# Geometries of Quantum States

**Dénes Petz* and Csaba Sudár****

Department of Mathematics, Faculty of Chemical Engineering
Technical University Budapest
H-1521 Budapest XI. Sztoczek u. 2, Hungary

**Abstract**

The quantum analogue of the Fisher information metric of a probability simplex is searched and several Riemannian metrics on the set of positive definite density matrices are studied. Some of them appeared in the literature in connection with Cramér-Rao type inequalities or the generalization of the Berry phase to mixed states. They are shown to be stochastically monotone here. All stochastically monotone Riemannian metrics are characterized by means of operator monotone functions and it is proven that there exist a maximal and a minimal among them. A class of metrics can be extended to pure states and the Fubini-Study metric shows up there.

* Also Mathematical Institute of the Hungarian Academy of Sciences, H-1364 Budapest, PF. 127, Hungary
* E-mail: PETZ@CH.BME.HU
** E-mail: SUDAR@CH.BME.HU

## I. Introduction

The state space of a classical system with $n$ alternatives is the simplex of probability distributions on the $n$-point-space. The probability simplex is an $n-1$ dimensional manifold with boundary and its affine structure is fairly trivial. The extreme boundary consists of $n$ discrete points. In quantum mechanics, the state space of an $n$ level system is identified with the set of all $n \times n$ positive semidefinite complex matrices of trace 1. (They are called density matrices.) The case $n = 2$ is easily visualized as the unit ball in the 3-space.

$$\frac{1}{2}\begin{pmatrix} 1+x & y-\mathrm{i}z \\ x-\mathrm{i}y & 1-x \end{pmatrix} \qquad \longleftrightarrow \qquad (x,y,z) \in \mathbb{R}^3 \qquad (x^2+y^2+z^2 \leq 1)$$

The boundary consists of noninvertible matrices and it is an infinite set. The case $n = 2$ is simple but for higher $n$ the structure of the topological boundary is rather complicated. The extreme boundary consists of the density matrices of rank one and for $n > 2$ it is much smaller than the topological boundary. As far as dimensionality concerned, the topological boundary is $n^2 - 2$ and the extreme one is $2n - 2$. The extreme states are usually called pure and they are described in the textbooks by nonzero vectors of a complex Hilbert space of linear dimension $n$. The same state is described by a vector $\psi$ as well as $\lambda\psi$, where $\lambda$ is any complex number different from 0. This means that pure states are in one-to-one correspondence to rays $\{\lambda\psi : 0 \neq \lambda \in \mathbb{C}\}$. The rays form a smooth manifold called complex projective space, $\mathbb{C}P^{(n-1)}$.

On the level of convex structure the difference between the classical and quantum state space is well-understood. The classical one is a Choquet simplex and different axiomatizations of the quantum one are available in the literature, the reader may be referred to the works [4, 5], for example. Our main concern here is the possible Riemannian structure in the quantum case. Before turning to that subject, we review briefly the classical case, that is, the Riemannian structure on the space of measures.

¿From the viewpoint of information geometry, the spherical representation of the probability simplex is adequate, because the squared length of the tangent vector of a curve equals the Fisher information. Indeed, introduce the parameters $z_i = 2\sqrt{p_i}$, where $1 \leq i \leq n$ and $\sum_i p_i = 1$. Then $\sum_i z_i^2 = 4$ and the probability simplex is parametrized with a portion of the $n$-sphere. Let $x(t)$ be a curve on the sphere. The square of the length of the tangent is

$$\langle \partial_t x, \partial_t x \rangle = \sum_i (\partial_t x_i)^2 = \sum_i p_i(t)(\partial_t \log p_i(t))^2 \,,$$

which is the Fisher information. The geodesic distance between two probability distributions $Q$ and $R$ can be computed along a great circle and it is a simple transform of the Hellinger distance. The lecture notes [2] contains further details as well as statistical applications of this geometric approach. To the best of our knowledge, Riemannian metric on quantum states was first considered by Helstrom in connection with state estimation theory [13]. Since Helstrom's work, several other metrics appeared in the literature, see for example [6], [7], [12], [23] and Uhlmann approached Helstrom's metric in a different way ([25], [26]).

The present paper is organized as follows. In Section II we survey the work of Chentsov both in the probabilistic and in the quantum case. We explain how he arrived at the study of invariant metrics on the space of probability measures, motivated by decision theory, and how far he could go towards the quantum generalization after his unicity result about the Fisher information in the probabilistic context. Section IV reviewes different approaches to Riemannian metric on the quantum state space. The relation of Uhlmann's and Helstrom's work to Chentsov's idea is enlightened and a concise description of the complex projective space is given. The main results are contained in Sections IV and V. We construct monotome metrics by means of operator monotone functions and prove that all monotone metrics are obtained in this way. Our result completes the program initiated by Chentsov. It turns out that the symmetric logarithmic derivative metric of Helstron (which is the same as the metric studied by Uhlmann) is monotone. Furthermore, this metric is minimal among all monotone metrics. The subject of Section V is the extension of monotone metrics to pure states. We prove that if the extension exists then it coincides with the standard metric of pure states up to a constant factor.

## II. The viewpoint of Chentsov

Chentsov was led by decision theory when he considered a category whose objects are probability spaces and whose morphisms are Markov kernels. Although he worked in [9] with arbitrary probability spaces, his idea can be demonstrated very well on finite ones. In this case a morphism from the probability $n$-simplex $\mathcal{S}_n$ to an $m$-simplex $\mathcal{S}_m$ is an $n \times m$ stochastic matrix. If $\Pi$ is such a matrix and $P \in \mathcal{S}_n$ then $P\Pi \in \mathcal{S}_m$ is considered more random than $P$. Generally speaking, the parametrized family $(Q_i)$ is more random than the parametrized family $(P_i)$ (with the same parameter set) if there exists a stochastic matrix $\Pi$ such that $P_i\Pi = Q_i$ for every value of the parameter $i$. Two parametric families $(P_i)$ and $(Q_i)$ are equivalent in the theory of statistical inferences if there are two stochastic matrices $\Pi^{(12)}$ and $\Pi^{(21)}$ such that

$$P_i\Pi^{(12)} = Q_i \quad \text{and} \quad Q_i\Pi^{(21)} = P_i \tag{2.1}$$

for every $i$. Chentsov said a numerical function $f$ defined on pairs of measures to be invariant if

$$(P_1, P_2) \sim (Q_1, Q_2) \quad \text{implies} \quad f(P_1, P_2) = f(Q_1, Q_2) \tag{2.2}$$

and monotone if

$$f(P_1, P_2) \geq f(P_1\Pi, P_2\Pi)\,. \tag{2.3}$$

for every stochastic matrix $\Pi$. A monotone function $f$ is obviously invariant. Statistics and information theory know a lot of monotone functions, relative entropy and its generalizations are so. If a Riemannian metric is given on all probability simplexes, then this family of metrics is called invariant (respectively, monotone) if the corresponding geodesic distance is an invariant (respectively, monotone) function. Chentsov's greate achievement was that up to a constant factor the Fisher information yields the only monotone family of Riemannian metrics on the class of finite probability simplexes ([9], see also [8]). A decade later Chentsov turned to the quantum case, where the probability simplex is replaced by the set of density matrices. A linear mapping between two matrix spaces sends a density matrix into a density if the mapping preserves trace and positivity (i.e., positive semidefinitness). By now it is well-understood that completely positivity is a natural and important requirement in the noncommutative case. Therefore, we call a trace preserving completely positive mapping stochastic. One of the equivalent forms of the completely positivity of a map $T$ is the following.

$$\sum_{i=1}^{n}\sum_{j=1}^{n} a_i^* T(b_i^* b_i) a_i \geq 0$$

for all possible choice of $a_i$, $b_i$ and $n$. A completely positive mapping $T$ satisfies the Schwarz inequality: $T(a^*a) \geq T(a)^*T(a)$.

Chentsov recognized that stochastic mappings are the appropriate morphisms in the category of quantum state spaces. (The monograph [1] contains more information about stochastic mappings, see also [18].) The above definitions of invariance and monotonicity make sense when stochastic matrices are replaced by stochastic mappings. Chentsov (with Morozova) aimed to find the invariant (or monotone) Riemannian metrics in quantum setting as well. They obtained the following result ([21]). Assume that a family of Riemannian metrics is given on all spaces of density matrices which is invariant, then there exist a function $c(x,y)$ and a constant $C$ such that the squared length of a tangent vector $A = (A_{ij})$ at a diagonal point $D = \mathbf{Diag}(p_1, p_2, \dots, p_n)$ is of the form

$$C \sum_{k=1}^{n} p_k^{-1} A_{kk}^2 + 2 \sum_{j<k} c(p_j, p_k) |A_{jk}|^2 . \tag{2.4}$$

Furthermore, the function $c(x,y)$ is symmetric and $c(\lambda x, \lambda y) = \lambda^{-1} c(x,y)$. This result of Morozova and Chentsov was not complete. Although they had proposals for the function $c(x,y)$, they did not prove monotonicity or invariance of any of the corresponding metrics. A complete result will be given here but now a few comments on (2.4) are in order.

Both the function $c(x,y)$ and the constant are independent from the matrix size $n$. Restricting ourselves to diagonal matrices, which is in some sense a step back to the probability simplex, we can see that there is no ambiguity of the metric. Loosely speaking, the unicity result in the simplex case survives along the diagonal and the offdiagonal provides new possibilities for the definition of a stochastically invariant metric.

## III. Riemannian metrics on quantum states

The demand for Riemannian structure on the whole quantum state space or on a parametrized family of density operators appeared in mathematical physics a long time ago and in rather different contexts.

In the parametric problem of quantum statistics a family $(D_\theta)$ of states of a systems is given and one has to decide between several alternative values of the parameter by using measurements. The set of outcomes of the applied measurements is the parameter set $\Theta$ and we assume that it is a region in $\mathbb{R}^m$. So an estimator measurement $M$ is a positive-operator valued measure on the Borel sets of $\Theta$ and its values are observables of the given quantum system. The probability measure $B \mapsto \mu_\theta(B) = \mathrm{Tr}\,(D_\theta M(B))$ $(B \subset \Theta)$ represents the result of the measurement $M$ when the "true" state is $D_\theta$. The choice of the estimators has to be made by taking into account the expected errors. The aim of an optimal desision process is to search estimators with small error. To an error one can attribute several sizes. For example, one can seek a measurement such that its value is "approximately" equal to the true parameter value. If this holds "in the mean" then the estimator is free of destorsion and such estimator is commonly called unbiased. The accuracy of an unbiased measurement is described by the total mean-square deviation which should be small on the parameter space if we want to choose an effective estimator measurement.

The quantum state estimation was initiated by Helstrom in the 1960's ([13], see also [15]). He followed the Cramér-Rao pattern of mathematical statistics and introduced the concept of symmetric logarithmic derivative. Let $M$ be a positive-operator valued measure on $\mathbb{R}^n$. The corresponding measurement is an unbiased estimator of the parameter $\theta = (\theta_1, \dots, \theta_m)$ if

$$\int_{\mathbb{R}^m} \theta_i \, d\mathrm{Tr}\,(D_t M)(\theta) = t_i \tag{3.1}$$

for every $1 \le i \le m$. (The integration is taken with respect to the measure $B \mapsto \mathrm{Tr}\,(D_t M(B))$.) The symmetric logarithmic derivatives $L^i_\theta$ are observables defined as

$$\frac{\partial \mathrm{Tr}\,(D_\theta A)}{\partial \theta_i} = \frac{1}{2}\mathrm{Tr}\,\big((L^i(\theta)D_\theta + D_\theta L^i(\theta))A\big) \tag{3.2}$$

for every observable $A$. The measurement has two characteristic matrices, the covariance matrix $C(\theta) = (C_{ij}(\theta))$ and the information matrix $J(\theta) = (J_{ij}(\theta))$. They are determined as follows.

$$\begin{aligned} C_{ij}(\theta)) &= \int_{\mathbb{R}^m} (t_i - \theta_i)(t_j - \theta_j)\, d\mathrm{Tr}\,(D_\theta M)(t)\,, \\ J_{ij}(\theta) &= \mathrm{Tr}\,(D_\theta L^i(\theta) L^j(\theta))\,. \end{aligned} \tag{3.3}$$

A quantum version of the Cramér-Rao inequality, due to Helstrom, says that

$$C(\theta) \ge J(\theta)^{-1} \tag{3.4}$$

for an unbiased measurement. (The inequality means that the difference is positive semidefinite.) The information matrix $J(\theta)$ may be regarded the metric tensor on the parameter space.

¿From the point of view of the statistical state estimation problem, the number $n$ of the real parameters is much smaller than the dimension of the whole state space. However, we can parametrize the whole state space as well. Assume that the parametrization is affine,

$$D_\theta = I/n + \sum_i \theta_i a_i\,, \tag{3.5}$$

where $a_i$ are traceless selfadjoint matrices. $D_\theta$ is positive definite if $\theta$ is in a certain open subset of $\mathbb{R}^{n^2-1}$ and the mapping $D_\theta \mapsto \theta \in \mathbb{R}^{n^2-1}$ yields atlas of a single chart. We refer to (3.5) as the affine parametrization of invertible density matrices $\mathcal{D}_n$.

The symmetric logarithmic derivative $L^i(\theta)$ is given by the equation

$$A_i = \tfrac{1}{2}\left(D_\theta L^i(\theta) + L^i(\theta) D_\theta\right). \tag{3.6}$$

When $A_i$ is regarded as a tangent vector at $D_\theta$, its squared length equals to

$$\operatorname{Tr}\left(D_\theta (L^i(\theta))^2\right) = \operatorname{Tr} L^i(\theta) A_i\,. \tag{3.7}$$

If $\sum_j \lambda_j(\theta) p_j(\theta)$ is the spectral decomposition of $D_\theta$, then the solution of (3.6) may be written in the form

$$L^i(\theta) = \sum_{k,j} \frac{2}{\lambda_k(\theta) + \lambda_j(\theta)} p_k(\theta) A_i p_j(\theta)\,. \tag{3.8}$$

To show an example, we consider the $2 \times 2$ case and choose $a_i = \sigma_i$ with the three Pauli spin matrices, that is,

$$\sigma_1 = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \sigma_2 = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 & -\mathrm{i} \\ \mathrm{i} & 0 \end{pmatrix}, \qquad \sigma_3 = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{3.9}$$

Then

$$\|\sigma_i\|_D^2 = \frac{2}{\mu + \nu} \qquad (i = 1, 2)$$

if the footpoint $D$ is diagonal $\mathbf{Diag}(\mu, \nu)$. A convenient affine coordinate system exists also for the whole set $\mathcal{M}_n$ of invertible $n \times n$ density matrices and the Riemannian metric of the symmmetric logarithmic derivative may be written in the form (2.4).

$$\|(A_{ij})\|_D^2 = \sum_{k=1}^{n} p_k^{-1} A_{kk}^2 + 2 \sum_{j<k} \frac{2}{p_j + p_k} |A_{jk}|^2\,, \tag{3.10}$$

where $D = \mathbf{Diag}(p_1, p_2, \ldots, p_n)$. So the value of the constant $C$ is 1 and the Morozova-Chentsov function of the metric of the symmetric logarithmic derivative is

$$c(x, y) = \frac{2}{x + y}\,. \tag{3.11}$$

Before we show how Uhlmann obtained essentially the same Riemannian metric in a completely different approach we review shortly the complex projective space $\mathbb{C}P^{(n-1)}$. It was explained in the introduction that the extreme boundary of the state space (of an $n$-level) quantum system is $\mathbb{C}P^{(n-1)}$.

$\mathbb{C}P^{(n-1)}$ is equipped with an atlas containing $n$ charts. Let $U_i$ be the set of the equivalence classes of all $n$-tuples $(z_1, z_2, \ldots, z_n)$ of complex numbers that $z_i \neq 0$ and set

$$\psi_i\big(p(z_1, z_2, \ldots, z_n)\big) = \Big(\frac{z_1}{z_i}, \ldots, \frac{z_{i-1}}{z_i}, \frac{z_{i+1}}{z_i}, \ldots, \frac{z_n}{z_i}\Big).$$

The standard Riemannian metric of $\mathbb{C}P^{(n-1)}$ is given by considering the $(2n-1)$-sphere

$$|z_1|^2 + |z_2|^2 + \ldots + |z_n|^2 = C > 0\,,$$

which is parametrized now by complex numbers. $S^1$ as the group of complex numbers of modulus one has a natural isometric action on the $(2n-1)$-sphere. The orbits are homeomorphic to circles and the space of orbits may be identified with $\mathbb{C}P^{(n-1)}$. The orbits may be given a metric by taking that obtained by projecting the metric on $S^{2n-1}$ orthogonally to the orbits. This metric is invariant under the natural action of the unitary group $U(n)$ and called sometimes the Fubini-Study metric. (Strictly speaking the Fubini-Study metric is a Kaehler metric on $\mathbb{C}P$ viewed as a Kaehler manifold.)

One of the key issues of quantum mechanics (compared with classical one) is that a subsystem of a system in pure state can be in a mixed state. More precisely, if $D$ is any density operator on the Hilbert space $\mathcal{H}$ then one can find a vector $\xi$ in the enlarged Hilbert space $\mathcal{H} \otimes \mathcal{H}$ such that

$$\mathrm{Tr}\, DA = \langle \xi, (A \otimes I)\xi \rangle \tag{3.12}$$

for every observable $A$ of the smaller system, i.e., acting on $\mathcal{H}$. The vector $\xi$ is not determined uniquely and called the purification of $D$. It is worthwile to regard $\mathcal{H} \otimes \mathcal{H}$ as the Hilbert-Schmidt operators acting on $\mathcal{H}$. Then the observable $A$ of the small system corresponds to the multiplication operator $L_A : X \mapsto AX$ on $\mathcal{H} \otimes \mathcal{H}$. So condition (3.12) reads as

$$\mathrm{Tr}\, DA = \langle W, L_A W \rangle \tag{3.13}$$

when $W$ is written instead of $\xi$. Since $\langle W, L_A W \rangle = \mathrm{Tr}\, W^* W A$, conditions (3.13) simply becomes

$$W^* W = D\,. \tag{3.14}$$

Among all lifts of $D$ into the fibration $W \mapsto W^* W$ there is a canonical one which satisfies the so-called parallelity (or horizontality) condition

$$W^* \dot{W} = \dot{W}^* W \tag{3.15}$$

Uhlmann arrived at this condition from the following minimization problem related to the generalization of the Berry phase to mixed states ([25, 26]). Let $D(t)$ be a smooth curve of density matrices with purification $W(t)$. If the arclength of $W(t)$ with respect to the standard Fubini-Study metric is minimal then the paralellity condition is satisfied. The vectors $Y \in T_W \mathbb{C}P^{(2n-1)}$ such that $W^* Y = Y W^*$ are called horizontal. Any vector $X \in T_D \mathcal{M}$ admits a horizontal lift $X' \in T_W \mathbb{C}P^{(2n-1)}$ and Uhlmann proposed the Riemannian metric

$$g_D^B(X, X) = g_W^{FS}(X', X') \tag{3.16}$$

for any $W$ with $W^* W = D$. If $DG + GD = \dot{D}$ then

$$g^B(\dot{D}, \dot{D}) = \tfrac{1}{2} \mathrm{Tr}\, G \dot{D}$$

In $G/2$ one recognize the symmetric logarithmic derivative and $g^B$ is the corresponding metric up to a factor one half. The letter $B$ in $g^B$ refers to Bures, because the geodesic distance in the metric $g^B$ coicides with the one introduced by Bures many years earlier. The Bures distance is

$$d_B(D_1, D_2) = \sqrt{2 - \mathrm{Tr}\,\left(D_1^{1/2} D_2 D_1^{1/2}\right)^{1/2}}\,. \tag{3.17}$$

It is worthwile to mention that Dittmann computed several geometric characteristics of the space of density matrices endowed with the above metric ([10]). For example this space is not locally symmetric and all sectional curvatures are greater than 1. Braunstein and Caves obtained recenly the same metric by optimizing over all generalized quantum measurements that can be used to distinguish neighboring quantum states $D$ and $D + dD$ ([7]).

## IV. Monotone metrics

If a distance between density matrices expresses statistical distinguishability then this distance must decrease under coarse-graining. A good example of coarse-graining arises when a density matrix is partitioned in the form of a $2 \times 2$ block matrix, and the coarse-graining forgets about the offdiagonal:

$$\begin{pmatrix} A & B \\ B^* & C \end{pmatrix} \longmapsto \begin{pmatrix} A & 0 \\ 0 & C \end{pmatrix}$$

In the mathematical formulation, a coarse-graining is a completely positive mapping which preserves the trace and hence sends density matrix into density matrix. Such mapping will be called stochastic below. A Riemannian metric is defined to be monotone if the

differential of any stochastic mapping is a contraction. If the affine parametrization is considered, then $D_t = D + tA$ is a curve for an invertible density $D$ and for a selfadjoint traceless $A$. Under a stochastic mapping $\mathbf{T}$ this curve is trasformed into $\mathbf{T}(D_t) = \mathbf{T}(D) + t\mathbf{T}(A)$ provided that $\mathbf{T}(D)$ is an invertible density and the real number $t$ is small enough. The monotonicity condition for the Riemannian metric $g$ on $\mathcal{M}_n$ reads as

$$g_{\mathbf{T}(D)}\big(\mathbf{T}(A), \mathbf{T}(A)\big) \le g_D(A, A)\,, \tag{4.1}$$

where $D$ is an invertible density, $A$ is traceless selfadjoint and $\mathbf{T}$ is stochastic. Our goal is to show many examples of monotone metrics and to give their characterization in terms of operator monotone functions.

Let us recall that a function $f : \mathbb{R}^+ \to \mathbb{R}$ is called operator monotone if the relation $0 \le K \le H$ implies $0 \le f(K) \le f(H)$ for any matrices $K$ and $H$ (of any order). The theory of operator monotone functions was established in the 1930's by Löwner and there are several reviews on the subject, for example [3], [11] are suggested.

Let us introduce some superoperators as

$$\mathbf{L}_D(A) = DA, \quad \mathbf{R}_D(A) = AD. \qquad (A \in M_n(\mathbb{C})) \tag{4.2}$$

**Theorem 4.1.** Let $f : \mathbb{R}^+ \to \mathbb{R}^+$ be an operator monotone function such that $f(t) = tf(t^{-1})$ for every $t > 0$ and set a superoperator

$$\mathbf{K}_D = \mathbf{R}_D^{1/2} f(\mathbf{L}_D \mathbf{R}_D^{-1}) \mathbf{R}_D^{1/2} \tag{4.3}$$

acting on matrices. Then the relation

$$g_D(A, B) = \mathrm{Tr}\left(\mathbf{K}_D^{-1}(A) B\right) \tag{4.4}$$

determines a monotone Riemannian metric on $\mathcal{M}_n$.

*Proof.* Since an operator monotone function is analytic, the bilinear form (4.4) is smooth in $D$. The condition $f(t) = tf(t^{-1})$ on $f$ makes sure that $\mathbf{K}_D^{-1}(A)$ is selfadjoint whenever $A$ is so. Hence the bilinear form (4.4) is real. For an invertible $D$ the superoperator $\mathbf{K}_D$ is invertible and positive definite. So (4.4) is really a nondegenerate metric and its monotonicity is to be checked.

In the paper [22] the following inequality was obtained:

$$\mathbf{T}\mathbf{R}_F^{1/2} f(\mathbf{L}_E \mathbf{R}_F^{-1}) \mathbf{R}_F^{1/2} \mathbf{T}^\star \le \mathbf{R}_{\mathbf{T}(F)}^{1/2} f(\mathbf{L}_{\mathbf{T}(E)} \mathbf{R}_{\mathbf{T}(F)}^{-1}) \mathbf{R}_{\mathbf{T}(F)}^{1/2} \tag{4.5}$$

if $E, F$ are positive definite matrices, $\mathbf{T}$ is a stochastic mapping and $\mathbf{T}^\star$ denotes its adjoint with respect to the Hilbert-Schmidt inner product. Putting $E = F = D$ (4.5) becomes

$$\mathbf{T}\mathbf{K}_D\mathbf{T}^\star \le \mathbf{K}_{\mathbf{T}(D)}\,,$$

which is equivalent to

$$\mathbf{T}^\star \mathbf{K}_{\mathbf{T}(D)}^{-1} \mathbf{T} \le \mathbf{K}_D^{-1}\,. \tag{4.6}$$

The latter condition is exactly the monotonicity of the metric (4.4). □

It is in order to make a comment on the relation of the function $f$ in Theorem 4.1 and the Morozova-Chentsov function $c(x, y)$ in (2.4). Given $f$, we have $c(x, y) = 1/yf(x/y)$ and conversely $f(t) = 1/c(t, 1)$. Some examples of functions $f$ satisfying the hypothesis of Theorem 4.1 are the following.

$$\frac{2x^{\alpha+1/2}}{1+x^{2\alpha}}, \quad \frac{x-1}{\log x}, \quad \frac{x-1}{\log x}\frac{2\sqrt{x}}{1+x}, \quad \left(\frac{x-1}{\log x}\right)^2 \frac{2}{1+x}, \quad \frac{1+x}{2} \tag{4.7}$$

where $0 \le \alpha \le 1/2$. The lattest function $f$ gives the Morozova-Chentsov function (3.11) and we obtain that the metric of the symmetric logarithmic derivative is monotone.

The metrics on $\mathcal{M}_2$ provided by Theorem 4.1 are rotation invariant, they depend only on $r = \sqrt{x^2 + y^2 + z^2}$ and split into radial and tangential components:

$$ds^2 = \frac{1}{1-r^2} dr^2 + \frac{1}{1+r} g\left(\frac{1-r}{1+r}\right) dn^2 \quad \text{where} \quad g(t) = \frac{1}{f(t)} \tag{4.8}$$

The radial component is independent of the function $f$. In case of the metric of the symmetric logarithmic derivative the tangential component is independent of $r$.

**Theorem 4.2.** Every monotonone metric is provided by Theorem 4.1.

*Proof.* A monotone metric is invariant in the sense of Section II and due to the result of Chentsov and Morozova the metric is of the form (2.4). Set a function $f$ as $f(t) = 1/c(t,1)$, where $c$ is the function of two variables from (2.4). By means of this function the monotone metric can be written in terms $f$ exactly in the form described in Theorem 4.1, see (4.3) and (4.2). What we have to prove is that $f$ is operator monotone. This will be shown following [24].

We choose a particular stochastic mapping $\mathbf{T}$:

$$\mathbf{T} : X \equiv \begin{pmatrix} X_1 & A \\ B & X_2 \end{pmatrix} \mapsto \frac{1}{2} \begin{pmatrix} X_1 + X_2 & A + B \\ A + B & X_1 + X_2 \end{pmatrix}.$$

With this choice the monotonicity condition yields that

$$Y \mapsto f(\mathbf{L}_Y \mathbf{R}_Y^{-1}) \mathbf{R}_Y$$

is a concave mapping, or equivalently

$$Y \mapsto f(Y \otimes (Y^{-1})^t)(I \otimes Y^t) \tag{4.9}$$

is concave for a positive definite density matrix $Y$. The concavity extends to all positive definite matrices obviously. We write (4.9) for a block matrix

$$Y = \begin{pmatrix} Y_1 & 0 \\ 0 & Y_2 \end{pmatrix}$$

then we observe that concavity of (4.9) implies the concavity of the mapping

$$(Y_1, Y_2) \mapsto f(Y_1 \otimes (Y_2^{-1})^t)(I \otimes Y_2^t)\,. \tag{4.10}$$

Now the choice $Y_2 = I$ gives that the mapping $Y_1 \mapsto f(Y_1)$ must be concave. What we have arrived at is the operator concavity of $f$ which is known to be equivalent to the operator monotonicity of $f$ (cf. [11]). □

Let $f_1$ and $f_2$ be functions satisfying the hypothesis of Theorem 4.1 and let $\mathbf{K}^1$ and $\mathbf{K}^2$ be the corresponding superoperators defined by (4.3). If $f_1 \le f_2$ then $\mathbf{K}_D^1 \le \mathbf{K}_D^1$. The inverse changes this ordering, hence $g_D^1(A,A) \ge g_D^2(A,A)$ for the corresponding metrics. The relation between operator monotone functions and monotone metrics established by Theorems 4.1 and 4.2 respects ordering in the sense that biger function gives a smaller metric. Comparison of different metrics is meaningful only under some normalization. The most natural is

$$g_D(A,A) = \operatorname{Tr} D^{-1} A^2 \quad \text{whenever} \quad DA = AD \tag{4.11}$$

which corresponds to $f(1) = 1$. It is known (see [19]) that among all operator monotone functions with $f(1) = 1$ and $f(t) = tf(t^{-1})$ there is a minimal and a maximal. They are

$$f_{\min}(t) = \frac{2t}{1+t}, \qquad f_{\max}(t) = \frac{1+t}{2}\,. \tag{4.12}$$

So we obtain

**Theorem 4.3.** Under the normalization (4.11), the metric of the symmetric logarithmic derivative is minimal among all monotone metrics.

*Proof.* One has to verify that the function $f_{\max}$ yields the stated metric. From (4.3) and (4.4) we have

$$g_D(A, A) = 2\langle (\mathbf{L}_D + \mathbf{R}_D)^{-1} A, A\rangle \tag{4.13}$$

and $L = 2(\mathbf{L}_D + \mathbf{R}_D)^{-1}$ is exactly the solution of equation (3.6). Hence (4.13) matches (3.7).

We have to emphesize that the theorem states the minimality of the logarithmic derivative metric only under the essential condition that the whole state space of a spin is parametrized. If this is not the case, then no information is provided by the theorem. The largest monotone metric is the metric of the so-called left logarithmic derivative. That appeared in the literature in connection with Cramér-Rao type inequalities. Its monotonicity was established in [23]. The fact that the left logarithmic metric is larger than the symmetric one is elementary and it has been known (for example, [15], p. 282).

The metric corresponding to the Morozova-Chentsov function

$$\frac{\log x - \log y}{x - y}$$

is the Kubo (or Mori, or Bogoliubov) inner product which showed up in [6] and was studied in [23]. In particular, it was proved that the Kubo product is monotone, under more general assumption than a finite spin, and a conjecture was made. Namely, the scalar curvature of the Kubo metric is monotone as well. Monotonicity of the Kubo metric is not surprising because this result is a kind of reformulation of the Lieb convexity theorem ([20]). However, the monotonicity of the scalar curvature seems to be an inequality of new type (provided that the cunjecture is really true). Concerning details we refer to [23] and [14].

In [12] Hasegawa introduced a family of metrics. They can be obtained by the above construction of monotone metrics, however, we are unable to prove that the auxiliary functions are operator monotone. (Numerical computations support the monotonicity of Hasegawa's metric.)

## V. Extension to pure states

The objective of this section is to discuss the extension of monotone metrics of $\mathcal{M}_n$ to pure states $\mathbb{C}P^{(n-1)}$. Since pure states form a low dimensional part of the topological boundary of $\mathcal{M}_n$, it should be well-specified how the extension is understood.

Let $\mathcal{M}_n^\circ$ denote the set of all elements of $\mathcal{M}_n$ whose eigenvalues are distinct and define a projection $\pi : \mathcal{M}_n^\circ \to \mathbb{C}P^{(n-1)}$ as follows. Let $\pi(D)$ be the one-dimensional eigenspace corresponding to the largest eigenvalue of $D \in \mathcal{M}_n^\circ$. This map is smooth (see [16], II.5.8) and $\mathcal{M}_n^\circ$ is a smooth fibre bundle over $\mathbb{C}P^{(n-1)}$ with projection $\pi$

(see [17] I.5.). (The structure group of this bundle is $U(1) \times U(n-1)$, where $U(k)$ is the group of $k \times k$ unitary matrices.) The fibre space is $\pi^{-1}(e)$, where $e$ is the ray generated by the vector $(1, 0, \ldots, 0) \in \mathbb{C}^n$.

Let $T_D\pi$ be the differential of $\pi$ at $D$ and let $H_D$ be the orthogonal complement of $\operatorname{Ker} T_D\pi$ in $T_D\mathcal{M}_n^\circ$ with respect to a fixed monotone Riemannian metric $g_D(\,\cdot\,,\,\cdot\,)$. Since $T_D\pi$ is surjective, the restriction of $T_D\pi$ gives a linear isomorphism between $H_D$ and $T_{\pi(D)}\mathbb{C}P^{(n-1)}$. If $v \in T_{\pi(D)}\mathbb{C}P^{(n-1)}$, then we can define a unique lift $\tilde{v} \in H_D$ of $v$ such that $T_D\pi(\tilde{v}) = v$. Using this lift we can define the following inner product $k^D_{\pi(D)}(\,\cdot\,,\,\cdot\,)$ on $T_{\pi(D)}\mathbb{C}P^{(n-1)}$:

$$k^D_{\pi(D)}(u, v) = g_D(\tilde{u}, \tilde{v}) \qquad (u, v \in T_{\pi(D)}\mathbb{C}P^{(n-1)}). \tag{5.1}$$

We say that a sequence $D_n \in \mathcal{M}_n^\circ$ is radial at $p \in \mathbb{C}P^{(n-1)}$ if $\pi(D_n) = p$ for every $n$ and $D_n$ is convergent to $p$ when $p$ is considered as a density matrix (that is, a one-dimensional projection operator). Now we can define the radial extension of $g(\,\cdot\,,\,\cdot\,)$. A smooth metric $k(\,\cdot\,,\,\cdot\,)$ on $\mathbb{C}P^{(n-1)}$ is called the radial extension of $g(\,\cdot\,,\,\cdot\,)$ if for every $p \in \mathbb{C}P^{(n-1)}$, $u, v \in T_p\mathbb{C}P^{(n-1)}$ and for every radial sequence $D_n$ at $p$

$$\lim_{n\to\infty} g_p^{D_n}(u, v) = k_p(u, v)$$

holds. In the next theorem we give a necessary and sufficient condition for the existence of the radial extension.

**Theorem 5.1.** Let $g(\,\cdot\,,\,\cdot\,)$ be a monotone Riemannian metric on $\mathcal{M}_n$ and let $f : \mathbb{R}^+ \to \mathbb{R}^+$ be the corresponding operator monotone function (described in Theorem 4.1). The radial extension $k(\,\cdot\,,\,\cdot\,)$ of the given metric $g(\,\cdot\,,\,\cdot\,)$ of $\mathcal{M}_n$ exist if and only if $f(0) \neq 0$. In the case of existence

$$k(\,\cdot\,,\,\cdot\,) = \frac{1}{2f(0)}\langle\,\cdot\,,\,\cdot\,\rangle,$$

where $\langle\,\cdot\,,\,\cdot\,\rangle$ is the standard Riemannian metric on $\mathbb{C}P^{(n-1)}$.

*Proof.* The proof is based on the direct computation of $k^D_{\pi(D)}(\,\cdot\,,\,\cdot\,)$. For any unitary matrix $U$ and $D \in \mathcal{M}_n^\circ$ we have

$$\pi(UDU^{-1}) = U\pi(D)$$

which implies

$$T_{UDU^{-1}}\pi(UXU^{-1}) = UT_D\pi(X) \qquad (X \in T_D\mathcal{M}_n^\circ)$$

by differentiation. Since $k(\,\cdot\,,\,\cdot\,)$ is unitary invariant,

$$U(\operatorname{Ker} T_D\pi)U^{-1} = \operatorname{Ker} T_{UDU^{-1}}\pi \quad \text{and} \quad UH_DU^{-1} = H_{UDU^{-1}}.$$

Moreover, $U\tilde{v}U^{-1} = \widetilde{Uv}$ for any $v \in T_{\pi(D)}\mathbb{C}P^{(n-1)}$, hence we get

$$g^D_{\pi(D)}(u, v) = g^{UDU^{-1}}_{U\pi(D)}(Uu, Uv)\,. \tag{5.2}$$

¿From this equality it follows that it is sufficient to compute $k^D(\cdot\,,\,\cdot)$ if $D$ is diagonal and $\pi(D)$ is the projection onto $e$. Assume these and let $X \in T_D\mathcal{M}_n^\circ$ and let $\lambda(t)$ and $v(t)$ be the largest eigenvalue and the unit eigenvector corresponding to $\lambda(t)$ of $D+tX$ where $t \in \mathbb{R}$. For sufficiently small $t$, $D+tX \in \mathcal{M}_n^\circ$ and $\lambda(t)$ and $v(t)$ are smooth functions of $t$. For $D(t) = D+tX$ we have

$$(D(t) - \lambda(t))v(t) = 0\,.$$

Differentiating this expression we obtain that $\lambda'(0) = x_{11}$ and

$$T_D\pi(X) = v'(0) = \big(0, \frac{x_{21}}{\lambda_1 - \lambda_2}, \ldots, \frac{x_{n1}}{\lambda_1 - \lambda_n}\big)\,, \tag{5.3}$$

where $\lambda_1, \ldots, \lambda_n$ are the eigenvalues of $D$, $\lambda_1 = \lambda(0)$ and $X = (x_{ij})$. If $X \in \operatorname{Ker} T_D\pi$ then the expression of $T_D\pi(X)$ gives

$$X = \begin{pmatrix} x_{11} & 0 & \ldots & 0 \\ 0 & x_{22} & \ldots & x_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & x_{n2} & \ldots & x_{nn} \end{pmatrix}$$

Let $\mathbf{K}_D^{-1} = f(\mathbf{L}_D\mathbf{R}_D^{-1})\mathbf{R}_D$ as in (4.3). Since D is diagonal,

$$K_D(X)_{ij} = \frac{x_{ij}}{f(\lambda_i/\lambda_j)\lambda_j} \tag{5.4}$$

hence we get $K_D(\operatorname{Ker} T_D\pi) = \operatorname{Ker} T_D\pi$. If $V \in H_D$ then the last equation gives

$$V = \begin{pmatrix} 0 & \bar{v}_2 & \ldots & \bar{v}_n \\ v_2 & 0 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ v_n & 0 & \ldots & 0 \end{pmatrix} \tag{5.5}$$

where $v_i \in \mathbb{C}$ for $i = 2, \ldots, n$. If $v = (0, v_2, \ldots, v_n) \in T_{[e]}\mathbb{C}P^{(n-1)}$ then (5.3) and (5.5) give

$$\tilde{v} = \begin{pmatrix} 0 & (\lambda_1 - \lambda_2)\bar{v}_2 & \ldots & (\lambda_1 - \lambda_n)\bar{v}_n \\ (\lambda_1 - \lambda_2)v_2 & 0 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ (\lambda_1 - \lambda_n)v_n & 0 & \ldots & 0 \end{pmatrix}.$$

Now we can express $g^D(\cdot\,,\,\cdot)$:

$$g^D(u,v) = \operatorname{Re} \sum_{i=2}^{n} \frac{(\lambda_1 - \lambda_i)^2}{f(\lambda_i/\lambda_1)\lambda_1} u^i \bar{v}^i, \tag{5.6}$$

where $u, v \in T_{[e]}\mathbb{C}P^{(n-1)}$.

Let us consider now the general case. Let $(D_m)$ be a radial sequence at $p$ and let $u, v \in T_p\mathbb{C}P^{(n-1)}$. Let $B_p^m$ be linear operators on $T_p\mathbb{C}P^{(n-1)}$ such that

$$g_p^{D_m}(u,v) = \langle B_p^m u, v\rangle_p\,,$$

where $\langle\,\cdot\,,\,\cdot\,\rangle_p$ is the inner product on $T_p\mathbb{C}P^{(n-1)}$ induced by the standard metric. Let $U_m$ be unitary operators such that, $D_m^0 = U_m D_m U_m^{-1}$ is diagonal and $\pi(D_m^0) = p_0$ with $p_0 = [e]$. Using (5.2) we have:

$$B_p^m = U_m^{-1} \cdot B_{p_0}^m \cdot U_m \tag{5.7}$$

Since $\lim_{m\to\infty} \lambda_1^m = 1$ and $\lim_{m\to\infty} \lambda_i^m = 0$ for $i = 2, \dots, n$, by (5.6)

$$\lim_{m\to\infty} \|B_{p_0}^m - cI_{p_0}\|_{p_0} = 0 \qquad (c = 1/2f(0)), \tag{5.8}$$

where $I_{p_0}$ is the identity map on $T_{p_0}\mathbb{C}P^{(n-1)}$ and $\|\cdot\|$ is the operator norm induced by $\langle\,\cdot\,,\,\cdot\,\rangle$. It follows from (5.7) that

$$\begin{aligned}&\|B_p^m - cI_p\| = \|U_m^{-1} \cdot B_{p_0}^m \cdot U_m - cU_m^{-1} \cdot I_{p_0} \cdot U_m\| \\ &= \|U_m^{-1} \cdot (B_{p_0}^m - cI_{p_0}) \cdot U_m\| \le \|U_m^{-1}\| \cdot \|B_{p_0}^m - cI_{p_0}\| \cdot \|U_m\|\end{aligned}$$

Since $U_m$ are isometries from $T_p\mathbb{C}P^{(n-1)}$ to $T_{p_0}\mathbb{C}P^{(n-1)}$, $\|U_m\| = 1$ and by (5.8) we obtain

$$\lim_{m\to\infty} \|B_p^m - cI_p\| = 0 \qquad (c = 1/2f(0))\,.$$

So we have proved that the radial extension exists if $f(0) \neq 0$.

The special case $n = 2$ is very transparent from (4.8) and it explains the terminology "radial extension". The $2 \times 2$ case shows also that the condition $f(0) \neq 0$ is necessary to speak about extension. □

## Acknowledgement

DP thanks to the Erwin Schrödinger Institut on Mathematical Physics (Vienna) for an invitation and CS acknowleges support of the Hungarian National Foundation for Scientific Research grant no. OTKA 1900.

## References

**[1]** P.M. Alberti, A. Uhlmann, *Stochasticity and partial order. Doubly stochastic maps and unitary mixing* (VEB Deutscher Verlag Wiss., Berlin, 1981)

**[2]** S. Amari, *Differential-geometrical methods in statistics*, Lecture Notes in Stat. **28** (Springer, Berlin, Heidelberg, New York, 1985)

**[3]** T. Ando, Concavity of certain maps and positive definite matrices and applications to Hadamard products, Linear Alg. Appl. **26**(1979), 203–241

**[4]** H. Araki, On the characterization of the state space in quantum mechanics, Commun. Math. Phys. **75**(1980), 1–25

[**5**] Sh.A. Ayupov, N.J. Yadgorov, Geometry of the state spaces in quantum probability, in *Probability Theory and Mathematical Statistics*, eds. B. Grigelionis et al., pp. 1–9

[**6**] R. Balian, Y. Alhassid, H. Reinhardt, Dissipation in many-body systems: A geometric approach based on information theory, Phys. Rep. **131**(1986), 1–146

[**7**] S.L. Braunstein, C.M. Caves, Statistical distance and the geometry of quantum states, Phys. Rev. Lett. **72**(1994), 3439–3443

[**8**] L.L. Campbell, An extended Čencov charecterization of the information metric, Proc. Amer. Math. Soc. **98**(1986), 135–141

[**9**] N.N. Cencov, *Statistical decision rules and optimal inferences*, Translation of Math. Monog. 53 (Amer. Math. Society, Providence, 1982)

[**10**] J. Dittmann, On the Riemannian geometry of finite dimensional mixed states, Seminar Sophus Lie, **3**(1993), 73–87

[**11**] F. Hansen, G.K. Pedersen, Jensen's inequality for operators and Löwner's theorem", Math. Ann. **258**(1982), 229–241

[**12**] H. Hasegawa, Non-commutative extension of the information geometry, to appear in *Quantum Communication and Measurement*, eds. V.P. Belavkin, O. Hirota, R.I. Hudson, Plenum

[**13**] C.W. Helstrom, *Quantum detection and estimation theory*, (Academic Press, New York, 1976)

[**14**] F. Hiai, D. Petz, G. Toth, Curvature in the geometry of canonical correlation, to appear in Studia Sci. Math. Hungar.

[**15**] A.S. Holevo: *Probabilistic and statistical aspects of quantum theory* (North-Holland, Amsterdam, 1982)

[**16**] T. Kato, *Perturbation Theory for Linear Operators*, (Springer, Berlin, Heidelberg, New York, 1980)

[**17**] S. Kobayashi, K. Nomizu, *Foundations of Differential Geometry, Volume I*, (Intersience, John Wiley & Sons, New York, London, 1963)

[**18**] K. Kraus, *States, Effects, and Operations*, Lecture Notes in Physics **190** (Springer, Berlin, Heidelberg, New York, 1983)

[**19**] F. Kubo, T. Ando, Means of positive linear operators, Math. Ann. **246**(1980), 205-224

[**20**] E.H. Lieb, Some convexity and subadditivity properties of entropy, Bull. Amer. Math. Soc. **81**(1975), 1–14

[**21**] E.A. Morozova, N.N. Chentsov, Markov invariant geometry on state manifolds (in Russian), Itogi Nauki i Tehniki **36**(1990), 69–102

[**22**] D. Petz, Quasi-entropies for finite quantum systems, Rep. Math. Phys. **21** (1986), 57–65

[**23**] D. Petz, Geometry of Canonical Correlation on the State Space of a Quantum System, J. Math. Phys. **35**(1994), 780–795

[**24**] D. Petz, Monotone metrics on matrix spaces, Linear Algebra Appl., to appear

[**25**] A. Uhlmann, The metric Bures and the geometric phase, in *Groups and Related Topics*, eds. R. Gielerak et al., 267–274, Kluwer Academic Publisher, 1992

[**26**] A. Uhlmann, Density operators as an arena for differential geometry, Rep. Math. Phys. **33**(1993), 253–263