Astérisque

ALESSIO FIGALLI

Regularity of optimal transport maps [after Ma-Trudinger-Wang and Loeper]

Astérisque, tome 332 (2010), Séminaire Bourbaki,
exp. nº 1009, p. 341-368

<http://www.numdam.org/item?id=AST_2010__332__341_0>

Séminaire BOURBAKI Juin 2009
61^{e} année, 2008-2009, n^{o} 1009, p. 341 à 368

REGULARITY OF OPTIMAL TRANSPORT MAPS
[after Ma–Trudinger–Wang and Loeper]

by **Alessio FIGALLI**

INTRODUCTION

In the field of optimal transportation, one important issue is the regularity of the optimal transport map. There are several motivations for the investigation of the smoothness of the optimal map:

- It is a typical PDE/analysis question.
- It is a step towards a qualitative understanding of the optimal transport map.
- If it is a general phenomenon, then non-smooth situations may be treated by regularization, instead of working directly on non-smooth objects.

In the special case "cost = squared distance" on $\mathbb{R}^n$, the problem was solved by Caffarelli [**4, 5, 6, 7**], who proved the smoothness of the map under suitable assumptions on the regularity of the densities and on the geometry of their support. However, a major open problem in the theory was the question of regularity for more general cost functions, or for the case "cost = squared distance" on a Riemannian manifold. A breakthrough in this problem has been achieved by Ma, Trudinger and Wang [**27**] and Loeper [**24**], who found a necessary and sufficient condition on the cost function in order to ensure regularity. This condition, now called MTW condition, involves a combination of derivatives of the cost, up to the fourth order. In the special case "cost = squared distance" on a Riemannian manifold, the MTW condition corresponds to the non-negativity of a new curvature tensor on the manifold (the so-called MTW tensor), which implies strong geometric consequences on the geometry of the manifold and on the structure of its cut-locus.

1. THE OPTIMAL TRANSPORTATION PROBLEM

The Monge transportation problem is more than 200 years old [**29**], and it has generated a huge amount of work in the last years.

Originally Monge wanted to move, in the Euclidean space $\mathbb{R}^3$, a rubble (*déblais*) to build up a mound or fortification (*remblais*) minimizing the cost. To explain this in a simple case, suppose that the rubble consists of masses, say $m_1, \dots, m_n$, at locations $\{x_1, \dots x_n\}$, and one is interested in moving them into another set of positions $\{y_1, \dots, y_n\}$ by minimizing the weighted travelled distance. Then, one tries to minimize

$$\sum_{i=1}^{n} m_i |x_i - T(x_i)|,$$

over all bijections $T : \{x_1, \dots x_n\} \to \{y_1, \dots, y_n\}$.

Nowadays, influenced by physics and geometry, one would be more interested in minimizing the energy cost rather than the distance. Therefore, one wants to minimize

$$\sum_{i=1}^{n} m_i |x_i - T(x_i)|^2.$$

Of course, it is desirable to generalize this problem to continuous, rather than just discrete, distributions of matter. Hence, the optimal transport problem is now formulated in the following general form: given two probability measures μ and ν, defined on the measurable spaces X and Y, find a measurable map $T : X \to Y$ with $T_\sharp \mu = \nu$, i.e.

$$\nu(A) = \mu\left(T^{-1}(A)\right) \qquad \forall\, A \subset Y \text{ measurable},$$

in such a way that T minimizes the transportation cost. This means

$$\int_X c(x, T(x))\, d\mu(x) = \min_{S_\# \mu = \nu} \left\{ \int_X c(x, S(x))\, d\mu(x) \right\},$$

where $c : X \times Y \to \mathbb{R}$ is some given cost function, and the minimum is taken over all measurable maps $S : X \to Y$ such that $S_\# \mu = \nu$. When the transport condition $T_\# \mu = \nu$ is satisfied, we say that T is a *transport map*, and if T also minimizes the cost we call it an *optimal transport map.*

Even in Euclidean spaces, with the cost c equal to the Euclidean distance or its square, the problem of the existence of an optimal transport map is far from being trivial. Moreover, it is easy to build examples where the Monge problem is ill-posed simply because there is no transport map: this happens for instance when μ is a Dirac mass while ν is not. This means that one needs some restrictions on the measures μ and ν.

We further remark that, if $\mu(dx) = f(x)dx$ and $\nu(dy) = g(y)dy$, the condition $T_\# \mu = \nu$ formally gives the Jacobian equation $\left|\det(\nabla T)\right| = f/(g \circ T)$.

1.1. Existence and uniqueness of optimal maps on Riemannian manifolds

In [**1, 2**], Brenier considered the case $X = Y = \mathbb{R}^n$, $c(x,y) = |x-y|^2/2$, and he proved the following theorem (the same result was also proven independently by Cuesta-Albertos and Matrán [**10**] and by Rachev and Rüschendorf [**30**]):

THEOREM 1.1 ([**1, 2**]). — *Let μ and ν be two compactly supported probability measures on $\mathbb{R}^n$. If μ is absolutely continuous with respect to the Lebesgue measure, then:*

(i) *There exists a unique solution T to the Monge problem.*
(ii) *The optimal map T is characterized by the structure $T(x) = \nabla\phi(x)$, for some convex function $\phi : \mathbb{R}^n \to \mathbb{R}$.*

Furthermore, if $\mu(dx) = f(x)dx$ and $\nu(dy) = g(y)dy$,

$$\big|\det(\nabla T(x))\big| = \frac{f(x)}{g(T(x))} \qquad \text{for } \mu\text{-a.e. } x \in \mathbb{R}^n.$$

After this result, many researchers started to work on the problem, showing existence of optimal maps with more general costs, both in an Euclidean setting, in the case of compact (Riemannian and sub-Riemannian) manifolds, and in some particular classes on non-compact manifolds. In particular, exploiting some ideas introduced by Cabré in [**3**] for studying elliptic equations on manifolds, McCann was able to generalize Brenier's theorem to (compact) Riemannian manifolds [**28**].

REMARK. — *From now on, we will always implicitly assume that all manifolds have no boundary.*

To explain McCann's result, let us first introduce a few definitions.

We recall that a function $\varphi : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ is convex and lower semicontinuous if and only if

$$\varphi(x) = \sup_{y\in\mathbb{R}^n} \big[x\cdot y - \varphi^*(y)\big],$$

where

$$\varphi^*(x) := \sup_{x\in\mathbb{R}^n} \big[x\cdot y - \varphi(x)\big].$$

This fact is the basis for the notion of *c-convexity*, where $c : X \times Y \to \mathbb{R}$ is an arbitrary function:

DEFINITION 1.2. — *A function $\psi : X \to \mathbb{R} \cup \{+\infty\}$ is* c-convex *if*

$$\psi(x) = \sup_{y\in Y}\big[\psi^c(y) - c(x,y)\big] \qquad \forall\, x \in X,$$

where

$$\psi^c(y) := \inf_{x\in X}\big[\psi(x) + c(x,y)\big] \qquad \forall\, y \in Y.$$

*Moreover, for a c-convex function ψ, we define its c-*subdifferential *at x as*

$$\partial^c\psi(x) := \{y \in Y \mid \psi(x) = \psi^c(y) - c(x,y)\}.$$

With this general definition, when $X = Y = \mathbb{R}^n$ and $c(x,y) = -x\cdot y$, the usual convexity coincides with the c-convexity, and the usual subdifferential coincides with the c-subdifferential.

In particular, in the case $X = Y = \mathbb{R}^n$ and $c(x,y) = |x-y|^2/2$, a function ψ is c-convex if and only if $\psi(x) + \frac{|x|^2}{2}$ is convex. The following result is the generalization of Brenier's Theorem to Riemannian manifolds:

THEOREM 1.3 ([**28**]). — *Let (M,g) be a Riemannian manifold, take μ and ν two compactly supported probability measures on M, and consider the optimal transport problem from μ to ν with cost $c(x,y) = d(x,y)^2/2$, where $d(x,y)$ denotes the Riemannian distance on M. If μ is absolutely continuous with respect to the volume measure, then:*

(i) *There exists a unique solution T to the Monge problem.*
(ii) *T is characterized by the structure $T(x) = \exp_x\big(\nabla\psi(x)\big) \in \partial^c\psi(x)$ for some c-convex function $\psi : M \to \mathbb{R}$.*
(iii) *For μ_0-a.e. $x \in M$, there exists a unique minimizing geodesic from x to $T(x)$, which is given by $[0,1] \ni t \mapsto \exp_x\big(t\nabla\psi(x)\big)$.*

Furthermore, if $\mu(dx) = f(x)\mathrm{vol}(dx)$ and $\nu(dy) = g(y)\mathrm{vol}(dy)$,

$$\big|\det(\nabla T(x))\big| = \frac{f(x)}{g(T(x))} \qquad \text{for } \mu\text{-a.e. } x \in M.$$

The last formula in the above theorem needs a comment: given a function $T : M \to M$, the determinant of its Jacobian is not intrinsically defined. Indeed, in order to compute the determinant of $\nabla T(x) : T_xM \to T_{T(x)}M$, one needs to identify the tangent spaces. On the other hand, $\big|\det(\nabla T(x))\big|$ is intrinsically defined as

$$\big|\det(\nabla T(x))\big| = \lim_{r\to 0} \frac{\mathrm{vol}(T(B_r(x)))}{\mathrm{vol}(B_r(x))},$$

whenever the above limit exists.

2. THE REGULARITY ISSUE: THE EUCLIDEAN CASE

Let Ω and Ω' be two bounded smooth open sets in $\mathbb{R}^n$, and let $\mu(dx) = f(x)dx$, $\nu(y) = g(y)dy$ be two probability measures, with f and g such that $f = 0$ in $\mathbb{R}^2 \setminus \Omega$, $g = 0$ in $\mathbb{R}^2 \setminus \Omega'$. We assume that f and g are C^∞ and bounded away from zero and infinity on Ω and Ω', respectively. By Brenier's Theorem, the optimal transport map T is given by the gradient of a convex function ϕ. Hence, at least formally, the Jacobian equation for T

$$\left|\det(\nabla T(x))\right| = \frac{f(x)}{g(T(x))},$$

gives a PDE for ϕ:

$$\det(D^2\phi(x)) = \frac{f(x)}{g(\nabla\phi(x))}. \tag{1}$$

This is a Monge-Ampère equation for ϕ, which is naturally coupled with the boundary condition

$$\nabla\phi(\Omega) = \Omega' \tag{2}$$

(which corresponds to the fact that T transports $f(x)dx$ onto $g(y)dy$).

As observed by Caffarelli [**6**], even for smooth densities, one cannot expect any general regularity result for ϕ without making some geometric assumptions on the support of the target measure. Indeed, suppose that $\Omega = B_1$ is the unit ball centered at the origin, and $\Omega' = \left(B_1^+ + e_n\right) \cup \left(B_1^- - e_n\right)$ is the union of two half-balls, where $(e_i)_{i=1,\dots,n}$ denote the canonical basis of $\mathbb{R}^n$, and

$$B_1^+ := \left(B_1 \cap \{x_n > 0\}\right), \qquad B_1^- := \left(B_1 \cap \{x_n < 0\}\right).$$

Then, if $f = g = \frac{1}{|B_1|}$ on Ω and Ω' respectively, it is easily seen that the optimal map T is given by

$$T(x) := \begin{cases} x + e_n & \text{if } x_n > 0, \\ x - e_n & \text{if } x_n < 0, \end{cases}$$

which corresponds to the gradient of the convex function $\phi(x) = |x_n| + |x|^2/2$.

Thus, as one could also show by an easy topological argument, in order to hope for a regularity result for ϕ, we need at least to assume the connectedness of Ω'. But, starting from the above construction and considering a sequence of domains Ω'_ε where one adds a small strip of width $\varepsilon > 0$ to glue together $\left(B_1^+ + e_n\right) \cup \left(B_1^- - e_n\right)$, one can also show that for $\varepsilon > 0$ small enough the optimal map will still be discontinuous (see [**6**]).

As proven by Caffarelli [**6**], the right geometric condition on Ω' which allows to prevent singularities of ϕ and to show the regularity of the optimal transport map is the convexity of the target: if Ω' is convex, and f and g are C^∞ and strictly positive on their respective support, then ϕ (and hence T) is C^∞ inside Ω [**4, 5, 6**]. Moreover, if one further assumes that both Ω and Ω' are smooth and uniformly convex, then $\phi \in C^\infty(\overline{\Omega})$, and $T : \overline{\Omega} \to \overline{\Omega'}$ is a smooth diffeomorphism [**7**] (the same result has been proven independently by Urbas [**33**]).

3. THE REGULARITY ISSUE: THE RIEMANNIAN CASE

The extension of Caffarelli's regularity theory to more general cost function or to the case of the squared distance function on Riemannian manifolds was for a long time a serious issue, not clear how to attack. To keep the exposition easier, we will focus on the case of the squared distance on Riemannian manifolds, although most of the arguments are exactly the same for a more general cost function. In what follows, we will use "smooth" as a synonymous of C^∞.

3.1. A PDE approach to the regularity issue

Let (M, g) be a (smooth) compact connected Riemannian manifold, let $\mu(dx) = f(x)\mathrm{vol}(dx)$ and $\nu(dy) = g(y)\mathrm{vol}(dy)$ be probability measures on M, and consider the cost $c(x,y) = d(x,y)^2/2$. Assume f and g to be C^∞ and strictly positive on M.

As before, we start from the Jacobian equation

$$\big|\det(\nabla T(x))\big| = \frac{f(x)}{g(T(x))}$$

to formally obtain an equation for ψ. It can be shown, by standard arguments of Riemannian geometry, that the relation $T(x) = \exp_x\big(\nabla\psi(x)\big)$ is equivalent to

$$\nabla\psi(x) + \nabla_x c(x, T(x)) = 0. \tag{3}$$

Writing everything in charts, we differentiate the above identity with respect to x, and by using the Jacobian equation we get

$$\begin{aligned}\det\big(D^2\psi(x) + D_x^2 c\big(x, \exp_x\big(\nabla\psi(x)\big)\big)\big) &= \frac{f(x)\mathrm{vol}_x}{g(T(x))\mathrm{vol}_{T(x)}\big|\det(d_{\nabla\psi(x)}\exp_x)\big|} \\ &=: h(x, \nabla\psi(x)),\end{aligned} \tag{4}$$

where vol_z denotes the volume density at a point $z \in M$ computed with respect to the chart. (Because ψ is c-convex (cf. Theorem 1.3(ii)), the matrix $D^2\psi(x) + D_x^2 c\big(x, \exp_x\big(\nabla\psi(x)\big)\big)$ is non-negative.) Hence ψ solves a Monge-Ampère type equation with a perturbation term $D_x^2 c\big(x, \exp_x\big(\nabla\psi(x)\big)\big)$, which is of first order in ψ. Unfortunately, for Monge-Ampère type equations lower order terms do matter, and it turns out that it is exactly the term $D_x^2 c\big(x, \exp_x\big(\nabla\psi(x)\big)\big)$ which can create obstructions to the smoothness.

The breakthrough in this problem came with the paper of Ma, Trudinger and Wang [**27**] (whose roots lie in an earlier work of Wang on the reflector antenna problem [**35**]), where the authors found a mysterious fourth-order condition on the cost functions, which turned out to be sufficient to prove the regularity of ψ. The idea was to differentiate twice Equation (4) in order to get a linear PDE for the second derivatives of ψ, and then to try to show an a priori estimate on the second derivatives of ψ. In this computation, one ends up at a certain moment with a term which needs to have a sign in order to conclude the desired a priori estimate. This term is what is now called the Ma–Trudinger–Wang tensor (in short MTW tensor):

$$(5)\quad \mathfrak{S}_{(x,y)}(\xi,\eta) := \frac{3}{2}\sum_{ijklrs}\left(c_{ij,r}c^{r,s}c_{s,kl} - c_{ij,kl}\right)\xi^i\xi^j\eta^k\eta^l, \qquad \xi \in T_xM,\ \eta \in T_yM.$$

In the above formula the cost function is evaluated at (x,y), and we used the notation $c_j = \frac{\partial c}{\partial x^j}$, $c_{jk} = \frac{\partial^2 c}{\partial x^j \partial x^k}$, $c_{i,j} = \frac{\partial^2 c}{\partial x^i \partial y^j}$, $c^{i,j} = (c_{i,j})^{-1}$, and so on. Moreover, all the derivatives are computed by introducing a system of coordinates $(x^1,\dots,x^n)$ around x, and a system $(y^1,\dots,y^n)$ around y. (We will discuss later on the independence of this expression on the choice of the system of coordinates, see Paragraph 3.4.) The condition to impose on $\mathfrak{S}_{(x,y)}(\xi,\eta)$ is

$$\mathfrak{S}_{(x,y)}(\xi,\eta) \geq 0 \qquad \text{whenever } \sum_{ij} c_{i,j}\xi^i\eta^j = 0$$

(this is called the MTW condition). Under this hypothesis, and a geometric condition on the supports of the measures (which is the analogous of the convexity assumption of Caffarelli), Ma, Trudinger and Wang could prove the following result:

THEOREM 3.1 ([**27, 31, 32**]). — *Let (M,g) be a compact Riemannian manifold. Assume that the* MTW *condition holds, that f and g are smooth and bounded away from zero and infinity on their respective supports Ω and Ω', and that the cost function $c = d^2/2$ is smooth on the set $\overline{\Omega} \times \overline{\Omega'}$. Finally, suppose that:*

(a) *Ω and Ω' are smooth;*
(b) *$(\exp_x)^{-1}(\Omega') \subset T_xM$ is uniformly convex for all $x \in \Omega$;*
(c) *$(\exp_y)^{-1}(\Omega) \subset T_xM$ is uniformly convex for all $y \in \Omega'$.*

Then $\psi \in C^\infty(\overline{\Omega})$, and $T : \overline{\Omega} \to \overline{\Omega'}$ is a smooth diffeomorphism.

Sketch of the proof. — As we already pointed out before, the key point is to show an a priori estimate on second derivatives of smooth solutions of (4). Indeed, once such an estimate is proven, Equation (4) becomes uniformly elliptic, and standard PDE methods based on approximation allow to show the desired regularity of ψ inside Ω. (The regularity up to the boundary is more complicated, and needs a barrier argument.)

We will assume for simplicity that a stronger MTW condition holds: there exists a constant $K > 0$ such that

$$\mathfrak{S}_{(x,y)}(\xi,\eta) \geq K\,|\xi|_x^2|\eta|_x^2 \qquad \text{whenever } \sum_{ij} c_{i,j}\xi^i\eta^j = 0. \text{ (1)} \tag{6}$$

Let us start from a smooth (say C^4) solution of (4), coupled with the boundary condition $T(\Omega) = \Omega'$, where $T(x) = \exp_x(\nabla\psi(x))$. The goal is to find a universal bound for the second derivatives of ψ.

We observe that, since $T(x) = \exp_x(\nabla\psi(x))$, we have

$$|\nabla\psi(x)| = d(x, T(x)) \leq \operatorname{diam}(M).$$

Hence ψ is globally Lipschitz, with a uniform Lipschitz bound. We define

$$w_{ij} := D^2_{x^ix^j}\psi + D^2_{x^ix^j}c\big(x, \exp_x(\nabla\psi(x))\big).$$

(Recall that by the c-convexity of ψ, (w_{ij}) is non-negative, and it is actually positive definite thanks to (4), as $h > 0$.) Then (4) can be written as

$$\det(w_{ij}) = h(x, \nabla\psi(x)), \tag{7}$$

or equivalently

$$\log\big(\det(w_{ij})\big) = \varphi,$$

with $\varphi(x) := \log\big(h(x, \nabla\psi(x))\big)$. By differentiating the above equation, and using the convention of summation over repeated indices, we get

$$w^{ij}w_{ij,k} = \varphi_k,$$

$$w^{ij}w_{ij,kk} = \varphi_{kk} + w^{is}w^{jt}w_{ij,k}w_{st,k} \geq \varphi_{kk},$$

where (w^{ij}) denotes the inverse of (w_{ij}). We use the notation $\psi_k = \frac{\partial}{\partial x^k}\psi$, $w_{ij,k} = \frac{\partial}{\partial x^k}w_{ij}$, $T_{s,k} = \frac{\partial}{\partial x^k}T_s$, and so on. Then the above equations become

$$w^{ij}\big[\psi_{ijk} + c_{ijk} + c_{ij,s}T_{s,k}\big] = \varphi_k, \tag{8}$$

$$w^{ij}\big[\psi_{ijkk} + c_{ijkk} + 2\,c_{ijk,s}T_{s,k} + c_{ij,s}T_{s,kk} + c_{ij,st}T_{s,k}T_{t,k}\big] \geq \varphi_{kk}. \tag{9}$$

(1) This stronger MTW condition is actually the one originally used in [**27, 31**]. The general case (i.e. $K = 0$) is treated in [**32**], where the authors relax the stronger assumption by applying a sort of barrier method, using a function $\tilde{u}$ which satisfies

$$\sum_{ij}\Big[D_{x^ix^j}\tilde{u} + \sum_k D_{p_k}A_{ij}(x, \nabla\psi(x))\,D_{x^k}\tilde{u}\Big]\xi^i\xi^j \geq \delta\,|\xi|^2, \qquad \delta > 0,$$

with $A_{ij}(x,p) := D^2_{x^ix^j}c\big(x, \exp_x(p)\big)$.

We fix now $\bar{x} \in \Omega$, we take η a cut-off function around $\bar{x}$, and define the function $G : \Omega \times \mathbb{S}^{n-1} \to \mathbb{R}$,

$$G(x,\xi) := \eta(x)^2\, w_{\xi\xi}, \qquad w_{\xi\xi} := \sum_{ij} w_{ij}\xi^i\xi^j.$$

We want to show that G is uniformly bounded by a universal constant C, depending only on $\operatorname{dist}(\bar{x},\partial\Omega)$, n, the cost function, and the function $h(x,p)$. (Observe that $G \geq 0$, since (w_{ij}) is positive definite.) In fact, this will imply that

$$\eta(x)^2 \Big| D^2\psi(x) + D_x^2 c\big(x, \exp_x(\nabla\psi(x))\big)\Big| \leq C,$$

and since $\nabla\psi(x)$ is bounded and c is smooth, the above equation gives that $|D^2\psi|$ is locally uniformly bounded by a universal constant, which is the desired a priori estimate.

To prove the bound on G, the strategy is the following: let $x_0 \in \Omega$ and $\xi_0 \in \mathbb{S}^{n-1}$ be a point where G attains its maximum. By a rotation of coordinates, one can assume $\xi_0 = e_1$. Then at x_0 we have

$$0 = (\log G)_i = \frac{w_{11,i}}{w_{11}} + 2\,\frac{\eta_i}{\eta}, \tag{10}$$

$$(\log G)_{ij} = \frac{w_{11,ij}}{w_{11}} + 2\,\frac{\eta_{ij}}{\eta} - 6\,\frac{\eta_i\eta_j}{\eta^2}.$$

Since the above matrix is non-positive, we get

$$0 \geq w_{11}w^{ij}(\log G)_{ij} = w^{ij}w_{11,ij} + 2\,\frac{w_{11}}{\eta}w^{ij}\eta_{ij} - 6\,w_{11}w^{ij}\,\frac{\eta_i\eta_j}{\eta^2}. \tag{11}$$

We further observe that, differentiating (3), we obtain the relation

$$w_{ij} = c_{i,k}T_{k,j}. \tag{12}$$

This gives in particular $T_{k,j} = c^{k,i}w_{ij}$ (which implies $|\nabla T| \leq C\,w_{11}$), and allows to write derivatives of T in terms of that of w and c.

The idea is now to start from (11), and to combine the information coming from (8), (9), (10), (12), to end up with an inequality of the form

$$0 \geq w^{ij}\big[c^{k,\ell}c_{ij,k}c_{\ell,st} - c_{ij,st}\big]c^{s,p}c^{t,q}w_{p1}w_{q1} - C_0,$$

for some universal constant C_0. (When doing the computations, one has to remember that the derivatives of φ depend on derivatives of $\nabla\psi$, or equivalently on derivatives of T.) By a rotation of coordinates, one can further assume that (w_{ij}) is diagonal at x_0. We then obtain

$$w^{ii}\big[c^{k,\ell}c_{ii,k}c_{\ell,st} - c_{ii,st}\big]c^{s,1}c^{t,1}w_{11}w_{11} \leq C_0.$$

Up to now, the MTW condition has not been used. So, we now apply (6) to get

$$K\,w_{11}^2\sum_i w^{ii} \leq C_0. \tag{13}$$

Observe that by the arithmetic-geometric inequality and by (7)

$$\sum_{i=1}^{n} w^{ii} \geq \sum_{i=2}^{n} w_{ii} \geq \left(\prod_{i=2}^{n} w^{ii}\right)^{1/(n-1)} \geq c_0\, w_{11}^{-1/(n-1)},$$

where $c_0 := \inf_{x\in\Omega} h(x, \nabla\psi(x))^{1/(n-1)} > 0$. Hence, combining the above estimate with (13) we finally obtain

$$c_0\, K \left[w_{11}(x_0)\right]^{2-1/(n-1)} \leq C_0,$$

which proves that $G(x,\xi) \leq G(x_0,\xi_0) \leq C_1$ for all $(x,\xi) \in \Omega \times \mathbb{S}^{n-1}$, as desired. □

3.2. A geometric interpretation of the MTW condition

Although the MTW condition seemed the right assumption to obtain regularity of optimal maps, it was only after Loeper's work [**24**] that people started to have a good understanding of this condition, and a more geometric insight. The idea of Loeper was the following: for the classical Monge-Ampère equation, a key property to prove regularity of convex solutions is that the subdifferential of a convex function is convex, and so in particular connected. Roughly speaking, this has the following consequence: whenever a convex function φ is not C^1 at a point x_0, there is at least a whole segment contained in the subdifferential of φ at x_0, and this fact combined with the Monge-Ampère equation provides a contradiction. (See also Theorem 3.6 below.) Hence, Loeper wanted to understand whether the c-subdifferential of a c-convex function is at least connected, believing that this fact had a link with the regularity. To explain all this in details, let us introduce some definitions.

Let $\varphi : \mathbb{R}^n \to \mathbb{R}$ be a convex function; its *subdifferential* $\partial\varphi(x)$ is given by

$$\begin{aligned}\partial\varphi(x) &= \left\{y \in \mathbb{R}^n \mid \varphi(x) + \varphi^*(y) = x\cdot y\right\}\\ &= \left\{y \in \mathbb{R}^n \mid \varphi(z) - z\cdot y \geq \varphi(x) - x\cdot y \quad \forall\, z \in \mathbb{R}^n\right\}.\end{aligned}$$

Then $\partial\varphi(x)$ is a convex set, a fortiori connected. More in general, given a semiconvex function $\phi : \mathbb{R}^n \to \mathbb{R}$ (i.e. ϕ can be locally written as the sum of a convex and a smooth function), its *subgradient* $\nabla^-\phi(\bar{x})$ is defined as

$$\nabla^-\phi(\bar{x}) := \left\{p \mid \phi(\bar{x}+v) \geq \phi(\bar{x}) + \langle p, v\rangle + o(|v|) \quad \forall\, v\right\}.$$

We remark that, by working in charts, this definition makes sense also for functions ϕ defined on manifolds.

If we now consider $\psi : M \to \mathbb{R}$ a c-convex function, $c = d^2/2$, then

$$\begin{aligned}\partial^c\psi(x) &= \left\{y \in M \mid \psi(x) = \psi^c(y) - c(x,y)\right\}\\ &= \left\{y \in M \mid \psi(z) + c(z,y) \geq \psi(x) + c(x,y) \quad \forall\, z \in M\right\}\end{aligned}$$

(see Definition 1.2). In this generality there is no reason for $\partial^c\psi(x)$ to be connected, and in fact in general this is not the case!

Conditions for the connectedness of $\partial^c\psi$. — We now wish to find some simple enough conditions implying the connectedness of sets $\partial^c\psi$. In all the following arguments, we will assume for simplicity that points $(x, y) \in M \times M$ vary in a compact subset where the cost function $c = d^2/2$ is smooth. In particular it is well know that, under this assumption, for any pair (x, y) there exists a unique minimizing geodesic $\gamma_{x,y}$ joining them, which is given by $[0, 1] \ni t \mapsto \exp_x(tv_{x,y})$, for some vector $v_{x,y} \in T_xM$. (See also Paragraph 3.5.1 below.) We will use the notation $(\exp_x)^{-1}(y) := v_{x,y}$.

First attempt to the connectedness. — Let us look first at the simplest c-convex functions:

$$\psi(x) := -c(x, y_0) + a_0.$$

Let $\bar{y} \in \partial^c\psi(\bar{x})$. Then the function $\psi(x) + c(x, \bar{y})$ achieves its minimum at $x = \bar{x}$, so that

$$-\nabla_x c(\bar{x}, y_0) + \nabla_x c(\bar{x}, \bar{y}) = 0.$$

This implies $(\exp_{\bar{x}})^{-1}(y_0) = (\exp_{\bar{x}})^{-1}(\bar{y})$, which gives $\bar{y} = y_0$. In conclusion $\partial^c\psi(\bar{x}) = \{y_0\}$ is a singleton, automatically connected, and so we do not get any information!

Second attempt to the connectedness. — The second simplest example of c-convex functions are

$$\psi(x) := \max\bigl\{-c(x, y_0) + a_0, -c(x, y_1) + a_1\bigr\}.$$

Take a point $\bar{x} \in \{x \mid -c(x, y_0) + a_0 = -c(x, y_1) + a_1\}$, and let $\bar{y} \in \partial^c\psi(\bar{x})$. Since $\psi(x) + c(x, \bar{y})$ attains its minimum at $x = \bar{x}$, we get

$$0 \in \nabla^-_{\bar{x}}\bigl(\psi + c(\cdot, \bar{y})\bigr),$$

or equivalently

$$-\nabla_x c(\bar{x}, \bar{y}) \in \nabla^-\psi(\bar{x}).$$

From the above inclusion, one can easily deduce that $\bar{y} \in \exp_{\bar{x}}\bigl(\nabla^-\psi(\bar{x})\bigr)$. Moreover, it is not difficult to see that

$$\nabla^-\psi(\bar{x}) = \{(1-t)v_0 + tv_1 \mid t \in [0, 1]\}, \qquad v_i := \nabla_x c(\bar{x}, y_i) = (\exp_{\bar{x}})^{-1}(y_i), \quad i = 0, 1.$$

Therefore, denoting by $[v_0, v_1]$ the segment joining v_0 and v_1, we obtain

$$\partial^c\psi(\bar{x}) \subset \exp_{\bar{x}}\bigl([v_0, v_1]\bigr).$$

The above formula suggests the following definition:

Definition 3.2. — *Let $\bar{x} \in M$, $y_0, y_1 \notin \mathrm{cut}(\bar{x})$. Then we define the c*-segment *from y_0 to y_1 with base $\bar{x}$ as*

$$[y_0, y_1]_{\bar{x}} := \bigl\{y_t = \exp_{\bar{x}}\bigl((1-t)(\exp_{\bar{x}})^{-1}(y_0) + t(\exp_{\bar{x}})^{-1}(y_1)\bigr) \mid t \in [0, 1]\bigr\}.$$

By slightly modifying some of the arguments in [**27**], Loeper showed that, under adequate assumptions, the connectedness of the c-subdifferential is a necessary condition for the smoothness of optimal transport (see also [**34**, Theorem 12.7]):

THEOREM 3.3 ([**24**]). — *Assume that there exist $\bar{x} \in M$ and $\psi : M \to \mathbb{R}$ c-convex such that $\partial^c\psi(\bar{x})$ is not (simply) connected. Then one can construct two probability densities f and g, C^∞ and strictly positive on M, such that the optimal map is discontinuous.*

While the above result was essentially contained in [**27**], Loeper's major contribution was to link the connectedness of the c-subdifferential to a differential condition on the cost function, which actually coincides with the MTW condition (see Paragraph 3.3). He proved (a slightly weaker version of) the following result, still assuming that the points (x, y) vary in a compact set where the cost function is smooth (see [**34**, Chapter 12] for a more general statement):

THEOREM 3.4 ([**24**]). — *The following conditions are equivalent:*

(i) *For any ψ c-convex, for all $\bar{x} \in M$, $\partial^c\psi(\bar{x})$ is connected.*

(ii) *For any ψ c-convex, for all $\bar{x} \in M$, $(\exp_{\bar{x}})^{-1}\big(\partial^c\psi(\bar{x})\big)$ is convex, and it coincides with $\nabla^-\psi(\bar{x})$.*

(iii) *For all $\bar{x} \in M$, for all y_0, y_1, if $[y_0, y_1]_{\bar{x}} = (y_t)_{t\in[0,1]}$, then*

$$d(x, y_t)^2 - d(\bar{x}, y_t)^2 \geq \min\big[d(x, y_0)^2 - d(\bar{x}, y_0)^2, d(x, y_1)^2 - d(\bar{x}, y_1)^2\big] \tag{14}$$

for all $x \in M$, $t \in [0, 1]$.

(iv) *For all $\bar{x}, y \in M$, for all $\eta, \xi \in T_{\bar{x}}M$ with $\xi \perp \eta$,*

$$\frac{d^2}{ds^2}\bigg|_{s=0} \frac{d^2}{dt^2}\bigg|_{t=0} d\big(\exp_{\bar{x}}(t\xi), \exp_{\bar{x}}(p + s\eta)\big)^2 \leq 0,$$

where $p = (\exp_{\bar{x}})^{-1}(y)$.

Moreover, if any of these conditions is not satisfied, C^1 c-convex functions are not dense in Lipschitz c-convex functions.

Sketch of the proof. — We give here only some elements of the proof.

(ii) $\Rightarrow$ (i): since $(\exp_{\bar{x}})^{-1}\big(\partial^c\psi(\bar{x})\big)$ is convex, it is connected, and so its image by $\exp_{\bar{x}}$ is connected too.

(i) $\Rightarrow$ (ii): for $\psi_{\bar{x},y_0,y_1} := \max\big\{-c(\cdot, y_0) + c(\bar{x}, y_0), -c(\cdot, y_1) + c(\bar{x}, y_1)\big\}$ we have $(\exp_{\bar{x}})^{-1}\big(\partial^c\psi_{\bar{x},y_0,y_1}(\bar{x})\big) \subset [(\exp_{\bar{x}})^{-1}(y_0), (\exp_{\bar{x}})^{-1}(y_1)]$, which is a segment. Since in this case connectedness is equivalent to convexity, if (i) holds we obtain $\partial^c\psi_{\bar{x},y_0,y_1}(\bar{x}) = [y_0, y_1]_{\bar{x}}$, and $\partial^c\psi_{\bar{x},y_0,y_1}(\bar{x}) = \exp_{\bar{x}}\big(\nabla^-\psi_{\bar{x},y_0,y_1}(\bar{x})\big)$

In the general case, we fix $y_0, y_1 \in \partial^c\psi(\bar{x})$. Then it is simple to see that

$$\partial^c\psi(\bar{x}) \supset \partial^c\psi_{\bar{x},y_0,y_1}(\bar{x}) = [y_0, y_1]_{\bar{x}},$$

and the result follows easily.

(ii) $\Leftrightarrow$ (iii): condition (14) is equivalent to $\partial^c\psi_{\bar{x},y_0,y_1} = [y_0, y_1]_{\bar{x}}$. Then the equivalence between (ii) and (iii) follows arguing as above.

(iii) $\Rightarrow$ (iv): fix $\bar{x} \in M$, and let $y := \exp_{\bar{x}}(p)$. Take ξ, η orthogonal and with unit norm, and define

$$y_0 := \exp_{\bar{x}}(p - \varepsilon\eta), \quad y_1 := \exp_{\bar{x}}(p + \varepsilon\eta) \qquad \text{for some } \varepsilon > 0 \text{ small.}$$

Moreover, let

$$h_0(x) := c(\bar{x}, y_0) - c(x, y_0), \;\; h_1(x) := c(\bar{x}, y_1) - c(x, y_1), \;\; \psi := \max\{h_0, h_1\} = \psi_{\bar{x},y_0,y_1}.$$

We now define $\gamma(t)$ as a curve contained in the set $\{h_0 = h_1\}$ such that $\gamma(0) = \bar{x}$, $\dot\gamma(0) = \xi$.

Since $y \in [y_0, y_1]_{\bar{x}}$, by (iii) we get $y \in \partial^c\psi(\bar{x})$, so that

$$\begin{aligned}\frac{1}{2}\big[h_0(\bar{x}) + h_1(\bar{x})\big] + c(\bar{x}, y) &= \psi(\bar{x}) + c(\bar{x}, y) \le \psi(\gamma(t)) + c(\gamma(t), y)\\ &= \frac{1}{2}\big[h_0(\gamma(t)) + h_1(\gamma(t))\big] + c(\gamma(t), y),\end{aligned}$$

where we used that $h_0 = h_1$ along γ. Recalling the definition of h_0 and h_1, we deduce

$$\frac{1}{2}\big[c(\gamma(t), y_0) + c(\gamma(t), y_1)\big] - c(\gamma(t), y) \le \frac{1}{2}\big[c(\bar{x}, y_0) + c(\bar{x}, y_1)\big] - c(\bar{x}, y),$$

so the function $t \mapsto \frac{1}{2}\big[c(\gamma(t), y_0) + c(\gamma(t), y_1)\big] - c(\gamma(t), y)$ achieves its maximum at $t = 0$. This implies

$$\frac{d^2}{dt^2}\bigg|_{t=0} \left[\frac{1}{2}\big(c(\gamma(t), y_0) + c(\gamma(t), y_1)\big) - c(\gamma(t), y)\right] \le 0,$$

i.e.

$$\left\langle \left[\frac{1}{2}\big(D_x^2 c(\bar{x}, y_0) + D_x^2 c(\bar{x}, y_1)\big) - D_x^2 c(\bar{x}, y)\right] \cdot \xi, \xi \right\rangle \le 0$$

(here we used that $\nabla_x c(\bar{x}, y) = \frac{1}{2}\big[\nabla_x c(\bar{x}, y_0) + \nabla_x c(\bar{x}, y_1)\big]$). Thus the function

$$\eta \mapsto \big\langle D_x^2 c\big(\bar{x}, \exp_{\bar{x}}(p + \eta)\big) \cdot \xi, \xi \big\rangle$$

is concave, and proves (iv). $\square$

The above theorem leads to the definition of the *regularity* property:

DEFINITION 3.5. — *The cost function $c = d^2/2$ is said to be* regular *if the properties listed in Theorem 3.4 are satisfied.*

To understand why the above properties are related to smoothness, consider Theorem 3.4(iii). It says that, if we take the function

$$\psi_{\bar{x},y_0,y_1} = \max\bigl\{-c(\cdot,y_0)+c(\bar{x},y_0),-c(\cdot,y_1)+c(\bar{x},y_1)\bigr\},$$

then we are able to touch the graph of this function from below at $\bar{x}$ with the family of functions $\{-c(\cdot,y_t)+c(\bar{x},y_t)\}_{t\in[0,1]}$. This suggests that we could use this family to regularize the cusp of $\psi_{\bar{x},y_0,y_1}$ at the point $\bar{x}$, by slightly moving above the graphs of the functions $-c(\cdot,y_t)+c(\bar{x},y_t)$. On the other hand, if (14) does not hold, it is not clear how to regularize the cusp preserving the condition of being c-convex.

By what we said above, the regularity property seems mandatory to develop a theory of smoothness of optimal transport. Indeed, if it is not satisfied, we can construct C^∞ strictly positive densities f, g such that the optimal map is not continuous. Hence the natural question is when it is satisfied, and what is the link with the MTW condition.

3.3. A unified point of view

As we have seen in Theorem 3.4, the regularity of $c = d^2/2$ is equivalent to

$$\frac{d^2}{ds^2}\bigg|_{s=0}\frac{d^2}{dt^2}\bigg|_{t=0} c\bigl(\exp_x(t\xi),\exp_x(p+s\eta)\bigr)\le 0, \tag{15}$$

for all $p,\xi,\eta\in T_xM$, with ξ and η orthogonal, $p=(\exp_x)^{-1}(y)$.

By introducing a local system of coordinates $(x^1,\dots,x^n)$ around x, and a system $(y^1,\dots,y^n)$ around y, it is not difficult to check (by some standard but tedious computations) that the above expression *coincides* up to the sign with the MTW tensor. Hence the MTW condition is *equivalent* to the connectedness of the c-subdifferential of a c-convex function, and by Theorems 3.1 and 3.3, it is a necessary and sufficient condition for the smoothness of the optimal transport map. (At least, as long as the cost function is smooth on the supports of the two densities!)

By exploiting (a variant of) Theorem 3.4, Loeper proved the following regularity result:

THEOREM 3.6 ([**24**]). — *Let (M,g) be a compact Riemannian manifold, and let Ω and Ω' denote the support of f and g respectively. Assume that* (6) *holds for some $K>0$, f is bounded from above on Ω, g is bounded away from zero on Ω', and the cost function $c=d^2/2$ is smooth on the set $\overline{\Omega}\times\overline{\Omega'}$. Finally, suppose that $(\exp_x)^{-1}(\Omega')\subset T_xM$ is convex for any $x\in\Omega$. Then $\psi\in C^{1,\alpha}(\overline{\Omega})$, with $\alpha=1/(4n-1)$, so that $T\in C^{0,\alpha}(\overline{\Omega},\overline{\Omega'})$.*

A remarkable fact of the above result is that the Hölder exponent found by Loeper is explicit. (For instance, for the classical Monge-Ampère equation one can prove $C^{1,\alpha}$ regularity of solutions under weak assumptions on the densities [**4, 5, 6**], but there is no explicit lower bound on the exponent.) As shown recently by Liu [**23**], the optimal exponent in the above theorem is $\alpha = 1/(2n-1)$.

Sketch of the proof. — We will just prove C^1 regularity of ψ. A very similar argument, with a slightly more refined analysis, gives the $C^{1,\alpha}$-estimate.

Assume that ψ is not C^1 at x_0. Being a c-convex function, ψ is defined as a supremum of smooth functions of the form $x \mapsto -c(x,y) + \psi^c(y)$, and in particular is semiconvex. Hence it is not difficult to see that being not differentiable at x_0 means that there exist two points $y_0, y_1 \in \partial^c\psi(x_0) \cap \Omega'$. Let now $(y_t)_{t\in[0,1]} = [y_0, y_1]_{x_0} \subset \Omega'$ be the c-segment from y_0 to y_1 with base x_0. (Here we are using the assumption that $\exp_{x_0}^{-1}(\Omega')$ is convex.) Thanks to the (stronger) MTW condition (6), one can prove an improved version of (14) (see also (20) below): writing everything in charts,

$$\psi(x) + c(x, y_t) \geq \psi(x_0) + c(x_0, y_t) + \delta_0\, |y_1 - y_0|^2\, |x - x_0|^2 + O(|x - x_0|^3)$$

for all $t \in [1/4, 3/4]$. The idea is now the following: let y belong to a ε-neighborhood of the curve $(y_t)_{t\in[1/4,3/4]}$, and consider the function $f_a := -c(\cdot, y) + a$ with $a \in \mathbb{R}$. If a is sufficiently negative, then this function is below ψ in the closed ball $\overline{B}_{c_0\varepsilon}(x_0)$, where $c_0 > 0$ has to be chosen. Now, let a increase until f_a touches ψ from below inside $\overline{B}_{c_0\varepsilon}(x_0)$. Thanks to the above inequality, if c_0 is chosen sufficiently large (but fixed once for all, independently of ε), then for all $\varepsilon > 0$ sufficiently small the contact point will belong to the open ball $B_{c_0\varepsilon}(x_0)$. By this fact and Theorem 3.4(ii), we easily obtain that

$$\partial^c\psi\big(B_{c_0\varepsilon}(x_0)\big) \supset N_\varepsilon\big((y_t)_{t\in[1/4,3/4]}\big),$$

where N_ε denotes the ε-neighborhood. In terms of the optimal transport problem, this means that any point y belonging to $N_\varepsilon\big((y_t)_{t\in[1/4,3/4]}\big)$ is the image through the optimal transport map T of a point in the ball $B_{c_0\varepsilon}(x_0)$ (see Theorem 1.3(ii)). By the transport condition $T_\#(f\mathrm{vol}) = g\mathrm{vol}$, this implies

$$\int_{B_{c_0\varepsilon}} f\, d\mathrm{vol} \geq \int_{N_\varepsilon\left((y_t)_{t\in[1/4,3/4]}\right)} g\, d\mathrm{vol}.$$

However, by the assumptions on f and g we have

$$\int_{N_\varepsilon\left((y_t)_{t\in[1/4,3/4]}\right)} g\, d\mathrm{vol} \gtrsim \mathrm{vol}\big(N_\varepsilon\big((y_t)_{t\in[1/4,3/4]}\big)\big) \sim \varepsilon^{n-1},$$

$$\int_{B_{c_0\varepsilon}} f\, d\mathrm{vol} \lesssim \varepsilon^n,$$

and all these conditions are not compatible if $\varepsilon > 0$ is sufficiently small. This contradiction proves the C^1 regularity of ψ. □

Let us now consider the following geometric example.

Absence of regularity in presence of negative sectional curvature. — We want to show how negative sectional curvature is an obstruction to regularity (indeed even to continuity) of optimal maps. We refer to [**34**, Theorem 12.4] for more details on the construction given below.

Let $M = \mathbb{H}^2$ be the hyperbolic plane (or a compact quotient thereof). Fix a point O as the origin, and fix a local system of coordinates in a neighborhood of O such that the maps $(x_1, x_2) \mapsto (\pm x_1, \pm x_2)$ are local isometries (it suffices for instance to consider the model of the Poincaré disk, with O equal to the origin in $\mathbb{R}^2$). Then define the points

$$A^{\pm} = (0, \pm\varepsilon), \quad B^{\pm} = (\pm\varepsilon, 0) \qquad \text{for some } \varepsilon > 0.$$

Take a measure μ symmetric with respect to 0 and concentrated near $\{A^+\} \cup \{A^-\}$ (say 3/4 of the total mass belongs to a small neighborhood of $\{A^+\} \cup \{A^-\}$), and a measure ν symmetric with respect to 0 and concentrated near $\{B^+\} \cup \{B^-\}$. Moreover assume that μ and ν are absolutely continuous, and have strictly positive densities everywhere. We denote by T the unique optimal transport map, and we assume by contradiction that T is continuous. By symmetry, we deduce that $T(O) = O$. Then, by counting the total mass, there exists a point A' close to A^+ which is sent to a point B' near, say, B^+.

But, by negative curvature (if A' and B' are close enough to A and B respectively), Pythagoras Theorem becomes an inequality: $d(O, A')^2 + d(O, B')^2 < d(A', B')^2$, and this contradicts the optimality of the transport map, as transporting A' onto O and O onto B' would be more convenient than transporting A' onto B' and letting O stay at rest.

Now, the natural question is: how does the above example fit into Ma, Trudinger and Wang's and Loeper's results? The answer is actually pretty simple: in [**24**] Loeper noticed that the MTW tensor satisfies the following remarkable identity:

$$\mathfrak{S}_{(x,x)}(\xi, \eta) = -\frac{3}{4} \frac{d^2}{ds^2}\bigg|_{s=0} \frac{d^2}{dt^2}\bigg|_{t=0} d\big(\exp_x(t\xi), \exp_x(s\eta)\big)^2 = \mathrm{Sect}_x([\xi, \eta]), \tag{16}$$

where $\xi, \eta \in T_xM$ are two orthogonal unit vectors, and $\mathrm{Sect}_x([\xi, \eta])$ denotes the sectional curvature of the plane generated by ξ and η.

In fact, as shown by Kim and McCann [**20**], $\mathfrak{S}$ is the sectional curvature of the manifold $M \times M$, endowed with the pseudo-metric $-d^2_{xy}c$. Combining (16) with Theorems 3.3 and 3.4, we get the following important negative result:

Theorem 3.7. — *Let (M, g) be a (compact) Riemannian manifold, and assume that there exist $x \in M$ and a plane $P \subset T_xM$ such that $\mathrm{Sect}_x(P) < 0$. Then there exist C^∞ strictly positive probability densities f and g such that the optimal map is discontinuous.*

After this negative result, one could still hope to develop a regularity theory on any manifold with non-negative sectional curvature. But this is not the case: as shown by Kim [**19**], the regularity condition is strictly stronger than the condition of non-negativity of sectional curvatures. In conclusion, except for some special cases (see Paragraphs 3.4 and 3.5.2 below), the optimal map is non-smooth!

3.4. More on the MTW condition

As shown above, the MTW tensor is a non-local version of the sectional curvature, and the MTW condition is a stronger condition than non-negative sectional curvature. We further remark that the MTW condition is intrinsic, and independent of the system of coordinates.

To see this, we first show that (5) can also be written as

$$\frac{\partial^2}{\partial p_\eta^2}\frac{\partial^2}{\partial x_\xi^2}\, c(x,y) \leq 0. \tag{17}$$

The meaning of the left-hand side in (17) is the following: first freeze y and differentiate $c(x,y)$ twice with respect to x in the direction $\xi \in T_xM$. Then, considering the result as a function of y, parameterize y by $p = -\nabla_x c(x,y)$, and differentiate twice with respect to p in the direction $\eta \in T_yM$. By the relation $p_i = -c_i(x,y)$ we get $\frac{\partial p_i}{\partial y^j} = -c_{i,j}$, which gives $\frac{\partial y^k}{\partial p_\ell} = -c^{k,\ell}$. Finally, using $-c_{i,j}$ and $-c^{i,j}$ to raise and lower indices ($\eta^k = -c^{k,l}\eta_i$, etc.), it is just a (tedious) exercise to show that the expression in (17) is equal to

$$\sum_{ijklrs} \left(c_{ij,kl} - c_{ij,r}c^{r,s}c_{s,kl}\right)\xi^i\xi^j\eta^k\eta^l,$$

where we used the formula $d(M^{-1})\cdot H = -M^{-1}HM^{-1}$.

Since the expression in (17) involves second derivatives (which are not intrinsic and depend on the choice of the coordinates), it is not a priori clear whether $\mathfrak{S}$ depends or not on the choice of coordinates. On the other hand, we can hope it does not, because of the (intrinsic) geometric interpretation of the regularity.

To see that $\mathfrak{S}$ is indeed independent of the choice of coordinates (so that one does not even need to use geodesic coordinates, as in (15)), we observe that, if we do a change of coordinates and compute first the second derivatives in x, we get some additional terms of the form

$$\Gamma_{ij}^k(x)c_k(x,y) = -\Gamma_{ij}^k(x)p_k(x,y) = \Gamma_{ij}^k(x)g_{k\ell}(x)p^\ell(x,y).$$

But when we differentiate twice with respect to p, this additional term disappears! This shows that the MTW tensor is *independent* of the system of coordinates.

Let us introduce the following:

DEFINITION 3.8. — *Given $K \geq 0$, we say that (M, g) satisfies the* MTW(K) *condition if, for all $(x, y) \in (M \times M)$ where $c = d^2/2$ is smooth, for all $\xi \in T_xM$, $\eta \in T_yM$,*

$$\mathfrak{S}_{(x,y)}(\xi, \eta) \geq K|\xi|_x^2|\tilde{\eta}|_x^2 \qquad \text{whenever } -c_{i,j}(x,y)\xi^i\eta^j = 0,$$

where $\tilde{\eta}^i = -g^{i,k}(x)c_{k,j}(x,y)\eta^j \in T_xM$.

Remark 3.9. — Observe that, thanks to (16), the MTW(K) condition implies in particular that all sectional curvatures are bounded from below by K. Therefore, if $K > 0$, by the Bonnet-Myers Theorem the diameter of the manifold is bounded, and the manifold is compact.

Some examples of manifolds satisfying the MTW condition are given in [**12, 20, 21, 24, 25**]:

- $\mathbb{R}^n$ and $\mathbb{T}^n$ satisfy MTW(0).
- $\mathbb{S}^n$, its quotients (like $\mathbb{RP}^n$), and its submersions (like $\mathbb{CP}^n$ or $\mathbb{HP}^n$), satisfy MTW(1).
- Products of any of the examples listed above (for instance, $\mathbb{S}^{n_1} \times \cdots \times \mathbb{S}^{n_k} \times \mathbb{R}^\ell$ or $\mathbb{S}^{n_1} \times \mathbb{CP}^{n_2} \times \mathbb{T}^{n_3}$) satisfy MTW(0).

We observe that the MTW condition is a non-standard curvature condition, as it is fourth order and nonlocal. Therefore an important open problem is whether this condition is stable under perturbation. More precisely, we ask for the following:

QUESTION. — *Assume that (M, g) satisfies the* MTW(K) *condition for $K > 0$, and let g_ε be a C^4-perturbation of g. Does (M, g_ε) satisfy the* MTW(K') *condition for some $K' > 0$?*

The answer is easily seen to be affirmative for manifolds with nonfocal cut-locus like the projective space $\mathbb{RP}^n$ (see [**12, 26**], and Theorem 3.12 below). Moreover, as proven by Figalli and Rifford [**12**], the answer is affirmative also for the 2-dimensional sphere $\mathbb{S}^2$ (see Theorem 3.14 below). The extension of this last result to arbitrary dimension has been recently achieved by Figalli, Rifford and Villani [**13**].

As a corollary of these facts, one can prove that regularity of optimal maps holds in all of these cases. In the next paragraph, we will explain the link between the MTW condition and the geometry of the cut-locus, and we will describe more in detail the aforementioned results.

3.5. Relation between the MTW condition and the geometry of the cut-locus

3.5.1. *The cut and focal-locus.* — We recall that, given a tangent vector $v \in T_xM$, the curve $(\exp_x(tv))_{t\geq 0}$ is a geodesic defined for all times, but in general is not minimizing for large times. On the other hand, it is possible to prove that $\exp_x(tv)$ is always minimizing between x and $\exp_x(\varepsilon v)$ for $\varepsilon > 0$ sufficiently small. We define the *cut-time* $t_C(x,v)$ as

$$t_C(x,v) := \inf\Big\{t > 0 \,|\, s \mapsto \exp_x(sv) \text{ is not minimizing between } x \text{ and } \exp_x(tv)\Big\}.$$

Given two points $x, y \in M$, whenever there exists a unique minimizing geodesic $(\exp_x(tv))_{0\leq t\leq 1}$ going from x to y in time 1, we write $(\exp_x)^{-1}(y) := v$.

Given $x \in M$, we define the *cut-locus* of x as

$$\mathrm{cut}(x) := \Big\{\exp_x\big(t_C(x,v)v\big) \,|\, v \in T_xM,\ |v|_x = 1\Big\}.$$

We further define

$$\mathrm{cut}(M) := \{(x,y) \in M \times M \ |\ y \in \mathrm{cut}(x)\}.$$

Example. — *On the sphere $\mathbb{S}^n$, the geodesics starting from a point x with unit speed describe great circles passing through its antipodal point $-x$. These geodesics are minimizing exactly until they reach $-x$ after a time π. Thus $t_C(x,v) = \pi$ for any $v \in T_xM$ with unit norm, and* $\mathrm{cut}(x) = \{-x\}$. *By time-rescaling, we get $t_C(x,v) = \frac{\pi}{|v|_x}$ for any $x \in \mathbb{S}^n$, $v \in T_xM \setminus \{0\}$.*

It is possible to prove that, if $y \notin \mathrm{cut}(x)$, then x and y are joined by a unique minimizing geodesic. The converse is close to be true: $y \notin \mathrm{cut}(x)$ if and only if there are neighborhoods U of x and V of y such that any two points $x' \in U$, $y' \in V$ are joined by a unique minimizing geodesic. In particular $y \notin \mathrm{cut}(x)$ if and only if $x \notin \mathrm{cut}(y)$.

Given now $x \in M$ and $v \in T_xM$, we define the *focal-time* $t_F(x,v)$ as

$$t_F(x,v) := \inf\Big\{t > 0 \,|\, d_{tv}\exp_x : T_xM \to T_{\exp_x(tv)}M \text{ is not invertible}\Big\}.$$

We further introduce the *tangent cut-locus* of x

$$\mathrm{TCL}(x) = \big\{t_C(x,v)v \,|\, v \in T_xM,\ |v|_x = 1\big\},$$

the *tangent focal-locus* of x

$$\mathrm{TFL}(x) = \big\{t_F(x,v)v \,|\, v \in T_xM,\ |v|_x = 1\big\},$$

the *injectivity domain* of the exponential map at x

$$\mathrm{I}(x) = \big\{tv \,|\, 0 \leq t < t_C(x,v),\ v \in T_xM,\ |v|_x = 1\big\},$$

and the *nonfocal domain* of the exponential map at x

$$\mathrm{NF}(x) = \big\{tv \,|\, 0 \le t < t_F(x,v),\ v \in T_xM,\ |v|_x = 1\big\}.$$

With these definitions, we have

$$\mathrm{cut}(x) = \exp_x\big(\mathrm{TCL}(x)\big), \qquad \mathrm{TCL}(x) = \partial\big(\mathrm{I}(x)\big), \qquad \mathrm{TFL}(x) = \partial\big(\mathrm{NF}(x)\big).$$

We finally define the *focal cut-locus* of x as

$$\mathrm{fcut}(x) := \exp_x\big(\mathrm{TCL}(x) \cap \mathrm{TFL}(x)\big).$$

It is a well-known fact of Riemannian geometry that $t_C \le t_F$ (see for instance [**17**, Corollary 3.77]). In the case of the sphere, $t_C \equiv t_F$, and $\mathrm{cut}(x) = \mathrm{fcut}(x)$ for all $x \in \mathbb{S}^n$.

The fact that a point $y \in M$ belongs to $\mathrm{cut}(x)$ is a phenomenon which is captured by the regularity of the distance function. Indeed, it can be proven that the following hold (see for instance [**9**, Proposition 2.5]):

(a) The function $d(x,\cdot)^2$ is smooth (i.e. C^∞) in a neighborhood of y if and only if $y \notin \mathrm{cut}(x)$.

(b.1) The function $d(x,\cdot)^2$ has an upward cusp at y if and only if $y \in \mathrm{cut}(x)$ and there are at least two minimizing geodesics between x and y.

(b.2) The function $d(x,\cdot)^2$ is C^1 at y and its Hessian has an eigenvalue $-\infty$ if and only if $y \in \mathrm{cut}(x)$ and there is a unique minimizing geodesic between x and y. (In this case, y necessarily belongs to $\mathrm{fcut}(x)$.)

In the above statement, having an "upward cusp" means that there exist two vectors $p_1 \neq p_2$ both belonging to the supergradient of $f := d(x,\cdot)^2$ at y: writing everything in charts, we have

$$\{p_1,p_2\} \subset \nabla^+ f(y) := \big\{p \,|\, f(y+v) \le f(y) + \langle p, v\rangle + o(|v|) \quad \forall\, v\big\},$$

that is f is locally below the function $v \mapsto f(y) + \min\{\langle p_1,v\rangle, \langle p_2,v\rangle\} + o(|v|)$ near y. Hence (b.1) corresponds to roughly say that the second derivative (along the direction $p_2 - p_1$) of $d(x,\cdot)^2$ at y is $-\infty$. (The fact that there is an upward cusp, means that one of the second directional derivatives is a negative delta measure!)

Furthermore, saying that "Hessian has an eigenvalue $-\infty$" means that (always working in charts)

$$\liminf_{|v|\to 0} \frac{f(y+v) - 2f(y) + f(y-v)}{|v|^2} = -\infty.$$

Thus, all the above description of the cut-locus in terms of the squared distance can be informally summarized as follows:

$$y \in \mathrm{cut}(x) \quad \Leftrightarrow \quad \big\langle D_y^2 d^2(x,y)\cdot v, v\big\rangle = -\infty \quad \text{for some } v \in T_yM. \tag{18}$$

This observation will be of key importance in what follows.

3.5.2. *The* MTW *condition and the convexity of the tangent cut-locus.* — In [**26**], Loeper and Villani noticed the existence of a deep connection between the MTW condition and the geometry of the cut-locus. The idea is the following: fix $x \in M$, and let $v_0, v_1 \in \mathrm{I}(x)$. Consider the segment $(v_t)_{t\in[0,1]}$, with $v_t := (1-t)v_0 + tv_1$. Set further $y_t := \exp_x(v_t)$. Since $v_0, v_1 \in \mathrm{I}(x)$, we have

$$y_0, y_1 \notin \mathrm{cut}(x).$$

In particular $c(x,\cdot) := d(x,\cdot)^2/2$ is smooth in a neighborhood of y_0 and y_1. Assume now that the MTW condition holds. Thanks to Theorem 3.4(iv), we know that the function

$$\eta \mapsto \big\langle D_x^2 c\big(\bar{x}, \exp_{\bar{x}}(p+\eta)\big) \cdot \xi, \xi \big\rangle$$

is concave for all $\eta \perp \xi$. (This is just a formal argument, as the theorem applies a priori only if $\exp_{\bar{x}}(p+\eta) \notin \mathrm{cut}(\bar{x})$.) Applying this fact along the segment $(v_t)_{t\in[0,1]}$, and exploiting the smoothness of $d(x,\cdot)^2$ near y_0 and y_1, we obtain, for $\xi \perp (v_1 - v_0)$,

$$\inf_{t\in[0,1]} \big\langle D_x^2 d^2(x,y_t) \cdot \xi, \xi \big\rangle \geq \min\big\{ \big\langle D_x^2 d^2(x,y_0) \cdot \xi, \xi \big\rangle, \big\langle D_x^2 d^2(x,y_1) \cdot \xi, \xi \big\rangle \big\} \geq C_0,$$

for some constant $C_0 \in \mathbb{R}$. Hence, if we forget for a moment about the orthogonality assumption between $v_1 - v_0$ and ξ, we see that the above equation implies that $x \notin \mathrm{cut}(y_t)$ for all $t \in [0,1]$ (compare with (18)), which by symmetry gives

$$y_t \notin \mathrm{cut}(x) \qquad \forall\, t \in [0,1],$$

or equivalently

$$v_t \notin \mathrm{TCL}(x) \qquad \forall\, t \in [0,1].$$

Since $v_0, v_1 \in \mathrm{I}(x)$, we have obtained

$$v_t \in \mathrm{I}(x) \qquad \forall\, t \in [0,1],$$

that is $\mathrm{I}(x)$ is convex! In conclusion, this formal argument suggests that the MTW condition (or a variant of it) should imply that all tangent injectivity loci $\mathrm{I}(x)$ are convex, for every $x \in M$. This would be a remarkable property. Indeed, usually the only regularity results available for $\mathrm{I}(x)$ say that $\mathrm{TCL}(x)$ is just Lipschitz [**8, 18, 22**]. Moreover, such a result would be of a global nature, and not just local like a semi-convexity property.

Unfortunately, the argument described above is just formal, and up to now there is no complete result in that direction. However, one can actually prove some rigorous results. To do this, we will need to introduce some variant of the MTW condition.

Convexity of the cut-loci: the nonfocal case

DEFINITION 3.10 (uniform MTW condition). — *If $K, C \geq 0$ are given, it is said that M satisfies the* $\mathrm{MTW}(K, C)$ *condition if, for all $(x, y) \in (M \times M) \setminus \mathrm{cut}(M)$, for all $(\xi, \eta) \in T_x M \times T_y M$,*

$$\mathfrak{S}_{(x,y)}(\xi, \eta) \geq K\, |\xi|_x^2\, |\tilde{\eta}|_x^2 - C\, \langle \xi, \tilde{\eta} \rangle_x^2, \tag{19}$$

where $v = (\exp_x)^{-1}(y)$, $\tilde{\eta} = (d_v \exp_x)^{-1}(\eta)$.

DEFINITION 3.11. — *We say that Riemannian manifold (M, g) has nonfocal cut-locus if* $\mathrm{fcut}(x) = \varnothing$ *for all $x \in M$.*

As shown in [**26**] by a compactness argument, as long as $y \notin \mathrm{cut}(x)$ stays uniformly away from $\mathrm{fcut}(x)$, the $\mathrm{MTW}(K, C)$ condition is actually equivalent to the $\mathrm{MTW}(K)$ condition. In particular, if (M, g) is a compact manifold with nonfocal cut-locus, and the $\mathrm{MTW}(K)$ condition holds for some $K \geq 0$, then there exists a constant $C > 0$ such that the $\mathrm{MTW}(K, C)$ condition is true. Thanks to this fact, the authors can prove a variant of Theorem 3.4(iii), where they exploit the information coming from the fact that now the vectors ξ and η do not need to be orthogonal, in order to get an improved version of that result: with the same notation as in Theorem 3.4(iii), then there exists $\lambda = \lambda(K, C) > 0$ such that, for any $t \in (0, 1)$,

$$\begin{aligned} d(x, y_t)^2 - d(\bar{x}, y_t)^2 \geq \min\Big(d(x, y_0)^2 - d(\bar{x}, y_0)^2,\, d(x, y_1)^2 - d(\bar{x}, y_1)^2\Big) \\ + 2\lambda\, t(1-t)\, d(\bar{x}, x)^2 |v_1 - v_0|_{\bar{x}}^2, \end{aligned} \tag{20}$$

where $v_0 = (\exp_{\bar{x}})^{-1}(y_0)$, $v_1 = (\exp_{\bar{x}})^{-1}(y_1)$. Moreover, they can even assume that y_t is not exactly a c-segment, but just a C^2-perturbation of it.

Thanks to this improved version of "regularity", Loeper and Villani showed the following result:

THEOREM 3.12 ([**26**]). — *Let (M, g) be a Riemannian manifold with nonfocal cut-locus, satisfying* $\mathrm{MTW}(K)$ *for some $K > 0$ (in particular, M is compact by Remark 3.9). Then there is $\kappa > 0$ such that all tangent injectivity domains* $\mathrm{I}(x)$ *are κ-uniformly convex.*

The (uniform) convexity of all injectivity loci is exactly what Ma, Trudinger and Wang needed as a geometric assumption in order to prove the regularity of the optimal map.

Hence, combining Theorem 3.1 with the strategy developed by Loeper in [**24**] (see Theorem 3.6), Loeper and Villani obtained the following theorem:

COROLLARY 3.13 ([**26**]). — *Let (M, g) be a Riemannian manifold with nonfocal cut-locus, satisfying* MTW(K) *for some $K > 0$. Assume that f and g are smooth probability densities, bounded away from zero and infinity on M. Then ψ (and hence T) is smooth.*

Sketch of the proof. — The first step of the proof consists in showing that ψ is C^1. This is done using the same strategy of Theorem 3.6, exploiting (20) and the convexity of all injectivity domains ensured by Theorem 3.12. We remark that the fact that (20) holds for C^2-perturbations of c-segments allows to simplify some technical parts of the original proof of Loeper, and to slightly relax some of his assumptions.

Then, one takes advantage of the nonfocality assumption to ensure the "stay-away property" $\mathrm{dist}(T(x), \mathrm{cut}(x)) \geq \sigma > 0$. To see how nonfocality plays a role in this estimate, we recall the description of the distance function given in Paragraph 3.5.1: roughly speaking

- $d(x, y)^2$ is smooth for $y \notin \mathrm{cut}(x)$.
- $d(x, y)^2$ is at most C^1 for $y \in \mathrm{fcut}(x)$.
- $d(x, y)^2$ is not C^1 for $y \in \mathrm{cut}(x) \setminus \mathrm{fcut}(x)$.

Hence, in presence of nonfocality, either $d(x, y)^2$ is smooth, or is not C^1, and in this last case there are at least two minimizing geodesics joining x to y. Now, when proving Theorem 1.3(iii), one actually shows that, whenever ψ is differentiable at x, there exists a unique minimizing geodesic from x to $T(x)$, given by $t \mapsto \exp_x(t\nabla\psi(x))$ [**28**]. Thus, if ψ is C^1, in the nonfocal case one immediately deduces that $T(x) \notin \mathrm{cut}(x)$ *for all $x \in M$*, and a simple compactness argument provides the existence of a positive $\sigma > 0$ such that $d(T(x), \mathrm{cut}(x)) \geq \sigma$.

Once the stay-away property is established, since all pairs $(x, T(x))$ belong to a set where d^2 is smooth, it is simple to localize the problem and apply the a priori estimates of Ma, Trudinger and Wang (see Theorem 3.1) to prove the smoothness of ψ. □

The above result applies for instance to the projective space $\mathbb{RP}^n$ and its perturbations. We also recall that the smoothness of optimal maps holds true in the case of the sphere $\mathbb{S}^n$, as shown by Loeper [**25**]. However, a non-trivial question is whether the regularity of optimal maps holds for perturbations of the sphere.

By imposing some uniform L^∞-bound on the logarithm of the densities (so that they are uniformly bounded away from zero and infinity), Delanoë and Ge showed that for small perturbations of the metric (the smallness depending on the L^∞-bound) the optimal map stays uniformly away from the cut-locus, in the sense that $\mathrm{dist}(T(x), \mathrm{cut}(x)) \geq \sigma$ for some $\sigma > 0$ [**11**], and in this case the regularity issue presents no real difficulties (see the last part of the proof of Corollary 3.13). However this stay-away property does not necessarily hold for general smooth densities, and the problem becomes much more complicated. The case of perturbations of $\mathbb{S}^2$ has been solved by Figalli and Rifford [**12**], and their result has been recently extended to arbitrary dimension by Figalli, Rifford and Villani [**14**].

The extended MTW condition. — We observe that, from the point of view of the structure of the cut-locus, the perturbations of the sphere are in some sense the worst case to treat. Indeed, since for $\mathbb{S}^n$ one has $\mathrm{cut}(x) = \mathrm{fcut}(x)$ for all $x \in M$ (which is completely the opposite of nonfocality), when one slightly perturbs the metric the structure of the cut-locus can be very wild. (The idea is that the cut-locus behaves nicely under perturbations of the metric away from focalization, while it is very difficult to control its behavior near the focal-locus [**8**]).

To overcome these difficulties, Figalli and Rifford introduced the following strategy: first of all, we observe that the MTW condition is defined only for $(x, y) \in M \times M$ with $y \notin \mathrm{cut}(x)$. Hence, we can write it as a condition on the pairs (x, v) instead of (x, y), where $v := (\exp_x)^{-1}(y) \in \mathrm{I}(x)$.

We fix now $\bar{x} \in M$, and we observe that the MTW tensor at $(\bar{x}, v)$ (or equivalently at $(\bar{x}, \exp_{\bar{x}}(v))$) is expressed in terms of derivatives of $d^2/2$ at $(\bar{x}, \exp_{\bar{x}}(v))$. Now, assume that v approaches $\mathrm{TCL}(\bar{x})$ but it is still far from $\mathrm{TFL}(\bar{x})$. This means that the map $(x, w) \mapsto (x, \exp_x(w))$ is a local diffeomorphism near $(\bar{x}, v)$. Hence, we can define a new cost function for (x, y) near $(\bar{x}, \exp_{\bar{x}}(v))$ as

$$\hat{c}(x, y) := \frac{\|(\exp_x)^{-1}(y)\|_x^2}{2},$$

where now $(\exp_x)^{-1}$ denotes the local smooth inverse of $\exp_x$, as explained above. This new cost function coincides with $d(x, y)^2/2$ as long as $y = \exp_x(w)$ with $w \in \mathrm{I}(x)$, and it provides a smooth extension of it up to the first conjugate time. This allows to define an *extended* MTW condition, which makes sense for all pairs (x, v) with $v \in \mathrm{NF}(x)$ (and not only for $v \in \mathrm{I}(x)$). The advantage of having extended the MTW condition up to the focal-locus is twofold: on the one hand, the extended MTW condition is more "local", as one can easily show that it only concerns the geodesic flow, and not the global topology of the manifold. On the other hand, the fact of being allowed to cross the cut-locus away from the focal points makes this extended condition more flexible than the usual one, and this strongly helps when trying to prove the convexity of all tangent injectivity domains. Exploiting these facts, Figalli and Rifford proved the following result:

THEOREM 3.14 ([**12**]). — *Let (M, g) be a Riemannian manifold which satisfies the extended* $\mathrm{MTW}(K, C)$ *condition for some $K, C > 0$, and assume that* $\mathrm{NF}(x)$ *is (strictly) convex for all $x \in M$. Then* $\mathrm{I}(x)$ *is (strictly) convex for all $x \in M$.*

We observe that, in the above result, the authors replace the nonfocality assumption as in Theorem 3.13 with the convexity of all tangent nonfocal domains. This hypothesis is satisfied for instance by any perturbation of the sphere $\mathbb{S}^n$ (see for example [**8**]).

The above theorem allows also to prove a regularity result for optimal maps:

Corollary 3.15 ([**12**]). — *Let (M,g) be a Riemannian manifold which satisfies the extended* $\mathrm{MTW}(K,C)$ *condition for some $K, C > 0$, and assume that* $\mathrm{NF}(x)$ *is (strictly) convex for all $x \in M$. Assume that f and g are two probability densities bounded away from zero and infinity on M. Then the optimal map is continuous.*

We remark that the statement of the above theorem does not say that if f and g are smooth, then T is smooth too. The difficulty to prove such a result comes again from focalization: if the cut-locus is nonfocal, as shown in the proof of Corollary 3.13 the continuity of the transport map implies the stay-away property $\operatorname{dist}(T(x), \operatorname{cut}(x)) \geq \sigma > 0$, and from this fact the higher regularity of T follows easily [**26**]. Unfortunately, without nonfocality (as in the above case), the continuity of T is not enough to ensure the stay-away property, and this is why the above statement is only about the continuity of the optimal map.

In [**12**] the authors show that the sphere $\mathbb{S}^n$ satisfies the (extended) $\mathrm{MTW}(K,C)$ condition for some $K = C > 0$, and they prove that this condition survives for perturbations of the two-dimensional sphere. In particular, they obtain as a corollary the following result:

Corollary 3.16 ([**12**]). — *Let $(M,g) = (\mathbb{S}^2, g_\varepsilon)$, where g_ε is a C^4-perturbation of canonical metric on $\mathbb{S}^2$. Then, for ε small enough,* $\mathrm{I}(x)$ *is strictly convex for all $x \in M$. Moreover, if f and g are two probability densities bounded away from zero and infinity on M, then the optimal map is continuous.*

Conclusions. — An interesting remark to the above result is the following: the first part of the statement of Corollary 3.16 is a statement on perturbations of the 2-sphere, which has nothing to do with optimal transport! Moreover, the same is true for many of the results stated above, which are just statements on the structure of the cut-locus. So, what happened can be summarized as follows: to prove regularity of optimal maps, Ma, Trudinger and Wang discovered a new tensor by purely PDE methods, starting from a Monge-Ampère type equation. Then it was realized that this tensor is intrinsic and has a geometric meaning, and now the MTW tensor is used as a tool (like the Ricci or the Riemann tensor) to prove geometric statements on manifolds. (For a recent account on other possible links between optimal transport and geometry, see [**16**].) This domain of research is new and extremely active, and there are still a lot of open problems. For instance, a complete understanding of the link between the MTW condition and the convexity of the tangent cut-loci is still missing (although in [**15**] the authors have a quite complete answer in the case of 2-dimensional manifolds). Another formidable challenge is for example the description of positively curved Riemannian manifolds which satisfy $\mathrm{MTW}(K, C)$, for some $K, C > 0$.

REFERENCES

[1] Y. Brenier – "Décomposition polaire et réarrangement monotone des champs de vecteurs", *C. R. Acad. Sci. Paris Sér. I Math.* **305** (1987), p. 805–808.

[2] ______, "Polar factorization and monotone rearrangement of vector-valued functions", *Comm. Pure Appl. Math.* **44** (1991), p. 375–417.

[3] X. Cabré – "Nondivergent elliptic equations on manifolds with nonnegative curvature", *Comm. Pure Appl. Math.* **50** (1997), p. 623–665.

[4] L. A. Caffarelli – "A localization property of viscosity solutions to the Monge-Ampère equation and their strict convexity", *Ann. of Math.* **131** (1990), p. 129–134.

[5] ______, "Some regularity properties of solutions of Monge Ampère equation", *Comm. Pure Appl. Math.* **44** (1991), p. 965–969.

[6] ______, "The regularity of mappings with a convex potential", *J. Amer. Math. Soc.* **5** (1992), p. 99–104.

[7] ______, "Boundary regularity of maps with convex potentials. II", *Ann. of Math.* **144** (1996), p. 453–496.

[8] M. Castelpietra & L. Rifford – "Regularity properties of the distance function to conjugate and cut loci for viscosity solutions of Hamilton-Jacobi equations and applications in Riemannian geometry", to appear in *ESAIM Control Optim. Calc. Var.*.

[9] D. Cordero-Erausquin, R. J. McCann & M. Schmuckenschläger – "A Riemannian interpolation inequality à la Borell, Brascamp and Lieb", *Invent. Math.* **146** (2001), p. 219–257.

[10] J. A. Cuesta & C. Matrán – "Notes on the Wasserstein metric in Hilbert spaces", *Ann. Probab.* **17** (1989), p. 1264–1276.

[11] P. Delanoë & Y. Ge – "Regularity of optimal transportation maps on compact, locally nearly spherical, manifolds", to appear in *J. reine angew. Math.*

[12] A. Figalli & L. Rifford – "Continuity of optimal transport maps and convexity of injectivity domains on small deformations of $\mathbb{S}^2$", *Comm. Pure Appl. Math.* **62** (2009), p. 1670–1706.

[13] A. Figalli, L. Rifford & C. Villani – "Nearly round spheres look convex", preprint, 2009.

[14] ______, "Necessary and sufficient conditions for continuity of optimal transport maps on Riemannian manifolds", in preparation.

[15] ______, "On the Ma–Trudinger–Wang curvature tensor on surfaces", to appear in *Calc. Var. Partial Differential Equations.*

[16] A. Figalli & C. Villani – "Optimal transport and curvature", Lecture notes, `http://cvgmt.sns.it/cgi/get.cgi/papers/figvil08/OTcurv-6.pdf`.

[17] S. Gallot, D. Hulin & J. Lafontaine – *Riemannian geometry*, third ed., Universitext, Springer, 2004.

[18] J.-i. Itoh & M. Tanaka – "The Lipschitz continuity of the distance function to the cut locus", *Trans. Amer. Math. Soc.* **353** (2001), p. 21–40.

[19] Y.-H. Kim – "Counterexamples to continuity of optimal transport maps on positively curved Riemannian manifolds", *Int. Math. Res. Not.* **2008** (2008), art. ID rnn120, 15.

[20] Y.-H. Kim & R. J. McCann – "Continuity, curvature, and the general covariance of optimal transportation", to appear in *J. Eur. Math. Soc.*

[21] ______, "Towards the smoothness of optimal maps on Riemannian submersions and Riemannian products (of round spheres in particular)", preprint, to appear in *J. reine angew. Math.*

[22] Y. Li & L. Nirenberg – "The distance function to the boundary, Finsler geometry, and the singular set of viscosity solutions of some Hamilton-Jacobi equations", *Comm. Pure Appl. Math.* **58** (2005), p. 85–146.

[23] J. Liu – "Hölder regularity of optimal mappings in optimal transportation", *Calc. Var. Partial Differential Equations* **34** (2009), p. 435–451.

[24] G. Loeper – "On the regularity of solutions of optimal transportation problems", *Acta Math.* **202** (2009), p. 241–283.

[25] ______, "Regularity of optimal maps on the sphere: The quadratic cost and the reflector antenna", to appear in *Arch. Ration. Mech. Anal.*

[26] G. Loeper & C. Villani – "Regularity of optimal transport in curved geometry: the nonfocal case", preprint, to appear in *Duke Math. J.*

[27] X.-N. Ma, N. S. Trudinger & X.-J. Wang – "Regularity of potential functions of the optimal transportation problem", *Arch. Ration. Mech. Anal.* **177** (2005), p. 151–183.

[28] R. J. McCann – "Polar factorization of maps on Riemannian manifolds", *Geom. Funct. Anal.* **11** (2001), p. 589–608.

[29] G. Monge – "Mémoire sur la théorie des déblais et des remblais", *Hist. de l'Acad. des Sciences de Paris* (1781), p. 666–704.

[30] S. T. RACHEV & L. RÜSCHENDORF – *Mass transportation problems. Vol. I*, Probability and its Applications (New York), Springer, 1998.

[31] N. S. TRUDINGER & X.-J. WANG – "On strict convexity and continuous differentiability of potential functions in optimal transportation", *Arch. Ration. Mech. Anal.* **192** (2009), p. 403–418.

[32] ______, "On the second boundary value problem for Monge-Ampère type equations and optimal transportation", *Ann. Sc. Norm. Super. Pisa Cl. Sci.* **8** (2009), p. 143–174.

[33] J. URBAS – "On the second boundary value problem for equations of Monge-Ampère type", *J. reine angew. Math.* **487** (1997), p. 115–124.

[34] C. VILLANI – *Optimal transport, old and new*, Grundl. Math. Wiss., vol. 338, Springer, 2009.

[35] X.-J. WANG – "On the design of a reflector antenna", *Inverse Problems* **12** (1996), p. 351–375.

Alessio FIGALLI

Department of Mathematics
The University of Texas at Austin
1 University Station C1200
Austin, TX 78712 – USA
E-mail : figalli@math.utexas.edu