On Hölder continuity in time of the optimal transport map towards measures along a curve

Nicola Gigli*
University of Bordeaux

01 November 2008

Abstract

We discuss the problem of the regularity in time of the map $t \mapsto T_t \in L^p(\mathbb{R}^d, \mathbb{R}^d; \sigma)$ where T_t is a transport map (optimal or not) from a reference measure σ to a measure μ_t which lies along an absolutely continuous curve $t \mapsto \mu_t$ in the space $(\mathscr{P}_p(\mathbb{R}^d), W_p)$. We prove that in most cases such a map is no more than $\frac{1}{p}$-Hölder continuous.

1 Introduction

Starting from the pioneering work of Otto [10], much is known today about the Riemannain structure of the Wasserstein space $(\mathscr{P}_2(\mathbb{R}^d), W_2)$. One of the basic facts of the theory is that for any probability measure σ with bounded second moment, there is a well defined 'exponential map' from $L^2(\mathbb{R}^d, \mathbb{R}^d; \sigma)$ to $\mathscr{P}_2(\mathbb{R}^d)$ given by:

$$v \mapsto (Id + v)_\# \sigma,$$

where Id is the identity map and $(Id + v)_\# \sigma$ the push forward of σ through $Id + v$. The trivial inequality

$$W_2\big((Id + v)_\# \sigma, (Id + w)_\# \sigma\big) \leq \sqrt{\int |v(x) - w(x)|^2 d\sigma(x)},$$

may be interpeted as the confirmation of the formal fact that $(\mathscr{P}_2(\mathbb{R}^d), W_2)$ has non-negative curvature, since the exponential map is non expansive. If the measure σ is absolutely continuous (this condition may be weakened, see for instance [2] or [12] for more general results), the exponential map has a natural right inverse: the function which associates to each $\mu \in \mathscr{P}_2(\mathbb{R}^d)$, the vector field $T_\sigma^\mu - Id$, where T_σ^μ is the optimal transport map from σ to μ. The existence of such map is given by the celebrated theorem of Brenier ([3]).

A natural question which arises is then: which kind of regularity should we expect from the map $\mu \mapsto T_\sigma^\mu$?

*Nicola.Gigli@math.u-bordeaux1.fr

A well known result in this direction is that, under the assumption $\sigma \ll \mathcal{L}^d$ which guarantees existence and uniqueness of the optimal transport map, from the so called 'stability of optimality' it follows that the function $\mathscr{P}_p(\mathbb{R}^d) \ni \mu \mapsto T_\sigma^\mu \in L^p(\mathbb{R}^d, \mathbb{R}^d; \sigma)$ is continuous.

It is then natural to ask whether there is more regularity or not. A typical question is the following: given an absolutely continuous curve $t \mapsto \mu_t \in \mathscr{P}_2(\mathbb{R}^d)$, which regularity does the map $t \mapsto T_\sigma^{\mu_t} \in L^2(\mathbb{R}^d, \mathbb{R}^d; \sigma)$ have?

This question has been investigated by several authors, among which Loeper and Ambrosio. Loeper published a work on the subject ([9]) where he obtained a result of the following kind: he assumed $\mu_t = (X(t,\cdot))_\# \sigma$, with $\sigma = \mathcal{L}^d{}_{|U}$ for some open set U, and $X(t,x) : [0,1] \times U \to \mathbb{R}^d$ with both X and $\partial_t X$ L^∞ in space and time, and he derived that the optimal transport maps T_t from σ to μ_t satisfies "$t \mapsto T_t$ is of bounded variation in $L^2(\mathbb{R}^d, \mathbb{R}^d, \sigma)$".

The results of Ambrosio are unpublished. With his permission, we report here his result, which shows that when $p = 2$, under certain conditions on σ and (μ_t) (similar to those of Caffarelli's regularity theory for the solutions of the Monge Ampere equation), the map $t \mapsto T_\sigma^{\mu_t} \in L^2(\mathbb{R}^d, \mathbb{R}^d; \sigma)$ is $\frac{1}{2}-$Hölder continuous.

The main result of this paper is that in the case $p = 2$, $\frac{1}{2}-$Hölder regularity is the most we can expect.

Actually, we prove much more: for every $1 < p < \infty$, and any geodesic of the kind $t \mapsto \mu_t := ((1-t)Id + tS)_\# \mu_0 \in \mathscr{P}_p(\mathbb{R}^d)$ for some optimal map S, and any family of maps T_t (not necessarily optimal) satisfying $(T_t)_\# \sigma = \mu_t$ and

$$T_1 \neq S \circ T_0,$$

the map $t \mapsto T_t \in L^p(\mathbb{R}^d, \mathbb{R}^d; \sigma)$ is at most $\frac{1}{p}-$Hölder continuous.

2 Preliminaries

For a given $1 < p < \infty$, we will denote by $\mathscr{P}_p(\mathbb{R}^d)$ the set of probability measures on $\mathbb{R}^d$ with bounded p moment, that is:

$$\mathscr{P}_p(\mathbb{R}^d) := \Big\{ \mu \in \mathscr{P}(\mathbb{R}^d) \ : \ \int |x|^p d\mu(x) < \infty \Big\}.$$

We endow $\mathscr{P}_p(\mathbb{R}^d)$ with the distance W_p defined as

$$W_p(\mu, \nu) := \inf \sqrt[p]{\int \int |x-y|^p d\boldsymbol{\gamma}},$$

where the infimum is taken among all *admissible* plans $\boldsymbol{\gamma} \in \mathscr{P}(\mathbb{R}^d \times \mathbb{R}^d)$ satisfying $\pi^1_\# \boldsymbol{\gamma} = \mu$ and $\pi^2_\# \boldsymbol{\gamma} = \nu$, where π^1, π^2 are the projection onto the first and second coordinate respectively. A plan which realizes the minimum is called *optimal.*

The following theorem is a well known generalization of Brenier's theorem to the case of general exponent p. It is not stated in his maximum generality, and the conclusion may be strenghtened by a characterization of the optimal map: for a detailed discussion of the topic and for the proof see [2] or [12].

Theorem 2.1 (Existence and uniqueness of optimal transport map) *Let $\mu, \nu \in \mathscr{P}_p(\mathbb{R}^d)$ and assume that μ is absolutely continuous w.r.t. the Lebesgue measure. Then there exists a unique optimal plan $\boldsymbol{\gamma}$ from μ to ν, and this plan is induced by a map, i.e.: there exists (a unique) $T \in L^p(\mathbb{R}^d, \mathbb{R}^d; \mu)$ such that $\boldsymbol{\gamma} = (Id, T)_\# \mu$, where Id is the identity map.*

There is a well known characterization of geodesics in $\mathscr{P}_p(\mathbb{R}^d)$; we recall the basics facts we will need in the following statement (see e.g. [2] or [12] for a proof).

Theorem 2.2 (Geodesics in $(\mathscr{P}_p(\mathbb{R}^d), W_p)$) *Let $(\mu_t) \subset \mathscr{P}_p(\mathbb{R}^d)$ be a geodesic on $[0,1]$. Then:*

i) there exists an optimal transport plan $\boldsymbol{\gamma}$ from μ_0 to μ_1 such that for any $t \in [0,1]$ it holds

$$\mu_t = \big((1-t)\pi^1 + t\pi^2\big)_\# \boldsymbol{\gamma},$$

ii) for any $t \in (0,1)$, $s \in [0,1]$ there exists only one optimal plan from μ_t to μ_s and such plan is induced by a Lipschitz map,

iii) for any $\varepsilon > 0$ there exists $C_\varepsilon \in \mathbb{R}$ such that for any $t \in [\varepsilon, 1-\varepsilon]$, $s \in [0,1]$ the Lipschitz constant of the optimal transport map from μ_t to μ_s is less than C_ε.

3 $\frac{1}{2}$−Hölder regularity is achievable

Here we report a proof, suggested to us by Ambrosio, that under appropriate hypothesis the $\frac{1}{2}$−Hölder regularity of $t \mapsto T_\sigma^{\mu_t} \in L^2(\mathbb{R}^d, \mathbb{R}^d; \sigma)$ is achievable when (μ_t) is an absolutely continuous curve in $\mathscr{P}_2(\mathbb{R}^d)$. The hypothesis we put on the measures involved ar far from being optimal: it is not the purpose here to look for maximum generality, but just to show that $\frac{1}{2}$−Hölder continuity of the optimal transport map is achievable. In particular, the regularity result due to Caffarelli, which is the key ingredient of the proof, is not recalled here in its maximum generality.

Theorem 3.1 (Caffarelli's regularity result) *Let $\mu, \sigma \in \mathscr{P}_2(\mathbb{R}^d)$. Assume that* $\operatorname{supp}(\mu), \operatorname{supp}(\sigma)$ *(i.e. the smallest closed sets on which μ and σ are concentrated) are both C^2 and uniformly convex. Also assume that both μ and σ are absolutely continuous with $C^{0,\alpha}$ densities on their supports, for some $\alpha \in (0,1)$, satisfying*

$$0 < c \le \left\| \frac{d\sigma}{d\mathcal{L}^d} \right\|_\infty \le C,$$
$$0 < \bar{c} \le \left\| \frac{d\mu}{d\mathcal{L}^d} \right\|_\infty \le \overline{C}.$$

Then the optimal transport map from μ to σ is the gradient of a $C^{2,\alpha}$ function on $\operatorname{supp}(\mu)$.

Corollary 3.2 (Uniform convexity of the optimal transport map) *With the same hypothesis of the previous theorem, let $\varphi \in C^{2,\alpha}(\operatorname{supp}(\mu))$ be the smooth function whose gradient is the optimal transport map from μ to σ. Then φ is strictly uniformly convex.*

Proof. From the bound on the densities of μ and σ and the well known formula

$$\frac{d\sigma}{d\mathcal{L}^d}\big(\nabla\varphi(x)\big)|\det(\nabla^2\varphi(x))| = \frac{d\mu}{d\mathcal{L}^d}(x),$$

we get

$$\frac{\bar{c}}{C} \leq |\det(\nabla^2\varphi(x))|, \qquad \forall x \in \operatorname{supp}(\mu).$$

By the Brenier's theorem, we know that φ is convex, thus the modulus in the above expression can be dropped. Also, by Caffarelli's regularity result we know that

$$\sup_{x\in\operatorname{supp}(\mu)} \|\nabla^2\varphi(x)\|_{\mathrm{op}} < \infty.$$

From this uniform upper bound on the eigenvalues of $\nabla^2\varphi(x)$ plus the uniform lower bound on $\det(\nabla^2\varphi(x))$ obtained before, we get the strict uniform convexity. $\square$

Proposition 3.3 *Let μ, σ be as in theorem 3.1, $\varphi \in C^{2,\alpha}(\operatorname{supp}(\mu))$ be the smooth function whose gradient is the optimal transport map from μ to σ, let $\lambda > 0$ be the modulus of uniform convexity of φ (i.e. λ is the supremum of λ' such that $x \mapsto \varphi(x) - \frac{\lambda'|x|^2}{2}$ is convex on $\operatorname{supp}(\mu)$) and $T := (\nabla\varphi)^{-1}$. Then for every transport map S from σ to μ it holds*

$$\|S - T\|_\sigma^2 \leq \frac{2}{\lambda}\left(\|S - Id\|_\sigma^2 - \|T - Id\|_\sigma^2\right).$$

Proof. We have

$$\begin{aligned} 0 &= \int \varphi(y)d\mu(y) - \int \varphi(y)d\mu(y) = \int \varphi(S(x)) - \varphi(T(x))d\sigma(x) \\ &\geq \int \langle\nabla\varphi(T(x)), S(x) - T(x)\rangle d\sigma(x) + \frac{\lambda}{2}\|S - T\|_\sigma^2. \end{aligned}$$

Now observe that $\nabla\varphi(T(x)) = x$ for every $x \in \operatorname{supp}(\sigma)$, thus it holds

$$\int \langle\nabla\varphi(T(x)), S(x) - T(x)\rangle d\sigma(x) = \int \langle x, S(x) - T(x)\rangle d\sigma(x) = -\frac{1}{2}\|S - Id\|_\sigma^2 + \frac{1}{2}\|T - Id\|_\sigma^2.$$

$\square$

Corollary 3.4 ($\frac{1}{2}$-Hölder regularity) *Let $\sigma \in \mathscr{P}_2(\mathbb{R}^d)$ and $(\mu_t) \subset \mathscr{P}_2(\mathbb{R}^d)$ a Lipschitz curve of absolutely continuous measures. Assume that σ and $\mu := \mu_0$ satisfy the assuptions of Caffarelli's theorem 3.1 and let, for every $t \in [0, 1]$, T_t be the optimal transport map from σ to μ_t. Then $t \mapsto T_t \in L^2(\mathbb{R}^d, \mathbb{R}^d; \sigma)$ satisfies*

$$\varlimsup_{t\to0^+} \frac{\|T_t - T_0\|_\sigma}{\sqrt{t}} < \infty.$$

Proof. Let L be the Lipschitz cosntant of the curve $t \mapsto \mu_t \in \mathscr{P}_2(\mathbb{R}^d)$. Apply Brenier's theorem to get the existence of optimal transport maps S_t from μ_t to μ_0. The map $S_t \circ T_t$ maps σ into μ_0, thus applying proposition 3.3 we get

$$\|S_t \circ T_t - T_0\|_\sigma^2 \leq C \left(\|S_t \circ T_t - Id\|_\sigma^2 - \|T_0 - Id\|_\sigma^2 \right), \tag{3.1}$$

for every $t \in [0,1]$ and some constant C independent on t.

Now observe that

$$\begin{aligned} \|S_t \circ T_t - Id\|_\sigma &\leq \|S_t \circ T_t - T_t\|_\sigma + \|T_t - Id\|_\sigma = \|S_t - Id\|_{\mu_t} + W_2(\mu_t, \sigma) \\ &\leq 2W_2(\mu_0, \mu_t) + W_2(\mu_0, \sigma) \leq 2Lt + W_2(\mu_0, \sigma), \end{aligned}$$

and similarly

$$\|S_t \circ T_t - T_0\|_\sigma \geq \|T_t - T_0\|_\sigma - \|S_t \circ T_t - T_t\|_\sigma \geq \|T_t - T_0\|_\sigma - Lt.$$

Using this two inequality in (3.1) and recalling that $\|T_0 - Id\|_\sigma = W_2(\mu_0, \sigma)$ we get the thesis. □

4 The basic idea of the result

Before turning to our main result on $\frac{1}{p}$-Hölder regularity, we discuss a simple example in dimension 2 and for $p = 2$, that shows the main idea of our argument.

Let $A := (-2, 1)$, $B := (2, 1)$, $C := (0, -2)$ and $O := (0, 0)$. Since the strict inequality

$$|A - O|^2 + |O - C|^2 = 5 + 4 < 13 + 0 = |A - C|^2 + |O - O|^2$$

holds, where $|\cdot|$ is the euclidean norm, we have that for $r > 0$ small enough it holds

$$|A - O'|^2 + |O - C'|^2 < |A - C'|^2 + |O - O'|^2, \quad \forall O' \in B_r(O),\ C' \in B_r(C). \tag{4.1}$$

Fix such an r and define the measures

$$\begin{aligned} \mu_0 :=& \tfrac{1}{2} \left(\delta_A + \delta_O \right), \\ \mu_1 :=& \tfrac{1}{2} \left(\delta_B + \delta_O \right), \\ \sigma :=& (2\pi r^2)^{-1} \left(\mathcal{L}^2|_{B_r(O) \cup B_r(C)} \right). \end{aligned}$$

Inequality (4.1) implies that the optimal transport map T_0 from σ to μ_0 satisfies $T_0(B_r(O)) = \{A\}$ and $T_0(B_r(C)) = \{O\}$. Symmetrically, for the optimal transport map T_1 from σ to μ_1 it holds $T_1(B_r(O)) = \{B\}$ and $T_1(B_r(C)) = \{O\}$.

Now observe that since

$$|A - O|^2 + |O - B|^2 = 5 + 5 < 16 + 0 = |A - B|^2 + |O - O|^2,$$

there is a unique optimal plan between μ_0 and μ_1 and this plan is induced by the map S, seen from μ_0, given by $S(A) = O$ and $S(O) = B$. Observe that it holds $S(T_0(B_r(O))) \neq T_1(B_r(O))$.

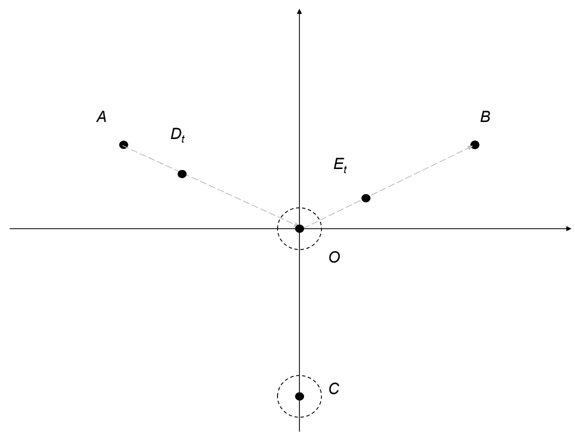

Figure 1: Position of the masses

Let $\mu_t := ((1-t)Id + tS)_{\#}\mu_0$ and T_t be the optimal transport map from σ to μ_t. Let $D_t := (1-t)A$ and $E_t := tB$, so that $\mathrm{supp}(\mu_t) = \{D_t, E_t\}$.

Here it comes the main idea of the example. We claim that the map $t \to T_t \in L^2(\mathbb{R}^2, \mathbb{R}^2; \sigma)$ is not C^α for $\alpha > 1/2$: we will argue by contradiction. Suppose that for some $\alpha > 1/2$ the map is C^α, let χ be the characteristic function of $B_r(0)$ (i.e. $\chi(B_r(0)) = \{1\}$ and $\chi(\mathbb{R}^2 \setminus B_r(0)) = \{0\}$) and observe that from the inequality

$$\int |T_t - T_s|^2 \chi d\sigma \leq \int |T_t - T_s|^2 d\sigma,$$

we get that 'any regularity of $t \mapsto T_t$ seen as curve in $L^2(\mathbb{R}^2, \mathbb{R}^2; \sigma)$ is inherited by the curve $t \mapsto T_t$ seen as curve with values in $L^2(\mathbb{R}^2, \mathbb{R}^2, 2\chi\sigma)$' (the factor 2 stands just for the renormalization of the mass). In particular the map $t \mapsto T_t \in L^2(\mathbb{R}^2, \mathbb{R}^2; 2\chi\sigma)$ is C^α, too. Therefore defining the measures

$$\nu_t := (T_t)_{\#}(2\chi\sigma),$$

and using the inequality

$$W^2(\nu_t, \nu_s) \leq \int |T_t - T_s|^2 d(2\chi\sigma),$$

we get that the curve $t \mapsto \nu_t \in (\mathscr{P}_2(\mathbb{R}^d), W_2)$ is C^α. The contradiction comes from the fact that the mass of ν_0 lies entirely on D_0, while the mass of ν_1 is on E_1. To make the contradiction evident, define the function $f : [0,1] \to [0,1]$ as $f(t) := \nu_t(D_t)$ and observe that it holds $f(0) = 1$ and $f(1) = 0$. Now we want to evaluate the distance $W(\nu_t, \nu_s)$: roughly speaking, the best way to move the mass from ν_t to ν_s is to move as much mass as poissible from D_t to D_s, as much mass as possible from E_t to E_s and then 'to adjust the rest'. More precisely, it can be easily checked that the optimal transport plan between ν_t and ν_s is given by

$$\min\{f(t), f(s)\}\delta_{(D_t,D_s)} + \min\{1 - f(t), 1 - f(s)\}\delta_{(E_t,E_s)}$$
$$+ (f(t) - f(s))^+\delta_{(D_t,E_s)} + (f(s) - f(t))^+\delta_{(E_t,D_s)},$$

as its support is either $\{(D_t, D_s), (E_t, E_s), (D_t, E_s)\}$ or $\{(D_t, D_s), (E_t, E_s), (E_t, D_s)\}$ (depending on whether $f(t) \geq f(s)$ or viceversa, respectively) and both of these sets are cyclically monotone. Therefore we get

$$\begin{aligned} W_2^2(\nu_t, \nu_s) = &\min\{f(t), f(s)\}|D_t - D_s|^2 + \min\{1 - f(t), 1 - f(s)\}|E_t - E_s|^2 \\ &+ (f(t) - f(s))^+|D_t - E_s|^2 + (f(s) - f(t))^+|E_t - D_s|^2. \end{aligned}$$

Considering only the last two terms of the expression on the right, and choosing $|s - t| < 1/2$ we get the bound

$$W_2(\nu_t, \nu_s) \geq \frac{\sqrt{5}}{2}\sqrt{f(t) - f(s)}.$$

From the fact that $t \mapsto \nu_t \in (\mathscr{P}_2(\mathbb{R}^d), W_2)$ is C^α we get

$$\sqrt{f(t) - f(s)} \leq c|t - s|^\alpha, \quad \forall t, s \text{ s.t. } |s - t| < 1/2,$$

for some constant c. The contradiction follows: indeed the above inequality and the fact that $\alpha > 1/2$ implies that f is constant on $[0, 1]$, while we know that $f(0) = 1$ and $f(1) = 0$.

5 The main result

Lemma 5.1 *Let $\sigma, \mu \in \mathscr{P}_p(\mathbb{R}^d)$ and T, S be two Borel transport maps from σ to μ. Assume that $T \neq S$ in L^2_σ. Then there exists a Borel set E such that $\sigma(E) > 0$ and $d(T(E), S(E)) > 0$.*

Proof. Since $T \neq S$ in L^2_σ we know that there exists $c > 0$ such that the Borel set

$$A := \Big\{x \in \mathbb{R}^d \ : \ |T(x) - S(x)| > c\Big\},$$

satisfies $\sigma(A) > 0$. Let $r := \frac{\sigma(A)}{4}$ and find $x_0 \in \mathbb{R}^d$ such that the ball $B_r(x_0)$ satisfies

$$\sigma\left(\Big\{x \in A \ : \ T(x) \in B_r(x_0)\Big\}\right) > 0,$$

(such a ball must exist, since a countable family of balls of radius r covers $\mathbb{R}^d$). We claim that the Borel set

$$E := \Big\{x \in A \ : \ T(x) \in B_r(x_0)\Big\},$$

satisfies the thesis. We know that $\sigma(E) > 0$, so we only have to prove that $\inf_{x,y\in E} |T(x) - S(y)| > 0$. This follows from

$$|T(x) - S(y)| > |T(y) - S(y)| - |T(x) - T(y)| > r - 2\frac{r}{4} = \frac{r}{2}.$$

$\square$

Theorem 5.2 *Let $(\mu_t) \subset \mathscr{P}_p(\mathbb{R}^d)$ be a geodesic induced by a Lipshitz map S with Lipschitz inverse*[1]*, $\sigma \in \mathscr{P}_p(\mathbb{R}^d)$ and, for any $t \in [0,1]$, $T_t \in L^p_\sigma$ a transport map from σ to μ_t. Assume that $T_1 \neq S \circ T_0$. Then the map $t \mapsto T_t \in L^p_\sigma$ has at most $\frac{1}{p}$-Hölder regularity.*

Proof. Use lemma 5.1 to obtain the existence of a Borel set E such that $\sigma(E) > 0$ such that $d(S(T_0(E)), T_1(E)) > 0$. Since S has Lipschitz inverse, it also holds $d(T_0(E), S^{-1}(T_1(E))) > 0$. Define the measures

$$\overline{\sigma} := \frac{1}{\sigma(E)} \sigma|_E,$$
$$\nu_t := (T_t)_\# \overline{\sigma},$$

the sets $A_0 := T_0(E)$, $B_0 := S^{-1}(T_1(E))$ and

$$A_t := \big((1-t)Id + tS\big)(A_0),$$
$$B_t := \big((1-t)Id + tS\big)(B_0),$$

so that ν_0 is concentrated on A_0 and ν_1 on B_1 and ν_t is concentrated on $A_t \cup B_t$ for any $t \in (0,1)$. From the choice of E, we know that $d(A_0, B_0) > 0$ and $d(A_1, B_1) > 0$; therefore, since S is Lipshitz, we get that $d := \inf_t d(A_t, B_t) > 0$. By continuity, for every t, there exists $\delta_t > 0$ such that $d(A_t, B_s) > d/2$ and $d(A_s, B_t) > d/2$ for any $s \in [t - \delta_t, t + \delta_t]$.

Define $f(t) := \overline{\sigma}(T_t^{-1}(A_t))$; observe that $f(0) = 1$, $f(1) = 0$. We claim that

$$W_p(\nu_t, \nu_s) \geq \frac{d}{2} \left(|f(s) - f(t)|\right)^{\frac{1}{p}}, \qquad \forall s \in [t - \delta_t, t + \delta_t] \tag{5.1}$$

for some constant $c > 0$.

To prove this, assume $f(s) \leq f(t)$ (the other inequality is similar) and let $\boldsymbol{\gamma}_t^s$ be an optimal transfer plan from ν_t to ν_s. From

$$\begin{aligned} f(t) = \nu_t(A_t) &= \boldsymbol{\gamma}_t^s(A_t \times B_s) + \boldsymbol{\gamma}_t^s(A_t \times A_s) \\ &\leq \boldsymbol{\gamma}_t^s(A_t \times B_s) + \boldsymbol{\gamma}_t^s(A_t \times A_s) + \boldsymbol{\gamma}_t^s(B_t \times A_s) = \boldsymbol{\gamma}_t^s(A_t \times B_s) + \nu_s(A_s) \end{aligned}$$

we get $\boldsymbol{\gamma}_t^s(A_t \times B_s) \geq f(t) - f(s)$. Arguing symmetrically for the case $f(t) < f(s)$ we can conclude that

$$\boldsymbol{\gamma}_t^s(A_t \times B_s \cup A_s \times B_t) \geq |f(s) - f(t)|.$$

Therefore if $s \in [t - \delta_t, t + \delta_t]$ it holds

$$W_p^p(\nu_t, \nu_s) = \int |x - y|^p d\boldsymbol{\gamma}_t^s(x,y) \geq \int_{A_t \times B_s \cup A_s \times B_t} |x - y|^p d\boldsymbol{\gamma}_t^s(x,y) \geq \left(\frac{d}{2}\right)^p |f(s) - f(t)|,$$

which is equation (5.1).

[1] by theorem 2.2, such a geodesic may be obtained by restriciton starting from any geodesic

To conclude the proof, we will argue by contradiction. Assume that $t \mapsto T_t \in L^2_\sigma$ is $\alpha-$Hölder continuous for some $\alpha > p^{-1}$. Coupling the inequality

$$W_p(\nu_t, \nu_s) \leq \left(\int |T_t - T_s|^p d\overline{\sigma} \right)^{\frac{1}{p}} \leq \frac{1}{(\sigma(E))^{\frac{1}{p}}} \left(\int |T_t - T_s|^p d\sigma \right)^{\frac{1}{p}} \leq C|s-t|^\alpha,$$

with (5.1), we get

$$\frac{d}{2} \left(|f(s) - f(t)| \right)^{\frac{1}{p}} \leq C|s-t|^\alpha, \qquad s \in [t - \delta_t, t + \delta_t],$$

which may be written as

$$\frac{|f(s) - f(t)|}{|s-t|} \leq \frac{2C}{d} |s-t|^{\alpha p - 1}, \qquad s \in [t - \delta_t, t + \delta_t].$$

Since we assumed $\alpha > p^{-1}$, this equation implies that f is constant. This is absurdum, as we know that $f(0) = 1$ and $f(1) = 0$. $\square$

We conclude with some comments on this result.

Remark 5.3 (Independence on the geometry) It is immediate to verify that the validity of theorem 5.2 does not rely on the fact that we are working on $\mathbb{R}^d$, rather than on a generic Riemannian manifold. A similar result holds when the geodesic (μ_t) is contained on $\mathscr{P}_c(M)$, i.e. on the set of probability measures with compact support on a Riemannian manifold M.

The only thing we should take care of, is the meaning of Hölder regularity for the transport map, as in this setting the transport maps do not belong anymore to an Hilbert space. The natural generalization is to define the set Tr_μ of all transport maps from $\mu \in \mathscr{P}_2(M)$ as

$$\mathrm{Tr}_\mu := \left\{ T : M \to M \ : \ T \text{ is Borel and } \int \mathrm{d}^2(x, T(x)) d\mu(x) < \infty \right\},$$

d being the Riemannian distance on M, to identify two maps in this set if they coincide $\mu-$a.e. and to endow this space with the distance D defined as

$$D^2(T, S) := \int \mathrm{d}^2(T(x), S(x))\, d\mu(x).$$

Then the space (Tr_μ, D) is a metric space, and it makes sense to say that a map $t \mapsto T_t \in \mathrm{Tr}_\mu$ is Hölder continuous.

It is known that theorem 2.2 is true also in this setting, thus it can be easily checked that the proof of theorem 5.2 can be generalized up to this level.

Remark 5.4 (Not only geodesics) The only step of the proof in which we used the fact that (μ_t) was (the restriction of) a geodesic, was the one in which we said that the optimal transport maps from μ_t to μ_s are uniformly Lipschitz. This was needed to be sure that the distance between the sets A_t, B_t was bounded from below by a positive constant.

The result can therefore be generalized to the class of curves (μ_t) for which it exists a family of transport maps $\{S_t^s\}_{t,s\in[0,1]}$ uniformly Lipschitz, satisfying $(S_t^s)_\#\mu_t = \mu_s$ and $S_t^s \circ S_r^s = S_r^t$ for any $t, r, s \in [0,1]$. In this case the hypothesis reads as

$$T_1 \neq S_0^1 \circ T_0,$$

and following step by step the proof it is simple to check that the conclusion still holds. (In [1] and [8] it is proved that the family of curves having this kind of Lipschitz property is dense in the class of absolutely continuous curves w.r.t. the uniform convergence).

References

[1] L. Ambrosio, N. Gigli, *Construction of parallel transport in the Wasserstein space*, accepted paper, Methods and Applications of Analysis, 2008.

[2] L. Ambrosio, N. Gigli and G. Savaré, *Gradient flows in metric spaces and in the space of probability measures*, Birkhäuser (2005).

[3] Y. Brenier, *Polar factorization and monotone rearrangement of vector-valued functions*, Comm. Pure Appl. Math., 44 (1991), pp. 375–417.

[4] R. J. DiPerna and P.-L. Lions, *Ordinary differential equations, transport theory and Sobolev spaces*, Invent. Math., 98 (1989), pp. 511–547.

[5] Caffarelli, L. A. *The regularity of mappings with a convex potential*, J. Amer. Math. Soc. 5, 1 (1992), 99–104.

[6] Caffarelli, L. A., *Boundary regularity of maps with convex potentials*, Comm. Pure Appl. Math. 45, 9 (1992), 1141–1151.

[7] Caffarelli, L. A., *Boundary regularity of maps with convex potentials II*, Ann. of Math. (2) 144, 3 (1996), 453–496.

[8] N. Gigli, *On the geometry of the space of measures in $\mathbb{R}^d$ endowed with the quadratic optimal transportation distance*, Scuola Normale Superiore, Pisa, 2008.

[9] G. Loeper, *On the regularity of the polar factorization for time dependent maps*, Calc. Var. Partial Differential Equations, (2004).

[10] F. Otto, *The geometry of dissipative evolution equations: the porous medium equation*, Comm. Partial Differential Equations, 26 (2001), pp. 101–174.

[11] C. Villani, *Topics in optimal transportation*, vol. 58 of Graduate Studies in Mathematics, American Mathematical Society, Providence, RI, 2003.

[12] C. Villani, *Optimal transport, old and new*, Springer Verlag, 2008.