Proceedings of the Edinburgh Mathematical Society (2011) **54**, 401–409
DOI:10.1017/S001309150800117X

ON HÖLDER CONTINUITY-IN-TIME OF THE OPTIMAL TRANSPORT MAP TOWARDS MEASURES ALONG A CURVE

NICOLA GIGLI*

UFR Mathématiques et Informatique, Université de Bordeaux, 351 cours de la Libération, 33405 Talence cedex, France

(Received 17 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 $(\mathcal{P}_p(\mathbb{R}^d), W_p)$. We prove that in most cases such a map is no more than $1/p$-Hölder continuous.

Keywords: optimal transport; Hölder continuity; regularity-in-time

2010 *Mathematics subject classification:* Primary 49Q20
Secondary 28A33

1. Introduction

Starting from the pioneering work of Otto [**8**], much is known today about the Riemannian structure of the Wasserstein space $(\mathcal{P}_2(\mathbb{R}^d), W_2)$. 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 $\mathcal{P}_2(\mathbb{R}^d)$ given by

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

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

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

may be interpreted as the confirmation of the formal fact that $(\mathcal{P}_2(\mathbb{R}^d), W_2)$ 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 example, [**1**] or [**9**] for more general results), the exponential map has a natural right inverse: the function which

* Present address: Laboratoire J. A. Dieudonné, Université Nice Sophia Antipolis, 28 avenue Valrose, BP 2135, 06103 Nice Cedex 2, France (gigli@unice.fr).

associates to each $\mu \in \mathcal{P}_2(\mathbb{R}^d)$ the vector field $T_\sigma^\mu - \mathrm{Id}$, where T_σ^μ is the optimal transport map from σ to μ. The existence of such a map is given by the celebrated theorem of Brenier [**2**].

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

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 $\mathcal{P}_2(\mathbb{R}^d) \ni \mu \mapsto T_\sigma^\mu \in L^2(\mathbb{R}^d, \mathbb{R}^d; \sigma)$ is continuous.

It is then natural to ask whether or not there is more regularity. A typical question is the following: given an absolutely continuous curve $t \mapsto \mu_t \in \mathcal{P}_2(\mathbb{R}^d)$, 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, including Loeper and Ambrosio. Loeper [**7**] obtained the following result: he assumed $\mu_t = (X(t, \cdot))_\# \sigma$, with $\sigma = \mathcal{L}^d|_U$ for some open set U, and $X(t, x)\colon [0, 1] \times U \to \mathbb{R}^d$ with both X and $\partial_t X$ L^∞ in space and time, and he derived that the optimal transport map T_t from σ to μ_t satisfies the condition

$$t \mapsto T_t \text{ is of bounded variation in } L^2(\mathbb{R}^d, \mathbb{R}^d, \sigma).$$

The results of Ambrosio are unpublished. With his permission, we report here his result, which shows that, under certain conditions on σ and (μ_t) (similar to those of Caffarelli's regularity theory for the solutions of the Monge–Ampère 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 $\frac{1}{2}$-Hölder regularity is the most we can expect (§ 3). We will also see that, for the same question asked in the space $(\mathcal{P}_p(\mathbb{R}^d), W_p)$, the maximum regularity is $1/p$-Hölderianity, and that in general we cannot gain greater regularity by dropping the requirement that the transport maps considered are optimal (Theorem 5.1).

2. Notation

For a given $1 < p < \infty$, the set $\mathcal{P}_p(\mathbb{R}^d) \subset \mathcal{P}(\mathbb{R}^d)$ is the set of Borel probability measures on $\mathbb{R}^d$ with bounded p moment, i.e.

$$\mathcal{P}_p(\mathbb{R}^d) := \left\{ \mu \in \mathcal{P}(\mathbb{R}^d)\colon \int |x|^p \,\mathrm{d}\mu(x) < \infty \right\}.$$

The Wasserstein distance W_p of order p is defined on $\mathcal{P}_p(\mathbb{R}^d)$ as

$$W_p(\mu, \nu) := \min \sqrt[p]{\int\!\!\int |x - y|^p \,\mathrm{d}\boldsymbol{\gamma}}, \tag{2.1}$$

where the minimum is taken among all *admissible* plans $\boldsymbol{\gamma} \in \mathcal{P}(\mathbb{R}^d \times \mathbb{R}^d)$ satisfying $\pi^1_\# \boldsymbol{\gamma} = \mu$ and $\pi^2_\# \boldsymbol{\gamma} = \nu$, where π^1 and π^2 are the projection onto the first and second coordinate, respectively. An admissible plan is called *optimal* if it realizes the minimum in (2.1).

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

Here we report a proof (L. Ambrosio, personal communication) that under appropriate hypotheses 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 $\mathcal{P}_2(\mathbb{R}^d)$. The hypotheses we make about the measures involved are far from being optimal: it is not our 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 [**3**–**5**], which is the key ingredient of the proof, is not recalled here in its maximum generality.

Theorem 3.1 (Caffarelli's regularity result). *Let $\mu, \sigma \in \mathcal{P}_2(\mathbb{R}^d)$. Assume that* $\operatorname{supp}(\mu)$ *and* $\operatorname{supp}(\sigma)$ *(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*

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

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 hypotheses as the previous theorem, let $\varphi \in C^{2,\alpha}(\operatorname{supp}(\mu))$ be a smooth function whose gradient is the optimal transport map from μ to σ (φ is uniquely defined up to a constant). Then φ is strictly uniformly convex.*

Proof. By Brenier's Theorem, we know that φ is convex; thus, the Monge–Ampère Equation
$$\frac{\mathrm{d}\sigma}{\mathrm{d}\mathcal{L}^d}(\nabla\varphi(x)) \det(\nabla^2\varphi(x)) = \frac{\mathrm{d}\mu}{\mathrm{d}\mathcal{L}^d}(x)$$
holds. From the bound on the densities of μ and σ we get

$$\frac{\bar{c}}{C} \leqslant \det(\nabla^2\varphi(x)) \quad \text{for all } x \in \operatorname{supp}(\mu).$$

Also, by Caffarelli's regularity result we know that

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

where $\|A\|_{\mathrm{op}}$ is the operatorial norm of the linear map $A\colon \mathbb{R}^d \to \mathbb{R}^D$. 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 and σ be as in Theorem 3.1, let $\varphi \in C^{2,\alpha}(\operatorname{supp}(\mu))$ 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 $x \mapsto \varphi(x) - \frac{1}{2}\lambda'|x|^2$ is convex on $\operatorname{supp}(\mu)$) and let $T := (\nabla\varphi)^{-1}$. Then for every transport map S from σ to μ the following holds:*

$$\|S - T\|_\sigma^2 \leqslant \frac{2}{\lambda}(\|S - \mathrm{Id}\|_\sigma^2 - \|T - \mathrm{Id}\|_\sigma^2).$$

Proof. We have

$$\begin{aligned} 0 &= \int \varphi(y)\,\mathrm{d}\mu(y) - \int \varphi(y)\,\mathrm{d}\mu(y) \\ &= \int \varphi(S(x)) - \varphi(T(x))\,\mathrm{d}\sigma(x) \\ &\geqslant \int \langle \nabla\varphi(T(x)), S(x) - T(x)\rangle\,\mathrm{d}\sigma(x) + \tfrac{1}{2}\lambda\|S - T\|_\sigma^2. \end{aligned}$$

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

$$\begin{aligned} \int \langle \nabla\varphi(T(x)), S(x) - T(x)\rangle\,\mathrm{d}\sigma(x) &= \int \langle x, S(x) - T(x)\rangle\,\mathrm{d}\sigma(x) \\ &= -\tfrac{1}{2}\|S - \mathrm{Id}\|_\sigma^2 + \tfrac{1}{2}\|T - \mathrm{Id}\|_\sigma^2. \end{aligned}$$

□

Corollary 3.4 ($\frac{1}{2}$-Hölder regularity). *Let $\sigma \in \mathcal{P}_2(\mathbb{R}^d)$ and let $(\mu_t) \subset \mathcal{P}_2(\mathbb{R}^d)$ be a Lipschitz curve of absolutely continuous measures. Assume that σ and $\mu := \mu_0$ satisfy the assumptions of Caffarelli's Theorem (Theorem 3.1) and, for every $t \in [0, 1]$, let 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*

$$\varlimsup_{t\to 0^+} \frac{\|T_t - T_0\|_\sigma}{\sqrt{t}} < \infty.$$

Proof. Let L be the Lipschitz constant of the curve $t \mapsto \mu_t \in \mathcal{P}_2(\mathbb{R}^d)$. 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

$$\|S_t \circ T_t - T_0\|_\sigma^2 \leqslant C(\|S_t \circ T_t - \mathrm{Id}\|_\sigma^2 - \|T_0 - \mathrm{Id}\|_\sigma^2) \tag{3.1}$$

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

Now observe that

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

and, similarly,

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

Using these two inequalities in (3.1) and recalling that $\|T_0 - \mathrm{Id}\|_\sigma = W_2(\mu_0, \sigma)$, we complete the proof. □

4. $\frac{1}{2}$-Hölder regularity is the best we can expect in general

Here we give an explicit example in $\mathcal{P}_2(\mathbb{R}^2)$ which shows that in general $\frac{1}{2}$-Hölder regularity is the best we can expect.

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 the following holds:

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

Fix such an r, with no loss generality assume $r < 1$, and define the measures

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

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 that $T_1(B_r(O)) = \{B\}$ 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 that

$$S(T_0(B_r(O))) \neq T_1(B_r(O)).$$

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

We now present 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 $\alpha > \frac{1}{2}$: we will argue by contradiction. Suppose that for some $\alpha > \frac{1}{2}$ the map is C^α. Let $\bar{\sigma} := (\pi r^2)^{-1}\mathcal{L}^2|_{B_r(O)}$ and observe that from $\bar{\sigma} \leqslant 2\sigma$ we deduce

$$\int |T_t - T_s|^2 \, \mathrm{d}\bar{\sigma} \leqslant 2 \int |T_t - T_s|^2 \, \mathrm{d}\sigma.$$

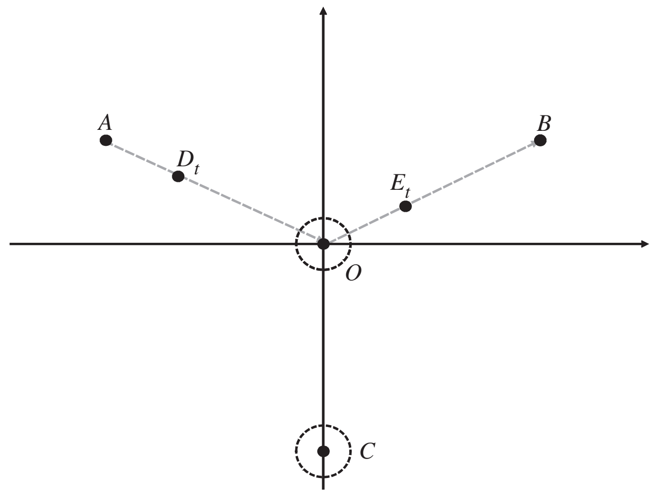

Figure 1. Position of the masses.

Thus, any 'regularity of $t \mapsto T_t$ seen as a curve in $L^2(\mathbb{R}^2, \mathbb{R}^2; \sigma)$ is inherited by the curve $t \mapsto T_t$ seen as a curve with values in $L^2(\mathbb{R}^2, \mathbb{R}^2, \bar{\sigma})$'. In particular, the map $t \mapsto T_t \in L^2(\mathbb{R}^2, \mathbb{R}^2; \bar{\sigma})$ is also C^α. Therefore, defining the measures

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

and using the inequality

$$W^2(\nu_t, \nu_s) \leqslant \int |T_t - T_s|^2 \, \mathrm{d}\bar{\sigma},$$

we get that the curve $t \mapsto \nu_t \in (\mathcal{P}_2(\mathbb{R}^d), W_2)$ 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 $f \colon [0,1] \to [0,1]$ as $f(t) := \nu_t(D_t)$ and observe that it holds that $f(0) = 1$ 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 possible 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 easily be checked that the optimal transport plan between ν_t and ν_s is given by

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

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 $f(t) \geqslant f(s)$ or vice versa, 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-hand side and choosing $|s-t| < \frac{1}{2}$, we get the bound

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

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

$$\sqrt{f(t) - f(s)} \leqslant c|t-s|^\alpha \quad \text{for all } t, s \text{ such that } |s-t| < \tfrac{1}{2},$$

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

5. Generalization of the previous example

For any open set $\Omega \subset \mathbb{R}^d \times [0,1]$, let $\Omega_t \subset \mathbb{R}^d$ be the section defined by

$$\Omega_t := \{x \in \mathbb{R}^d \colon (x,t) \in \Omega\}.$$

Theorem 5.1. *Let $(\mu_t) \subset \mathcal{P}_p(\mathbb{R}^d)$ be an absolutely continuous curve and assume that there exist two open sets $\Omega^1, \Omega^2 \subset \mathbb{R}^d \times [0,1]$ such that*

(i) *$\mu_t(\Omega_t^1 \cup \Omega_t^2) = 1$ and $\mu_t(\Omega_t^1), \mu_t(\Omega_t^2) > 0$ for any $t \in [0,1]$,*

(ii) *for every $t \in [0,1]$ there exists $\delta_t > 0$ such that $d_t := \inf_{s \in [t-\delta_t, t+\delta_t]} d(\Omega_t^1, \Omega_t^2) > 0$.*

Also let $\sigma \in \mathcal{P}_p(\mathbb{R}^d)$ and $(T_t) \in L^p(\mathbb{R}^d, \mathbb{R}^d; \sigma)$ be any choice of Borel maps satisfying $(T_t)_\# \nu = \mu_t$ for any $t \in [0,1]$. Assume that for some $t_0, t_1 \in [0,1]$ the following holds:

$$\sigma((T_{t_0}^{-1}(\Omega_{t_0}^1)) \Delta (T_{t_1}^{-1}(\Omega_{t_1}^1))) \neq 0, \tag{5.1}$$

where Δ stands for the symmetric difference. Then the map $t \mapsto T_t \in L^p(\mathbb{R}^d, \mathbb{R}^d; \sigma)$ is not α-Hölder continuous for any $\alpha > 1/p$.

Proof. Without loss of generality we may assume $t_0 = 0$ and $t_1 = 1$. From $\mu_0(\Omega_0^1) > 0$ and $(T_0)_\# \sigma = \mu_0$ we derive $\sigma(T_0^{-1}(\Omega_0^1)) > 0$. Define

$$\bar{\sigma} := c\sigma|_{T_0^{-1}(\Omega_0^1)},$$
$$\nu_t := (T_t)_\# \bar{\sigma},$$

where c is a normalization constant, and

$$f(t) := \nu_t(\Omega_t^1).$$

Observe that ν_t is concentrated on $\Omega_t^1 \cup \Omega_t^2$ for any $t \in [0,1]$. From the absolute continuity of the curve (μ_t) and the hypotheses (i) and (ii) it may immediately be verified that $\mu_t(\Omega_t^1)$ does not depend on t. Therefore, (5.1) implies $f(1) < f(0)$.

https://doi.org/10.1017/S001309150800117X Published online by Cambridge University Press

For any $t, s \in [0,1]$, let $\boldsymbol{\gamma}_t^s$ be any optimal plan from ν_t to ν_s; $\boldsymbol{\gamma}_t^s$ must be concentrated on $(\Omega_t^1 \cup \Omega_t^2) \times (\Omega_s^1 \cup \Omega_s^2)$. From

$$\begin{aligned}
\boldsymbol{\gamma}_t^s(\Omega_t^2 \times \Omega_s^1) &= \boldsymbol{\gamma}_t^s(\mathbb{R}^d \times \Omega_s^1) - \boldsymbol{\gamma}_t^s(\Omega_t^1 \times \Omega_s^1) \\
&= f(s) - (\boldsymbol{\gamma}_t^s(\Omega_t^1 \times \mathbb{R}^d) - \boldsymbol{\gamma}_t^s(\Omega_t^1 \times \Omega_s^2)) \\
&= f(s) - f(t) + \boldsymbol{\gamma}_t^s(\Omega_t^1 \times \Omega_s^2)
\end{aligned}$$

and the positivity of $\boldsymbol{\gamma}_t^s$ we deduce

$$\begin{aligned}
\boldsymbol{\gamma}_t^s(\Omega_t^2 \times \Omega_s^1) &\geqslant \max\{f(s) - f(t), 0\}, \\
\boldsymbol{\gamma}_t^s(\Omega_t^1 \times \Omega_s^2) &\geqslant \max\{f(t) - f(s), 0\}.
\end{aligned}$$

Thus, recalling (ii) we get

$$\begin{aligned}
W_p^p(\nu_s, \nu_t) &= \int |x-y|^p \,\mathrm{d}\boldsymbol{\gamma}_t^s(x,y) \\
&\geqslant \int_{\Omega_t^1 \times \Omega_s^2 \cup \Omega_t^2 \times \Omega_s^1} |x-y|^p \,\mathrm{d}\boldsymbol{\gamma}_t^s(x,y) \\
&\geqslant d_t^p |f(t) - f(s)| \qquad (5.2)
\end{aligned}$$

for every $s \in [t - \delta_t, t + \delta_t]$.

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

$$W_p(\nu_t, \nu_s) \leqslant \left(\int |T_t - T_s|^p \,\mathrm{d}\bar{\sigma} \right)^{1/p} \leqslant c^{1/p} \left(\int |T_t - T_s|^p \,\mathrm{d}\sigma \right)^{1/p} \leqslant C|s-t|^\alpha$$

with (5.2), we get

$$d_t |f(s) - f(t)|^{1/p} \leqslant C|s-t|^\alpha, \quad s \in [t - \delta_t, t + \delta_t],$$

which may be written as

$$\frac{|f(s) - f(t)|}{|s-t|} \leqslant \frac{C^p}{d_t^p} |s-t|^{\alpha p - 1}, \quad s \in [t - \delta_t, t + \delta_t].$$

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

We conclude with some comments on this result.

Remark 5.2. An example of curve (μ_t) satisfying the assumptions of the theorem is the restriction to $[0,1]$ of some geodesic defined in a larger interval $(-\varepsilon, 1+\varepsilon)$, such that μ_0 is concentrated on two distant sets and gives positive mass to each of them. Indeed, in this case it is known that there exists a bi-Lipschitz map S such that

$$\mu_t := ((1-t)\mathrm{Id} + tS)_{\#}\mu_0,$$

and from this fact it is easy to build open sets Ω^1, Ω^2 for which (i) and (ii) of the theorem are satisfied.

Remark 5.3. The fact that the maps T_t are optimal transport maps for the Wasserstein distance W_p is not one of the assumptions of the above theorem.

Remark 5.4 (independence of the geometry). It can immediately be verified that the validity of Theorem 5.1 does not rely on the fact that we are working on $\mathbb{R}^d$ rather than on a generic Polish space (X, d). A similar result holds when the curve (μ_t) is contained on $\mathcal{P}_p(X)$, i.e. on the set of Borel probability measures μ on X such that

$$\int d^p(x, x_0)\,\mathrm{d}\mu(x) < \infty \quad \text{for some } x_0 \in X.$$

The only thing we should address is the meaning of Hölder regularity for a time-dependent transport map, as in this setting the transport maps no longer belong to a Hilbert space. The natural generalization is to define the set Tr_μ of all transport maps from $\mu \in \mathcal{P}_p(X)$ as

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

to identify two maps in this set if they coincide μ-a.e. and to endow this space with the distance D defined as

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

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.

Acknowledgements. I thank the referee for a valuable contribution towards fixing an error present in the preliminary version of this paper.

References

1. L. Ambrosio, N. Gigli and G. Savaré, *Gradient flows in metric spaces and in the space of probability measures* (Birkhäuser, Basel, 2005).
2. Y. Brenier, Polar factorization and monotone rearrangement of vector-valued functions, *Commun. Pure Appl. Math.* **44** (1991), 375–417.
3. L. A. Caffarelli, The regularity of mappings with a convex potential, *J. Am. Math. Soc.* **5** (1992), 99–104.
4. L. A. Caffarelli, Boundary regularity of maps with convex potentials, *Commun. Pure Appl. Math.* **45** (1992), 1141–1151.
5. L. A. Caffarelli, Boundary regularity of maps with convex potentials, II, *Annals Math. (2)* **144** (1996), 453–496.
6. N. Gigli, On the geometry of the space of measures in $\mathbb{R}^d$ endowed with the quadratic optimal transportation distance, PhD thesis, Scuola Normale Superiore, Pisa (2008; available at http://cvgmt.sns.it/people/gigli/).
7. G. Loeper, On the regularity of the polar factorization for time dependent maps, *Calc. Var. PDEs* **22** (2005), 343–374.
8. F. Otto, The geometry of dissipative evolution equations: the porous medium equation, *Commun. PDEs* **26** (2001), pp. 101–174.
9. C. Villani, *Optimal transport, old and new* (Springer, 2008).