Information and metrics in Hilbert space

M. Raviculé,[1] M. Casas,[2] and A. Plastino[1]
[1]Departamento de Física, Universidad Nacional de La Plata, Casilla de Correo 67, 1900 La Plata, Argentina
[2]Departament de Física, Universitat de les Illes Balears, E-07071, Palma de Mallorca, Spain
(Received 13 June 1996)

The concept of distance in Hilbert space is relevant in a variety of scenarios, in particular for investigating the quality of different approximations. In this work we study the relations between (i) statistical distances (SD) on a probability space, on the one hand, and (ii) different metrics on Hilbert space (MHS), on the other hand. As a result, we are able to establish some universal relations between SD and MHS and to apply them to one-dimensional problems. [S1050-2947(97)05003-8]

PACS number(s): 03.65.−w, 05.30.−d

I. INTRODUCTION

The concept of distance between different rays in (the same) Hilbert space is certainly a very important one. It is related, for instance, to different preparations of the same quantum system [1] and to the geometric properties of the quantum evolution submanifold [2]. It becomes relevant in discussing squeezed coherent states, displaced number states, generalized coherent spin states, etc. [2]. It can be used also, of course, in order to study the quality of different approximations, as very few problems are amenable to an exact treatment. From a rather different point of view, a connection can be found with the problem of detecting weak signals that cause small changes in the state of a system. This problem can be regarded as being equivalent to that of distinguishing ''neighboring'' states along a suitably parametrized path, which involves the concept of distance.

The present effort is inspired by the pioneer work of Wootters [1] and is concerned with the relationship between statistical and geometric distances. The former are distances on probability space and are determined by the size of statistical fluctuations [1]. The latter refer, of course, to metrics on Hilbert space.

Given two probability distributions f_1 and f_2 one can define a distance between them [1] (to be herefrom called the Wootters one),

$$D_W(f_1,f_2)=\lim_{n\to\infty}\frac{1}{\sqrt{n}}Y_n, \tag{1}$$

where Y_n is the maximum number of mutually distinguishable (in n trials) intermediate probabilities. The Wootters distance is *not* (in principle, at least) the same as the Euclidean distance, which reads

$$D_E(f_1,f_2)=\sqrt{\int |f_1-f_2|^2 d\tau}, \tag{2}$$

with $d\tau$ an appropriate volume element in that function space to which f_1,f_2 belongs [1]. Indeed, the distance function (1) was introduced by Fisher in his analysis on genetic drift [3]. Wootters was able to exhibit a notable finding: although D_W is not *a priori* related to the usual distance (or angle) between rays, these two kinds of distance *coincide*, a result that depends of quantum mechanics peculiarities elucidated by him [1] and by an illuminating study of Braunstein and Caves, which extended Wootters' work to the realm of density matrices [4].

In the present work we shall pursue Wootters' ideas and study a *different*, information-theory-related statistical distance. We shall be concerned with (i) its connection to Wootters' one, and (ii) the precise fashion in which these statistical distances can be related to ordinary Hilbert space metrics. It will be seen that some new, interesting physical insight is gained by recourse to this type of investigation. We begin in Sec. II by reviewing diverse distances that can be employed with reference to wave functions in Hilbert space. In Sec. III we analyze the different distances and are able to establish some universal relations between them, which supplement the original one discovered by Wootters. We apply our results to some one-dimensional soluble problems in Sec. IV and, finally, some conclusions are drawn in Sec. V.

II. DISTANCES BETWEEN WAVE FUNCTIONS

A. The Fubini-study metric

Suppose that a quantum state $\psi(\mathbf{x})$ is parametrized by n real parameters that we collectively denote by α, i.e., we write $\psi_\alpha(\mathbf{x})$. A set of rays corresponding to the states with all possible α values constitutes an n-dimensional manifold $\mathcal{K}$ of the Hilbert space $\mathcal{H}$. The geometric structure of the submanifold can be studied with the help of the Fubini-Study metric induced on it [5,6].

$$dS_F^2=1-|\langle\psi(\alpha)|\psi(\alpha+d\alpha)\rangle|^2, \tag{3}$$

where, for the sake of a lighter notation, we use herefrom

$$\psi_\alpha(\mathbf{x})\equiv\psi(\alpha), \tag{4}$$

and one assumes that states that differ by a change $\alpha\to\alpha+d\alpha$ can be related via the expansion

$$|\psi(\alpha+d\alpha)\rangle=|\psi(\alpha)\rangle+d\alpha\frac{d}{d\alpha}|\psi(\alpha)\rangle+\frac{1}{2}(d\alpha)^2\frac{d^2}{d\alpha^2}|\psi(\alpha)\rangle+\cdots. \tag{5}$$

This metric has been discussed by several authors [7–10]. The metric structure of the manifold is completely expressed by physical quantities, which are the uncertainties and correlations of Hermitian operators generating various evolutions of a given quantum state [2]. It is of importance to mention the fact that, as Braunstein and Caves [4] have pointed out, dS_F^2 is the maximum value of Wootters' distance, which means that the Fubini metric can be regarded as the statistical distance between neighboring pure states [4].

B. The Euclidean metric

This is included here for reasons of completeness. It reads, of course, for two wave functions ψ_1, ψ_2,

$$dS_E^2 = \int |(\psi_1 - \psi_2)|^2 d\tau. \tag{6}$$

Notice that, for the sake of an easier notation, we shall herefrom denote volume elements simply as $d\tau = dx$.

C. The Wootters distance

This is *not* the Euclidean distance. Wootters uses, instead [1],

$$dS_W^2 = (\arccos|\langle \psi_1|\psi_2\rangle|)^2 \equiv \gamma^2, \tag{7}$$

and with very good reasons: the angle (γ in our case) in Hilbert space is the only Riemannian metric on the set of rays, up to a constant factor, which is invariant under all unitary transformations [1]. Notice that, for very "close" rays ("neighboring states") the overlap $\langle \psi_1|\psi_2\rangle \approx 1$ and we can omit the "absolute value" symbols in Eq. (7), i.e.,

$$dS_W^2 = (\arccos\langle \psi_1|\psi_2\rangle)^2. \tag{8}$$

Interestingly enough, recourse to estimation theory ([11–16], and references therein) has allowed Braunstein and Caves [4] to show that the Wootters metric is identical with the so-called Cramer-Rao bound [12].

D. Statistical distributions in Hilbert space

The concept of statistical distance is quite independent of quantum mechanics and can be defined in any probability space [1]. In order to find an explicit expression for Eq. (1), Wootters [1] regards probability distributions as points, say, $\vec{p}_1$, $\vec{p}_2$, belonging to an N-dimensional probability space $[\vec{p}_1 = (p_1^1, p_1^2, \ldots, p_1^N);\ \vec{p}_2 = (p_2^1, p_2^2, \ldots, p_2^N)]$. Let us connect two of these points by a smooth curve $\vec{p}(t)$ parametrized by the variable t ($0 \leqslant t \leqslant 1$). Thus, $\vec{p}(0) = \vec{p}_1$ and $\vec{p}(1) = \vec{p}_2$. Performing a variational calculation in order to find the shortest curve between $\vec{p}_1$ and $\vec{p}_2$ and thereby determine the statistical distance, Wootters finds that, in the limit $N \to \infty$,

$$D_W = \arccos\left[\sum_{i=1}^{\infty} (p_1^i)^{1/2} (p_2^i)^{1/2}\right], \tag{9}$$

and, as anticipated above, reaches (and explains) a notable result, namely that [cf. Eq. (1)]

$$dS_W = D_W. \tag{10}$$

E. Kullback's distance

In the present considerations we intend to underline the importance of a different statistical distance that arises out of the celebrated maximum entropy principle [17]. We shall call this distance the Kullback one.

Kullback's minimum cross-entropy principle (MinEnt) [18] is an entropy optimization principle that rivals Jaynes' celebrated maximum entropy principle (MEP) [17] both in its range of applications and in its theoretical outreach. Indeed, MEP and MinEnt are related in the fashion to be made explicit below. We define now the Kullback distance. Let $\vec{p} = (p_1, p_2, \ldots, p_n)$ and $\vec{q} = (q_1, q_2, \ldots, q_n)$ be two normalized (to unity) probability distributions. The Kullback-Leibler measure is defined as [18,19]

$$D_{\mathrm{KL}}(\vec{p}:\vec{q}) = \sum_{i=1}^{n} p_i \ln\frac{p_i}{q_i}, \tag{11}$$

where we assume that whenever $q_i = 0$, the corresponding p_i is also zero. We define, as usual [17], $0\ln(0/0) = 0$. Some important properties of this measure are as follows.

(i) $D_{\mathrm{KL}}(\vec{p}:\vec{q})$ is a continuous function of $p_1, \ldots, p_n$ and of $q_1, \ldots, q_n$.

(ii) $D_{\mathrm{KL}}(\vec{p}:\vec{q})$ is permutationally symmetric [we can permute among themselves pairs (p_j, q_j)].

(iii) $D_{\mathrm{KL}}(\vec{p}:\vec{q}) \geqslant 0$ and vanishes if $\vec{p} = \vec{q}$. This property will play a central role in our considerations. It may be helpful to have the proof at hand, so that we give it in the Appendix.

(iv) The minimum value of D_{KL} is zero.

(v) D_{KL} is a convex function of both $\vec{p}$ and $\vec{q}$. This property is important in establishing the properties of global minimum.

(vi) $D_{\mathrm{KL}}(\vec{p}:\vec{q}) \neq D_{\mathrm{KL}}(\vec{q}:\vec{p})$.

The non-negativity $[D_{\mathrm{KL}}(\vec{p}:\vec{q}) \geqslant 0]$ and the identity $[D_{\mathrm{KL}}(\vec{p}:\vec{p}) = 0]$ are essential for any measure of *discrepancy*. On the other hand, the Kullback-Leibler distance does not satisfy the symmetry and triangle inequality, conditions that apply for metric distances. The last one is not essential for our present purposes because (a) we are interested in discrepancies or deviations from [cf. Eq. (3)] $\psi(\alpha)$ when one sightly varies the parameter(s) α; (b) we are interested in considering only two wave functions at a time, so that the triangle inequality is not required.

Nevertheless in order to preserve the symmetry property we can also use the (related) measure [18,19]

$$2\,D_K = D_{\mathrm{KL}}(\vec{p}:\vec{q}) + D_{\mathrm{KL}}(\vec{q}:\vec{p}) = \sum_{i=1}^{n} p_i \ln\frac{p_i}{q_i} + \sum_{i=1}^{n} q_i \ln\frac{q_i}{p_i}, \tag{12}$$

so that, obviously, $D_K(\vec{p}:\vec{q}) = D_K(\vec{q}:\vec{p})$.

Now, if we take $\vec{q}$ to be the *uniform* ($q_1 = q_2 = \cdots = q_n$) probability distribution $\vec{u}$, then we immediately find

$$D_{\mathrm{KL}}(\vec{p}:\vec{u}) = \ln n - S(p_1, p_2, \ldots, p_n), \tag{13}$$

where S stands for Shannon's measure [17]. Thus minimizing $D_{\mathrm{KL}}(\vec{p}:\vec{u})$ is tantamount to maximizing S.

Of course, for the present purposes we take

$$\vec{q}=|\psi(\alpha)|^2, \tag{14}$$

$$\vec{p}=|\psi(\alpha+\Delta\alpha)|^2 \tag{15}$$

(or vice versa) so that

$$2\,D_K=\int dx|\psi(\alpha)|^2\ln\left[\frac{|\psi(\alpha)|^2}{|\psi(\alpha+\Delta\alpha)|^2}\right]+\int dx|\psi(\alpha+\Delta\alpha)|^2\ln\left[\frac{|\psi(\alpha+\Delta\alpha)|^2}{|\psi(\alpha)|^2}\right], \tag{16}$$

and, in order to keep dimensions properly accounted for, we need to regard D_K as a metric in the sense (self-explanatory notation)

$$dS_K^2\equiv D_K. \tag{17}$$

III. COMPARING DISTANCES IN HILBERT SPACE

The idea is that we have a vector $\psi(\alpha)$ that depends upon a set of parameters collectively denoted by α. We slightly change α to $\alpha+\Delta\alpha$, so that our wave function is now $\psi(\alpha+\Delta\alpha)$. We wish to compare these two wave functions, according to the different criteria that arise out of the considerations of the preceding section, by recourse to evaluating the diverse distances between $\psi(\alpha+\Delta\alpha)$ and $\psi(\alpha)$. To this end, and following Eq. (5), we expand $\psi(\alpha+\Delta\alpha)$ up to second order and assume that both $\psi(\alpha)$ and $\psi(\alpha+\Delta\alpha)$ are properly normalized to unity. The goal is, of course, that of establishing universal relationships among the various distances between neighboring states (NS). We begin, however, with a quite general relationship between Euclidean and Wootters distances.

Before embarking into our discussion we first note the following.

(i) On account of normalization

$$\mathcal{J}=\left(\frac{\partial}{\partial\alpha}\right)\langle\psi(\alpha)|\psi(\alpha)\rangle=\left\langle\frac{\partial\psi}{\partial\alpha}\middle|\psi(\alpha)\right\rangle+\text{c.c.}=0 \tag{18}$$

implies

$$\left\langle\frac{\partial\psi}{\partial\alpha}\middle|\psi(\alpha)\right\rangle=0,$$

$$\left\langle\psi(\alpha)\middle|\frac{\partial\psi}{\partial\alpha}\right\rangle=0. \tag{19}$$

(ii) Up to second order in $\Delta\alpha$ [we set $\Delta^2\alpha=(\Delta\alpha)^2$],

$$\mathcal{F}=|\psi(\alpha+\Delta\alpha)-\psi(\alpha)|^2=\left|\Delta\alpha\frac{\partial\psi}{\partial\alpha}+(1/2)\Delta^2\alpha\frac{\partial^2\psi}{\partial\alpha^2}\right|^2\approx\Delta^2\alpha\left|\frac{\partial\psi}{\partial\alpha}\right|^2. \tag{20}$$

(iii) Up to this order, and using Eqs. (19) (the asterisk indicates, as usual, complex conjugation),

$$\mathcal{O}\equiv\langle\psi(\alpha)|\psi(\alpha+\Delta\alpha)\rangle\approx1+(1/2)\Delta^2\alpha\int dx\,\psi^*\frac{\partial^2\psi}{\partial\alpha^2}. \tag{21}$$

To make further progress we notice that

$$\int dx\,\psi^*\frac{\partial^2\psi}{\partial\alpha^2}=\frac{\partial}{\partial\alpha}\int dx\left(\psi^*\frac{\partial\psi}{\partial\alpha}\right)-\int dx\frac{\partial\psi^*}{\partial\alpha}\,\frac{\partial\psi}{\partial\alpha}, \tag{22}$$

which, since the first integral on the right-hand side vanishes [see Eqs. (19)], finally gives

$$1-\mathcal{O}=(1/2)\Delta^2\alpha\int dx\left|\frac{\partial\psi}{\partial\alpha}\right|^2. \tag{23}$$

The results (i)–(iii) are quite general. The next one is restricted to real wave functions, so that it becomes specially useful in the case of one-dimensional problems, where (for stationary states) one always can assume, without loss of generality, that wave functions are of such a nature [20,21].

(iv) Up to second order ($\psi'=\partial\psi/\partial\alpha$),

$$\begin{aligned}[\psi(\alpha+\Delta\alpha)]^2&=\left[\psi(\alpha)+\Delta\alpha\frac{\partial\psi}{\partial\alpha}+(\Delta^2\alpha/2)\frac{\partial^2\psi}{\partial\alpha^2}\right]^2\\&=\psi^2\left[1+\Delta\alpha\frac{\partial\ln\psi^2}{\partial\alpha}+\Delta^2\alpha[\psi''/\psi+(\psi'/\psi)^2]\right].\end{aligned} \tag{24}$$

A. Euclidean distance between neighboring states

By recourse to Eq. (20) we evaluate now, for the sake of later reference, the Euclidean distance

$$dS_E^2=\int dx[\psi(\alpha+\Delta\alpha)-\psi(\alpha)]^2 \tag{25}$$

up to second order in $\Delta\alpha$,

$$\begin{aligned}dS_E^2&=\int dx\mathcal{F}\\&=\Delta^2\alpha\int dx\left|\frac{\partial\psi}{\partial\alpha}\right|^2\\&\equiv\Delta^2\alpha I_1,\end{aligned} \tag{26}$$

so that

$$dS_E^2\equiv\Delta^2\alpha I_1. \tag{27}$$

B. Euclidean distance and Wootters distance

We show here that they are intimately related. Just consider two wave functions ψ_1,ψ_2 $[\psi_1=\psi(\alpha+\Delta\alpha),\psi_2=\psi(\alpha)]$ and compare Eqs. (6) and (8).

Since $\int dx|\psi_1|^2=\int dx|\psi_2|^2=1$, we have

$$dS_E^2=2(1-\langle\psi_1|\psi_2\rangle)=2(1-\cos\gamma)=2(1-\mathcal{O}), \quad (28)$$

where γ is, of course, a small angle. Expansion of the trigonometric function gives

$$dS_E^2=2\left(1-1+\frac{\gamma^2}{2}-\frac{\gamma^4}{24}+\cdots\right)=\gamma^2\left(1-\frac{\gamma^2}{12}\right)\simeq\gamma^2$$
$$=[\arccos(\langle\psi_1|\psi_2\rangle)]^2, \quad (29)$$

i.e., the Wootters distance reappears in the last equality. From Eq. (23) we get

$$dS_E^2=\Delta^2\alpha\int dx\left|\frac{\partial\psi}{\partial\alpha}\right|^2, \quad (30)$$

which coincides, of course, with our previous result. Thus,

$$dS_E^2=dS_W^2, \quad (31)$$

up to second order in α.

C. Fubini distance between NS

We start with the definition

$$dS_F^2=1-|\langle\psi(\alpha)|\psi(\alpha+\Delta\alpha)\rangle|^2, \quad (32)$$

and use Eq. (21),

$$dS_F^2=[1-\mathcal{O}^2], \quad (33)$$

which, after expansion up to second order, and on account of Eq. (23), gives

$$dS_F^2=\Delta^2\alpha\int dx\left|\frac{\partial\psi}{\partial\alpha}\right|^2, \quad (34)$$

so that the interesting relation ensues

$$dS_F^2=(\Delta^2\alpha)I_1=dS_E^2, \quad (35)$$

up to second order in $\Delta\alpha$.

Next let us suppose that the α evolution of $\psi(\alpha)$ is generated by a unitary operator $\hat{U}$, i.e., that for an arbitrary initial state $\psi(\alpha_0)$ we can write

$$\psi(\alpha)=\hat{U}(\alpha,\alpha_0)\psi(\alpha_0). \quad (36)$$

Correspondingly, the infinitesimal evolutions we are interested in here can be ascribed to the action of a set of m Hermitian operators [22],

$$-i\partial_k\psi(\alpha)=\hat{A}_k\psi(\alpha), \quad (37)$$

with

$$\hat{A}_k(\alpha)=i\hat{U}(\alpha,\alpha_0)\partial_k\hat{U}^{-1}(\alpha,\alpha_0) \quad (k=1\cdots m). \quad (38)$$

As shown in [2], the Fubini metric neatly captures the essentials of the uncertainties and correlations of the operators $\hat{A}_k$ generating the α evolution of $\psi(\alpha)$. This has been best demonstrated by using squeezed states [22].

Our result Eq. (35) tells us that, up to second order in $\Delta\alpha$*, the Euclidean metric is also able to account (in the sense of [2]) for the above mentioned uncertainties and correlations.*

D. MinEnt distances between NS

We address now out leitmotiv topic: i.e., concerning ourselves with statistical distances of a quite different origin: information theory (IT) [17], whose main tenet asserts that to any probability distribution one can associate, in unique fashion, an information measure [17]. $D_{\rm KL}(\vec{p}:\vec{q})$ is the measure of a *relative* information [23,24]. Assume that you have an *a priori* estimation $\vec{q}$ of how the pertinent probability distribution (pd) should look. You are provided now with some additional information concerning the system of interest and conclude that the pd $\vec{p}$ is the one that best reflects what you know *now* about the system. $D_{\rm KL}(\vec{p}:\vec{q})$ measures the amount of information associated to $\vec{p}$ relative to that contained in $\vec{q}$ [23]. Let us rephrase this in a slightly different form. Suppose that the maximum amount of information you can gather concerning your system is called $\mathcal{M}$. $D_{\rm KL}(\vec{p}:\vec{q})$ can also be said to represent the additional amount of information that is still required, in going from $\vec{q}$ to $\vec{p}$, to attain the ideal amount $\mathcal{M}$. Of course, if starting from $\vec{q}$ you were free to choose $\vec{p}$, you would select it so as to *minimize* $D_{\rm KL}(\vec{p}:\vec{q})$ [23,24].

Within the present context we have

$$\vec{q}(\alpha)=[\psi(\alpha)]^2,$$
$$\vec{p}(\alpha)\equiv\vec{q}(\alpha+\Delta\alpha)=[\psi(\alpha+\Delta\alpha)]^2, \quad (39)$$

so that we can write, if we restrict ourselves to the one-dimensional instance, and are thus allowed to use Eq. (24) (see the comment made before deriving that relationship),

$$\vec{p}(\alpha)=\vec{q}(\alpha)\left[1+\Delta\alpha\frac{\partial\ln\vec{q}(\alpha)}{\partial\alpha}+(\Delta^2\alpha/2)(\vec{q}''/\vec{q})+\cdots\right]. \quad (40)$$

$D_{\rm KL}(\vec{p}:\vec{q})$ measures how much information is gained in going from α to $\alpha+\Delta\alpha$ relative to that already contained in $\psi(\alpha)$, the so-called ''quantal entropy'' S_Q [25], that has been the subject of much recent work [25–32]. One writes

$$S_Q=-\int dx|\psi(\alpha)|^2\ln|\psi(\alpha)|^2 \quad (41)$$

(for details see, for instance, Refs. [25,29] and references therein).

It is clear that, in principle, $D_{\rm KL}$ is not related in an obvious way to the Wootters distance, which refers to the maximum number $\mathcal{N}$ of mutually distinguishable (in n trials) intermediate probability distributions $|\psi(\gamma)|^2$ $(\alpha\leqslant\gamma\leqslant\alpha+\Delta\alpha)$.

Notice that, because both $\psi(\alpha)$ and $\psi(\alpha+\Delta\alpha)$ are properly normalized wave functions, Eq. (40) implies (we drop the "$\vec{q}$" notation for the sake of a lighter notation)

$$\Delta\alpha\int dxq(\alpha)\frac{\partial\ \ln q(\alpha)}{\partial\alpha}=0 \tag{42}$$

and

$$(\Delta^2\alpha/2)\int dxq''(\alpha)=0, \tag{43}$$

so that we can recast our KL measure in the fashion

$$D_{\mathrm{KL}}(q:p)=\int dxq(\alpha)\ln\frac{q(\alpha)}{p(\alpha)}$$
$$=-\int dxq\left(\ln\left[1+\Delta\alpha\frac{\partial\ \ln q}{\partial\alpha}+(\Delta^2\alpha/2)(q''/q)\right]\right) \tag{44}$$

$$=\int dx\left[q\Delta\alpha\frac{\partial\ \ln q}{\partial\alpha}+(\Delta^2\alpha/2)q''(\alpha)\right]$$
$$-\int dxq\left(\ln\left[1+\Delta\alpha\frac{\partial\ \ln q}{\partial\alpha}+(\Delta^2\alpha/2)(q''/q)\right]\right) \tag{45}$$

$$=\int dxq\left\{\Delta\alpha\frac{\partial\ \ln q}{\partial\alpha}+(\Delta^2\alpha/2)(q''/q)\right.$$
$$\left.-\left(\ln\left[1+\Delta\alpha\frac{\partial\ \ln q}{\partial\alpha}+(\Delta^2\alpha/2)(q''/q)\right]\right)\right\}. \tag{46}$$

1. First-order relationships

If we expand now

$$\ln[1+y]\equiv\ln\left[1+\Delta\alpha\frac{\partial\ln q}{\partial\alpha}+(\Delta^2\alpha/2)(q''/q)\right], \tag{47}$$

up to first order in y we immediately obtain from Eq. (46)

$$D_{\mathrm{KL}}(q:p)=0, \tag{48}$$

which tells us that the KL measure is stable against first-order changes in $\Delta\alpha$. *This constitutes one of our main results*: up to first order in $\Delta\alpha$, no information is gained in going from α to $\alpha+\Delta\alpha$ (or vice versa).

2. A second-order relation

If we expand now up to second order in $\Delta\alpha$ in Eq. (47) we find from Eq. (46), taking care also of Eq. (43),

$$D_{\mathrm{KL}}(q:p)=(1/2)\Delta^2\alpha\int dxq\left(\frac{\partial\ln q}{\partial\alpha}\right)^2, \tag{49}$$

which, in terms of $\psi(\alpha)$ reads

$$D_{\mathrm{KL}}(q:p)=2\Delta^2\alpha\int dx\left[\frac{\partial\psi(\alpha)}{\partial\alpha}\right]^2, \tag{50}$$

i.e.,

$$dS^2_{\mathrm{KL}}=2dS^2_E, \tag{51}$$

and, for the symmetrized Kullback distance,

$$dS^2_K=4dS^2_E, \tag{52}$$

which tell us that the both the Kullback and the KL distance become proportional to the Euclidean distance (6) and thus to the Wootters one.

IV. APPLICATIONS

The concept of distance between rays can be used in order to discuss the quality of different approximate treatments. We apply now some of the previous considerations to some one-dimensional soluble problems.

A. Sextic anharmonic potential

The special sextic anharmonic oscillator has been extensively studied by Dutta and Wiley [33] and by Leach *et al.* [34]. It reads

$$V(x)=a_0+a_1x+a_2x^2+a_3x^3+a_4x^4+a_5x^5+a_6x^6, \tag{53}$$

and we tackle here the special case

$$a_0=a_1=a_3=a_5=0. \tag{54}$$

It is well known that the associated, exact ground state (g.s.) wave function (w.f.) can be written in the fashion

$$\Psi(x)=\exp\left(-\frac{1}{2}\sum_{i=0}^{4}\lambda_ix^i\right), \tag{55}$$

where the relations between the λ's and the a_n coefficients are given by (see Ref. [24])

$$a_6=2\lambda_4^2, \tag{56}$$

$$a_4=2\lambda_2\lambda_4, \tag{57}$$

$$\lambda_3=0, \tag{58}$$

$$\lambda_1=0, \tag{59}$$

$$a_2=\frac{1}{2}\lambda_2^2-3\lambda_4. \tag{60}$$

The energy of the g.s. is $E=\frac{1}{2}\lambda_2$. It is also well known that maximization of the quantal entropy S_Q (MEP), or, equivalently, minimization of the KL one, yields the exact g.s. wave function.

In order to test the accuracy of the MEP technique, and thus use the distance concept to numerically measure it, we add now a perturbation a_8x^8 to the potential (53) and compare the exact g.s. wave function for this *new* potential with

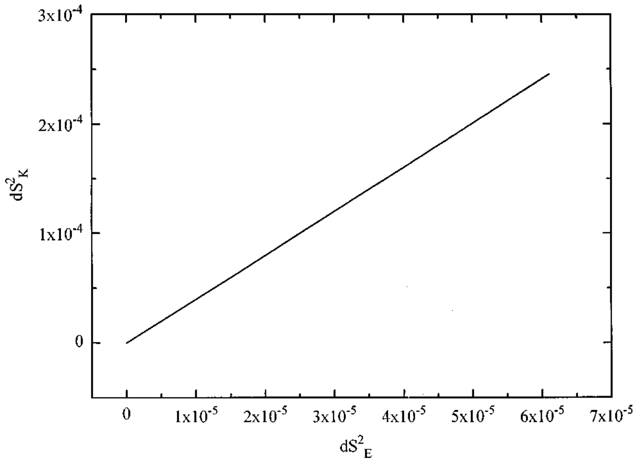

FIG. 1. Euclidean vs symmetrized Kullback distances between two ground state wave functions, namely, (i) that obtained with the approximate MEP technique (62) with $\lambda_2 = 1$ and $\lambda_4 = 0.5$, and (ii) the exact one, for the case of the potential (53) to which an additional perturbative term $a_8 x^8$ has been added. In the graph, the a_8 coefficient varies between 0 and 1.

the one obtained, following the approximate MEP technique described in [27], by recourse to the ansatz

$$\Psi_{\mathrm{MEP}}(x) = \exp\left[-\frac{1}{2}(\lambda_0 + \lambda_2 x^2 + \lambda_4 x^4)\right]. \tag{61}$$

We compute the different distances between the exact and the MEP wave functions and display in Fig. 1 the (symmetrized) Kullback vs the Euclidean one. Of course, a straight line is obtained (a_8 has been allowed to vary between 0 and 1). Numerically, we have

$$dS_K^2 = 4.01 dS_E^2 - 5.410^{-8}, \tag{62}$$

and the coefficient of correlation equals 1.0. In Fig. 2 the different distances are plotted against a_8. The solid line represents the Euclidean (or, equivalently, the Fubini or the Wootters) distance (all three coincide within the scale of the figure). The dashed curve gives the (symmetrized) Kullback

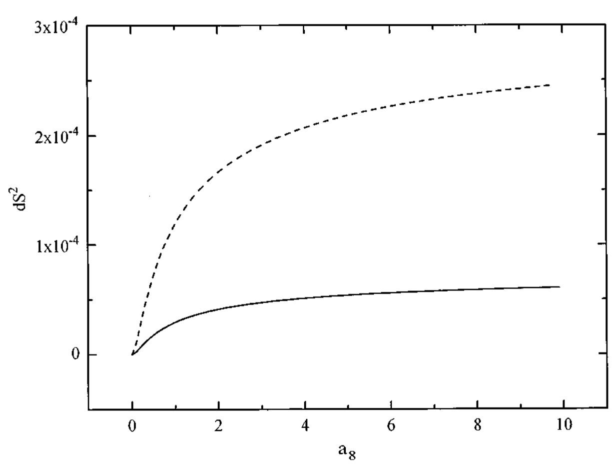

FIG. 2. Same as Fig. 1, but here distances are plotted against the coefficient a_8. The dashed curve corresponds to the symmetrized Kullback's distance, and the solid one to the Euclidian distance.

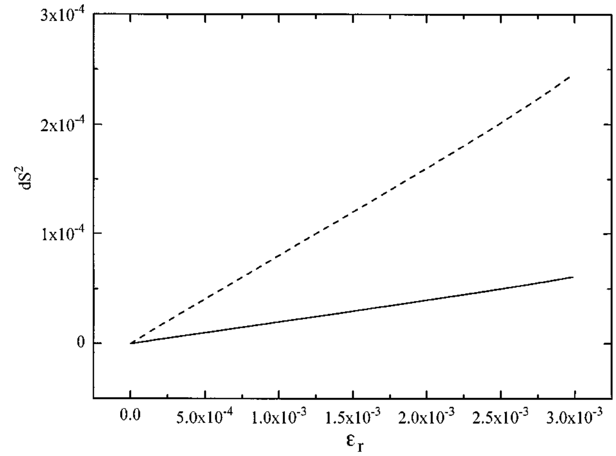

FIG. 3. Relative ground state energy error ϵ_r vs the symmetrized Kullback distance (dashed) and Euclidian one (solid). The error is that associated to the approximate MEP wave function referred to in Fig. 1.

distance. Figure 3 depicts the same distances as Fig. 2, but here against the relative energy error

$$\epsilon_r = \frac{E_{\mathrm{MEP}} - E_{\mathrm{exact}}}{E_{\mathrm{exact}}}, \tag{63}$$

which establishes a direct connection between the quality of an approximate quantal treatment and the concept of distance.

1. Harmonic oscillator

We use the distance concept here in order to compare two different approximate techniques: (i) the MEP approach and (ii) the perturbative one. To this end we add a perturbation $a_4 x^4$ to the harmonic oscillator potential.

We examine the MEP approximation by recourse to the ansatz (61), on the one hand, and employ results obtained by recourse to second order perturbation theory, on the other hand. The Euclidean (solid lines) and the symmetrized Kullback's distance (dotted-dashed curves) of Figs. 4 (MEP vs exact) and 5 (perturbative vs exact) that separate approximate from exact results neatly illustrates the fact that the MEP method is clearly superior to second order perturbation theory, even in the case of very small anharmonicities $a_4 x^4$.

V. CONCLUSIONS

In the present work we have established some universal relations between two distances on probability space (Refs. [1,4,18,19]) and different metrics on Hilbert space. The angle γ in Hilbert space [cf. Eq. (7)] is the only Riemannian metric (dS_W^2) on the set of rays, up to a constant factor, which is invariant under all possible time evolutions (more generally, under all unitary transformations), being thus, in a sense, the ''natural'' metric on the set of quantum states [1].

Wootters has shown [1] that the same metric arises from a quite different starting point: the analysis of statistical fluctuations in a finite sequence of measurements, with the result that distance between two states becomes tantamount to

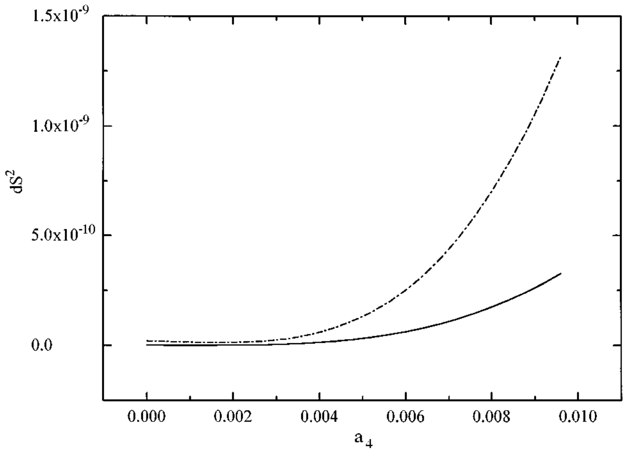

FIG. 4. Euclidean (solid) and symmetrized Kullback (dotted-dashed) distances between two ground state wave functions, namely, (i) that obtained with the approximate MEP technique and (ii) the exact one, vs a_4, for a quartic anharmonic oscillator. a_4 is the coefficient of the x^4 term in the associated potential $V(x)=\frac{1}{2}x^2+a_4x^4$. The a_4 parameter varies between 0 and 1. Due to the very small figures one finds near the origin that some ''imperfections'' in the drawings are unavoidable.

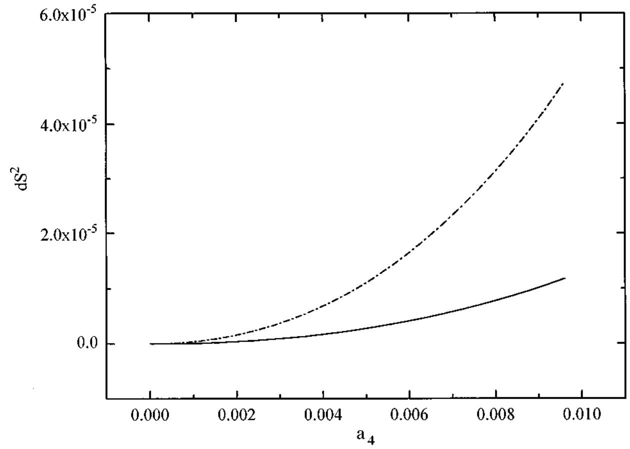

FIG. 5. Same as Fig. 4, but here the approximate wave function obtained by recourse to second order perturbation theory is compared to the exact one. Notice the difference in the vertical scales of Fig. 4 and this figure.

counting the number of distinguishable intermediate ones [1]. A connection between statistics and geometry is thus established [1] that has been considerably strengthened by the study of Ref. [4]. Although the intricacies of this connection are not at this point *totally* understood, the present work allows one to obtain some additional insights into the matter that, in a sense, allow one to advance a few steps into the road inaugurated by Wootters and considerably widened by Braunstein and Caves.

Thus, with reference to neighboring states, we have established that, up to second order in a suitable parameter α,[1] the following identities obtain:

$$dS_W^2=dS_E^2 \quad \text{(Euclidean metric)},$$

$$dS_W^2=dS_F^2 \quad \text{(Fubini -Study metric)},$$

$$dS_W^2=dS_{\mathrm{KL}}^2/2 \quad \text{(Kullback -Leibler metric)}. \tag{64}$$

The five metrics here discussed[2] become essentially identical up to second order in $\Delta\alpha$. Two of them were already analyzed in the pioneer work of Wootters [1] and carefully discussed in Ref. [4]. The three ones on the right-hand sides of (64) are of two types: the first two have a geometrical origin, while the last one is of an statistical one. *The connection between geometry and statistics referred to above become now entrenched in a firmer manner*.

Wootters has suggested [1] that statistical fluctuations in the outcomes of measurements might be partly responsibly for the Hilbert-space structure of quantum mechanics. These statistical fluctuations are intertwined with *the uncertainties and correlations* of the operators associated to the measurements, *whose essentials are neatly captured by the Fubini-Study metric* [2,4]. Thus, our second equality above reinforces the plausibility of Wootters' suggestion.

The KL metric is associated to an optimization principle, Shannon's maximum entropy one. In a sense, it reflects efficient management of the available information (always related to the expectation values at our disposal in building up the concomitant wave function [25]). The last equality above connects this management to the number of intermediate distinguishable states between $\psi(\alpha)$ and $\psi(\alpha+\Delta\alpha)$.

The application of our results to NS states in some one-dimensional soluble problems numerically illustrates the fact that the Euclidean, the Wootters, and the Fubini distances are equivalent. The Kullback's distance is, as explained, proportional to the Euclidean one [Eq. (6)]. The different relations provide one with a quantitative measure of the quality of different approximate treatments of Schrödinger's equation.

ACKNOWLEDGMENTS

This work has been performed under the auspices of the General Agreement between the Universidad de les Illes Balears and the Universidad Nacional de La Plata, and partially supported by the DGICyT (Spain), Grant No. PB95-0492-A, and PROTEM programme of CONICET (Argentina).

APPENDIX: POSITIVITY OF THE KULLBACK-LEIBLER MEASURE

Within the present context the KL measure is given by

$$D_{\mathrm{KL}}(p:q)=\int dx p(x,\mu)\ln\frac{p(x,\mu)}{q(x,\mu)}, \tag{A1}$$

with

$$q(x,\mu)=|\psi_\mu(x)|^2 \tag{A2}$$

and

[1] This parameter allows one to distinguish a given state from a neighboring one (from α to $\alpha+\Delta\alpha$).

[2] That is, Euclidean, Fubini-Study, Wootters, Kullback-Leibler, and symmetrized KL.

$$p(x,\mu)=|\psi_{\mu+\Delta\mu}(x)|^2, \tag{A3}$$

μ being, of course, a real quantity that parametrizes the wave function ψ. Now, we deal with situations such that

$$\begin{aligned} p(x,\mu)&=q(x,\mu+\Delta\mu)\\ &=q(x,\mu)\left(1+\Delta\mu\frac{\partial\ \ln q(x,\mu)}{\partial\mu}+\cdots\right), \end{aligned} \tag{A4}$$

with

$$\left(1+\Delta\mu\frac{\partial\ \ln q(x,\mu)}{\partial\mu}\right)\geqslant 0. \tag{A5}$$

On account of normalization we have

$$\int dx p(x,\mu)=\int dx q(x,\mu)=1, \tag{A6}$$

which implies

$$\int dx q(x,\mu)\Delta\mu\frac{\partial\ \ln q(x,\mu)}{\partial\mu}=0, \tag{A7}$$

so that we can recast Eq. (A1) in the fashion

$$\begin{aligned} D_{\mathrm{KL}}(q:p)&=\int dx q(x,\mu)\ln\frac{q(x,\mu)}{p(x,\mu)}\\ &=-\int dx q\left(\ln\left[1+\Delta\mu\frac{\partial\ \ln q}{\partial\mu}\right]\right)\\ &=\int dx q\Delta\mu\frac{\partial\ \ln q}{\partial\mu}-\int dx q\left(\ln\left[1+\Delta\mu\frac{\partial\ \ln q}{\partial\mu}\right]\right)\\ &=\int dx q\left\{\Delta\mu\frac{\partial\ \ln q}{\partial\mu}-\left(\ln\left[1+\Delta\mu\frac{\partial\ \ln q}{\partial\mu}\right]\right)\right\}\geqslant 0. \end{aligned} \tag{A8}$$

The last inequality is a result of the fact that

$$f(y)=y-\ln(1+y), \tag{A9}$$

for $1+y>0$, is of a non-negative character. Indeed, one has

$$\begin{aligned} f(0)&=0,\\ f'(y)&\equiv\frac{df}{dy}=\frac{y}{1+y},\\ f'(0)&=0,\\ f''(y)&=(1+y)^{-2}. \end{aligned} \tag{A10}$$

Thus, the KL measure is positive $[f'(y)<0,\ y<0$, and $f'(y)>0,\ y>0]$, as it should be.

[1] W. K. Wootters, Phys. Rev. D **23**, 357 (1981).
[2] S. Abe, Phys. Rev. A **48**, 4102 (1993).
[3] R. A. Fisher, Proc. R. Soc. Edinburgh **42**, 321 (1922).
[4] S. M. Braunstein and C. M. Caves, Phys. Rev. Lett. **72**, 3439 (1994).
[5] D. N. Page, Phys. Rev. A **36**, 3479 (1987).
[6] A. Gibbons, J. Geom. Phys. **8**, 147 (1992).
[7] J. Provost and G. Vallée, Commun. Math. Phys. **76**, 289 (1980).
[8] E. R. Caianiello and W. Guz, Phys. Lett. A **126**, 223 (1988).
[9] V. Bargmann, Ann. Math. **59**, 1 (1954).
[10] A. K. Pati, Phys. Lett. A **159**, 105 (1991).
[11] L. L. Campbell, Inform. Sci. **35**, 147 (1985).
[12] H. Cramer, *Mathematical Methods of Statistics* (Princeton University Press, Princeton, 1946).
[13] B. R. Frieden and B. H. Soffer, Phys. Rev. E **52**, 2274 (1995).
[14] B. R. Frieden, Phys. Rev. A **41**, 4265 (1990).
[15] B. R. Frieden, Am. J. Phys. **57**, 1004 (1989).
[16] B. Nikolov and B. R. Frieden, Phys. Rev. E **49**, 4815 (1994).
[17] A. Katz, *Principles of Statistical Mechanics* (Freeman, San Francisco, 1967).
[18] S. Kullback, *Information Theory and Statistics* (J. Wiley, New York, 1959).
[19] S. Kullback and R. A. Leibler, Ann. Math. Stat. **22**, 79 (1951).
[20] A. Messiah, *Quantum Mechanics* (North-Holland, Amsterdam, 1961), Vol. I.
[21] D. R. Bates, *Quantum Theory* (Academic Press, New York, 1961), Vol. I.
[22] S. Abe, Phys. Rev. A **46**, 1667 (1992).
[23] E. T. Jaynes, in *Maximum Entropy and Bayesian Methods in Science and Engineering*, Vol. 1: Foundations, edited by G. J. Erikson and C. Ray Smith (Kluwer, Dordrecht, 1988), pp. 25–30; M. Tribus, *ibid.*, pp. 31–52; S. F. Gull, *ibid.*, pp. 53–74.
[24] M. Casas, F. Garcias, A. Plastino, and Ll. Serra, Physica A **213**, 376 (1995).
[25] A. R. Plastino and A. Plastino, Phys. Lett. A **181**, 446 (1993).
[26] M. Casas, A. Plastino, and A. Puente, Phys. Rev. A **49**, 2312 (1994).
[27] M. Casas, A. Plastino, A. Puente, N. Canosa, and R. Rossignoli, Phys. Rev. A **47**, 3530 (1993).
[28] M. Casas, A. Plastino, and A. Puente, Phys. Rev. A **52**, 949 (1995).
[29] A. R. Plastino, M. Casas, A. Plastino, and A. Puente, Phys. Rev. A **52**, 2601 (1995).
[30] N. Canosa, A. Plastino, and R. Rossignoli, Phys. Rev. A **40**, 519 (1989).
[31] N. Canosa, A. Plastino, and R. Rossignoli, Nucl. Phys. A **512**, 520 (1990).
[32] N. Canosa, A. Plastino, and R. Rossignoli, Nucl. Phys. A **550**, 453 (1992).
[33] A. K. Dutta and R. S. Wiley, J. Math. Phys. **29**, 892 (1988).
[34] P. G. L. Leach, G. P. Flessas, and V. M. Gorringe, J. Math. Phys. **30**, 406 (1989).