Letters in Mathematical Physics **27**: 205–216, 1993.

The Bogoliubov Inner Product in Quantum Statistics*

Dedicated to J. Merza on his 60th birthday

DÉNES PETZ
Department of Mathematics, Faculty of Chemical Engineering, Technical University Budapest, Sztoczek u. 2. H ép. II. 25, H-1521 Budapest XI, Hungary

and

GABOR TOTH
Department of Mathematical Sciences, Rutgers University, Campus at Camden, Camden, NJ 0812, U.S.A.

(Received: 28 September 1992; revised: 25 January 1993)

Abstract. A natural Riemannian geometry is defined on the state space of a finite quantum system by means of the Bogoliubov scalar product which is infinitesimally induced by the (nonsymmetric) relative entropy functional. The basic geometrical quantities, including sectional curvatures, are computed for a two-level quantum system. It is found that the real density matrices form a totally geodesic submanifold and the von Neumann entropy is a monotone function of the scalar curvature. Furthermore, we establish information inequalities extending the Cramér–Rao inequality of classical statistics. These are based on a very general new form of the logarithmic derivative.

Mathematics Subject Classification (1991). 82B10.

0. Introduction

Let us consider a density matrix $D \equiv e^H$ of a finite quantum system $\mathscr{A}$. Then the Bogoliubov inner product is defined for the observables, $A, B \in \mathscr{A}^{sa}$ by the formula

$$\langle\!\langle A, B \rangle\!\rangle_D = \int_0^1 \mathrm{Tr}\, e^{uH} A\, e^{(1-u)H} B \, du. \tag{1}$$

This quantity appears in the second-order term of a perturbation expansion of the partition function,

$$\langle\!\langle A, B \rangle\!\rangle_D = \frac{\partial^2}{\partial t\, \partial s} \mathrm{Tr}\, e^{H + tA + sB}. \tag{2}$$

In quantum statistical mechanics, the definition of the Bogoliubov inner product involves the inverse temperature β which is chosen to be 1 in this Letter. The term

* This work was supported by the Hungarian National Foundation for Scientific Research, grant No. 1900. Authors' e-mail addresses are: H1128PET@ella.hu and TOTH@zodiac.rutgers.edu.

Kubo–Mori scalar product is also used [2, 13]. This scalar product is important in linear response theory [15] and it was generalized to KMS states of arbitrary quantum systems. In fact, the KMS property may be characterized by auto-correlation upper bounds which seem to be informational inequalities (see [2, 4, 5]).

In this Letter, we restrict our treatment to finite quantum systems. We argue that the Bogoliubov scalar product defines a natural Riemannian geometry on the state space. (This was previously proposed in [11].) The basic geomerical quantities, including sectional curvatures, are computed for a two-level quantum system. It is found that the real density matrices form a totally geodesic submanifold and the von Neumann entropy is a monotone function of the scalar curvature. (Differential geometrical methods in classical statistics have a huge literature, we mention only the books [1] and [3] but in the quantum case there are many things to be done.) Furthermore, we establish information inequalities extending the Cramér–Rao inequality of classical statistics. These are based on a new form of logarithmic derivative and show that the quantity

$$\sqrt{\langle\!\langle A - I \operatorname{Tr} DA, A - I \operatorname{Tr} DA \rangle\!\rangle_D} \tag{3}$$

is more appropriate for being the dispersion of the observable A in the state D than the popular variance

$$\sqrt{\operatorname{Tr} D(A - I \operatorname{Tr} A)^2}. \tag{4}$$

Concerning information inequalities in the quantum setting, we refer to the book [10].

1. Parameter Estimation

In the parametric problem of quantum statistics, a family (φ_θ) of states of a system is given and one has to decide among several alternative values of the parameter by using measurements. The set of outcomes of the applied measurements is the parameter set Θ and we assume that it is a region in $\mathbb{R}^n$. So an estimator measurement M is a positive-operator-valued measure on the Borel sets of Θ and its values are observables of the given quantum system [7, 9, 14]. The probability measure

$$B \mapsto \mu_\theta(B) = \varphi_\theta(M(B))(B \subset \Theta)$$

represents the result of the measurement M when the 'true' state is φ_θ. The choice of the estimators has to be made by taking into account the expected errors.

First we concentrate on the one-dimensional case and follow the Cramér–Rao pattern of mathematical statistics. Let (φ_θ) be a family of states of the quantum system $\mathscr{M}$ and let the parameter set Θ be an interval in $\mathbb{R}$ with $0 \in \Theta$. For the sake of simplicity, let the estimator be an observable M of the finite quantum system $M_n(\mathbb{C})$ and let the family (φ_θ) be given by density matrices $D_\theta \in M_n(\mathbb{C})$. We assume that $\theta \mapsto D_\theta$ is sufficiently differentiable. The estimator M is called unbiased if

$$\operatorname{Tr} D_\theta M = \theta. \tag{1.1}$$

Since we are interested in the local behaviour of the estimation, the less restrictive condition

$$\left.\frac{\partial \operatorname{Tr} D_\theta M}{\partial \theta}\right|_{\theta=0} = 1 \tag{1.2}$$

will also work. If M satisfies (1.2), then it is called locally unbiased at 0 (being unbiased means locally unbiased at each point).

Our approach, will be closely related to the so-called Bogoliubov inner product. (Note that Kubo–Mori scalar product and Duhamel two-point function are also frequently used terms.) For $A, B, H \in M_n(\mathbb{C})^{\mathrm{sa}}$, we set

$$\langle\!\langle A, B \rangle\!\rangle_{\exp H} = \int_0^1 \operatorname{Tr} \mathrm{e}^{uH} A\, \mathrm{e}^{(1-u)H} B \,\mathrm{d}u. \tag{1.3}$$

It is known that

$$\operatorname{Tr} XYX^* f(Y) \leqslant \operatorname{Tr} XX^* Y f(Y)$$

holds provided that $Y = Y^*$ and $f\colon \mathbb{R} \to \mathbb{R}$ is an increasing function [12]. In particular,

$$\operatorname{Tr} \mathrm{e}^{uH} A\, \mathrm{e}^{(1-u)H} A \leqslant \operatorname{Tr} \mathrm{e}^{H} A^2 \tag{1.4}$$

and we obtain

$$\langle\!\langle A, A \rangle\!\rangle_{\exp H} \leqslant \operatorname{Tr} \mathrm{e}^{H} A^2 \quad \text{for every } A \in M_n(\mathbb{C})^{\mathrm{sa}}. \tag{1.5}$$

(This inequality is a very special case of the Bogoliubov inequality but we prefer the above elementary proof, see also [2] and [5].) To make clear the relevance of the Bogoliubov inner product to the parameter estimation problem, we choose two differentiable curves of density matrices

$$D_t^1 = \exp(H + tA + O(t^2)) \quad \text{and} \quad D_t^2 = \exp(H + tB + O(t^2)).$$

(Note that the condition $\operatorname{Tr} D_t^1 = 1$ implies $\operatorname{Tr} \mathrm{e}^H A = 0$ and, similarly, $\operatorname{Tr} \mathrm{e}^H B = 0$.) Recall that the relative entropy

$$S(D_1, D_2) \equiv \operatorname{Tr} D_1(\log D_1 - \log D_2)$$

of the densities D_1 and D_2 measures the information between the corresponding states (see [16, 18, 19]). Using the perturbation series for the exponential function we arrive at

$$\left.\frac{\partial^2}{\partial t\, \partial s}\right|_{t=s=0} S(D_t^1, D_s^2) = -\langle\!\langle A, B \rangle\!\rangle_{\exp H}. \tag{1.6}$$

Hence, the Bogoliubov norm in quantum statistics can take the role of the Fischer information from classical statistics.

The Cramér–Rao pattern consists of finding a lower bound for $\operatorname{Tr} D_\theta(M-\theta)^2$, which is the variance of the estimator. To state our version of the Cramér–Rao inequality, we introduce the concept of a logarithmic derivative (see also Section 3). If $\theta \mapsto \varphi_\theta$ is a differentiable family of states, then its logarithmic derivative (with respect to the Bogoliubov inner product and taken at 0) is an operator $L \in \mathcal{M}$ which is characterized by the condition

$$\frac{\partial}{\partial\theta}\bigg|_{\theta=0} \varphi_\theta(A) = \langle\!\langle L^*, A\rangle\!\rangle_{\varphi_0} \quad \text{for every } A \in \mathcal{M}. \tag{1.7}$$

PROPOSITION 1.1. *Let the observable $M \in M_n(\mathbb{C})$ be a locally unbiased estimator for the differentiable family (D_θ) of the density matrices at $\theta = 0$. Then*

$$\varphi_0(M^2) \geqslant \langle\!\langle L^*, L^*\rangle\!\rangle_{\varphi_0}^{-1},$$

where L is the logarithmic derivative of the states $\varphi_\theta(\,\cdot\,) \equiv \mathrm{Tr}(\,\cdot\, D_\theta)$.

Proof. As a direct consequence of (1.2) and (1.7) we have $1 = \langle\!\langle L^*, M\rangle\!\rangle_{\varphi_0}$. Now applying first the Schwarz inequality and then (1.5), we infer

$$1 = \langle\!\langle L^*, M\rangle\!\rangle_{\varphi_0}^2 \leqslant \langle\!\langle L^*, L^*\rangle\!\rangle_{\varphi_0} \times \langle\!\langle M, M\rangle\!\rangle_{\varphi_0} \leqslant \langle\!\langle L^*, L^*\rangle\!\rangle_{\varphi_0} \times \varphi_0(M^2),$$

which is equivalent to the stated inequality. □

This proposition and formula (1.6) suggest that Bogoliubov inner product may be used to define a statistically relevant Riemannian metric on the state space of a quantum system, as was proposed in [11]. The definition and the detailed study of the geometry obtained for a 2-level (or spin $\frac{1}{2}$) quantum system is the content of the next section.

2. Riemannian Geometry of the State Space

The observables of a 2-level quantum system are the 2×2 complex self-adjoint matrices; their space is denoted by $M_2(\mathbb{C})^{\mathrm{sa}}$. The real linear space $M_2(\mathbb{C})^{\mathrm{sa}}$ is conveniently parametrized by the Pauli matrices:

$$\sigma_0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix},$$
$$\sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{2.1}$$

Any element of $M_2(\mathbb{C})^{\mathrm{sa}}$ is of the form

$$\begin{pmatrix} t + x_3 & x_1 - x_2 i \\ x_1 + x_2 i & t - x_3 \end{pmatrix} = ts_0 + x_1\sigma_1 + x_2\sigma_2 + x_3\sigma_3 \equiv (t, x) \cdot \sigma. \tag{2.2}$$

In fact, $\{\sigma_0, \sigma_1, \sigma_2, \sigma_3\}$ is an orthonormal basis for the Hilbert–Schmidt inner product

$$(A, B) \mapsto \tfrac{1}{2}\mathrm{Tr}\, AB \quad (A, B \in M_2(\mathbb{C})^{\mathrm{sa}}).$$

The inverse of the mapping

$$(t, x_1, x_2, x_3) \mapsto (t, x_1, x_2, x_3) \cdot \sigma \tag{2.3}$$

gives the affine (or Stokes) coordinates on $M_2(\mathbb{C})^{sa}$. The affine coordinate system establishes an isometry between the Euclidean 4-space $\mathbb{R}^4$ and $M_2(\mathbb{C})^{sa}$ endowed with the Hilbert–Schmidt scalar product. It is useful to write elements of $\mathbb{R}^4$ as (t, x) with $t \in \mathbb{R}$ and $x \in \mathbb{R}^3$ because of the following formulas:

$$\begin{aligned} &\mathrm{Tr}(t, x) \cdot \sigma = t, \\ &\mathrm{Det}(t, x) \cdot \sigma = t^2 - |x|^2, \quad \text{where } |x|^2 = x_1^2 + x_2^2 + x_3^2. \end{aligned} \tag{2.4}$$

Hence

$$M_2(\mathbb{C})^+ = \{(t, x) \cdot \sigma \in M_2(\mathbb{C})^{sa} : |x|^2 < t^2\} \tag{2.5}$$

is the set of all positive definite matrices of $M_2(\mathbb{C})^{sa}$. Geometrically, $M_2(\mathbb{C})^+$ is a circular cone of opening half-angle $\pi/4$. (Actually, this is nothing but the positive light cone in $\mathbb{R}^4$ considered as the Minkowski 4-space.) The slice

$$S = \{D \in M_2(\mathbb{C})^+ : \mathrm{Tr}\, D = 1, D \text{ is invertible}\}$$

is a ball with center $\sigma_0/2$ and radius $1/2$.

In the customary formalism of quantum mechanics, the states of the 2-level system correspond to density matrices and the closure of S is identified with the state space. The extremal (surface) points are the pure states and S is the set of mixed states.

The exponential map is a diffeomorphism of $M_2(\mathbb{C})^{sa}$ to $M_2(\mathbb{C})^+$, so that any $p \in M_2(\mathbb{C})^+$ may be represented as e^H, $H \in M_2(\mathbb{C})^{sa}$. Explicitly,

$$\exp((t, x) \cdot \sigma) = \left(e^t \cosh |x|, \frac{e^t \sinh |x|}{|x|} x \right) \cdot \sigma. \tag{2.6}$$

The inverse of the mapping

$$(t, x) \mapsto \exp((t, x) \cdot \sigma)$$

gives the logarithmic coordinate system on $M_2(\mathbb{C})^+$. At each $p = e^H \in M_2(\mathbb{C})^+$, the tangent space $\mathbf{T}_p(M_2(\mathbb{C})^+)$ is identified with $\mathbb{R}^4$ by translating the tangent vectors to the origin. Hence, $\sigma_0, \sigma_1, \sigma_2, \sigma_3$ can be thought of as globally defined vector fields on $M_2(\mathbb{C})^+$. Another possibility is to regard $\mathbf{T}_p(M_2(\mathbb{C})^+)$ as $M_2(\mathbb{C})^{sa}$. For example,

$$s \mapsto \exp(H + (st, sx) \cdot \sigma) \tag{2.7}$$

is an integral curve of the vector field to $t\sigma_0 + x_1\sigma_1 + x_2\sigma_2 + x_3\sigma_3$. However, this curve does not remain in S, even if $\mathrm{Tr}\, e^H = 1$. To get around this, we introduce three vector fields Σ^i, $1 \leqslant i \leqslant 3$ on the state space S by

$$(\Sigma^i)_\rho \equiv \sigma_i - (\mathrm{Tr}\, \sigma_i \rho)\sigma_0 \quad (1 \leqslant i \leqslant 3). \tag{2.8}$$

Now

$$s \mapsto \frac{\exp(H + s\sigma_i)}{\mathrm{Tr}(\exp(H + s\sigma_i))} \tag{2.9}$$

is an integral curve of Σ^i $(1 \leqslant i \leqslant 3)$ and belongs to S for all value of the real parameter s. Hence, Σ^i, $i = 1, 2, 3$ give a parallelization of S and, at each point $\rho \in S$, the tangent space $\mathbf{T}_\rho(S)$ can be identified with

$$\{A \in M_2(\mathbb{C})^{sa} : \operatorname{Tr} \rho A = 0\}.$$

The inverse of the mapping

$$(y_1, y_2, y_3) \mapsto \exp((t, y) \cdot \sigma) \quad \text{with } t = -\log(\exp(|y|) + \exp(-|y|)), \tag{2.10}$$

will be called the *logarithmic coordinate system on S*. In these coordinates, the vector field Σ^i is nothing but the differentiation with respect to y_i $(1 \leqslant i \leqslant 3)$.

As suggested by (1.6), the Bogoliubov inner product gives rise to a Riemannian metric on S. Explicitly, for $A, B \in \mathbf{T}_\rho(S)$ we define the scalar product of the tangent vectors A and B as $\langle\!\langle A, B \rangle\!\rangle_{\exp H}$ given in (1.3). (Note that this definition may be extended to the whole $M_2(\mathbb{C})^+$ but we are mainly interested in the geometry of the state space.)

The Riemannian metric is invariant under rotations. We let SO(3) act on $M_2(\mathbb{C})^{sa}$ by rotation around the σ_0-axis:

$$U((t, x) \cdot \sigma) = (t, Ux) \cdot \sigma, \tag{2.11}$$

where Ux is the ordinary action of SO(3) on $\mathbb{R}^3$. Note that this is nothing but the action of SU(2) on $M_2(\mathbb{C})^{sa}$ by conjugation after we factored out the kernel of the canonical epimorphism SU(2) $\to$ SO(3). Hence, we have the following properties.

(i) $\operatorname{Tr} U(A)U(B) = \operatorname{Tr} AB$ for $A, B \in M_2(\mathbb{C})^{sa}$, in particular $\operatorname{Tr} U(A) = \operatorname{Tr} A$.
(ii) $U(\exp H) = \exp U$ for $H \in M_2(\mathbb{C})^{sa}$.

LEMMA 2.1. *For $A, B \in M_2(\mathbb{C})^{sa}$ the invariance property*

$$\langle\!\langle U(A), U(B) \rangle\!\rangle_{\exp U(H)} = \langle\!\langle A, B \rangle\!\rangle_{\exp H} \tag{2.12}$$

holds for any $U \in$ SO(3).

The proof follows from the invariance properties (i) and (ii) through (2).

Let $A, B \in \mathbf{T}_\rho(S)$ be given by Stokes coordinates

$$A = (t_A, a) \cdot \sigma \quad \text{and} \quad B = (t_B, b) \cdot \sigma$$

and let $\rho \in S$ be given by logarithmic coordinates as

$$\rho = \exp((t, y) \cdot \sigma).$$

Then we have the following convenient formula for the Riemannian metric

$$\langle\!\langle A, B \rangle\!\rangle_\rho = \frac{\tanh r}{r} \langle a, b \rangle + \frac{r - \tanh r - r \tanh^2 r}{r} \cdot \frac{\langle a, y \rangle \langle b, y \rangle}{r^2}, \tag{2.13}$$

where $r = (y_1^2 + y_2^2 + y_3^2)^{1/2}$ is the length of y. To prove this, we first observe that the right-hand side of (2.13) is invariant under the action of SO(3), while invariance of the left-hand side is guaranteed by Lemma 2.1. Hence, we may assume that ρ is diagonal, that is, $y_1 = y_2 = 0$ and $r = |y_3|$. On the other hand, as easy computation shows, at a diagonal point ρ with $r > 0$, the Riemannian metric in logarithmic coordinates takes the form

$$\begin{pmatrix} r^{-1}\tanh r & 0 & 0 \\ 0 & r^{-1}\tanh r & 0 \\ 0 & 0 & \cosh^{-2} r \end{pmatrix}. \tag{2.14}$$

Formula (2.13) now follows for $r > 0$ by noting that the logarithmic coordinate fields are Σ^i, $1 \leqslant i \leqslant 3$. At the center $r = 0$, we take the limit $r \to 0$ and observe that the Riemannian metric becomes Euclidean, i.e.

$$\langle\!\langle A, B \rangle\!\rangle_I = \langle a, b \rangle. \tag{2.15}$$

Introducing the projection operator $\mathrm{Proj}_{[y]}$ that projects the whole space orthogonally to the line spanned by y, i.e. $\mathrm{Proj}_{[y]}(b) = \langle b, y \rangle r^{-2} y$, the matrix of the Riemannian metric (2.13) can be conveniently written

$$\mathbf{G} = \frac{\tanh r}{r}\,\mathrm{Proj}_{[y]}^{\perp} + \frac{1}{\cosh^2 r}\,\mathrm{Proj}_{[y]}. \tag{2.16}$$

Now we are going to compute some basic geometrical quantities for the geometry of S. Our basic reference on Riemannian geometry is [8]. We start with the covariant derivation ∇ induced by the metric and use the Kostant formula

$$X \cdot \langle\!\langle Y, Z \rangle\!\rangle + Z \cdot \langle\!\langle X, Y \rangle\!\rangle - Y \cdot \langle\!\langle X, Z \rangle\!\rangle$$
$$= \langle\!\langle [X, Y], Z \rangle\!\rangle + \langle\!\langle [Z, Y], X \rangle\!\rangle + \langle\!\langle [X, Z], Y \rangle\!\rangle + 2\langle\!\langle Y, \nabla_Z X \rangle\!\rangle.$$

Setting

$$X = \Sigma^j, \qquad Y = \Sigma^k, \quad \text{and} \quad Z = \Sigma^i$$

we have

$$2\langle\!\langle \nabla_{\Sigma^i}\Sigma^j, \Sigma^k \rangle\!\rangle = \Sigma^j \cdot \langle\!\langle \Sigma^i, \Sigma^k \rangle\!\rangle + \Sigma^i \cdot \langle\!\langle \Sigma^j, \Sigma^k \rangle\!\rangle - \Sigma^k \langle\!\langle \Sigma^i, \Sigma^j \rangle\!\rangle, \tag{2.17}$$

since the Lie brackets vanish. Let us recall briefly that the right-hand side terms have the following meaning. $\langle\!\langle \Sigma^j, \Sigma^k \rangle\!\rangle$ is a function on S, that is, a function of the coordinates (y_1, y_2, y_3) and the vector field Σ^i acts on this function as derivation with respect to y_i. Using this, we compute

$$\Sigma^k \langle\!\langle \Sigma^i, \Sigma^j \rangle\!\rangle$$
$$= \left(\frac{1}{\cosh^2 r} - \frac{\tanh r}{r}\right) \times$$
$$\times \left(\frac{\delta_{ij} y_k}{r^2} + \frac{\delta_{ki} y_j}{r^2} + \frac{\delta_{jk} y_i}{r^2} - \frac{y_k y_i y_j}{r^4}\right) - \frac{2\tanh r}{r^3 \cosh^2 r} y_i y_j y_k \tag{2.18}$$

and we observe that this expression is symmetric in all indices. Therefore,

$$2\langle\!\langle \nabla_{\Sigma^i}\Sigma^j, \Sigma^k\rangle\!\rangle = \Sigma^i \cdot \langle\!\langle \Sigma^j, \Sigma^k\rangle\!\rangle$$

from (2.17). We now expand

$$\nabla_{\Sigma^i}\Sigma^j = \Gamma^1_{ij}\Sigma^1 + \Gamma^2_{ij}\Sigma^2 + \Gamma^3_{ij}\Sigma^3,$$

with coefficients Γ^k_{ij} being the Christoffel symbols relative to the parallelization Σ^i, $1 \leqslant i \leqslant 3$. Equivalently,

$$\begin{pmatrix} \Gamma^1_{ij} \\ \Gamma^2_{ij} \\ \Gamma^3_{ij} \end{pmatrix} = \mathbf{G}^{-1} \begin{pmatrix} \langle\!\langle \nabla_{\Sigma^i}\Sigma^j, \Sigma^1\rangle\!\rangle \\ \langle\!\langle \nabla_{\Sigma^i}\Sigma^j, \Sigma^2\rangle\!\rangle \\ \langle\!\langle \nabla_{\Sigma^i}\Sigma^j, \Sigma^3\rangle\!\rangle \end{pmatrix}. \tag{2.19}$$

Fortunately, $\mathbf{G}^{-1}$ is easy to obtain from (2.16) and we have

$$\mathbf{G}^{-1} = \frac{r}{\tanh r} \operatorname{Proj}^{\perp}_{[y]} + \cosh^2 r \operatorname{Proj}_{[y]}. \tag{2.20}$$

After a lengthy but elementary computation, we obtain

$$\begin{aligned} 2\Gamma^k_{ij} &= \frac{1}{r^2}\left(\frac{r}{\sinh r \cosh r} - 1\right)\left(y_i\delta_{jk} + y_j\delta_{ik} - \frac{2y_iy_jy_k}{r^2}\right) + \\ &\quad + \frac{1}{r^2}\left(1 - \frac{\sinh r \cosh r}{r}\right)\left(y_k\delta_{ij} - \frac{y_iy_jy_k}{r^2}\right) - 2\,\frac{\tanh r}{r}\,\frac{y_iy_jy_k}{r^2}. \end{aligned} \tag{2.21}$$

We now determine the sectional curvatures. By the very definition of Riemannian curvature R, we have

$$\langle\!\langle R(\Sigma^i, \Sigma^j)\Sigma^j, \Sigma^i\rangle\!\rangle = \langle\!\langle \nabla_{\Sigma^i}\Sigma^j, \nabla_{\Sigma^i}\Sigma^j\rangle\!\rangle - \langle\!\langle \nabla_{\Sigma^i}\Sigma_{\Sigma^j}\Sigma^j\rangle\!\rangle.$$

Here, we used that

$$\Sigma^i\langle\!\langle \nabla_{\Sigma^j}\Sigma^j, \Sigma^i\rangle\!\rangle = \Sigma^j\langle\!\langle \nabla_{\Sigma^i}\Sigma^i, \Sigma^j\rangle\!\rangle,$$

which follows from the full symmetry of (2.18). Taking expansions with respect to the coordinate fields, we arrive at

$$\langle\!\langle R(\Sigma^i, \Sigma^j)\Sigma^j, \Sigma^i\rangle\!\rangle = \sum_{k,l=1}^{3} (\Gamma^k_{ij}\Gamma^l_{ij} - \Gamma^k_{ii}\Gamma^l_{jj})\langle\!\langle \Sigma^k, \Sigma^l\rangle\!\rangle.$$

Due to rotation invariance we may assume that $y_1 = y_2 = 0$ and $y_3 = r$. Then

$$\langle\!\langle R(\Sigma^i, \Sigma^j)\Sigma^j, \Sigma^i\rangle\!\rangle = \sum_{k=1}^{3} ((\Gamma^k_{ij})^2 - \Gamma^k_{ii}\Gamma^k_{jj})\langle\!\langle \Sigma^k, \Sigma^k\rangle\!\rangle. \tag{2.22}$$

For $i \neq j$, we denote by

$$K(\Sigma^i, \Sigma^j) = \frac{\langle\!\langle R(\Sigma^i, \Sigma^j)\Sigma^j, \Sigma^i\rangle\!\rangle}{\langle\!\langle \Sigma^i, \Sigma^i\rangle\!\rangle\langle\!\langle \Sigma^j, \Sigma^j\rangle\!\rangle - \langle\!\langle \Sigma^i, \Sigma^j\rangle\!\rangle^2}$$

the sectional curvature with respect to the plane spanned by Σ^i and Σ^j. Substituting (2.21) into (2.22) we obtain

$$K(\Sigma^1, \Sigma^2) = -\frac{1}{4\tanh^2 r}\left(\frac{1}{\cosh r} - \frac{\sinh r}{r}\right)^2 \tag{2.23}$$

and

$$\begin{aligned} K(\Sigma^1, \Sigma^3) &= K(\Sigma^2, \Sigma^3) \\ &= \frac{1}{4}\left(1 - \frac{\sinh r \cosh r}{r}\right)\left(1 + \coth^2 r - \frac{\coth r}{r}\right). \end{aligned} \tag{2.24}$$

These are zero at $r = 0$ and strictly decreasing to $-\infty$ as $r \to \infty$. It follows that the scalar curvature is a strictly decreasing function of r. Similarly to the curvature, the von Neumann entropy is rotation invariant as well. We have

$$S(r) = \log(e^r + e^{-r}) - r \tanh r$$

and observe that scalar curvature and entropy are a monotone function of each other.

We claim that the submanifold S_r ($\{(y_1, y_2, y_3) : y_2 = 0\}$ in logarithmic coordinates) of real density matrices is totally geodesic. In fact, (2.21) implies

$$\langle\!\langle \nabla_{\Sigma^1}\Sigma^2, \Sigma^2 \rangle\!\rangle = \langle\!\langle \nabla_{\Sigma^2}\Sigma^2, \Sigma^2 \rangle\!\rangle = \langle\!\langle \nabla_{\Sigma^3}\Sigma^2, \Sigma^2 \rangle\!\rangle = 0$$

and the claim follows.

The study of the geometry of higher spin systems is in progress [17].

3. Logarithmic Derivatives

Let (φ_θ) be a family of states of the quantum system $\mathscr{M}$ and let the parameter set Θ be an interval in $\mathbb{R}$. We assume that $\theta \mapsto \varphi_\theta$ is differentiable in a certain sense and give a very general definition of the logarithmic derivative with respect to a real or complex inner product $\alpha(\cdot\,,\cdot)$ on $\mathscr{M}$. We define the operator $L \in \mathscr{M}$ to be the logarithmic derivative of the family (φ_θ) at a fixed point θ if

$$\frac{\partial}{\partial\theta}\varphi_\theta(a) = \alpha(L^*, a) \tag{3.1}$$

for every $a \in \mathscr{M}$. The derivation is understood at a fixed point $\theta \in \Theta$ and the form α may depend on this point. To show some very concrete examples of logarithmic derivatives, let ν be a measure on [0, 1] and set

$$\alpha_\nu(a, b) = \int_0^1 \operatorname{Tr} D_\theta^t a^* D_\theta^{1-t} b \,\mathrm{d}\nu(t), \tag{3.2}$$

if the state φ_θ is given by a density D_θ (on a finite quantum system, for example). When $\nu = (\delta_0 + \delta_1)/2$ is a convex combination of two Dirac measures, the corresponding symmetric logarithmic derivative was considered by Helstrom ([6], see also [7]) and it is given by the equation

$$\frac{\partial}{\partial\theta}\operatorname{Tr}(D_\theta a) = \tfrac{1}{2}\operatorname{Tr} D_\theta(L_s a + aL_s) \tag{3.3}$$

or equivalently

$$2\frac{\partial}{\partial\theta} D_\theta = D_\theta L_s + L_s D_\theta. \tag{3.4}$$

In a similar way, the right logarithmic derivative L_r is the solution of the equation

$$\frac{\partial D_\theta}{\partial\theta} = D_\theta L_r \tag{3.5}$$

and it appeared in [20]. While L_s is self-adjoint, in general L_r is not. The Lebesgue measure in place of v supplies us with the Bogoliubov inner product (1). The Bogoliubov logarithmic derivative L_B is the solution of the equation

$$\frac{\partial}{\partial\theta}\operatorname{Tr}(D_\theta a) = \int_0^1 \operatorname{Tr} D_\theta^t L_B D_\theta^{1-t} b \, \mathrm{d}t \tag{3.6}$$

and it may be written in an integral form:

$$L_B = \int_0^\infty (t + D_\theta)^{-1}\frac{\partial D_\theta}{\partial\theta}(t + D_\theta)^{-1}\,\mathrm{d}t. \tag{3.7}$$

PROPOSITION 3.1. *Let the observable $M \in M_n(\mathbb{C})$ be a locally unbiased estimator for the differentiable family (D_θ) of the density matrices at $\theta = 0$. Then*

$$\varphi_0(M^2) \geqslant \alpha_v(L_v, L_v)_{\varphi_0}^{-1},$$

where L_v is the logarithmic derivative of the states $\varphi_\theta(\,\cdot\,) \equiv \operatorname{Tr}(\,\cdot\, D_\theta)$ with respect to the inner product

$$\alpha_v(a, b)_{\varphi_0} = \int_0^1 \operatorname{Tr} D_0^t a^* D_0^{1-t} b \, \mathrm{d}v(t).$$

To prove this statement, we can proceed as in the proof of Proposition 1.1 but we have to use the inequality

$$\alpha_v(a, a)_{\varphi_0} \leqslant \varphi_0(a^2),$$

which is a consequence of (1.4). Next, we compare the lower bounds in the previous proposition when the right, the symmetric, and the Bogoliubov logarithmic derivatives are used. Denote by C_r, C_s, and C_B the corresponding bounds, respectively. Let (D_θ) be a family of the densities such that

$$\left.\frac{\partial}{\partial\theta} D_\theta\right|_{\theta=0} = R$$

and let $\Sigma_i \lambda_i p_i$ be the spectral decomposition of D_0. We have

$$C_r^{-1} = \operatorname{Tr} D_0 L_r L_r^* = \operatorname{Tr} R^2 D_0^{-1} = \sum_i \lambda_i^{-1} \operatorname{Tr} R^2 p_i = \sum_{i,j} \lambda_i^{-1} \operatorname{Tr} R p_j R p_i$$

$$= \tfrac{1}{2} \sum_{i,j} (\lambda_i^{-1} + \lambda_j^{-1}) \operatorname{Tr} R p_j R p_i$$

in the first case. L_s is the solution of (3.4), hence,

$$L_s = 2 \sum_{i,j} \frac{1}{\lambda_i + \lambda_j} p_i R p_j$$

and

$$C_s^{-1} = 2 \sum_{i,j} \frac{1}{\lambda_i + \lambda_j} \operatorname{Tr} R p_j R p_i .$$

Finally, one obtains from (3.7) that

$$C_B^{-1} = \sum_{i,j} \frac{\log \lambda_i - \log \lambda_j}{\lambda_i - \lambda_j} \operatorname{Tr} R p_j R p_i ,$$

and elementary inequalities give that

$$C_r \leqslant C_B \leqslant C_s .$$

As far as estimation of $\varphi_0(M^2)$ is concerned, the bound C_B is less informative than C_s and it is more informative than C_r. By means of the Bogoliubov inner product, the following Cramér–Rao-type inequality may be stated:

$$\langle\langle M, M \rangle\rangle_{\varphi_0} \geqslant \frac{1}{\langle\langle L_B, L_B \rangle\rangle_{\varphi_0}} \tag{3.8}$$

under the conditions of Proposition 3.1. Here the left-hand side is a natural noncommutative analogue of the standard deviation and L_B is the logarithmic derivative with respect to the Bogoliubov inner product.

References

1. Amari, S., *Differential-Geometrical Methods in Statistics*, Lecture Notes in Stat. 28, Springer, Berlin, Heidelberg, New York, 1985.
2. Bratteli, O. and Robinson, D. W., *Operator Algebras and Quantum Statistical Mechanics II*, Springer, New York, Berlin, Heidelberg, 1981.
3. Cencov, N. N., *Statistical Decision Rules and Optimal Inferences*, Amer. Math. Soc., Providence, 1982.
4. Fannes, M. and Verbeure, A., Correlation inequalities and equilibrium states, *Comm. Math. Phys.* **55**, 125–131 (1977).
5. Garrison, J. C. and Wong, J., Bogoliubov inequalities for infinite systems, *Comm. Math. Phys.* **26**, 1–5 (1970).

6. Helstrom, C. W., Minimum mean-square error of estimates in quantum statistics, *Phys. Lett.* **25**, 101–102 (1967).
7. Helstrom, C. W., *Quantum Detection and Estimation Theory*, Academic Press, New York, 1976.
8. Hicks, N. J., *Notes on Differential Geometry*, Van Nostrand, Princeton, 1965.
9. Holevo, A. S., An analogue of the theory of statistical decisions in non-commutative probability theory, *Trans. Moscow Math. Soc.* **26**, 133–149 (1972).
10. Holevo, A. S., *Probabilistic and Statistical Aspects of Quantum Theory*, North-Holland, Amsterdam, 1982.
11. Ingarden, R. S., Janyszek, H., Kossakowski, A., and Kawaguchi, T., Information geometry of quantum statistical systems, *Tensor* **37**, 105–111 (1982).
12. Marshall, A. W. and Olkin, I., Inequalities for the trace function, *Aequationes Math.* **29**, 36–39 (1985).
13. Mori, H., Transport, collective motion and Brownian motion, *Progr. Theoret. Phys.* **33**, 423–455 (1965).
14. Nagaoka, H., A new approach to Cramér–Rao bounds for quantum state estimation, IEICE Technical Report, 89, No. 228, IT89-42, pp. 9–14.
15. Naudts, J., Verbeure, A., and Weder, R., Linear response theory and the KMS condition, *Comm. Math. Phys.* **44**, 87–99 (1975).
16. Petz, D., Entropy in quantum probability I, *Quantum Probability and Related Topics VII*, World Scientific, Singapore, pp. 275–297.
17. Petz, D. and Toth, G.: work in progress.
18. Umegaki, H., Conditional expectations in an operator algebra IV (entropy and information), *Kodai Math. Sem. Rep.* **14**, 59–85 (1962).
19. Wehrl, A., General properties of entropy, *Rev. Modern Phys.* **50**, 221–260 (1978).
20. Yuen, H. P. and Lax, M., Multiple-parameter quantum estimation and measurement of nonself-adjoint observables, *Trans. IEEE* **IT-19**, 740–750 (1973).