Information theory and Riemann spaces: An invariant reformulation

E. Duering,* D. Otero, A. Plastino,† and A. N. Proto
Departamento de Física, Comisión Nacional de Energía Atómica, 1429 Buenos Aires, Argentina
(Received 7 November 1984)

A geometric representation for information theory is introduced by recourse to a covariant formulation, according to the customary procedure employed in connection with Riemann spaces. The central tool of this representation is the metric tensor that characterizes the particular dynamics of a given system and yields the corresponding quantal invariants.

I. INTRODUCTION

The information-theoretic (IT) approach to statistical mechanics has been pioneered by Jaynes[1] more than twenty years ago. Since then, numerous applications of his ideas can be found in the literature,[2–4] especially in connection with approaches to irreversible thermodynamics. Recently, several studies[5–7] have provided a way to insert IT concepts into quantum dynamics, thereby giving rise to an interesting and suggestive combination of the microscopic and macroscopic conceptions of nature. For example, a connection between IT and the Ehrenfest theorem[5] has been outlined and expressions for usual thermodynamical relationships in terms of the expectation values of quantal operators have been derived.[8] In particular, "Onsager-like" coefficients can easily be derived using a quantal linear-response theory.

These developments are based in the idea of regarding the entropy as a constant of motion. This well-known fact provides us with a dynamical conception of the entropy, a quantity broadly used within thermodynamical contexts, although not often seen within quantum dynamical ones.[1–8] Quantal friction is also a good example of this type of approach, as shown in Ref. 6.

In the light of the above-mentioned works,[5,6,8] it may be of interest to look into the possibility of extracting from IT new insights into well-known concepts, and regard it as a tool with which one can do much more than reformulate statistical mechanics.

In this sense, it should be useful (and this is the aim of the present effort) to carefully study the physical context into which the two main ingredients of IT, the Lagrange multipliers $\{\lambda_i\}$ and the operator set $\{O_i\}$, are inserted. This purpose is achieved by describing them as the geometric elements of a Riemann space, each of them having a covariant or contravariant behavior. This covariant formulation of IT results in the interpretation of λ and O as dual spaces. Within this geometric representation, the Massieu-Planck function λ_O is the potential function of the vectorial field of operators $\{O_i\}$ and the sources are the usual quantal dispersions.

As a bonus, we have outlined the procedure needed in order to construct all the invariants of motion for a given system, valid also for time-dependent Hamiltonians. As is well known, these invariants provide us with two main results: the identification of symmetries and the solution of the time-dependent Schrödinger equation.

This work is organized as follows: In Sec. II we give a brief review of IT basic concepts, in Sec. III we outline the covariant formulation of IT, Sec. IV is devoted to some illustrative examples, and in Sec. V some suggestions and conclusions are drawn.

II. THE MAXIMUM-ENTROPY FORMALISM

We shall present here a brief summary concerning information theory and the least biased (probability) assignment criterium, which is usually referred to as the maximum-entropy principle (MEP). Within the IT context, the statistical operator (or density matrix) $\hat{\rho}$ is constructed[1–7] starting from the knowledge of the expectation values, of, say, $\hat{M}$ operators $\hat{O}_j$ ($\hat{O}_0=\hat{I}=$identity operator),

$$\langle \hat{O}_j/\hat{\rho}\rangle=\mathrm{Tr}[\hat{\rho}(t)\hat{O}_j]=o_j,\quad j=0,1,\ldots,M\ . \tag{2.1}$$

The subindex "0" refers to the normalization condition $\mathrm{Tr}\hat{\rho}=1$. The operator is given, within the IT framework, by

$$\hat{\rho}=\exp\left[-\lambda_0-\sum_{j=1}^{M}\lambda_j(t)\hat{O}_j\right], \tag{2.2}$$

$$\ln\hat{\rho}=-\lambda_0-\sum_{j=1}^{M}\lambda_j(t)\hat{O}_j \tag{2.3}$$

in terms of the $M+1$ Lagrange multipliers, λ_i, $i=0,1,\ldots,M$, determined so as to fulfill Eq. (2.1). The density matrix $\hat{\rho}$ maximizes the entropy, $S[\hat{\rho}]$ given (in units of the Boltzmann constant) by

$$S[\hat{\rho}]=-\mathrm{Tr}(\hat{\rho}\ln\hat{\rho})=\lambda_0+\sum_{j=1}^{M}\lambda_j\langle\hat{O}_j/\hat{\rho}\rangle\ , \tag{2.4}$$

subject to the constraints given by Eq. (2.1). The operators $\hat{\rho}(t)$ and $\ln\hat{\rho}(t)$ obey, respectively,[2] the equations of motion

$$i\hbar\frac{\partial\hat{\rho}}{\partial t}=[\hat{H}(t),\hat{\rho}(t)] \tag{2.5}$$

and

$$i\hbar\frac{\partial}{\partial t}\ln\hat{\rho}(t)=[\hat{H}(t),\ln\hat{\rho}(t)] . \tag{2.6}$$

It is well known[9] that if $\hat{\rho}$ is constructed so as to obey Eq. (2.5), then S is a constant of the motion.[5–8] Consequently, one should endeavor to find those (relevant) operators entering Eq. (2.2) so as to satisfy Eq. (2.5) [or Eq. (2.6)] in order to guarantee that S is a constant of the motion. Using Eqs. (2.2) and (2.5) it is easy to verify that the relevant operators are those that close a partial Lie algebra under commutation with the Hamiltonian $\hat{H}$,

$$[\hat{H}(t),\hat{O}_j]=\sum_{j=0}^{q} g_{ij}\hat{O}_i , \tag{2.7}$$

where the g_{ji} are the elements (c-numbers) of a $q\times q$ matrix, G (which may depend upon the time if $\hat{H}$ is time dependent). Then, the number of observables we need is q [see Eq. (2.1)]. The case $M<q$ is discussed in Ref. 5. For our present purposes we shall assume $M=q$.

Equation (2.7) constitutes the central requirement to be fulfilled by the operators entering the density matrix. Moreover, the closure condition (2.7) on the $\hat{O}_j$ leads to the fact that the time-dependent Schrödinger equation [or equivalently Eq. (2.6)] can be replaced by a set of coupled equations for the λ_i's,[5–7]

$$\frac{d}{dt}\lambda_i=\sum_{l=0}^{q} g_{il}\lambda_l , \tag{2.8}$$

which is easily obtained using Eqs. (2.2), (2.5), and (2.7).

The Lagrange multipliers are related to $\langle O_r/\hat{\rho}\rangle$ by[1,2]

$$-\frac{\partial\lambda_0}{\partial\lambda_j}=\langle\hat{O}_j/\hat{\rho}\rangle . \tag{2.9}$$

The temporal evolution of the expectation values of the operators [Eq. (2.1)] can be easily obtained by recourse to the condition

$$S(t_0)=S(t) , \tag{2.10}$$

as shown in Ref. 5 or using sum rules.[2] Of course, in both cases, the result is (using a vectorial notation)

$$\langle\hat{O}\rangle_t=\langle O\rangle_{t_0}\underline{F}(t,t_0) , \tag{2.11}$$

where $\langle O\rangle_t,\langle O\rangle_{t_0}$ are row matrices and $\underline{F}(t,t_0)$ is a square matrix defined by[2]

$$-\frac{\partial\underline{F}}{\partial t}=\underline{F}\cdot\underline{g} . \tag{2.12}$$

In the Heisenberg representation we can also write

$$\hat{O}_t=\hat{O}_{t_0}\underline{F}(t,t_0) , \tag{2.13}$$

which, in respect to the Lagrange multipliers λ, can also be arranged as a column matrix,

$$\lambda^t=F^{-1}(t,t_0)\lambda^{t_0} . \tag{2.14}$$

The subscript or superscript t or t_0 in Eqs. (2.11) and (2.14) indicates whether we are working with a covariant or a contravariant vector, respectively.

As we have pointed out before, the operators entering $\hat{\rho}$ are those which fulfill Eq. (2.7); besides, the $\underline{F}(t,t_0)$ matrix (or $\underline{G}$) contains all the information about the dynamics of the Hamiltonian. Within this context, Eqs. (2.11) and (2.13) define covariant vectors with respect to the transformation characterized by $F(t,t_0)$. This matrix is not necessarily a unitary one.[9] Instead, λ^t is defined as a contravariant vector by Eq. (2.9).

III. INFORMATION THEORY IN COVARIANT FORM

A. The vectorial equation

As demonstrated in the preceding section, the $\{\lambda_i\}$ and $\{\hat{O}_i\}$ sets can be recast in contravariant (λ^t) and covariant ($\hat{O}_t$ or $\langle\hat{O}\rangle_t$) forms, respectively. In doing so, the entropy [Eq. (2.4)] reads

$$S=\langle\hat{O}\rangle_t\lambda^t , \tag{3.1}$$

which clearly allows for a geometrical interpretation of the invariance of S.[2] In the Heisenberg representation, and using Eqs. (2.13) and (2.14), the surprisal $\ln\hat{\rho}$ [Eq. (2.3)] becomes

$$-\ln\hat{\rho}=\hat{O}_t\lambda^t . \tag{3.2}$$

However, in order to be able to define vectorial Riemann spaces with $\langle\hat{O}\rangle_t$ (or $\hat{O}_t$) and λ^t as elements, we need to define scalar products of the form

$$\langle\hat{O}\rangle_t\langle\hat{O}\rangle^t=\langle\hat{O}\rangle_0\langle\hat{O}\rangle^0 , \tag{3.3a}$$

or

$$\hat{O}_t\hat{O}^t=\hat{O}_0\hat{O}^0 \tag{3.3b}$$

and

$$\lambda_t\lambda^t=\lambda_0\lambda^0 , \tag{3.3c}$$

where the "zero" index refers to the "original" [before the $\underline{F}(t,t_0)$ temporal transformation is applied] vector. We now need to find the metric tensor of the space, e (with $ee^{-1}=1$), for which

$$\langle\hat{O}\rangle^t=e\langle\hat{O}\rangle_t , \tag{3.4}$$

where on the right-hand side (rhs) we are referring to a transposed vector, or in other words, we need to find the transformation which allows us to change the character of the vector from covariant to contravariant class or vice versa. Additionally (and equivalently),

$$\hat{O}^t=e\hat{O}_t , \tag{3.5}$$

and, correspondingly,

$$\lambda^t=\underline{e}'\lambda_t , \tag{3.6}$$

where, of course, $\underline{e}$ and $\underline{e}'$ are determined so as to satisfy the relations (3.3). Thus, using Eqs. (3.3) and (3.4) we find ($\tilde{\underline{F}}$ indicates transposed matrix)

$$\underline{F}\,\underline{e}\,\tilde{\underline{F}}=\underline{e} , \tag{3.7a}$$

and, equivalently,

$$\tilde{\underline{F}}^{-1}\underline{e}'\underline{F}^{-1}=\underline{e}' . \tag{3.7b}$$

In particular, if $\underline{F}$ is a unitary matrix (which is not always the case) we can write

$$[\underline{F},\underline{e}]=\underline{0} \text{ and } [\underline{F},\underline{e}']=\underline{0} . \tag{3.8}$$

The metric tensor will be, in general,[9] of the block form

$$\begin{bmatrix} \underline{e} & \mathbb{1} \\ \mathbb{1} & \underline{e}' \end{bmatrix} \tag{3.9}$$

and will contain $2(q+1)\times 2(q+1)$ elements, with q defined in Eq. (2.7).

We may ask ourselves whether there is a unique choice of the metric $\underline{e}$. The answer is a qualified *no*, and, moreover, we argue that this constitutes an advantageous facet of our formalism. Indeed, we are not tied to any specific basis. Were we to be so tied, the choice would be certainly a unique one (cf. Ref. 9, Chap. 2.III). In physical terms, this would entail selecting specific numerical values for our invariants, which we have not seen fit to do, as one of the main advantages of resorting to tensor calculus is to free ourselves from any basis. An obvious way to make our choice of $\underline{e}$ a unique one would be to work with a canonical basis defined by, for example,

$$\begin{bmatrix} \langle \hat{x}_0 \rangle_1 \\ \langle \hat{p}_0 \rangle_1 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} , \qquad \begin{bmatrix} \langle \hat{x}_0 \rangle_2 \\ \langle \hat{p}_0 \rangle_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} . \tag{3.10}$$

In any case, as stated above, we wish to keep our formalism as general as possible, and will illustrate below, with reference to some specific examples, how to select $\underline{e}$. It will be clearly shown that the lack of uniqueness does not pose any difficulty whatsoever.

Until now we have dealt with the mathematical tools of our approach, of which the metric tensor $\underline{e}$ is the most important. Now we shall insert the elements of IT. Let us begin with Eq. (2.9), from which we can write

$$\frac{\partial^2 \lambda_0}{\partial \lambda\, \partial \lambda_n} = -\frac{\partial \langle O_n \rangle}{\partial \lambda_m} = -\frac{\partial \langle O_m \rangle}{\partial \lambda_n} = K_{nm} = K_{mn} , \tag{3.11a}$$

with

$$K_{mn} = \tfrac{1}{2}\langle [\hat{O}_m,\hat{O}_n]_+ \rangle - \langle \hat{O}_m \rangle \langle \hat{O}_n \rangle , \tag{3.11b}$$

where $[,]_+$ denotes anticommutation.

For the sake of brevity, we leave the demonstration of Eqs. (3.11) to the Appendix. The K space is the direct Kronecker or tensor product of the $\langle \hat{O} \rangle_t$, having, then, a covariant character. The corresponding tensor $\underset{\sim}{K}$ is thus expressed as a (covariant) second-order tensor ($\hat{O}_t$ is referred to the transposed vector)

$$K_t = \tfrac{1}{2}\langle [\tilde{\hat{O}}_t,\hat{O}_t]_+ \rangle - \langle \tilde{\hat{O}}_t \rangle \langle \hat{O}_t \rangle . \tag{3.11c}$$

In operator form, Eq. (3.11c) reads

$$K_t = \tfrac{1}{2}[\tilde{\hat{O}}_t,\hat{O}_t]_+ - \langle \tilde{\hat{O}}_t \rangle \langle \hat{O}_t \rangle , \tag{3.11d}$$

although this is not the only possible way of arranging things, any other expression compatible with (3.11) being equally acceptable.

We are now in a position to define the divergence and the rotor of $\langle O \rangle_t$ with respect to λ^t, if we are willing to consider the latter as a "coordinate," remembering that we are working with an $(N+1)$-dimensional "space." We obtain

$$\nabla_\lambda \langle \hat{O} \rangle = -\sum_n K_{nn} \tag{3.12}$$

and

$$\nabla_\lambda \times \langle \hat{O} \rangle = 0 . \tag{3.13}$$

The results (3.12) and (3.13) are a consequence of Eqs. (3.7) and can be traced to the fact that the tensor metric $\underline{e}$ is independent of the λ's.[9] Moreover, Eq. (3.11a) can be rewritten as a Poisson equation,

$$\Delta \lambda_0 = -\sum_n K_{nn} . \tag{3.14}$$

Equations (3.13) and (3.14) show that the Massieu-Planck function λ_0 (Ref. 10) is the potential function of an irrotational field, whose sources are the quantal dispersions. This is in agreement with Feynman's philosophy (path-integral method).[11] An elementary example of Eq. (3.14) is given by the specific heat of a canonical system.[8]

If all the dispersions vanish, then Eq. (3.14) is Laplacian, with the trivial solution

$$\lambda_0 = -\sum_{j=1}^{N} \langle \hat{O}_j \rangle \lambda_j , \tag{3.15}$$

where by hypothesis $\langle O_j \rangle$ is independent of the λ's and the entropy is equal to zero. Of course, this situation obtains for a complete set of commuting operators.

Summarizing, we have defined the gradient and rotor of the $\langle \hat{O} \rangle$ vector, by resorting to using the operators which fulfill Eq. (2.7). Besides, we established in Eq. (3.14) the relation between the quantal fluctuations and λ_0 (which is equivalent to the partition function).

In the following subsection we shall delve further into the physical meaning of the (direct-product) K space.

B. The K space

As we have seen in the preceding subsection, the quantal correlations K_{nm} play an important role in our formalism (see Appendix). They can be conveniently expressed by introducing the Kronecker product[9]

$$\hat{K} = \hat{O} \otimes \hat{O} \tag{3.16a}$$

or

$$K = \langle \hat{O} \rangle \otimes \langle \hat{O} \rangle . \tag{3.16b}$$

The temporal evolution is given, using Eq. (2.11) and the usual rules of the direct product,[9] by

$$K(t)=[\underline{F}\otimes\underline{F}]K(t_0)=[\underline{F}]_2\underline{K}(t_0) \tag{3.17a}$$

for column vectors, or, equivalently, for square matrices, by

$$\underline{K}(t)=\tilde{\underline{F}}\underline{K}(t_0)\underline{F} \tag{3.17b}$$

or

$$\underline{K}_t=\tilde{\underline{F}}\underline{K}_{t_0}\underline{F}\ . \tag{3.17c}$$

For the sake of simplicity, in the following we shall work with square matrices.

The expressions we have derived above can be used in combination with Eq. (2.4). In particular, taking the derivative of Eq. (2.4) with respect to λ_i, and using Eq. (2.9), we obtain

$$\frac{\partial S}{\partial \lambda_j}=\sum_{r=1}^{q}\lambda_r\frac{\partial\langle \hat{O}/\hat{\rho}\rangle}{\partial\lambda_j} \tag{3.18a}$$

and, using Eq. (3.11a) and the covariant notation

$$\nabla_\lambda S=-\underline{K}_t\lambda^t\ , \tag{3.18b}$$

so that multiplying by λ^t we obtain the directional derivative of S,

$$-\lambda_t\nabla_\lambda S=\tilde{\lambda}^t\underline{K}_t\lambda^t=\sigma\ . \tag{3.19}$$

σ is an invariant since, using Eqs. (3.17) and (2.14), we immediately obtain

$$\tilde{\lambda}^t\underline{K}_t\lambda^t=\tilde{\lambda}^t\tilde{\underline{F}}\underline{K}_0\underline{F}\lambda^t=\tilde{\lambda}^0\underline{K}_0\lambda^0\ , \tag{3.20}$$

which is formally analogous to the entropy source (see, for example, the illuminating report by Nicolis[12]), with a positive definite sign. However, Eqs. (2.5)—(2.7) imply no entropy variation and a continuity equation cannot be defined within the $(\lambda,\hat{O})$ space.[13]

We can extend our formalism further by studying temporal evolutions in the K space, and, using Eqs. (3.4), (3.5), and (3.11d), we can write (in covariant notation)

$$\underline{K}^t=\underline{e}\,\underline{K}_t\tilde{\underline{e}}\ , \tag{3.21}$$

and multiplying from the right by K_t,

$$\underline{K}^t\underline{K}_t=\underline{e}\,\underline{K}_t\tilde{\underline{e}}\,\underline{K}_t\ . \tag{3.22}$$

Using Eq. (3.17c), Eq. (3.19) can be rewritten as

$$\begin{aligned}\underline{K}^t\underline{K}_t&=\underline{e}\,\underline{K}_t\tilde{\underline{e}}\,\tilde{\underline{F}}\,\underline{K}_{t_0}\underline{F}\\&=\underline{e}\,\tilde{\underline{F}}\,\underline{K}_{t_0}\underline{F}\,\tilde{\underline{e}}\,\tilde{\underline{F}}\,\underline{K}_{t_0}\underline{F}\\&=\underline{F}^{-1}\underline{e}\,\underline{K}_t\tilde{\underline{e}}\,\underline{K}_{t_0}\underline{F}\\&=\underline{F}^{-1}\underline{K}^0\underline{K}_0\underline{F}\ .\end{aligned} \tag{3.23}$$

IV. EXAMPLES

In this section we shall illustrate our formalism with reference to some specific and simple examples. The main objective here is that of explicitly showing how to evaluate the metric $\underline{e}$. The general procedure can be concisely described as follows:

(a) Construct the $\underline{g}$ matrix by implementing Eq. (2.7).

(b) Construct the $\tilde{\underline{F}}$ matrix, following Eq. (2.12).

(c) Finally, find the $\underline{e}$ matrix by solving the linear system of q equations and q unknowns [cf. Eq. (2.7)] that arises out of Eqs. (3.7).

A. The free particle

For a free particle $H=\hat{p}^2/2m$ with the operator set $\{\hat{1},\hat{x},\hat{p}\}$ [which fulfills Eq. (2.7) (see Ref. 5)] we find, using Eqs. (2.7) and (2.12),

$$F=\begin{bmatrix}1&0&0\\0&1&t\\0&0&1\end{bmatrix} \tag{4.1a}$$

and, after solving the system (3.7), we find

$$e=\begin{bmatrix}e_{11}&0&e_{13}\\0&0&e_{23}\\e_{31}&-e_{23}&e_{33}\end{bmatrix}, \tag{4.1b}$$

where the e_{ij} are c-numbers to be determined. The indices 1,2,3 refer to the operators $\{\hat{1},\hat{x},\hat{p}\}$, respectively. Then

$$\hat{O}_t\hat{O}^t=\mathbb{1}e_{11}+\hat{p}(e_{13}+e_{31})+[\hat{x},\hat{p}]e_{23}+e_{33}\hat{p}^2\ , \tag{4.2}$$

which is a linear combination of quantal invariants. [Notice that it is not necessary to fix an *a priori* value to the lhs of (4.2), as explained above.] Heisenberg's uncertainty principle is seen to provide us with a seldom referred to invariant. If we calculate $\langle\hat{O}_t\rangle\langle\hat{O}^t\rangle$, then only classical invariants are involved, as expected. For the vector λ we obtain a similar structure.

If, instead, we use the operator set

$$\{\hat{1},\hat{x}^2,\hat{L}=(\hat{p}\hat{x}+\hat{x}\hat{p})/2,\hat{p}^2\}\ ,$$

which also fulfills Eq. (2.7), then we obtain a less trivial invariant combination (the indices 4,5,6 refer to the operators $\hat{x}^2,\hat{L},\hat{p}^2$, respectively)

$$\begin{aligned}\hat{O}^t\hat{O}_t=e_{11}\mathbb{1}+(e_{16}+e_{61})\hat{p}^2+e_{46}(\hat{x}^2\hat{p}^2+\hat{p}^2\hat{x}^2-2\hat{L}^2)\\+e_{56}(Lp^2-p^2L)+e_{66}p^4\ ,\end{aligned} \tag{4.3a}$$

with

$$\begin{bmatrix}e_{11}&0&0&e_{61}\\0&0&0&e_{46}\\0&0&-2e_{46}&e_{56}\\e_{16}&e_{46}&-e_{56}&e_{66}\end{bmatrix} \tag{4.3b}$$

and, for the mean values, $(\langle O_t\rangle\langle O^t\rangle)$, only the first, second, and fifth terms survive.

B. The harmonic oscillator

For a one-dimensional harmonic oscillator of frequency ω (mass m) we obtain,[5] following the same procedure as before,

$$H=\frac{p^2}{2m}+m\omega^2x^2\ , \tag{4.4a}$$

$$L=\tfrac{1}{2}(\hat{p}\hat{x}+\hat{x}\hat{p})\ , \tag{4.4b}$$

$$\begin{aligned}\hat{O}_t\hat{O}^t=&\,e_{11}\mathbb{1}+(e_{21}+e_{12})(m\omega\hat{x}^2+\hat{p}^2)\\&+e_{22}[(m\omega\hat{x})^4+2(m\omega\hat{L})^2+\hat{p}^4]\\&+e_{23}[(m\omega)^3[\hat{x}^2,\hat{L}]+m\omega[\hat{L},\hat{p}^2]]\\&+m\omega^2e_{24}(\hat{x}^2\hat{p}^2-2\hat{L}^2+\hat{p}^2\hat{x}^2)\ ,\end{aligned} \tag{4.4c}$$

if the relevant set is chosen to be $\{\hat{1},(m\omega\hat{x})^2,\hat{p}^2,(m\omega\hat{L})^2\}$, and only the first three terms contribute to the mean value.

C. Larmor precession

The Hamiltonian is

$$H=\omega_L S_Z\ , \tag{4.5}$$

where S_Z is an operator corresponding to the Z component of the spin ($\hat{S}^2=\hat{S}_x^2+\hat{S}_y^2+\hat{S}_z^2$) of a particle of mass $\underline{m}$ and charge $\underline{e}$, such that

$$\omega_L=-\frac{eB}{2mc} \tag{4.6}$$

is the Larmor frequency in an external magnetic field of strength B.

The set of relevant operators results $\{\hat{1},\hat{S}_x,\hat{S}_y\}$ and the $\underline{e}$ matrix is

$$e=\begin{bmatrix}e_{11}&0&0\\0&e_{22}&e_{23}\\0&-e_{23}&e_{22}\end{bmatrix}, \tag{4.7}$$

so that

$$O_tO^t=e_{11}\hat{1}+e_{22}(\hat{S}_x^2+\hat{S}_y^2)+e_{23}[\hat{S}_x,\hat{S}_y]\ , \tag{4.8}$$

which is a sum of invariants of motion as both $\hat{S}_x^2+\hat{S}_y^2$ and $[\hat{S}_x,\hat{S}_y]$ commute with the Hamiltonian. As for the mean values, only the first and second terms contribute.

D. Sets of identical systems

An interesting feature of this formalism lies in the fact that two systems described by the same H and O_r set also generate invariants of the form

$$\hat{O}_t{}^1\hat{O}_2{}^t=\hat{O}_0{}^1\hat{O}_2{}^0\ , \tag{4.9}$$

where the indices 1,2 indicate to which system one is referring. Also,

$$\underline{e}\,\underline{K}_t^{11}\underline{\tilde{e}}\,\underline{K}_t^{22}=\underline{F}^{-1}\underline{e}\,\underline{K}_{t_0}^{11}\underline{\tilde{e}}\,\underline{K}_{t_0}^{22}\underline{F} \tag{4.10}$$

or

$$\underline{e}\,\underline{K}_t^{12}\underline{\tilde{e}}\,\underline{K}_t^{12}=\underline{F}^{-1}\underline{e}\,\underline{K}_{t_0}^{12}\underline{\tilde{e}}\,\underline{K}_{t_0}^{12}\underline{F} \tag{4.11}$$

(with $\hat{K}_{12}=\hat{O}_1\times\hat{O}_2$, etc.), and where $\hat{O}_t{}^i\hat{O}^t{}_i$ can be considered as a sort of generalized "pressure" and $O_t{}^jO^t{}_i$ are the nondiagonal elements of a generalized "stress tensor."

For a number of systems greater than two, the corresponding generalization is not simple, but in working it out correlations of the form $K_{ijk},K_{ijkl},\ldots$ will appear. Such elements do not appear to be useful in physics.

V. CONCLUSIONS

The formalism we have outlined deserves some additional comments. The covariant formulation of IT is a compact representation of the dual spaces of the λ's and $\hat{O}$'s. The quantum invariants that we have found are the scalar products defined in a space whose metric tensor is a characteristic both of $\hat{H}$ and the set $\{\hat{O}_r\}$. The procedure can be easily extended to the product "space" $\hat{K}$ $(=\hat{O}\otimes\hat{O})$; Eq. (27) and the possibility of applying it to generate an infinite number of independent invariants may be regarded as a basic property of the covariant products.

Our central tool here is the metric $\underline{e}$ that characterizes the particular dynamics of a given system and automatically yields the invariant combinations of the operator set defined by Eq. (2.1).

Within this geometric representation, the Massieu-Planck function λ_0 is the potential function of a vectorial field of operators [Eq. (2.1)] (or observables [Eq. (2.4)]), where the sources are the quantal dispersions [Eq. (2.5)].

Up to this point we have always assumed that the set of operators defined in Eq. (2.1) closes an algebra. If this is not the case, or if a "new" operator not contained in the original set is taken into account, the entropy and the K's are not well defined.

However it is possible, by means of a straightforward calculation, to obtain

$$\frac{\partial\langle O_m\rangle}{\partial\lambda_n}=K'_{m,n} \tag{5.1}$$

and ($\hat{A}$ is a "new" operator)

$$\frac{\partial\langle A\rangle}{\partial\lambda_n}=K'_{An}\ , \tag{5.2}$$

with

$$K'_{m,n}=\langle[\hat{O}_m,\bar{\hat{O}}_n]_+\rangle-\langle\hat{O}_m\rangle\langle\hat{O}_n\rangle\ , \tag{5.3}$$

$$K'_{Am}=\langle[\hat{A},\bar{\hat{O}}_m]_+\rangle-\langle\hat{A}\rangle\langle\hat{O}_m\rangle \tag{5.4}$$

(where $\bar{\hat{O}}$ is the Kubo transform of $\hat{O}$).[14,15]

Then (see Ref. 8 for further details)

$$K'_{mA}\neq K'_{Am},\ \ K'_{mn}\neq K'_{nm}\ , \tag{5.5}$$

except if $[\hat{A},\hat{O}_m]=0$, $[\hat{O}_n,\hat{O}_m]=0$.

Besides, as a consequence of Eq. (5.2) we can write

$$\nabla_\lambda\wedge\langle\hat{O}\rangle=J\ , \tag{5.6}$$

where the components of J are

$$J_\rho=K'_{mn}-K'_{nm} \tag{5.7}$$

(J represents a hyperplane with $\lambda=N-2$ dimensions).

We see that in this case the $\langle O\rangle$-field is not a conservative one. In other words, the change from λ_0^1 to λ_0^2 is "path dependent," and, of course, the system follows that one that increases the entropy.

As a concluding comment we can say that this covariant formulation provides us with a rather powerful methodology for dealing with the basic elements of IT, the λ's and $\hat{O}$'s, and may shed some light upon their dual-space characteristics.

ACKNOWLEDGMENTS

Two of us (E.D. and A.P.) acknowledge support by Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Argentina.

APPENDIX

To analyze the meaning of the correlation coefficient K_{mn}, defined in Eq. (3.11a), we begin taking derivatives in Eq. (2.4) with respect to a particular λ_j,

$$\frac{\partial S}{\partial \lambda_\delta} = +\frac{\partial \lambda_0}{\partial \lambda_j} + \sum_{r=1}^{q} \left[\delta_{rj} \langle \hat{O}_r \rangle + \lambda_r \frac{\partial(\langle \hat{O}_r / \hat{\rho} \rangle)}{\partial \lambda_j} \right] , \tag{A1}$$

where δ_{rj} takes into account the $\partial \lambda_r / \partial \lambda_j$ derivative. Remembering Eq. (2.9),

$$\frac{\partial \lambda_0}{\partial \lambda_j} = -\langle \hat{O}_j / \hat{\rho} \rangle , \tag{A2}$$

we find

$$\frac{\partial S}{\partial \lambda_j} = \sum_r \lambda_r \frac{\partial(\langle O_r / \rho \rangle)}{\partial \lambda_j} . \tag{A3}$$

The next step is to find an adequate expression for the derivative of the expectation value with respect to λ_j. It is readily verified that

$$\langle \hat{O}_r / \hat{\rho} \rangle = \tfrac{1}{2} \mathrm{Tr}(\hat{\rho} \hat{O}_r + \hat{O}_r \hat{\rho}) . \tag{A4}$$

In taking the derivative of Eq. (A4) with respect to λ_j, we factorize in the following way,

$$\hat{\rho} = e^{-\lambda_0} \exp \left[-\sum_{r=1}^{q} \lambda_r \hat{O}_r \right] , \tag{A5}$$

and in dealing with the exponential function on the rhs we apply the well-known expression

$$e^{A+\epsilon} = e^{A}(1+\bar{\bar{\epsilon}}) , \tag{A6}$$

with (derivatives taken only with respect to λ_j)

$$\hat{A} = -\left[\sum_{\substack{r=0 \\ r \neq j}} \lambda_r \hat{O}_r + \lambda_j \hat{O}_j \right] , \tag{A7}$$

$$\hat{\epsilon} = -\left[\hat{O}_j \delta \lambda_j + \frac{\partial \lambda_0}{\partial \lambda_j} \delta \lambda_j \right] , \tag{A8}$$

$$\bar{\bar{\epsilon}} = -(\bar{\bar{O}}_j \delta \lambda_j + \langle \hat{O}_j \rangle \delta \lambda_j) .$$

Then, it is readily verified that

$$-\frac{\partial \rho}{\partial \lambda_{j'}} = +\hat{\rho} \langle \hat{O}_j \rangle - \hat{\rho} \hat{O}_j , \tag{A9}$$

where $\bar{\bar{O}}_j$ is the Kubo transform of $\hat{O}_j$.

The desired expression is then

$$\frac{\partial \langle \hat{O}_r / \hat{\rho} \rangle}{\partial \lambda_j} = \tfrac{1}{2} \mathrm{Tr}[(\hat{\rho} \bar{\bar{O}}_j \hat{O}_r + \hat{O}_r \hat{\rho} \bar{\bar{O}}_j) - \langle \hat{O}_j \rangle \langle \hat{O}_r \rangle] . \tag{A10}$$

Consequently, if Eq. (A10) is inserted into Eq. (A3), we obtain a more useful expression. After using

$$\sum_{r=1}^{q} \lambda_r \hat{O}_r = \ln \hat{\rho} - \lambda_0 , \tag{A11}$$

the following equation obtains:

$$-\frac{\partial S}{\partial \lambda_j} = \sum_{r=1}^{q} \lambda_r \langle \hat{O}_r \rangle \langle \hat{O}_j \rangle - \mathrm{Tr}[\hat{\rho} \bar{\bar{O}}_j (\ln \hat{\rho} - \lambda_0) + \hat{\rho}(\ln \hat{\rho} - \lambda_0) \bar{\bar{O}}_j] . \tag{A12}$$

The Kubo transform can be easily eliminated from Eq. (A12) by introducing $\ln \hat{\rho}$ into the Kubo transform of $\hat{O}_j$ and noting that $\hat{\rho}$ and $\ln \hat{\rho}$ commute.

Finally, we arrive at

$$-\frac{\partial S}{\partial \lambda_j} = -\sum_{r=1}^{q} \lambda_r (\tfrac{1}{2} \langle [\hat{O}_r, \hat{O}_j]_+ \rangle - \langle \hat{O}_r \rangle \langle \hat{O}_j \rangle)$$
$$= -\sum_{r=1}^{q} \lambda_r K_{rj}^{(2)} , \tag{A13}$$

and, comparing Eqs. (A3) and (A13), we obtain

$$\frac{\partial \langle \hat{O}_r / \hat{\rho} \rangle}{\partial \lambda_j} = +K_{rj}^{(2)} . \tag{A14}$$

We have thus seen that the variation of the expectation value of one of the O_r with respect to a given λ_j Lagrange multiplier is given by the correlation between the corresponding $\hat{O}_r$ and $\hat{O}_j$ operators. When $\delta \lambda_j$ values can be measured (i.e., the temperature, chemical potential, etc.), we obtain, through Eq. (A14), the quantal-statistical correlation between the involved operators. Equation (A14) differs from Eq. (A10) in that it does not contain the Kubo transform.

*Permanent address: Universidad Nacional del Centro de la Provincia de Buenos Aires, Pinto 399, 700 Tandil, Argentina.
†Permanent address: Departamento de Física, Universidad Nacional de La Plata, Casilla de Correo 67, 1900 La Plata, Argentina.

[1]E. T. Jaynes, Phys. Rev. **106**, 620 (1957); **108**, 171 (1957).
[2]Y. Alhassid and R. D. Levine, J. Chem. Phys. **67**, 4321 (1977).
[3]W. T. Grandy, Jr., Phys. Rep. **62**, 175 (1980).
[4]O. Penrose, Rep. Prog. Phys. **42**, 1937 (1979); A. Katz, *Principles of Statistical Mechanics* (Freeman, San Francisco, 1967); A. Hobson, *Concepts of Statistical Mechanics* (Gordon and Breach, New York, 1971).
[5]D. Otero, A. Plastino, A. N. Proto, and G. Zannoli, Phys. Rev. A **26**, 1208 (1982).
[6]D. Otero, A. Plastino, A. N. Proto, and G. Zannoli, Z. Phys. A **316**, 323 (1984).
[7]Y. Alhassid and R. D. Levine, Phys. Rev. A **18**, 89 (1978).
[8]E. Duering, D. Otero, A. Plastino, and A. Proto, Phys. Rev. A (to be published).
[9]R. Gilmore, *Lie Groups, Lie Algebras and Some of Their Applications* (Wiley, New York, 1974).
[10]D. N. Zubarev, *Non-equilibrium Statistical Thermodynamics* (Consultants Bureau, New York, 1974).
[11]R. P. Feynman and A. R. Hibbs, *Quantum Mechanics and Path Integrals* (McGraw-Hill, New York, 1965).
[12]G. Nicolis, Rep. Prog. Phys. **42**, 236 (1979).
[13]E. S. Hernández and B. Remaud, Phys. Lett. **75A**, 269 (1980); H. R. Lewis, Phys. Rev. Lett. **18**, 510 (1967).
[14]R. P. Feynman, Phys. Rev. **84**, 108 (1951).
[15]R. Kubo, J. Phys. Soc. Jpn. **12**, 570 (1957).