a mode indicates chemical oscillation. In the special case that the mode has also a vanishing imaginary part of the frequency, it leads to the formation of a dissipative spatial structure.

PACS numbers: 82.20. – w, 05.70.Ln

*Notation*: If $\delta\varphi$ denotes the first variation of a generic quantity $\varphi$ and it is given by $\delta\varphi = \Sigma_j \delta A_j \delta B_j$, then $\delta_m\varphi$ will denote the mixed first differential which is given by the expression

$$2\delta_m\varphi = \sum_j \delta A_j^* \delta B_j + \delta A_j \delta B_j^*.$$

An analogous definition applies for the case of the second variation of a quantity $\theta$: If

$$\delta^2\theta = \sum_j \delta A_j \delta B_j,$$

then

$$\delta_m^2\theta \overset{\text{def}}{=} \frac{1}{2}\sum_j \delta A_j^* \delta B_j + \delta A_j \delta B_j^*.$$

## I. INTRODUCTION

In this paper open reaction systems with a Fick's law of diffusion will be considered. It is assumed that the system obeys the laws of linear thermodynamics of irreversible processes in a neighborhood of equilibrium. The system is driven far from equilibrium (outside the linear domain[1]) by increasing the values of the overall affinity.[2] No convective effects are considered. The purely periodical normal modes of the system around a steady state $\mathbf{x}_0$ which lies outside the linear thermodynamics domain are characterized by the equations

$$\delta_m P = 0, \qquad \delta_m \pi \neq 0, \tag{1}$$

where $P$ represents the entropy production

$$P = \int_V \left(\sum_\alpha J_\alpha X_\alpha\right) dV \tag{1'}$$

($J_\alpha \equiv$ flows, $X_\alpha \equiv$ conjugate thermodynamic forces). Here $\delta_m P$ is the entropy production corresponding to the first variation of the entropy $\delta_m s$ along the normal mode, that is, the excess entropy production

$$\delta_m P = P(\delta_m s) = \left[\sum_\alpha \int \delta J_\alpha \delta X_\alpha \, dV\right]_{\text{mixed}}$$

$$= \frac{1}{2}\sum_\alpha \int \delta J_\alpha^* \delta X_\alpha + \delta J_\alpha \delta X_\alpha^* \, dV. \tag{1''}$$

Here $\delta_m \pi$ is given by the expression

$$\delta_m \pi = -\frac{i}{2}\int \sum_\alpha (\delta J_\alpha \delta X_\alpha^* - \delta J_\alpha^* \delta X_\alpha) dV, \tag{1'''}$$

where "*" means complex conjugate. The statement corresponding to Eq. (1) is proved in Glansdorff and Prigogine.[1] The increments of the thermodynamic flows $\delta J_\alpha$ and the thermodynamic forces $\delta X_\alpha$ should be evaluated along the normal mode.

In this paper we shall prove that the manifold of periodic motion $\Omega$, along which $\delta_m P = 0$, is locally attractive and locally invariant. (Rigorous definitions of these properties are given in the next section.)

This approach provides also a useful method of calculating the frequencies for purely periodic modes without evaluating the equation $\delta_m P = 0$ with the kinetic restrictions in the increments of the forces and flows. Since $\Omega$ is locally attractive, the purely periodic modes "absorb" other modes of excitation. Consider, for example, the case in which a time-independent normal mode exists, *this mere fact does not imply that the system will adopt a spatially periodic organization unless we prove in addition that this mode is locally attractive.*

## II. THE MANIFOLD

The space variable will be denoted $r$, $r \in [0,1]$ (we assume one-dimensional diffusion). The boundary conditions here imposed are of the Neumann type although we could also develop a theory when Dirichlet boundary conditions are imposed.

The Neumann boundary conditions correspond to zero flux at the boundary, that is,

$$\frac{\partial}{\partial r}\mathbf{x}(1,t) = \frac{\partial}{\partial r}\mathbf{x}(0,t) = 0, \quad t \geqslant 0, \tag{2}$$

$\mathbf{x}(r,t)$ is the concentration vector, and

$$\dot{\mathbf{x}} = \frac{d\mathbf{x}}{dt} = \varphi(\mathbf{x}), \quad \mathbf{x} \in \mathbb{R}^2$$

 0022-2488/84/082576-06$02.50  

is the kinetic equation. The generalization of the theory to the case of $n$ intermediate species is straightforward; $\mathbf{x}_0$ is our steady state of interest:

$$\varphi(\mathbf{x}_0) = 0. \tag{3}$$

Here $\mathbf{x}_0$ is assumed to lie outside the linear thermodynamics domain.[3]

$J(\varphi)|\mathbf{x}_0$ is the Jacobian matrix of $\varphi$ at the steady state. Assuming a Fick's law of diffusion the problem reads

$$\frac{d\mathbf{u}}{dt} = \dot{\mathbf{u}} = \mathbf{D}\frac{\partial^2 \mathbf{u}}{\partial r^2} + \mathbf{N}(\mathbf{u}), \tag{4}$$

where $\mathbf{D}$ is the diffusion matrix which is assumed to be diagonal.

We made in (4) the translation

$$\mathbf{u} = \mathbf{x} - \mathbf{x}_0, \quad \mathbf{N}(\mathbf{u}) = \varphi(\mathbf{u} + \mathbf{x}_0) = \varphi(\mathbf{x}). \tag{5}$$

We also impose the following restriction:

$$\mathbf{N}(\mathbf{u}) = J(\varphi)|_{\mathbf{x}=\mathbf{x}_0}\mathbf{u} + \mathbf{F}(\mathbf{u}), \tag{6}$$

with

$$\mathbf{F}(0) = 0, \quad \frac{\partial \mathbf{F}}{\partial \mathbf{u}}(0) = 0. \tag{7}$$

For operator $\partial^2/\partial r^2$ can be extended to a self-adjoint operator $\hat{\mathscr{B}}$, $\hat{\mathscr{B}}$ has domain

$\mathscr{D}\hat{\mathscr{B}}$ = closure of the set

$$\left\{u(r)\in C^2([0,1])\,\frac{\partial u}{\partial r}(0) = \frac{\partial u}{\partial r}(1) = 0\right\}$$

in the second Sobolev space $H^2([0,1])$. (For the theory of Sobolev spaces see, for example, Ref. 4.) Hence, $D\hat{\mathscr{B}}$:

$$H^2([0,1])\times H^2([0,1]) \to L^2([0,1])\times L^2([0,1]), \tag{8}$$

and $J(\varphi)|_{\mathbf{x}_0}$ will be regarded as a perturbation of $\mathbf{D}\hat{\mathscr{B}}$. The operator $\mathbf{D}\hat{\mathscr{B}} + J(\varphi)|_{\mathbf{x}_0} = \mathscr{L}$ generates a compact analytic semigroup

$$\{e^{\mathscr{L}t}\}_{t>0}$$

(see Kato[4]). From this, it follows (see, for example, Holmes and Mardsen[5]), that the Neumann problem considered above (4) defines a flow $\mathscr{F}_t$ on $S = \mathscr{D}_{\hat{\mathscr{B}}} \times \mathscr{D}_{\hat{\mathscr{B}}}$, for $t\geqslant 0$,

$$\mathscr{F}_t\mathbf{x}(r) = \mathbf{x}(r,t). \tag{9}$$

Then $\mathscr{F}_t$ is a monoparametric family of diffeomorphisms representing the evolution of the system. If $(\partial F/\partial u)(0) = 0$, this flow verifies

$$D_S\mathscr{F}_t|_{\mathbf{u}=0} = e^{t\mathscr{L}}, \tag{10}$$

where $D_S$ denotes the Fréchet derivative of $\mathscr{F}_t$ in $S$. From now on, the symbol $E$ denotes the eigenspace of $\mathscr{L}$ and $E[\mathrm{Re}\,\sigma(\mathscr{L}) = 0]$ is then the eigenspace associated to the purely imaginary eigenvalues of $\mathscr{L}$.

If the conditions

$$\begin{aligned} &\text{(a) } \dim_R E[\mathrm{Re}\,\sigma(\mathscr{L}) = 0] < \infty, \\ &\text{(b) } \mathrm{Re}\,\sigma(\mathscr{L}) \leqslant 0 \end{aligned} \tag{11}$$

are fulfilled, then there exists a center manifold $\Omega \subseteq \mathscr{D}_{\hat{\mathscr{B}}} \times \mathscr{D}_{\hat{\mathscr{B}}}$, (see, for example, Marsden[6]) such that the following hold true.

(a') $\Omega$ contains $\mathbf{x}_0$ (or $\mathbf{u} = 0$).

(b') $\Omega$ is tangent to $E[\mathrm{Re}\,\sigma(\mathscr{L}) = 0]$ at $\mathbf{x}_0$ (or at $\mathbf{u} = 0$): $\Omega = T_0E[\mathrm{Re}\,\sigma(\mathscr{L}) = 0]$.

(c') $\Omega$ is locally invariant: If $\mathbf{x}(r)\in\Omega$, there exists $T = T(\mathbf{x})$ such that $F_t(\mathbf{x})\in\Omega$ for $t$ in the interval $0\leqslant t\leqslant T = T(\mathbf{x})$, $T > 0$.

(d') $\Omega$ is locally attractive: There exists an open neighborhood $U$ of $u = 0$ such that if $\mathscr{F}_t(\mathbf{u})\in U$ $t\geqslant 0$, then

$$\lim_{t\to\infty}\inf_{v\in\Omega}\sup_{r\in[0,1]}\|\mathscr{F}_t(u(r)) - v(r)\| = 0$$

($\|\ \|$ is simply the Euclidean norm.)

*Observation*: The eigenvalues of $\hat{\mathscr{B}}$ for the Neumann problem are

$$-n^2\pi^2 = \lambda_n (n = 0,1,2,\ldots),$$

therefore if $\mathrm{Tr}\,J(\varphi)|_{\mathbf{x}_0} = 0$, $\Omega$ is contained in the manifold $T_0E[\sigma\hat{\mathscr{B}} = 0]$.

## III. EXCESS OF ENTROPY PRODUCTION ALONG $\Omega$

For the normal modes given by

$$\mathbf{x} - \mathbf{x}_0 = \mathbf{M}e^{wt + ir/\lambda}, \tag{12}$$

where $\lambda$ is the wavelength of the inhomogeneity and

$$\|\mathbf{M}\|/\|\mathbf{x}_0\| \ll 1 \quad \text{and} \quad w = w_1 + w_2 i,$$

we have from the last section that $w_1 = 0$ along $\Omega$. If the invariants of flows are written around the steady state $\mathbf{x}_0$ as

$$\delta J_\alpha = \sum_\beta \Lambda_{\alpha\beta}\delta X_\beta, \quad \Lambda_{\alpha\beta} = \frac{\partial J_\alpha}{\partial X_\beta}, \tag{13}$$

we get along $\Omega$,

$$\begin{aligned} \delta_m P &= \frac{1}{2}\int\left(\sum_\alpha \delta J_\alpha \delta X_\alpha^* + \delta J_\alpha^* \delta X_\alpha\right)dV \\ &= \int\left(\sum_{\alpha\beta}\left(\frac{\Lambda_{\alpha\beta} + \Lambda_{\beta\alpha}}{2}\right)\delta X_\alpha^* \delta X_\beta\right)dV \\ &= \int\left(\sum_\alpha \Lambda^s_{\alpha\beta}\delta X_\alpha^* \delta X_\beta\right)dV = 0 \end{aligned} \tag{14}$$

[defining the symmetric component of $\Lambda_{\alpha\beta}$: $\Lambda^s_{\alpha\beta} = (\Lambda_{\alpha\beta} + \Lambda_{\beta\alpha})/2$] since the $X_\alpha$ are independent forces, we have proven the following theorem.

**Theorem I**: $\Lambda^s_{\alpha\beta} = 0$ along $\Omega$. Hence, the tensor $\Lambda_{\alpha\beta}$ is purely antisymmetric along $\Omega$.

We have also proved the following theorems.

**Theorem II**: The number of normal modes of excitation along which $\delta_m P = 0$ is the number

$$\dim\Omega = \dim E[\mathrm{Re}\,(\sigma(\mathscr{L})) = 0]. \tag{15}$$

**Theorem III**: Along $\Omega$ the increment $\delta_m\pi$ is valuated as

$$i\int\sum_{\alpha,\beta}\Lambda_{\alpha\beta}\delta X_\alpha^*\delta X_\beta\, dV = \delta\pi. \tag{16}$$

*Proof*:

$$\begin{aligned} \delta_m\pi &= -\frac{i}{2}\int\left(\sum_\alpha \delta J_\alpha \delta X_\alpha^* - \delta J_\alpha^*\delta X_\alpha\right)dV \\ &= \int\frac{i}{2}\sum_{\alpha,\beta}\left(\frac{\Lambda_{\alpha\beta} - \Lambda_{\beta\alpha}}{2}\right)(\delta X_\alpha^*\delta X_\beta)\,dV. \end{aligned} \tag{17}$$

In the last equality Eq. (13) has been used.

Since, from Theorem I, along $\Omega$: $\Lambda^{s}_{\alpha\beta}=0$, we have

$$\delta_m \pi = i\int \sum_{\alpha,\beta} \Lambda_{\alpha\beta}\delta X^{*}_{\alpha}\delta X_{\beta}\, dV, \tag{18}$$

we can also state the following theorem as a summary.

**Theorem IV**: The manifold of modes of excitation along which $\delta_m P=0$ is locally attractive and locally invariant provided conditions given by Eqs. (6) and (7) are fulfilled.

If stationary boundary conditions are imposed instead of no flux boundary conditions, the existence of a center manifold is similarly developed. Consider the Dirichlet boundary conditions

$$\frac{d}{dt}\mathbf{x}(0,t)=\frac{d}{dt}\mathbf{x}(1,t)=0. \tag{19}$$

It is assumed that $\mathbf{x}=0$ is the steady state for the problem (4). Then the operator $\nabla^2$ can naturally be extended to a self-adjoint operator $\hat{\mathscr{A}}$ defined in the closure in $H^2([0,1])$ of the set

$$\{u(r)\in C^2([0,1])|u(0)=u(1)=0\}.$$

Since the eigenvalues of $\hat{\mathscr{A}}$ are now $\lambda_n=-n^2\pi^2$, $n\geqslant 1$, we can state that our Dirichlet problem is locally equivalent to the restriction of the Neumann problem to the manifold

$$\bigcup_{j>1} T_0 E\,[\sigma(\nabla^2)=-j^2\pi^2].$$

In general, the range of validity of relations (1) is restricted to the time-independent boundary conditions $(d/dt)\,\mu_j=0$ along the boundary ($\mu_j$ is the chemical potential of intermediate $X_j$), $j=1,2,\ldots,N$ for an isothermal system in mechanical equilibrium (see Ref. 1).

In this paper it is assumed that the only thermodynamic forces are the $T^{-1}A_j$'s ($A_j$ is the affinity of the $j$th reaction). Then, the key relation

$$(w_1^2+w_2^2)\delta_m^2 S = w_1\delta_m P + w_2\delta_m\pi \leqslant 0,$$

which implies relation (1), can be derived for both Dirichlet and Neumann boundary conditions:

$$-\sum_j \frac{d(\delta\mu_\gamma T^{-1})}{dt}\frac{d\delta\rho_j}{dt}=\frac{\rho}{T}\sum_{jj'}-\mu_{jj'}\frac{d\delta n_j}{dt}\frac{d\delta n_{j'}}{dt}, \tag{20}$$

where $\rho_j$ is the local density of intermediate $X_j$, $\rho$ is the total local density,

$$\mu_{jj'}=\frac{\partial}{\partial n_{j'}}\mu_j,$$

and $n_j$ is the mass fraction of the intermediate $X_j$.

From the law of conservation of mass, we get

$$-\sum_j\sum_z \upsilon_{jz}\mu_j\delta J_z\frac{d}{dt}(\delta\mu_j T^{-1}) = -\sum_j \frac{d(\delta\rho)_j}{dt}\frac{d}{dt}(\delta\mu_j T^{-1}), \tag{21}$$

$\upsilon_{jz}$ is the stoichiometric coefficient of species $X_j$ in the $z$th reaction, $M_j$ is the molecular mass of the species $X_j$, and

$$A_z=-\sum_j \upsilon_{jz}\mu_j T^{-1}.$$

Also

$$-\sum_{jj'}\frac{\rho}{T}\mu_{jj'}\frac{d\delta n_j}{dt}\frac{d\delta n_{j'}}{dt}=\sum_z \delta J_z\frac{d}{dt}(\delta A_z T^{-1})\leqslant 0. \tag{22}$$

The rhs of Eq. (22) is the response of the excess entropy production to the changes in the thermodynamic forces.

Taking mixed differentials in (22), we get for a normal mode of frequency $w=w_1+iw_2$,

$$-(w_1^2+w_2^2)\left[\left(+\sum_j(\delta\mu_j T^{-1})\delta\rho_j^{*}\right)+\left(\sum_j(\delta\mu_j T^{-1})^{*}\delta\rho_j\right)\right] = (w_1^2+w_2^2)\delta^2 s\leqslant 0. \tag{23}$$

Again from (22) and the definitions of $\delta P$ and $\delta\pi$ we get the desired relation. Therefore, the independence of relation (1) from the boundary conditions, makes the basic Theorems I–IV valid for the Dirichlet as well as the Neumann problem. If $\hat{\mathscr{O}}$ denotes here the operator $\hat{\mathscr{B}}$ for the Neumann problem or the operator $\hat{\mathscr{A}}$ for the Dirichlet problem, we get the following theorem.

**Theorem V**: Assuming $\mathscr{D}_{\hat{\theta}}\times\mathscr{D}_{\hat{\theta}}$ can be written as a sum of two $\mathscr{L}$-invariant subspaces: $\mathscr{D}_{\hat{\theta}}\times\mathscr{D}_{\hat{\theta}}=V_1\oplus V_2$ ($V_1$ and $V_2$ are $\mathscr{L}$-invariant, therefore also invariant under the compact semigroup generated by $\mathscr{L}$, that is, $e^{\mathscr{L}t}$, $t\geqslant 0$) such that $\delta_m^2 S$ is a constant of motion along $T_0V_1$, and $\delta_m P>0$ along $T_0V_2$, then $T_0V_1$ is locally attractive.

Our problem (4) then splits into the equations

$$\left.\begin{aligned}\dot{\mathbf{x}}_1&=\mathscr{L}_1\mathbf{x}_1+\mathbf{F}_1(\mathbf{x}_1,\mathbf{x}_2)\\ \dot{\mathbf{x}}_2&=\mathscr{L}_2\mathbf{x}_2+\mathbf{F}_2(\mathbf{x}_1,\mathbf{x}_2)\end{aligned}\right\}\quad \mathbf{x}_1\in V_1,\quad \mathbf{x}_2\in V_2, \tag{24}$$

where $\mathbf{x}_1+\mathbf{x}_2$ is the (unique) decomposition of a generic element of $\mathscr{D}_{\hat{\theta}}\times\mathscr{D}_{\hat{\theta}}$, $\mathrm{Re}\,\sigma(\mathscr{L}_1)=0$ and $\mathrm{Re}\,\sigma(\mathscr{L}_2)<0$. The manifold $T_0V_1$ is given by a smooth function $\mathbf{x}_2=\psi(\mathbf{x}_1)$, where $\psi$ has the following properties:

$$\left.\begin{aligned}&\psi(0)=0\\ &D\psi(0)=0\\ &\|D\psi(x_1)\|<1\end{aligned}\right\}\ D=\text{Fréchet derivative.} \tag{25}$$

The restriction of the system (4) to $T_0V_1$ is

$$\dot{\mathbf{x}}_1=\mathscr{L}\mathbf{x}_1+\mathbf{F}(\mathbf{x}_1). \tag{26}$$

If $\mathbf{x}_1(t)$ is the solution of (26), then $\mathbf{x}_1(t)+\psi(\mathbf{x}_1(t))$ is the solution of (4) contained in $T_0V_1$.

## IV. EXAMPLES

(a) $\Omega$ for the Lotka–Volterra model with Neuman boundary conditions for diffusion. If the kinetic laws corresponding to infinite total affinity are given by

$$\left[\begin{aligned}\frac{dx}{dt}&=k_1Ax-k_2xy\\ \frac{dy}{dt}&=k_2xy-k_3y\end{aligned}\right]\left[\begin{aligned}A+X&\xrightarrow{k_1}2X\\ X+Y&\xrightarrow{k_2}2Y\\ Y&\xrightarrow{k_3}P\end{aligned}\right] \tag{27}$$

($x$ is the concentration of species $X$, $y$ is the concentration of species $Y$). The steady state is $x_0=k_3/k_2$, $y_0=(k_1/k_2)A$. The concentration of $A$ is considered constant. We have

$$\mathscr{L}=\begin{bmatrix}0 & -k_3\\ k_1A & 0\end{bmatrix}+\mathbf{D}\nabla^2,\quad \nabla^2=\frac{\partial^2}{\partial r^2}, \tag{28}$$

$$\Omega=T_0\left[E\left(\pm i\sqrt{k_3k_1A}\right)\right]. \tag{29}$$

Considering only the relaxation process for the 0 eigenvalue of $\nabla^2$, we get

$$\delta_m P = \left[ T^{-1} \int [(\delta J_1)_x (\delta A_1)_x + (\delta J_2)_x (\delta A_2)_x + (\delta J_2)_y (\delta A_2)_y + (\delta J_3)_y (\delta A_3)_y ] dV \right]_{\text{mixed}}. \quad (30)$$

Here $A_j$ indicates the affinity of the $j$th reaction and the subindex $x$ or $y$ indicates that $\delta x$ or $\delta y$ is responsible for the increment of the quantity in brackets. We shall also assume $RT = 1$ in an appropriate scale. Then we have

$$(\delta J_1)_x = k_1 A \delta x = (\delta k_1 A x)_x, \quad (31)$$

$$(\delta A_1)_x = \left( \delta \ln \frac{Ax}{x^2} \right)_x = -\frac{\delta x}{x_0}, \quad (32)$$

$$(\delta J_2)_x = k_2 y_0 x = (\delta k_2 x y)_x, \quad (33)$$

$$(\delta A_2)_x = \left( \delta \ln \frac{xy}{y^2} \right)_x = \frac{\delta x}{x_0}, \quad (34)$$

$$(\delta J_2)_y = k_2 x_0 \delta y = (\delta k_2 x y)_y, \quad (35)$$

$$(\delta A_2)_y = \left( d \ln \frac{xy}{y_2} \right)_y = -\frac{\delta y}{y_0}, \quad (36)$$

$$(\delta J_3)_y = k_2 \delta y, \quad (37)$$

$$(\delta A_3)_y = \left( \delta \ln \frac{y}{p} \right)_y = \frac{\delta y}{y_0}. \quad (38)$$

But the relations (31)–(38) imply $\delta P \equiv 0$ (identically zero) for any point around the steady state. This implies that the manifold of the phase space locally tangent to $E[\sigma(\nabla^2) = 0]$ is a locally attractive manifold. (Actually, in this case, $\delta P = \delta_m P = 0$.)

Since $\delta P$ vanishes identically around the steady state we cannot calculate the relaxation frequencies from the equation $\delta_m P = 0$.

(b) Consider the model introduced by Glansdorff and Prigogine[1]:

$$\begin{bmatrix} A \to X & (1) \\ 2X + Y \to 3X & (2) \\ B + X \to Y + D & (3) \\ X \to E & (4) \end{bmatrix} \quad \begin{bmatrix} \dfrac{dx}{dt} = A + x^2 y - Bx - x \\ \dfrac{dy}{dt} = Bx - x^2 y \end{bmatrix}. \quad (39)$$

The concentration of species $A,B,D,E$ are considered constant; $x_0 = A$, $y_0 = B/A$. Neumann (no flux) boundary conditions are applied.

We assume a Fourier expansion of the arbitrary perturbation in eigenfunctions of the operator $\hat{\mathscr{B}}$, that is to say, $\mathbf{x}(r,t) - \mathbf{x}_0 = \Sigma_{n=0}^{\infty} \mathbf{a}_n(t) \cos(n\pi r)$. The part of the spectrum of $\mathscr{L}$ satisfying $\operatorname{Re} \sigma(\mathscr{L}) = 0$ could be contained in any of the slices $T_0 E[\sigma(\hat{\mathscr{B}}) = -n^2\pi^2]$, $n \geqslant 0$ ("$E$" represents the eigenspace with respect to the operator $\mathscr{L}$). Since $\operatorname{Re} \sigma(\mathscr{L}) = 0$ if and only if for a fixed $n$ the trace of $J(\varphi) - n^2\pi^2$ **D** vanishes (that is, $B - 1 - A^2 - D_x n^2\pi^2 - D_y n^2\pi^2 = 0$), we get for $D_x, D_y \gg B$ that the only possible local attractor should be contained in

$$T_0 E[\sigma(\mathscr{B}) = 0] : \{\delta_m P = 0\} \subset T_0 E[\sigma(\nabla^2) = 0]. \quad (40)$$

The normal modes, along which $\delta_m = 0$, are contained in $\Omega = T_0 E[\operatorname{Re} \sigma(\mathscr{L}) = 0] = T_0 E[\operatorname{Re} \sigma(J(\varphi)) = 0]$,

$$J(\varphi) = \begin{bmatrix} B - 1 & A^2 \\ -B & -A^2 \end{bmatrix}. \quad (41)$$

The frequencies for which $\delta_m P = 0$ satisfy

$$\operatorname{Det}(J(\varphi) - wI) = \Delta(J((\varphi) - wI)) = 0, \operatorname{Re} w = 0, \quad (42)$$

$w$ is purely imaginary only for $B_0 = 1 + A^2$:

$$w = iw_2 = \pm iA. \quad (43)$$

We shall get back these eigenfrequencies if we solve the equation $[\delta_m P = 0]$ along $T_0 E[\sigma(\nabla^2) = 0]$ for $B = B_0$:

$$\delta_m P = \frac{1}{2} T^{-1} \int dV [(\delta J_2)_x (\delta A_2)_x^* + (\delta J_2)_x^* (\delta A_2)_x + (\delta J_2)_y (\delta A_2)_y^* + (\delta J_2)_y^* (\delta A_2)_y + (\delta J_3)_x (\delta A_3)_x^* + (\delta J_3)_x^* (\delta A_3)_x + (\delta J_4)_x (\delta A_4)_x^* + (\delta J_4)_x^* (\delta A_4)_x ] \Big|_{T_0 E[\sigma(\nabla^2) = 0]} = 0.$$

A computation analogous to the one done for the Lotka–Volterra system gives

$$\delta_m P|_{T_0 E[\sigma(\nabla^2) = 0]} = -A \delta x \delta x^* + [A^3/(1 + A^3)] \delta y \delta y^*|_{T_0 E[\sigma(\nabla^2) = 0]} = 0. \quad (44)$$

The restriction of the increments $\delta x$ and $\delta y$ to $T_0 E[\sigma(\nabla^2) = 0]$ satisfies

$$J(\varphi) \begin{bmatrix} \delta x_1 + i \delta x_2 \\ \delta y_1 + i \delta y_2 \end{bmatrix} = J(\varphi) \cdot \begin{pmatrix} \delta x \\ \delta y \end{pmatrix} = \begin{pmatrix} \dot{\delta x} \\ \dot{\delta y} \end{pmatrix} \quad (45)$$

(since the restriction of $\hat{\mathscr{B}}$ is zero) which gives

$$\begin{bmatrix} \delta y_2 = w_2 \delta y_1 \\ \delta x_2 = -w_2 \delta y_1 \\ \delta x_1 = 0 \end{bmatrix}. \quad (46)$$

This implies from (44) that

$$\delta_m P|_{T_0 E[\sigma(\nabla^2) = 0]} = 0 \Leftrightarrow w_2 = \pm A. \quad (47)$$

But this is precisely the result obtained from equation (42) using the center manifold theory. The locally attractive manifold for this system is then given by the equations

$$\begin{bmatrix} \delta_m P = 0 \\ J(\varphi) \cdot \begin{pmatrix} \delta x \\ \delta y \end{pmatrix} = \begin{pmatrix} \dot{\delta x} \\ \dot{\delta y} \end{pmatrix} \\ \sigma(\nabla^2) = 0 \end{bmatrix}. \quad (48)$$

The kinetic laws in both examples obey the conditions given by Eqs. (6) and (7), therefore Theorems I–IV apply.

## V. THE LOCALLY UNSTABLE MANIFOLD

Provided $\dim E[\operatorname{Re} \sigma(\mathscr{L}) > 0] < \infty$, condition (11)(b) could be lifted and the validity of the results regarding the center manifold remains unaltered. In this case we introduce the new unstable manifold

$$M = T_0 E[\operatorname{Re} \sigma(\mathscr{L}) > 0]. \quad (49)$$

Since along $M$, $w_1 > 0$, we get from

$$\frac{1}{2} \frac{d\delta_m^2 S}{dt} = w_1 \delta_m^2 S = \delta_m P = P[\delta_m S] \quad (50)$$

and $\delta_m^2 S < 0$, that $\delta_m P < 0$. Here "$E[\cdots]$" indicates the eigenspace of $\mathscr{L}$ associated with the part of the spectrum of $\mathscr{L}$ satisfying the condition in square brackets.

We now have the following theorem.

**Theorem VI**: $\delta_m P < 0$ along $M$ if $M$ is finite dimensional.

Let us consider again the example (b) of the last section. If instead of assuming that the center manifold lies in the manifold $T_0E\,[\mathrm{Re}\,\sigma(\nabla^2) = 0]$ we now assume that it is contained in $T_0E\,[\mathrm{Re}\,\sigma\nabla^2 < 0]$ (we recall that "$E$" denotes eigenspace with respect to $\mathscr{L}$), then an unstable manifold exists for this problem. To illustrate this point, the following assumptions will be made: $D_x = D_y = D$, and the center manifold lies in $T_0E\,[\sigma\nabla^2 = -\pi^2]$. The critical value of the control parameter $B$ for which purely periodical modes will appear is now

$$B_0 = 1 + A^2 + 2D\pi^2 \quad (A > \pi^2 D). \tag{51}$$

It is therefore clear that the manifold of negative excess entropy production is then $T_0E\,[\sigma\nabla^2 = 0] = M$. Since the manifold $T_0E\,[\sigma\nabla^2 < -\pi^2] = T_0E\,[\sigma(\mathscr{L}) < 0]$, this manifold contains the normal modes along which the excess entropy production is positive (the stable manifold.)

## VI. THE STRUCTURAL STABILITY PROBLEM

The idea of topological equivalence of dynamical systems was introduced through the concept of structural stability by Adronov and Pontriaguin (see Ref. 1). Consider the problem

$$\dot{\mathbf{x}} = \boldsymbol{\varphi}(\mathbf{x}), \quad \mathbf{x} \in R^2 \tag{52}$$

defined on the disk $D$. The system (52) is called structurally stable if given any arbitrary $\epsilon > 0$, there exists $\delta > 0$ such that for any $C^1$-perturbation $\boldsymbol{\mu}(\mathbf{x})$ of $\boldsymbol{\varphi}(\mathbf{x})$ that obeys

$$\|\boldsymbol{\mu}(\mathbf{x})\| < \delta, \tag{53}$$

$$\left|\frac{\partial}{\partial x i}\mu_j(x)\right| < \delta, \quad i,j = 1,2, \tag{54}$$

there exists an $\epsilon$-homeomorphism $H: D \to D$ that maps the trajectories of $\boldsymbol{\varphi}(\mathbf{x})$ onto the trajectories of $\boldsymbol{\varphi}(\mathbf{x}) + \boldsymbol{\mu}(\mathbf{x})$. $H$ obeys

$$\|\mathbf{x} - H(\mathbf{x})\| < \epsilon. \tag{55}$$

From the Adronov and Pontriaguin characterization of structurally stable systems follows that problem (52) is structurally stable if the singularities of $\boldsymbol{\varphi}(\mathbf{x})$ are hyperbolic.[7]

The following result holds.

**Theorem VII**: The system with the kinetics given by (52) is structurally unstable if and only if $\Lambda^S_{\alpha\beta} = 0$ along a certain normal mode (cf. Sec. III).

*Proof*: $\Lambda^S_{\alpha\beta} = 0$ along a normal mode implies that the singularity $\mathbf{x}_0$ of $\boldsymbol{\varphi}(\mathbf{x})$ is not hyperbolic, reciprocally, if there is a normal mode of perturbation for which $w_1 = 0$, then $\Lambda^S_{\alpha\beta} = 0$. Consider again the restriction of the reaction-diffusion problem (4) to the manifolds $T_0E\,[\sigma(\mathscr{B}) = -n^2\pi^2] = W_n (n = 0,1,2...)$. These restrictions are two-dimensional problems of the form (52). The problem (4) is structurally unstable if and only if there exists a $W_n$ such that the restriction to $W_n$ is structurally unstable. Therefore the following theorem holds.

**Theorem VIII**: If $\delta^2_m S$ is a constant of motion of the restriction of system (4) to a $W_n$, then the system is structurally unstable.

*Proof*: Since $(d/dt)\delta^2_m S = 0$, we get $P[\delta_m S] = \delta_m P = 0$ but this implies that $W_n$ is locally attractive, so the singularity $\mathbf{u} = \mathbf{0}$ is a center, hence the system is structurally unstable.

The system discussed in Sec. IV(b) becomes structurally unstable for the following values of the parameter $B$:

$$B_n = 1 + A^2 + 2n^2\pi^2 D \quad (D_x = D_y). \tag{56}$$

The center manifolds become, respectively,

$$W_n = T_0E\,[\sigma\hat{\mathscr{B}} = -n^2\pi^2] \quad (A > n^2\pi^2 D).$$

## VII. SPATIALLY PERIODIC CENTER MANIFOLDS

So far we have not considered the case in which simultaneously both excesses $\delta_m P$ and $\delta_m \pi$ vanish. Let us first state the following theorem.

**Theorem IX**: If a locally attractive normal mode obeys $\delta_m \pi = 0$, then it is a time-independent mode.

*Proof*: Since the mode is locally attractive, it must be contained in $T_0E\,[\mathrm{Re}\,\sigma(\mathscr{L}) = 0]$ and since $\delta_m \pi = 0$, also the imaginary part of $w$ in Eq. (12) vanishes, so the mode is time independent.

Obviously the purely spatial modes are all contained in the manifold $T_0E\,[\sigma\mathscr{L} = 0]$. This implies that the operator

$$U_n = J(\boldsymbol{\varphi}) + \mathbf{D}(-n^2\pi^2), \quad n = 0,1,2,..., \tag{57}$$

becomes singular for a certain value of $n$. Equivalently, there exists a normal mode contained in $W_n$ such that $w = 0$ along it. This mode is then a spatially periodic attractor along which $\delta^2_m S$ becomes (trivially) a constant of motion. The existence of a spatial organization indicates that the system is structurally unstable. (This is the point of exchange of stabilities[1].) As an example consider again the system defined in Sec. IV(b): since $J(\boldsymbol{\varphi})$ is nonsingular, the purely spatial mode should be contained in the manifold $\bigcup_{j>1} W_j$.

If we assume that the spatial attractor is contained in $W_n$, then

$$B = 1 + A^2 + n^2\pi^2 D + A^2/n^2\pi^2 D \quad (D = D_x = D_y). \tag{58}$$

Regarding this familiar example, one can therefore state the following theorem.

**Theorem X**: If any of the relations (56) or (58) (the last ones, only when $n \geqslant 1$) holds, then there exists a $C^1$-perturbation of the kinetic equations (52) that will lead to a topologically nonequivalent phase portrait, that is, the homeomorphism $H$ fails to exist (cf. Sec. V).

This $C^1$-perturbation could be, for example, a new depleting step introduced in the reactor for one of the intermediates $X$ or $Y$.

## VIII. FLUCTUATIONS

The probability of small fluctuations around a far from equilibrium steady state was confirmed to obey the relation

$$\bar{P} = \text{probability} \sim \exp(\tfrac{1}{2}\delta^2 S_{ss}/k). \tag{59}$$

Here $\bar{P}$ is the probability of fluctuations, $\delta^2 S_{ss}$ is the second variation of the entropy starting at the steady state, and $k$ is the Boltzmann's constant. This is an extrapolation of the

Einstein's formula valid for isolated systems at equilibrium. (See Nicolis and Babloyantz.[8])

Let us consider, first, the restriction of our system to the manifold $T_0E\,[\mathrm{Re}\,\sigma(\mathscr{L})<0]$. Since along this manifold, we have asymptotic stability, we obtain

$$\frac{1}{2}\frac{d}{dt}\delta^2 S = P\,[\delta S\,] > 0, \tag{60}$$

and also

$$\frac{d}{dt}\bar{P}\Big|_{T_0E\,[\mathrm{Re}\,\sigma(\mathscr{L})<0\,]} > 0. \tag{61}$$

Equation (61) implies the regression of the system to the local maximum of the function $\bar{P}$ (obtained at the steady state) along the manifold $T_0E\,[\mathrm{Re}\,\sigma(\mathscr{L})<0]$. The relation implying the regression of the system to the local maximum is not fulfilled if the manifold $T_0E\,[\mathrm{Re}\,\sigma(\mathscr{L})=0]$ is nontrivial.

Consider the case in which $T_0E\,[\sigma(\mathscr{L})=0]$ is nontrivial. Clearly, $\bar{P}$ is constant along this mode since $\delta^2 S$ and $\delta_m^2 S$ are constants of motion. In this case, although along $T_0E\,[\mathrm{Re}\,\sigma(\mathscr{L})<0]$ there is regression to the steady state, since $T_0E\,[\sigma(\mathscr{L})=0]$ is locally attractive, the fluctuations in $H^2([0,1])\times H^2([0,1])$ grow till the system reaches the local maximum in $\bar{P}$ at $T_0E\,[\sigma(\mathscr{L})=0]$. This allows us to give a new criteria for structural instability.

**Theorem XI**: The following three conditions are equivalent: (a) the system is structurally unstable; (b) there exists a normal mode along which the regression condition $(d\bar{P}/dt)>0$ is violated; and (c) the locally attractive manifold $\Omega$ is nontrivial.


[1] P. Glansdorff and I. Prigogine, *Structure, Stability and Fluctuations* (Wiley–Interscience, New York, 1971), Chap. VI, VII, and XIV.
[2] R. Lefever, G. Nicolis, and I. Prigogine, J. Chem. Phys. **47**, 1045 (1967).
[3] S. R. De Groot and P. Mazur, *Non-Equilibrium Thermodynamics* (North-Holland, Amsterdam, 1962), Chaps. III and IV.
[4] T. Kato, *Perturbation Theory for Linear Operators* (Springer, New York, 1966).
[5] P. Holmes and J. Mardsen, "Bifurcation to divergence and flutter in flow-induced oscillations: an infinite dimensional analysis," *Control of Distributed Parameter Systems* (Pergamon, Oxford, 1978), pp. 133.
[6] J. Mardsen, Bull. Am. Math. Soc. **84**, 1139 (1978).
[7] A. Adronov and L. S. Pontriaguin, Dokl. Akad. Nauk. SSSR **14**, 247 (1937).
[8] G. Nicolis and A. Babloyantz, J. Chem. Phys. **51**, 2632 (1969).