SOME OPEN PROBLEMS IN INFORMATION GEOMETRY

P. GIBILISCO
Dipartimento SEFEMEQ and Centro V. Volterra,
Facoltà di Economia, Università di Roma "Tor Vergata",
Via Columbia 2, 00133 Rome, Italy
E-mail: gibilisco@volterra.mat.uniroma2.it

T. ISOLA
Dipartimento di Matematica,
Università di Roma "Tor Vergata",
Via della Ricerca Scientifica, 00133 Rome, Italy
E-mail: isola@mat.uniroma2.it

In this paper we recall some open problems in Information Geometry.

1. Orlicz geometry and statistical manifolds

Let $(X, \mathcal{F}, \mu)$ be a measure space and let

$$\mathcal{M}_\mu := \{\rho : X \to \mathbb{R} | \rho > 0\,, \int \rho = 1\}$$

be the associated maximal statistical model. To have an infinite-dimensional version of Information Geometry one has to solve the following problems: i) to give a differentiable manifold structure to $\mathcal{M}_\mu$; ii) to equip $\mathcal{M}_\mu$ with the α-geometries, namely the family of geometries containing exponential, mixture, and Fisher-Rao geometry. These two problems where solved in Ref. 14, 7, 2. The Pistone-Sempi solution is particularly appealing also from a physical point of view. Indeed in the $\mathcal{M}_\mu$ manifold points are close if the Kullback-Leibler relative entropy is not too big, while in L^p topologies the situation is quite different; therefore if one has to take into account entropy then the right topology should be an Orlicz one (see Ref. 16).

In Ref. 14, 7, 2 there are two main ideas: i) use Orlicz geometry to give a manifold structure to $\mathcal{M}_\mu$; ii) use natural geometry (Levi-Civita connection) of L^p spheres to costruct α-connections. The quantum version of these ideas do not share the same fate. Indeed the second idea has an immediate quantum version because of the similarity between commutative and non-commutative L^p spaces (see Ref. 2). The Pistone-Sempi costruction does not appear so easy to "quantize", because of intricacy of the theory of quantum Orlicz spaces. Some important steps in this direction have been done by Grasselli, Streater and recently also by Jencova.

Maybe the situation may change if we are able to present the Pistone-Sempi costruction in a more natural way using an embedding (similarly to the α-connection case). Actually the exponential geometry is a limit case of α-geometries, namely L^p geometries. Nevertheless, while the exponential parallel transport is simple, the exponential geometry does not come from a natural embedding. Maybe a new look to the commutative case can be of help in the quantum one.

We suggest to consider an idea already contained in Refs. 7, 8.

Given a Young function Φ, we denote by $L^\Phi = L^\Phi(\mu)$ the associated Orlicz space. Suppose the Young function Φ is invertible when restricted to the positive axis (this is not always the case: consider the Φ relative to L^∞).

We define the *Amari Φ–embedding* $A^\Phi : \mathcal{M}_\mu \to L^\Phi(\mu)$ by

$$A^\Phi(\rho) := \Phi^{-1}(\rho) \ .$$

The standard example is $\rho \to p\rho^{\frac{1}{p}}$ (we would like to give a similar treatment to $\rho \to \log(\rho)$). In Ref. 7 we proved the following result.

Proposition 1.1. *Let $S^\Phi = \{v \in L^\Phi : \|v\|_\Phi = 1\}$ be the unit sphere of the Banach space L^Φ. Then*

$$A^\Phi(\mathcal{M}_\mu) \subset S^\Phi \ .$$

In general, the space L^Φ is not uniformly convex, so the L^p approach cannot be imitated directly. Nevertheless, is quite possible that for a suitable Young function Φ the sphere S^Φ has a "good" behaviour in the region $A^\Phi(\mathcal{M}_\mu)$. In this case we may try to deduce the Pistone-Sempi structure as a pulback of the natural geometry of the S^Φ sphere by the Φ–embedding.

2. Quantum Fisher Information and Uncertainty Principle

The Heisenberg uncertainty principle

$$\mathrm{Var}_\rho(A) \cdot \mathrm{Var}_\rho(B) \geq \frac{1}{4}|\mathrm{Tr}(\rho[A,B])|^2$$

is a immediate consequence of a stronger inequality involving covariance proved by Schrödinger that is

$$\mathcal{S}_\rho(A,B) \geq \frac{1}{4}|\mathrm{Tr}(\rho[A,B])|^2,$$

where

$$\mathcal{S}_\rho(A,B) := \mathrm{Var}_\rho(A)\cdot\mathrm{Var}_\rho(B) - |\mathrm{Re}\{\mathrm{Cov}_\rho(A,B)\}|^2,$$

with ρ density matrix and A, B self-adjoint matrices.

It is natural to ask if a similar bound exists as a function of the commutators $[A,\rho],[B,\rho]$ instead of the commutator $[A,B]$; to this respect one should consider the Wigner-Araki-Yanase theorem for quantum measurement, which states that observables not commuting with a conserved quantity cannot be measured exactly (see Ref. 11).

It can be surprising that only in very recent times an inequality of this type has been proved. To present it in an expressive form we need to introduce the machinery of monotone metrics (the quantum counterpart of Fisher information).

Let M_n be the space of complex $n\times n$ matrices and let $\mathcal{D}_n^1$ be the set of density matrices namely

$$\mathcal{D}_n^1 = \{\rho \in M_n | \mathrm{Tr}\rho = 1,\ \rho > 0\}.$$

A monotone metric (or quantum Fisher information) is a family of riemannian metrics $g = \{g^n\}$ on $\{\mathcal{D}_n^1\}$, $n \in \mathbb{N}$, such that

$$g^m_{T(\rho)}(TX, TX) \leq g^n_\rho(X,X)$$

holds for every Markov morphism $T : M_n \to M_m$ (completely positive, trace preserving map) and all $\rho \in \mathcal{D}_n^1$ and $X \in T_\rho\mathcal{D}_n^1$.

With each operator monotone function f one associates the so-called Chentsov–Morotzova function

$$c_f(x,y) := \frac{1}{yf(xy^{-1})} \qquad \text{for} \qquad x, y > 0.$$

Define $L_\rho(A) := \rho A$, and $R_\rho(A) := A\rho$.

Now we can state the fundamental theorem about monotone metrics (classification is up to scalars).

Theorem 2.1. *(Petz 1996) There exists a bijective correspondence between monotone metrics on $\mathcal{D}_n^1$ and symmetric operator monotone functions. This correspondence is given by the formula*

$$g_f(A,B) := g_{f,\rho}(A,B) := \mathrm{Tr}(A\cdot c_f(L_\rho, R_\rho)(B)).$$

Let V be a finite dimensional real vector space with a scalar product $g(\cdot,\cdot)$. We define, for $v, w \in V$,

$$\mathrm{Area}_g(v,w) := \sqrt{g(v,v)\cdot g(w,w) - |g(v,w)|^2}.$$

Let

$$f_\beta(x) := \frac{(x-1)^2}{(x^\beta - 1)(x^{1-\beta}-1)} \qquad \beta \in [-1,2] \setminus \{0,1\}\,.$$

One can prove that the functions f_β are operator monotone.

Since $i[\rho, A]$ is traceless and selfadjoint, then $i[\rho, A] \in T_\rho \mathcal{D}_n^1$. Let $g_\beta(\cdot,\cdot) := g_{f_\beta}(\cdot,\cdot)$ be the quantum Fisher information associated to f_β, and $\mathrm{Area}_\beta(\cdot,\cdot)$ the corresponding area functional on the tangent space.

The monotone metric g_β is known as Wigner-Yanase-Dyson monotone metric of parameter β.

We are ready for the main result

Theorem 2.2. *The inequality*

$$\mathcal{S}_\rho(A,B) \geq \frac{1}{4}\mathrm{Area}_\beta(i[A,\rho], i[B,\rho])^2$$

is true for $\beta \in (0,1)$ *and is false for* $\beta \in [-1,0) \cup (1,2]$.

The case $\beta = \frac{1}{2}$ has been conjecture in Ref. 11 by S. Luo and Z. Zhang and proved in Ref. 10 by S. Luo himself and Q. Zhang. The general case $\beta \in (0,1)$ has been proved independently by H. Kosaki in Ref. 9 and by K.Yanagi, S. Furuichi and K. Kuriyama in Ref. 17. The counterexample for $\beta \in [-1,0) \cup (1,2]$ is due to the present authors and can be found in Ref. 6.

Let us make some more comments on this result. The proof by Yanagi-Furuichi-Kuriyama appears simpler than Kosaki proof. Nevertheless Kosaki was also able to establish necessary and sufficient conditions to have equality and moreover he showed that the function

$$A(\beta) := \mathrm{Area}_\beta(i[A,\rho], i[B,\rho])^2$$

is increasing on $(0, \frac{1}{2})$. Note that the inequality respects the ordering of the associated operator monotone functions. Let us underline that the WYD-metrics come from L^p-geometry (see Ref. 4) where $p = \frac{1}{\beta}$. Therefore the inequality of theorem 2.2 is true for $p \in (1, +\infty)$ that is when L^p spaces are well behaved Banach spaces.

A natural question is the following: are there other operator monotone functions f such that for the associated quantum Fisher information $g := g_f$

it is true that

$$\mathcal{S}_\rho(A,B) \geq \frac{1}{4}\text{Area}_g(i[A,\rho], i[B,\rho])^2$$

for any ρ, A, B?

Another problem has been suggested by Kosaki in Ref. 9: can Theorem 2.2 be generalized in the setting of arbitrary von Neumann algebras?

3. Schur-convexity of curvature for statistical models

Let us recall that a function $f : \mathcal{D}_n^1 \to \mathbb{R}$ is Schur-convex (Schur-increasing) if $A \succ B \Rightarrow f(A) \geq f(B)$, where the symbol $\succ$ stays for the "more mixed" relation. A Schur-convex function behaves like entropy, namely it increases with mixing. Let $(\mathcal{M}, g)$ be a riemannian manifold. The scalar curvature at the point ρ, denoted by $\text{Scal}_g(\rho)$, is (up to normalizing factor) the "average curvature" at ρ.

Because of its relation with volume of geodesic balls it has been suggested by Petz that, in Information Geometry, the scalar curvature should have the meaning of *average statistical uncertainty* and therefore the Schur-convexity of scalar curvature would be a desirable property (see Ref. 13).

In what follows $\tilde{p}$ is defined by $1/p + 1/\tilde{p} = 1$. Recall that the function $f(x) = (x-1)/\log(x)$ is operator monotone and the associated monotone metric is known as the Bogoliubov-Kubo-Mori metric. The $WYD(p)$-metric is the quantum Fisher information associated to

$$f_p(x) := \frac{1}{p\tilde{p}} \frac{(x-1)^2}{(x^{\frac{1}{p}}-1)(x^{\frac{1}{\tilde{p}}}-1)}$$

(here we use a different parameter, $p = 1/\beta$, and a different normalization with respect to Section 2). The L^p-geometries on the state space $\mathcal{D}_n^1$ are the geometries given by pull-back of the embeddings $\rho \mapsto p\rho^{\frac{1}{p}}$ for $p \in [1, +\infty)$ and $\rho \mapsto \log(\rho)$ for $p = +\infty$ (these are simply the α-geometries quoted in Section 1). A number of conjectures have been formulated in this field.

Conjecture 3.1 *(Petz conjecture)*
The scalar curvature of BKM-metric is Schur-convex.

Conjecture 3.2 *(WYD-conjecture)*
The scalar curvature of $WYD(p)$-metric is Schur-convex for p near 1 .

Conjecture 3.3 *(L^p-conjecture)*
The scalar curvature of L^p-geometry is
i) Schur-convex for $p \in (2, +\infty]$;
ii) Schur-concave for $p \in (1, 2)$.

Using a continuity argument one can prove that if the WYD-conjecture is true then the Petz conjecture is true (see Ref. 5). At first sight the WYD-conjecture does not seem any easier to prove than Petz conjecture. But as explained in Ref. 5 the L^p-conjecture "almost" implies the WYD-conjecture and the geometric content of the L^p-conjecture is self-evident when one looks at a picture of the unit sphere in L^p spaces. Up to now the Petz conjecture has been proved only in the 2×2 case (see Ref. 12) while the L^p-conjecture has been proved in the commutative case for $n = 2$ (Ref. 5).

Further interest in this area derives from the geometrical approach to statistical mechanics where it is postulated that the scalar curvature is proportional to free energy density: a recent account of this subject can be found in Refs. 15, 1.

References

1. D.Brody and A. Ritz. Information geometry of finite Ising models, *J. Geom Phys.*, **47** 207-220, (2003).
2. P.Gibilisco and T. Isola. Connections on Statistical manifolds of Density Operators by Geometry of Noncommutative L^p-Spaces, *Inf. Dim. Anal. Quant Prob. & Rel. Top.*, **2** 169-178, (1999).
3. P.Gibilisco and T. Isola. Wigner-Yanase information on quantum state space: the geometric approach, *J. Math. Phys.*, **44**(9): 3752-3762, (2003).
4. P.Gibilisco and T. Isola. , On the characterization of paired monotone metrics, *Ann. Ins. Stat. Math.*, **56**(2): 369-381, (2004).
5. P.Gibilisco and T. Isola. On the monotonicity of scalar curvature in classical and quantum information geometry, *J. Math. Phys.*, **46**(2): 023501,14, (2005).
6. P.Gibilisco and T. Isola. Uncertainty Principle and Quantum Fisher Information, Preprint arXiv:math-ph/0509046v1, (2005).
7. P.Gibilisco and G. Pistone. Connections on nonparametric statistical manifolds by Orlicz space geometry, *Inf. Dim. Anal. Quant Prob. & Rel. Top.*, **1** 325-347, (1998).
8. P.Gibilisco and G. Pistone. Analytical and geometrical properties of statistical connections in Information Geometry, in "Mathematical Theory of Networks and Systems", A. Beghi, L. Finesso, G. Picci (eds.), p. 811-814, Il Poligrafo, Padova (1999).
9. H. Kosaki. Matrix trace inequalities related to uncertainty principle, *Inter. Jour. Math.*, **6** 629-645, (2005).
10. S. Luo and Q. Zhang. On skew information, *IEEE Trans. Infor. Theory*, **50**(8), 1778-1782, (2004).
11. S. Luo and Z. Zhang. An informational characterization of Schrödinger's uncertainty relations, *J. Stat. Phys.*, **114**, 1557-1576, (2004).
12. D. Petz. Geometry of canonical correlation on the state space of a quantum system, *J. Math. Phys.*, **35**, 780-795, (1994).

13. D. Petz. Covariance and Fisher information in quantum mechanics, *J. Phys. A: Math. Gen*, **35**, 929-939, (2002).
14. G. Pistone and C. Sempi. An infinite-dimensional geometric structure on the space of all probability measures equivalent to a given one, *Ann. Stat.*, **33**, 1543-1561, (1995).
15. G. Ruppeiner, Riemannian geometry of thermodynamics and systems with repulsive power-law interactions, *Phys. Rev. E* **72**, 016120, (2005).
16. R.F. Streater, Quantum Orlicz Spaces in Information Geometry, *Open. Sys. & Inf. Dyn.* **11**, 359-375, (2004).
17. K. Yanagi, S. Furuichi and K. Kuryama. A generalized skew information and uncertainty relation, to apper on *IEEE Trans. Infor. Theory*, Preprint arXiv:quant-ph/0501152v2, (2005).