Connes distance and optimal transport

Pierre Martinetti
Dipartimento di matematica, università di Genova, via Dy odecaneso 35, 16146 Genova
E-mail: martinetti@dima.unige.it

Abstract. We give a brief overview on the relation between Connes spectral distance in noncommutative geometry and the Wasserstein distance of order 1 in optimal transport. We first recall how these two distances coincide on the space of probability measures on a riemannian manifold. Then we work out a simple example on a discrete space, showing that the spectral distance between arbitrary states does not coincide with the Wasserstein distance with cost the spectral distance between pure states.

1. The metric aspect of noncommutative geometry

Topology is the minimal structure required for a set to be called *space*, and for its elements to become *points*. Topology gives sense to the notion of neighbourhood (and more generally to the notion of open sets). The algebraic dual notion is that of continuity (a continuous function being, by definition, such that the inverse image of any open set is an open set). Actually this duality goes quite far, since all the information of a topological space is contained within the algebra of continuous functions defined on it. More precisely, Gelfand's duality states that any complex *commutative* C^*-algebra $\mathcal{A}$ is isomorphic to the algebra of continuous functions vanishing at infinity on some locally compact topological space - the space $\mathcal{P}(\mathcal{A})$ of pure states of $\mathcal{A}$ - and conversely any locally compact topological space $\mathcal{X}$ is homeomorphic to the space of pure states of the *commutative* algebra $C_0(\mathcal{X})$ of continuous functions on $\mathcal{X}$ vanishing at infinity:

$$\mathcal{A} \simeq C_0(\mathcal{P}(\mathcal{A})), \quad \mathcal{X} \simeq \mathcal{P}(C_0(\mathcal{X})). \tag{1}$$

Recall that the pure states of an involutive algebra $\mathcal{A}$ are the extremal points of the set of states, the latter being the linear maps φ on $\mathcal{A}$ wich are positive - $\varphi(a^*a) \in \mathbb{R}^+$ - and of norm 1 (where $||\varphi|| = \sup_{a\in\mathcal{A}} \frac{|\varphi(a)|}{||a||}$). In particular, a state of the commutative algebra $C_0(\mathcal{X})$ is the integration with respect to a probability measure μ, with pure states given by Dirac δ measures (i.e. evaluation at a point):

$$\mathcal{P}(C_0(\mathcal{X})) \ni \delta_x : x \to \int_{\mathcal{X}} f\delta_x = f(x)\,, \qquad S(C_0(\mathcal{X})) \ni \varphi : f \to \int_{\mathcal{X}} f\, \mathrm{d}\mu \quad \forall f \in C_0(\mathcal{X}).$$

Connes' noncommutative geometry extends Gelfand's duality beyond topology, so that to encompass all the aspects of riemannian geometry, in particular the metric. To do so, one needs more than an algebra: a *spectral triple* [4] consists in an involutive algebra $\mathcal{A}$ acting faithfully on an Hilbert space $\mathcal{H}$, with D a selfadjoint operator on $\mathcal{H}$ such that the commutator $[D, a]$ is bounded and $a[D - \lambda\mathbb{I}]^{-1}$ is compact for any $a \in \mathcal{A}$ and $\lambda \notin$ Sp D. When a set of conditions (dimension, regularity, finitude, first order, orientability) is satisfied, then one is able to characterize a riemannian manifold by purely spectral data [5]:

- For $\mathcal{M}$ a compact Riemann manifold, then

$$(C^\infty(\mathcal{M}), \Omega^\bullet(\mathcal{M}), d + d^\dagger) \tag{2}$$

is a spectral triple, where $\Omega^\bullet(\mathcal{M})$ is the Hilbert space of square integrable differential forms on $\mathcal{M}$ and $d{+}d^\dagger$ is the Hodge-Dirac operator (d the exterior derivative, $d^\dagger$ its Hodge-adjoint).
- When $(\mathcal{A}, \mathcal{H}, D)$ is a spectral triple with $\mathcal{A}$ unital commutative, then there exists a compact Riemannian manifold $\mathcal{M}$ such that $\mathcal{A} = C^\infty(\mathcal{M})$.

By adding two extra-conditions (real structure and Poincaré duality), the result is extended to spin manifolds.

Why is such a spectral characterization of manifolds interesting ? Because the properties defining a spectral triple still make sense for a noncommutative $\mathcal{A}$ [2]. A *noncommutative geometry* is thus a spectral triple where the algebra $\mathcal{A}$ is noncommutative. At the light of Gelfand duality, this is the geometrical object whose algebra of functions defined on it is non commutative. As such it cannot be a usual topological space (otherwise its algebra of continuous functions would be commutative), but it rather appears as a "space without points".

$$\begin{array}{rcl} \text{commutative spectral triple} & \rightarrow & \text{noncommutative spectral triple} \\ \updownarrow & & \downarrow \\ \text{Riemannian geometry} & & \text{non-commutative geometry} \end{array}$$

However, always at the light of Gelfand duality, it is tempting to consider the pure states of the noncommutative algebra as the equivalent of points in the noncommutative context. This is all the more appealing that the same formula that allows to retrieve the riemannian geodesic distance in Connes reconstruction theorem, also provides the space of states with a distance.

Explicitly, given a spectral triple $(\mathcal{A}, \mathcal{H}, D)$, with $\mathcal{A}$ commutative or not, one defines on its state space $\mathcal{S}(\mathcal{A})$ the *spectral distance* [3]

$$d_D(\varphi, \psi) = \sup_{a \in \mathcal{A}} \{ |\varphi(a) - \psi(a)| \,/\, \|[D, a]\| \leq 1 \} \qquad \forall \varphi, \psi \in \mathcal{S}(\mathcal{A}). \tag{3}$$

It is not difficult to check that d_D has all the properties of a distance (zero if and only if $\varphi = \psi$, $d_D(\varphi, \psi) = d_D(\psi, \varphi)$, triangle inequality), except that it may be infinite. By a slight abuse of language, we still call it *distance.* For an overview of explicit computation of this distance in various examples of commutative and noncommutative spectral triples, see [11].

2. Rieffel's remark and Wasserstein distance of order 1

Rieffel noticed in [12] that formula (3), applied to the spectral triple of a riemannian manifold (2), was nothing but the Wasserstein distance of order 1 in the theory of optimal transport, or more exactly a reformulation of Kantorovich dual of the Wasserstein distance.

To see it, let us first remind what the Wasserstein (or Monge Kantorovich distance) is. Let $\mathcal{X}$ be a locally compact Polish space, $c(x, y)$ a positive real function, the "cost". The minimal work W required to transport the probability measure μ_1 to μ_2 is

$$W(\mu_1, \mu_2) \doteq \inf_{\pi} \int_{\mathcal{X} \times \mathcal{X}} c(x, y) \, d\pi \tag{4}$$

where the infimum is over all *transportation plans*, i.e. measures π on $\mathcal{X} \times \mathcal{X}$ with marginals μ_1, μ_2. When the cost function c is a distance d, then

$$W(\mu_1, \mu_2) \doteq \inf_{\pi} \int_{\mathcal{X}\times\mathcal{X}} d(x, y)\, d\pi \tag{5}$$

is a distance (possibly infinite) on the space of probability measures on $\mathcal{X}$, called the Monge-Kantorovich (or Wasserstein) distance of order 1.

In [9], Kantorovich showed that Monge problem of minimizing the cost (eq. (4)) had an equivalent dual formulation (interpreted as maximizing a profit). Namely, $W(\mu_1, \mu_2)$ is equal to

$$W(\varphi_1, \varphi_2) = \sup_{\|f\|_{\mathrm{Lip}}\leq 1} \left(\int_{\mathcal{X}} f d\mu_1 - \int_{\mathcal{X}} f d\mu_2 \right) \tag{6}$$

where φ_1, φ_2 are the states of $C(\mathcal{X})$ defined by the measure μ_1, μ_2:

$$\varphi_i(f) = \int_{\mathcal{X}} f d\mu_i \quad \forall f \in C(\mathcal{X}),\ i = 1, 2; \tag{7}$$

and the supremum in (6) is on all the functions 1-Lipschitz with respect to the cost, that is

$$f(x, y) \leq c(x, y) \quad \forall x, y \in \mathcal{X}. \tag{8}$$

Let $\mathcal{X} = \mathcal{M}$ be a complete, connected, without boundary, Riemannian manifold. For any $\varphi, \tilde{\varphi} \in \mathcal{S}(C_0(\mathcal{M}))$,

$$W(\varphi, \tilde{\varphi}) = d_D(\varphi, \tilde{\varphi})$$

where W is the Wasserstein distance associated to the cost d_{geo}, while d_D is the spectral distance associated to $\left(C_0^\infty(\mathcal{M}), \Omega^\bullet(\mathcal{M}), D = d + d^\dagger\right)$. That Kantorovich dual (6) coincides with the spectral distance (3) then follows from the observation that the supremum in the latter can be searched equivalently on selfadjoint elements, for which one has

$$\left\| [d + d^\dagger, f]^2 \right\| = \|f\|^2_{\mathrm{Lip}}\,. \tag{9}$$

As pointed out in [7], one has to be careful that the manifold is complete,otherwise there is no guarantee that the supremum on the 1-Lipschitz functions in (6) coincides with the supremum on 1-Lipschitz functions vanishing at infinity in (3).

3. Towards a theory of optimal transport in noncommutative geometry ?

Connes spectral distance on a manifold coincides with Kantorovich dual formulation of the Wasserstein distance of order 1. It is quite natural to wonder if the same is true is the noncommutative setting. But there does not exist any "noncommutative Wasserstein distance" of whom the spectral distance would be the dual. Is it possible to build one ? More specifically, given a spectral triple $(\mathcal{A}, \mathcal{H}, D)$ with noncommutative $\mathcal{A}$, is there a Wasserstein distance W_D on $\mathcal{S}(\mathcal{A})$ such that its Kantorovich dual is the spectral distance d_D ?

$$\begin{array}{rcl}
\textbf{Commutative case:} & & \textbf{Noncommutative case:} \\
\text{Connes distance } d_D & \rightarrow & \text{Connes distance } d_D \\
\uparrow & & | \\
\text{Kantorovich duality} & & \text{Kantorovich duality ?} \\
\downarrow & & \downarrow \\
\text{Wassertein distance } W & & W_D \text{ for some noncommutative cost ?} \\
\text{with } d_D(\delta_x, \delta_y) \text{ as a cost function} & &
\end{array}$$

In the commutative case $\mathcal{A} = C^\infty(\mathcal{M})$, one retrieves the cost function as the Wasserstein distance between pure states:

$$W(\delta_x, \delta_y) = c(x, y).$$

So it is tempting to define W_D on the whole space of states $\mathcal{S}(\mathcal{A})$ as the Wasserstein distance associated with the cost c defined by the spectral distance on the space of pure states $\mathcal{P}(\mathcal{A})$, that is

$$c(\omega_1, \omega_2) := d_D(\omega_1, \omega_2) \tag{10}$$

We worked out this construction in [10], restricting to unital separable C^*-algebras, for which it is known that a state is a probability measure on $\mathcal{P}(\mathcal{A})$ [1, p.144] [1], namely to any $\varphi \in \mathcal{S}(\mathcal{A})$, there exists a (non-necessarily unique) probability measure $\mu \in \mathrm{Prob}(\mathcal{P}(\mathcal{A}))$ such that

$$\varphi(a) = \int_{\mathcal{P}(\mathcal{A})} \hat{a}(\omega)\, d\mu(\omega) \tag{11}$$

where $\hat{a}(\omega) \doteq \omega(a)$ denotes the evaluation at $\omega \in \mathcal{P}(\mathcal{A})$ of an element a of $\mathcal{A}$, viewed as a function on $\mathcal{P}(\mathcal{A})$. The Wasserstein distance on $\mathcal{S}(\mathcal{A})$ associated with the cost (10) has Kantorovich-dual formulation

$$W_D(\varphi, \tilde{\varphi}) \doteq \sup_{a \in \mathrm{Lip}_D(\mathcal{A})} \left\{ \left| \int_{\mathcal{P}(\mathcal{A})} \hat{a}(\omega)\, d\mu(\omega) - \int_{\mathcal{P}(\mathcal{A})} \hat{a}(\omega)\, d\tilde{\mu}(\omega) \right| \right\}, \tag{12}$$

where

$$\mathrm{Lip}_D(\mathcal{A}) \doteq \{a \in \mathcal{A} \text{ such that } |\omega_1(a) - \omega_2(a)| \leq d_D(\omega_1, \omega_2)\ \forall \omega_1, \omega_2 \in \mathcal{P}(\mathcal{A})\} \tag{13}$$

is the set of element of $\mathcal{A}$ that are 1-Lipschitz with respect to the cost (10).

It is not difficult to show that the Wasserstein distance provides a upper bound to the spectral distance [10, Prop. III.1],

$$d_D(\varphi, \tilde{\varphi}) \leq W_D(\varphi, \tilde{\varphi}) \quad \forall \varphi, \tilde{\varphi} \in \mathcal{S}(\mathcal{A}). \tag{14}$$

The equality holds on any subset of $\mathcal{S}(\mathcal{A})$ given by a convex linear combination of two pure states: fixed $\omega_1, \omega_2 \in \mathcal{P}(\mathcal{A})$, one denotes $\varphi_\lambda := \lambda\omega_1 + (1 - \lambda)\omega_2$. Then

$$d_D(\varphi_{\lambda_1}, \varphi_{\lambda_2}) = W_D(\varphi_{\lambda_1}, \varphi_{\lambda_2}) \quad \forall \lambda_1, \lambda_2 \in \mathbb{R}. \tag{15}$$

This shows in particular that $W_D = d_D$ on the whole of $\mathcal{S}(\mathcal{A})$ if $\mathcal{A} = M_2(\mathbb{C})$, since the pure state space of the algebra of 2×2 matrices is homeomorphic to the 2-sphere, so that the space of states is the 2-ball, and any two states φ_1, φ_2 can always be decomposed as two convex linear combinations $\varphi_{\lambda_1}, \varphi_{\lambda_2}$ of the same two pure states.

However the two distances are not equal in general, as can be seen in the following counter-example, taken from [13, §7]. Consider $\mathcal{A} = \mathbb{C}^3$ acting on $\mathcal{H} = \mathbb{C}^3$ as a diagonal matrix,

$$\pi(z_1, z_2, z_3) := \begin{pmatrix} z_1 & 0 & 0 \\ 0 & z_2 & 0 \\ 0 & 0 & z_3 \end{pmatrix} \quad \forall\, (z_1, z_2, z_3) \in \mathbb{C}^3, \tag{16}$$

[1] This may be true in general, but for safety we restrict to this well known case.

and take as Dirac operator

$$D = \begin{pmatrix} 0 & 0 & \alpha \\ 0 & 0 & \beta \\ \alpha & \beta & 0 \end{pmatrix} \qquad \alpha, \beta \in \mathbb{R}^+. \tag{17}$$

There are three pure states δ_i for $\mathcal{A}$, defined as

$$\delta_i(z_1, z_2, z_3) = z_i \quad i = 1, 2, 3. \tag{18}$$

So the space of states is the plain triangle with summit $\delta_1, \delta_2, \delta_3$.

By (15), one has that W_D coincides with d_D on each edge of the triangle. But the two distances do not agree on the whole triangle.

Proposition 3.1. *Let φ, φ' be states in $\mathcal{S}(\mathbb{C}^3)$,*

$$\varphi = \lambda_1\delta_1 + \lambda_2\delta_2 + (1 - \lambda_1 - \lambda_2)\delta_3, \qquad \varphi' = \lambda_1'\delta_1 + \lambda_2'\delta_2 + (1 - \lambda_1' - \lambda_2')\delta_3 \tag{19}$$

where $\lambda_i, \lambda_i' \in \mathbb{R}$, $i = 1, 2$ are such that $\Lambda_1 := \lambda_1 - \lambda_1'$ and $\Lambda_2 := \lambda_2 - \lambda_2'$ have the same sign. Then

$$W_D(\varphi, \varphi') = \frac{|\Lambda_1|}{\alpha} + \frac{|\Lambda_2|}{\beta} \tag{20}$$

while

$$d_D(\varphi, \varphi') = \sqrt{\frac{\Lambda_1^2}{\alpha^2} + \frac{\Lambda^2}{\beta^2}} \tag{21}$$

Proof. The cost function (10) is obtained computing explicitly the spectral distance between pure states (see e.g. [8, Prop. 7]):

$$d_D(\delta_1, \delta_2) = \sqrt{\frac{1}{\alpha^2} + \frac{1}{\beta^2}}, \qquad d_D(\delta_1, \delta_3) = \frac{1}{\alpha}, \qquad d_D(\delta_2, \delta_3) = \frac{1}{\beta}. \tag{22}$$

The Lipschitz ball (13) is thus

$$\mathrm{Lip}_D(\mathbb{C}^3) = \left\{ (z_1, z_2, z_3) \in \mathbb{C}^3 \,;\, |z_1 - z_2| \leq \sqrt{\frac{1}{\alpha^2} + \frac{1}{\beta^2}},\, |z_1 - z_3| \leq \frac{1}{\beta},\, |z_2 - z_3| \leq \frac{1}{\alpha} \right\}. \tag{23}$$

For any $a = (z_1, z_2, z_3) \in \mathbb{C}^3$, one has

$$\varphi(a) - \varphi'(a) = \Lambda_1(z_1 - z_3) + \Lambda_2(z_2 - z_3). \tag{24}$$

Therefore

$$W_D(\varphi, \varphi') \leq \frac{|\Lambda_1|}{\alpha} + \frac{|\Lambda_2|}{\beta}. \tag{25}$$

Since Λ_1 e Λ_2 have the same sign, this upper bound is attained by a_0 in $\mathrm{Lip}_D(\mathbb{C}^3)$ defined by

$$z_1 = \frac{1}{\alpha}, \quad z_2 = \frac{1}{\beta}, \quad z_3 = 0. \tag{26}$$

To prove (21), one computes the commutator of $a = (z_1, z_2, z_3)$ with D (see e.g. [8])

$$||[D, a]|| = \sqrt{\alpha^2|z_3 - z_1|^2 + \beta^2|z_3 - z_2|^2}. \tag{27}$$

By subtracting $z_1\mathbb{I}$, one can always assume that $z_3 = 0$. The commutator condition $||[D,a]|| \leq 1$ becomes $\alpha^2|z_1|^2 + \beta^2|z_2|^2 \leq 1$, which is equivalent to

$$|z_2| \leq \sqrt{\frac{1-\alpha^2|z_1|^2}{\beta^2}}. \tag{28}$$

For such a, one obtains from (24)

$$|\varphi(a) - \varphi'(a)| \leq \Lambda_1|z_1| + \Lambda_2|z_2| \tag{29}$$

$$\leq \Lambda_1|z_1| + \Lambda_2\sqrt{\frac{1-\alpha^2|z_1|^2}{\beta^2}}. \tag{30}$$

On $[1, \frac{1}{\alpha}]$, the function $f(x) = \Lambda_1 x + \Lambda_2\sqrt{\frac{1-\alpha^2x^2}{\beta^2}}$ reaches its maximum when f' vanishes, that is for

$$x_0 = \frac{\Lambda_1}{\alpha^2\sqrt{\frac{\Lambda_1^2}{\alpha^2} + \frac{\Lambda_2^2}{\beta^2}}}. \tag{31}$$

This maximum,

$$f(x_0) = \sqrt{\frac{\Lambda_1^2}{\alpha^2} + \frac{\Lambda^2}{\beta^2}}, \tag{32}$$

is an upper bound for the spectral distance, reached by the element $a = (x_0, f(x_0), 0)$. $\square$

4. Conclusion and outlook

On a manifold, Connes spectral distance between arbitrary states coincides with the Wasserstein distance of order 1 with cost the geodesic distance. On an arbitrary spectral triple, the spectral distance d_D on the space of states $\mathcal{S}(\mathcal{A})$ (viewed as the convex hull of the pure states space $\mathcal{P}(\mathcal{A})$) does not coincide with the Wasserstein distance W_D with cost function d_D (on $\mathcal{P}(\mathcal{A})$). However, W_D always provides an upper bound to the spectral distance, and the two distances coincides on any convex combination of two fixed pure states.

This shows that the interpretation of the spectral distance as a Wasserstein distance is more involved that could be initially thought, although the intriguing example worked out in this paper (W_D is the sum of the opposite and adjacent sides of a right a triangle, d_D is the length of the hypothenuse) suggests that there might exist a simple relation between the two distances.

Let us also mention that there do exists a formulation of the spectral distance as an infimum rather than a supremum (an analogue to "dual of the dual" formula of the Wasserstein distance in optimal transport), whose possible interpretation as a noncommutative cost still has to be elucidated [7].

References

[1] O. Bratteli and D. W. Robinson, *Operator algebras and quantum statistical mechanics 1*, Springer, 1987.

[2] A. Connes, *Gravity coupled with matter and the foundations of noncommutative geometry*, Commun. Math. Phys. **182** (1996), 155–176.

[3] A. Connes and J. Lott, *The metric aspect of noncommutative geometry*, Nato ASI series B Physics **295** (1992), 53–93.

[4] Alain Connes, *Noncommutative geometry*, Academic Press, 1994.

[5] Alain Connes, , *On the spectral characterization of manifolds*, J. Noncom. Geom. **7** (2013), no. 1, 1–82.

[6] Francesco D'Andrea and Pierre Martinetti, *A view on optimal transport from noncommutative geometry*, SIGMA **6** (2010), no. 057, 24 pages.

[7] Francesco D'Andrea and Pierre Martinetti, in preparation.

[8] Bruno Iochum, Thomas Krajewski, and Pierre Martinetti, *Distances in finite spaces from noncommutative geometry*, J. Geom. Phy. **31** (2001), 100–125.
[9] L. V. Kantorovich, *On the transfer of masses*, Dokl. Akad. Nauk. SSSR **37** (1942), 227–229.
[10] P. Martinetti, *Towards a Monge-Kantorovich distance in noncommutative geometry*, Zap. Nauch. Semin. POMI **411** (2013).
[11] P. Martinetti, *From Monge to Higgs: a survey of distance computation in noncommutative geometry*, Contemporary Mathematics **676** (2016), 1–46.
[10]
[12] Marc A. Rieffel, *Metrics on states from actions of compact groups*, Documenta Math. **3** (1998), 215–229.
[13] Marc A. Rieffel, *Metric on state spaces*, Documenta Math. **4** (1999), 559–600.