Published in
J. Eur. Math. Soc. 12 (2010), no. 6, pp. 1355–1369.

THE MONGE PROBLEM FOR STRICTLY CONVEX NORMS IN $\mathbb{R}^d$

THIERRY CHAMPION AND LUIGI DE PASCALE

ABSTRACT. We prove the existence of an optimal transport map for the Monge problem in a convex bounded subset of $\mathbb{R}^d$ under the assumptions that the first marginal is absolutely continuous with respect to the Lebesgue measure and that the cost is given by a strictly convex norm. We propose a new approach which does not use disintegration of measures.

1. INTRODUCTION

The Monge problem has origin in the *Mémoire sur la théorie des déblais et remblais* written by G. Monge [22], and may be stated as follows:

$$\inf\left\{\int_\Omega |x-T(x)|d\mu(x)\ :\ T\in\mathcal{T}(\mu,\nu)\right\},\tag{1.1}$$

where Ω is the closure of a convex open subset of $\mathbb{R}^d$, $|\cdot|$ denotes the usual Euclidean norm of $\mathbb{R}^d$, μ,ν are Borel probabilities on Ω and $\mathcal{T}(\mu,\nu)$ denotes the set of transport maps from μ to ν, i.e. the class of Borel maps T such that $T_\sharp\mu=\nu$ (i.e. $T_\sharp\mu(B):=\mu(T^{-1}(B))$ for each Borel set B).

The main result of this paper is to prove the following existence result for a generalization of this problem:

Theorem 1.1. *Let $\|\cdot\|$ be a strictly convex norm on $\mathbb{R}^d$ and assume that μ is absolutely continuous with respect to the Lebesgue measure $\mathcal{L}^d$, then the problem*

$$\min\left\{\int_\Omega \|x-T(x)\|d\mu(x)\ :\ T\in\mathcal{T}(\mu,\nu)\right\}\tag{1.2}$$

has at least one solution.

Before describing the previous results that we know on this problem and our contribution on the subject, we make a brief introduction on the Kantorovich relaxation for (1.2). For general probability measures the set of transport maps $\mathcal{T}(\mu,\nu)$ may be empty, for example if $\mu=\delta_0$ and $\nu=\frac{1}{2}(\delta_1+\delta_{-1})$. But even when $\mathcal{T}(\mu,\nu)$ is non-empty it may happen that problem (1.1) does not admit a minimizer in $\mathcal{T}(\mu,\nu)$: for example take $\mu:=\mathcal{H}^1_{\lfloor\{0\}\times[0,1]}$ and $\nu:=\frac{1}{2}(\mathcal{H}^1_{\lfloor\{-1\}\times[0,1]}+\mathcal{H}^1_{\lfloor\{1\}\times[0,1]})$. Moreover, the objective functional of problem (1.2) is non-linear in T and the set $\mathcal{T}(\mu,\nu)$ does not possess the right compactness properties to deal with the direct methods of the Calculus of Variations. A suitable relaxation was introduced by Kantorovich [20, 21] and it proved to be a strong,

---

Date: Version May 5 2008.
2000 *Mathematics Subject Classification.* 49Q20, 49K30, 49J45.
Key words and phrases. Monge-Kantorovich problem, optimal transport problem, cyclical monotonicity.
money used.

decisive tool to deal with this problem. This relaxation is defined as follows. The set of transport plans from μ to ν is defined as

$$\Pi(\mu,\nu) := \{\gamma \in \mathcal{P}(\Omega\times\Omega) \mid \pi^1_\sharp\gamma = \mu,\ \pi^2_\sharp\gamma = \nu\},$$

where π^i denotes the standard projection in the Cartesian product. The set $\Pi(\mu,\nu)$ is always non-empty as it contains at least $\mu\otimes\nu$. Then Kantorovich proposed to study the problem

$$\min\left\{\int_{\Omega\times\Omega}\|x-y\|d\gamma(x,y)\ :\ \gamma\in\Pi(\mu,\nu)\right\}. \tag{1.3}$$

Problem (1.3) is convex and linear in γ then the existence of a minimizer may be obtained by the direct method of the Calculus of Variations. At this point, to obtain the existence of a minimizer for (1.2) it is sufficient to prove that some solution $\gamma\in\Pi(\mu,\nu)$ of (1.3) is in fact induced by a transport $T\in\mathcal{T}(\mu,\nu)$, i.e. may be written as $\gamma = (id\times T)_\sharp\mu$.

In [29], Sudakov devised an efficient strategy to solve (1.2) for a general norm $\|\cdot\|$ on $\mathbb{R}^d$. However this strategy involved a crucial step on the disintegration of an optimal measure γ for (1.3) which was not completed correctly at that time. In more recent years the problem (1.1) has been solved first by Evans *et al.* [18] with additional regularity assumptions on μ and ν, and then by Ambrosio [1] and Trudinger *et al.* [30] for μ and ν with integrable density. For C^2 uniformly convex norms the problem (1.2) has been solved by Caffarelli *et al.* [11] and Ambrosio *et al.* [3], and finally for crystalline norms in $\mathbb{R}^d$ and general norms in $\mathbb{R}^2$ by Ambrosio *et al.* [2]. The original proof of Sudakov was based on the reduction of the transport problems to affine regions of smaller dimension, and all the proof we listed above are based on the reduction of the problem to a 1-dimensional problem via a change of variable or area-formula.

In this paper, we prove the existence of a solution to (1.2) for a general strictly convex norm $\|\cdot\|$ on $\mathbb{R}^d$, without any regularity assumption on the norm $\|\cdot\|$. The originality of our method for the proof of Theorem 1.1 above is that it does not require disintegration of measures and relies on a simple but powerful regularity result (see Lemma 4.3 below) which has been used in some transport problem with cost functional in non-integral form [12]. In section §2 we recall some well known results on duality and optimality conditions for problem (1.3). In section §3, we introduce a secondary transport problem in order to select solutions (1.3) that have a particular regularity property. Section §4 is devoted to the notion of regular points of a transport γ and in particular to Lemma 4.3, which states that a transport map $\gamma\in\Pi(\mu,\nu)$ is concentrated on a set of regular points. In the following section §5, we take advantage of this fact to prove a regularity result on the transport set associated to a solution of (1.3). The proof of our main result Theorem 1.1 is finally derived in §6, while a possible extension to the case of a general norm $\|\cdot\|$ is discussed in §7.

2. Preliminary on optimal transportation: duality and necessary conditions

The content of this section is classical (for example see [1, 31]). Problem (1.3) is convex and linear, then classical convex duality brings useful information on its minimizers. In particular, the following duality theorem holds (for example we refer to Theorems 3.1. and 3.3 in [3]).

Theorem 2.1. *The minimum in problem (1.3) is equal to*

$$\max\left\{\int_\Omega v(x)d\mu(x) - \int_\Omega v(y)d\nu(y) \ : \ v \in Lip_1(\Omega, \|\cdot\|)\right\} \tag{2.1}$$

where $Lip_1(\Omega, \|\cdot\|)$ is the set of functions $v : \Omega \to \mathbb{R}$ which are 1-Lipschitz with respect to the norm $\|\cdot\|$, i.e.

$$\forall x, y \in \Omega, \qquad |v(x) - v(y)| \le \|x - y\|.$$

Moreover if $u \in Lip_1(\Omega, \|\cdot\|)$ is a maximizer for problem (2.1) then $\gamma \in \Pi(\mu, \nu)$ is a minimizer of problem (1.3) if and only if

$$\forall (x, y) \in \operatorname{supp}\gamma, \qquad u(x) - u(y) = \|x - y\|.$$

In the following, maximizers of (2.1) are referred to as *Kantorovich transport potentials for (2.1)*. If we follow the interpretation of γ as a plan of transport we may deduce from this last theorem that in order to realize an optimal transport the mass should follow the direction of maximal slope of a Kantorovich transport potential u. We give a more precise statement of this classical fact in Lemma 2.2 below, and give a short proof to underline the role of the strict convexity of the norm.

Lemma 2.2. *Assume that $\|\cdot\|$ is a strictly convex norm. Let γ be an optimal transport plan for (1.3), let $u \in Lip_1(\Omega, \|\cdot\|)$ be a Kantorovitch potential for (2.1) and let (x, y) belong to* $\operatorname{supp}(\gamma)$ *with $x \neq y$. If u is differentiable at x and $z \in \Omega$ is such that $u(x) = u(z) + \|z - x\|$ and $z \neq x$ then*

$$\frac{z - x}{\|z - x\|} \ = \ \frac{y - x}{\|y - x\|}.$$

Remark 2.3. In particular x, y and z are on the same line and $z \in [x, y]$ or $y \in [x, z]$.

Proof. Without loss of generality we may assume that $x = 0$. Since $u \in Lip_1(\Omega, \|\cdot\|)$, we infer that

$$\forall t \in [0, 1], \qquad u(0) \ = \ u(tz) + t\,\|z\|.$$

Since u is differentiable at 0, we then get $\nabla u(0) \cdot z = -\|z\|$. On the other hand, for any $z' \neq 0$ one also has $\nabla u(0) \cdot z' \ge -\|z'\|$. As a consequence, $-\nabla u(0)$ belongs to the normal cone of the closed convex set $K := \{z' : \|z'\| \le 1\}$ at $\frac{z}{\|z\|}$.

Since $(x, y) \in \operatorname{supp}(\gamma)$ and u is a Kantorovich potential, $-\nabla u(0)$ also belongs to the normal cone of K at $\frac{y}{\|y\|}$. Since K is strictly convex and has nonempty interior, the intersection of the normal cones to two of its boundary points is empty unless they coincide, so that we get $\frac{z}{\|z\|} = \frac{y}{\|y\|}$. □

Another crucial property of optimal transport plans is the cyclical monotonicity relative to the cost under consideration: we shall state this in a more general setting to handle the secondary transport problem of the next section.

Definition 2.4. Let $c : \Omega^2 \to [0, +\infty]$. A transport plan $\gamma \in \Pi(\mu, \nu)$ is cyclically monotone for the cost c (or c-cyclically) monotone if it is concentrated on a set C such that

$$\sum_{i=1}^{n} c(x_i, y_i) \leq \sum_{i=1}^{n} c(x_i, y_{\sigma(i)})$$

for all $n \geq 2$, $(x_1, y_1), \ldots, (x_n, y_n) \in C$ and any permutation σ of $\{1, \ldots, n\}$.

The following proposition gives a necessary condition for optimality in terms of cyclical monotonicity; for a proof, we refer to Theorem 3.2 in [3].

Theorem 2.5. *Let $c : \Omega^2 \to [0, +\infty]$ be a lower semicontinuous cost function, and assume that the infimum of the corresponding transport problem is finite:*

$$\inf \left\{ \int_{\Omega \times \Omega} c(x, y) d\lambda \, : \, \lambda \in \Pi(\mu, \nu) \right\} < +\infty$$

If γ is an optimal transport plan for this problem, then there exists a c-cyclically monotone Borel set C on which c is finite and γ is concentrated.

Remark 2.6. Duality and sufficiency of cyclical monotonicity may be pursued in very general settings [24, 3, 28, 23, 7], however for the purpose of this paper duality may be obtained more easily and we refer the reader to [1, 31].

3. Secondary transport problem to select monotone transport plans

Following the line of [2], we introduce a secondary transport problem to select optimal transport plans for (1.3) which have some more regularity: in the next sections, we shall prove that these particular optimal transport plans are induced by transport maps. The idea that a secondary variational problem may help to choose "more regular" or particular minimizers is the root of the idea of asymptotic development by Γ-convergence (see [4] and [5]) .

We denote by $\mathcal{O}_1(\mu, \nu)$ the set of optimal transport plans for (1.3), and fix a Kantorovich transport potential $\overline{u}$, i.e. a maximizer of (2.1). Let us define the new cost function

$$\beta(x, y) := \begin{cases} |x - y|^2 & \text{if } \overline{u}(x) = \overline{u}(y) + \|x - y\|, \\ +\infty & \text{otherwise.} \end{cases} \tag{3.1}$$

We then consider the following transport problem:

$$\min \left\{ \int_{\Omega \times \Omega} \beta(x, y) d\lambda(x, y) \ : \ \lambda \in \Pi(\mu, \nu) \right\}. \tag{3.2}$$

Because of the characterization of the minimizers for (1.3) given in Theorem 2.1, it appears that the above problem may be rewritten as

$$\min \left\{ \int_{\Omega \times \Omega} \beta(x, y) d\lambda(x, y) \ : \ \lambda \in \mathcal{O}_1(\mu, \nu) \right\}.$$

In other words, the problem (3.2) consists in minimizing the new cost functional $\lambda \mapsto \int \beta d\lambda$ among the minimizers of problem (1.3), and in this sense it may be considered as a secondary variational problem.

Definition 3.1. We shall denote by $\mathcal{O}_2(\mu,\nu)$ the minimizers for (3.2).

By Theorem 2.5, the set $\mathcal{O}_2(\mu,\nu)$ is non-empty and any of its elements enjoy the additional property of being concentrated on a set which is also β-cyclically monotone. This implies the following monotonicity, whose proof is derived from that of Lemma 4.1 in [2].

Proposition 3.2. *Let γ be a minimizer of problem (3.2). Then γ is concentrated on a σ-finite set Γ with the following property:*

$$\forall (x,y),(x',y') \in \Gamma, \qquad x \in [x',y'] \ \Rightarrow\ (x-x')\cdot(y-y') \geq 0 \tag{3.3}$$

where $\cdot$ denotes the usual scalar product on $\mathbb{R}^d$.

Proof. Applying Theorem 2.5, we get that γ is concentrated on a β-cyclically monotone Borel set Γ on which β is finite. Up to removing a γ-negligible set from Γ, we may assume that Γ is σ-finite.

Let $(x,y),(x',y') \in \Gamma$ be such that $x \in [x',y']$. Since γ is optimal for (1.3) and $\overline{u}$ is a Kantorovich potential for (2.1) we deduce that

$$\overline{u}(x) = \overline{u}(y) + \|x-y\| \qquad \textit{and} \qquad \overline{u}(x') = \overline{u}(y') + \|x'-y'\|.$$

Since $x \in [x',y']$ we also have $\|x'-y'\| = \|x-x'\| + \|x-y'\|$, and then using the fact that $\overline{u} \in Lip_1(\Omega, \|\cdot\|)$ we have

$$\overline{u}(x') = \overline{u}(y') + \|x-x'\| + \|x-y'\| \geq \overline{u}(x) + \|x-x'\|$$

and then again since $\overline{u} \in Lip_1(\Omega, \|\cdot\|)$ we infer that the above inequality is an equality, so that

$$\overline{u}(x) = \overline{u}(y') + \|x-y'\| \qquad \textit{and} \qquad \overline{u}(x') = \overline{u}(x) + \|x-x'\|.$$

But then we also have $\overline{u}(x') = \overline{u}(y) + \|x-y\| + \|x-x'\|$ so that $\overline{u}(x') = \overline{u}(y) + \|y-x'\|$. It then follows that $\beta(x',y) = |x-y|^2$ and $\beta(x,y') = |x-y'|^2$. Since Γ is β-cyclically monotone, we conclude

$$|x-y|^2 + |x'-y'|^2 \ \leq\ |x-y'|^2 + |x'-y|^2$$

which is equivalent to $(x-x')\cdot(y-y') \geq 0$. □

Remark 3.3. The reason to deal with σ-compact sets Γ, in the above proposition as well as in the following, is that the projection $\pi^1(\Gamma)$ is also σ-compact, and in particular is a Borel set.

4. A property of transport plans

We begin by considering some general properties of transport plans. This section is independent of the transport problem (1.3), and the definitions and techniques detailed below are refinements of similar ones which were first applied in [12] in the framework of non-classical transportation problems involving cost functionals not in integral form.

Definition 4.1. Let $\gamma \in \Pi(\mu,\nu)$ be a transport plan and Γ a σ-compact set on which it is concentrated. For $y \in \Omega$ and $r > 0$ we define

$$\Gamma^{-1}(\overline{B(y,r)}) := \pi^1(\Gamma \cap (\Omega \times \overline{B(y,r)})).$$

In other words, $\Gamma^{-1}(\overline{B(y,r)})$ is the set of points whose mass is partially or completely transported to $\overline{B(y,r)}$ by the restriction of γ to Γ. We may justify this slight abuse of notations by the fact that γ should be thought of as a device that transports mass. Notice also that $\Gamma^{-1}(\overline{B(y,r)})$ is a σ-compact set.

Since this notion is important in the sequel, we recall that when A is $\mathcal{L}^d$-measurable, one has

$$\lim_{r\to 0} \frac{\mathcal{L}^d(A\cap B(x,r))}{\mathcal{L}^d(B(x,r))} \;=\; 1$$

for almost every x in A: we shall call such a point x a Lebesgue point of A, this terminology deriving from the fact that such a point may also be considered as a Lebesgue point of χ_A. In the following, we shall denote by $\mathrm{Leb}(A)$ the set of points $x \in A$ which are Lebesgue points of A.

Remark 4.2. In the definition of Lebesgue points, one may replace the open ball $B(x,r)$ by the set $x + rC$, where C is a convex neighborhood of 0.

The following Lemma, although quite simple, is an important step in the proof of Proposition 5.2 and Theorem 6.1 below. Its proof is a straightforward adaptation of that of Lemma 5.2 from [12] and we detail it for the convenience of the reader.

Lemma 4.3. *Let $\gamma \in \Pi(\mu,\nu)$ and Γ a σ-compact set on which it is concentrated. If we assume that $\mu << \mathcal{L}^d$, then γ is concentrated on a σ-compact set $R(\Gamma)$ such that for all $(x,y) \in R(\Gamma)$ the point x is a Lebesgue point of $\Gamma^{-1}(\overline{B(y,r)})$ for all $r > 0$.*

Proof. Let

$$A := \{(x,y)\in\Gamma \;:\; x \notin \mathrm{Leb}(\Gamma^{-1}(\overline{B(y,r)})) \text{ for some } r>0\};$$

we first intend to show that $\gamma(A) = 0$. To this end, for each positive integer n we consider a finite covering $\Omega \subset \bigcup_{i\in I(n)} B(y_i^n, r_n)$ by balls of radius $r_n := \frac{1}{2n}$. We notice that if $(x,y) \in \Gamma$ and x is not a Lebesgue point of $\Gamma^{-1}(\overline{B(y,r)})$ for some $r > 0$, then for any $n \geq \frac{1}{r}$ and y_i^n such that $|y_i^n - y| < r_n$ the point x belongs to $\Gamma^{-1}(\overline{B(y_i^n, r_n)})$ but is not a Lebesgue point of this set. Then

$$\pi^1(A) \;\subset\; \bigcup_{n\geq 1}\;\bigcup_{i\in I(n)} \Big(\Gamma^{-1}(\overline{B(y_i^n,r_n)}) \setminus \mathrm{Leb}(\Gamma^{-1}(\overline{B(y_i^n,r_n)}))\Big)\,.$$

Notice that the set on the right hand side has Lebesgue measure 0, and thus μ-measure 0. It follows that $\gamma(A) \leq \gamma(\pi^1(A)\times\Omega) = \mu(\pi^1(A)) = 0$.

Finally, since $\mathcal{L}^d(\pi^1(A)) = 0$, there exists a sequence $(U_k)_{k\geq 0}$ of open sets such that

$$\forall k\geq 0, \qquad \pi^1(A)\subset U_k \quad \text{and} \quad \lim_{k\to\infty} \mathcal{L}^d(U_k) = 0.$$

Then the set $R(\Gamma) \;:=\; \Gamma \bigcap (\bigcup_{k\geq 0} (\Omega\setminus U_k)\times\Omega)$ has the desired properties. □

The above Lemma yields us to introduce the following notion:

Definition 4.4. The couple $(x, y) \in \Gamma$ is a Γ-*regular point* if x is a Lebesgue point of $\Gamma^{-1}(\overline{B(y,r)})$ for any positive r.

Notice that any element of the set $R(\Gamma)$ of Lemma 4.3 is a Γ-regular point. Lemma 4.3 above therefore states that any transport plan Γ is concentrated on a Borel set consisting of regular points: this regularity property turns out to be a powerful tool in the study of the support of optimal transport plans for problem (1.3), as the proof of Proposition 5.2 below illustrates.

5. A property of optimal transport plans

In this section, we obtain a regularity result on the transport plans that are optimal for problem (1.3). Following the formalism of [3], we first introduce the notions of transport set related to a subset Γ of $\mathbb{R}^d \times \mathbb{R}^d$.

Definition 5.1. Let Γ be a subset of $\mathbb{R}^d \times \mathbb{R}^d$, the transport set $T(\Gamma)$ of Γ is

$$\{(1-t)x + ty \mid (x, y) \in \Gamma,\ t \in (0,1)\}.$$

Notice that if Γ is σ-compact then $T(\Gamma)$ is also σ-compact.

The following Proposition 5.2 gives a regularity property for optimal transport plans for (1.3) in the case where $\|\cdot\|$ is a strictly convex norm. This property is obtained using two principle ingredients. The first is the fact that an optimal transport plan is concentrated on a set of regular points (see Lemma 4.3). The second ingredient relies on the property of the Kantorovich potentials stated in Lemma 2.2 which allow to derive a density estimate on the transport rays. This estimate is close to that stated in Lemma 5.4 of [6] (see also [8]) for the transport potential of the variational problem studied therein.

Let us introduce some notations: let $x, y \in \mathbb{R}^d$ with $x \neq y$, we then denote by P_{xy} the orthogonal projection on the line xy passing through x and y with respect to the Euclidean norm. For $\Delta, t_1, t_2 \in \mathbb{R}$ with $\Delta > 0$ and $t_1 < t_2$ we then define the following portion of cylinder with axis xy:

$$Q(x,y,t_1,t_2,\Delta) \ := \ \left\{ z \in \mathbb{R}^d \ : \ (P_{xy}(z) - x) \cdot \frac{(y-x)}{|y-x|} \in [t_1, t_2] \ \textit{and} \ |z - P_{xy}(z)| \leq \Delta \right\}.$$

We can now state the following regularity result.

Proposition 5.2. *Assume that $\|\cdot\|$ is a strictly convex norm and $\mu << \mathcal{L}^d$. Let also $\gamma \in \Pi(\mu,\nu)$ be an optimal transport plan for problem (1.3) and Γ a σ-compact set on which γ is concentrated. Then γ is concentrated on a σ-compact subset $R_T(\Gamma)$ of $R(\Gamma)$ such that for any $(x,y) \in R_T(\Gamma)$ with $x \neq y$ and for $r > 0$ small enough it holds*

$$\liminf_{\delta \to 0^+} \frac{\mathcal{L}^d\left(T\left(\Gamma \cap Q_{-\delta,r}(x,y) \times \overline{B(y,r)}\right) \cap Q_{+\delta,r}(x,y)\right)}{\mathcal{L}^d(Q_{+\delta,r}(x,y))} \ > \ 0 \qquad (5.1)$$

where for any $\delta > 0$ we set

$$Q_{-\delta,r}(x,y) \ := \ Q(x,y,-\delta,-\frac{\delta}{2}, r\delta) \quad \textit{and} \quad Q_{+\delta,r}(x,y) \ := \ Q\left(x,y,0,\delta,r\delta\Delta_r\right).$$

with $\Delta_r := 1 + \frac{2}{|y-x|}$.

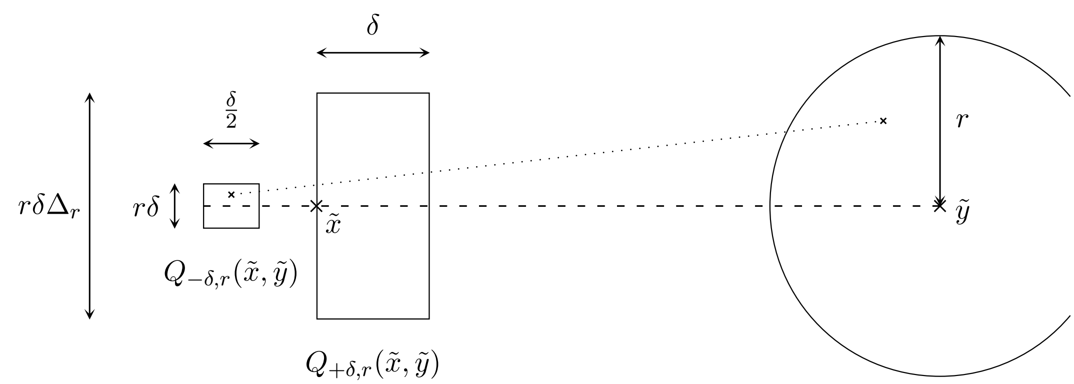

Figure 1

Proof. Step 1: definition of $R_T(\Gamma)$. Let $u \in Lip_1(\Omega, \|\cdot\|)$ be a Kantorovich potential for problem (1.3), and denote by $\mathrm{Diff}(u)$ the set of points of differentiability of u. Since u is Lipschitz continuous in Ω, $\mathrm{Diff}(u)$ has full Lebesgue measure in Ω, so that there exists a sequence $(U_k)_{k\geq 0}$ of open subsets of Ω such that

$$\forall k \geq 0, \qquad (\Omega \setminus U_k) \subset \mathrm{Diff}(u) \quad \text{and} \quad \lim_{k\to\infty} \mathcal{L}^d(U_k) = 0.$$

We set

$$A \ := \ R(\Gamma) \cap \bigcup_{k\geq 0} (\Omega \setminus U_k) \times \Omega.$$

and notice that A is a σ-compact set which has full measure for γ. In particular, $\pi^1(A)$ is also σ-compact and it has full measure for μ. Since $\mathcal{L}^d(\pi^1(A) \setminus \mathrm{Leb}(\pi^1(A))) = 0$, there exists a sequence $(V_k)_{k\geq 0}$ of open subsets of Ω such that

$$\forall k \geq 0, \qquad (\pi^1(A) \setminus \mathrm{Leb}(\pi^1(A))) \subset V_k \quad \text{and} \quad \lim_{k\to\infty} \mathcal{L}^d(V_k) = 0.$$

We may now define

$$R_T(\Gamma) \ := \ A \cap \bigcup_{k\geq 0} (\Omega \setminus V_k) \times \Omega.$$

Then $R_T(\Gamma)$ is a σ-compact set which is included in $R(\Gamma)$ and has full measure for γ. Moreover, notice that if $(x, y) \in R_T(\Gamma)$ then $x \in \mathrm{Diff}(u)$ and x is a Lebesgue point of $\pi^1(R_T(\Gamma))$.

We shall prove that the set $R_T(\Gamma)$ has the desired property.

Step 2: reduction of the proof. In the following, $(\tilde{x}, \tilde{y})$ is an element of $R_T(\Gamma)$ with $\tilde{x} \neq \tilde{y}$, and we aim to show that for $r > 0$ small enough it holds

$$\liminf_{\delta\to 0^+} \frac{\mathcal{L}^d\left(T\left(\Gamma \cap Q_{-\delta,r}(\tilde{x},\tilde{y}) \times \overline{B(\tilde{y},r)}\right) \cap Q_{+\delta,r}(\tilde{x},\tilde{y})\right)}{\mathcal{L}^d(Q_{+\delta,r}(\tilde{x},\tilde{y}))} \ > \ 0 \tag{5.2}$$

Without loss of generality we may assume that $\tilde{x} = 0$ and $\frac{\tilde{y}-\tilde{x}}{|\tilde{y}-\tilde{x}|} = \frac{\tilde{y}}{|\tilde{y}|} = e_1$ is the first vector of the canonical Euclidean basis of $\mathbb{R}^d$. If for $s > 0$ we denote by $\overline{B^{d-1}(0,s)}$ the closed ball of $\mathbb{R}^{d-1}$ of center 0 and radius s, we can rewrite

$$Q_{-\delta,r}(\tilde{x},\tilde{y}) \;=\; [-\delta,-\frac{\delta}{2}]\times\overline{B^{d-1}(0,r\delta)} \quad\text{and}\quad Q_{+\delta,r}(\tilde{x},\tilde{y}) \;=\; [0,\delta]\times\overline{B^{d-1}(0,r\delta\Delta_r)}.$$

Fix $r>0$ and $s>0$ small enough so that

$$\eta := \inf\{|y-x| \;:\; x\in[-s,s]\times\overline{B^{d-1}(0,rs\Delta_r)},\, y\in\overline{B(\tilde{y},r)}\} \;>\; 0. \tag{5.3}$$

Since $(0,\tilde{y})\in\Gamma$, 0 is a Lebesgue point of $\Gamma^{-1}(\overline{B(\tilde{y},r)})$. Since 0 is also a Lebesgue point of $\pi^1(R_T(\Gamma))$, we infer that it is a Lebesgue point of the σ-compact set $\mathcal{R} := \Gamma^{-1}(\overline{B(\tilde{y},r)})\cap\pi^1(R_T(\Gamma))$. It then follows from the Fubini theorem, the definition of Lebesgue points and remark 4.2 that for $\delta\in]0,s[$ small enough one has

$$\mathcal{L}^1\left(\left\{t\in[-\delta,\delta] \;:\; \mathcal{H}^{d-1}(\mathcal{R}\cap\{t\}\times B^{d-1}(0,r\delta))\geq\frac{1}{2}(r\delta)^{d-1}\omega_{d-1}\right\}\right)\geq\frac{8}{5}\delta$$

where $\omega_{d-1} = \mathcal{L}^{d-1}(B^{d-1}(0,1))$. We fix such a small enough $\delta\in]0,s[$, and choose $t_\delta\in[-\delta,-\frac{\delta}{2}]$ such that

$$\mathcal{H}^{d-1}(\mathcal{R}\cap\{t_\delta\}\times B^{d-1}(0,r\delta))\geq\frac{1}{2}(r\delta)^{d-1}\omega_{d-1}.$$

We finally take a compact subset $\mathcal{R}_\delta$ of $\mathcal{R}\cap\{t_\delta\}\times B^{d-1}(0,r\delta)$ such that $\mathcal{H}^{d-1}(\mathcal{R}_\delta)\geq\frac{1}{4}(r\delta)^{d-1}\omega_{d-1}$ and we shall now obtain a lower bound for

$$\mathcal{L}^d\left(T(\Gamma\cap\mathcal{R}_\delta\times\overline{B(\tilde{y},r)})\cap Q_{+\delta,r}(0,\tilde{y})\right).$$

Step 3: an approximation for $T(\Gamma\cap\mathcal{R}_\delta\times\overline{B(\tilde{y},r)})$ on $Q_{+\delta,r}(0,\tilde{y})$. Let $\{y_k\}_{k\geq0}$ be a dense sequence in $\overline{B(\tilde{y},r)}$, then for $x\in\Omega$ and $N\geq0$ we set

$$M_N(x) := \left\{k\in\{0,\ldots,N\} \;:\; u(y_k)+\|x-y_k\| = \min_{0\leq j\leq N}\{u(y_j)+\|x-y_j\|\}\right\}.$$

We now consider

$$C_{\delta,N} := \bigcup_{k=0}^{N}\{(x,y_k) \;:\; x\in\mathcal{R}_\delta \text{ and } k\in M_N(x)\}.$$

Notice that $C_{\delta,N}$ is a compact set and that $\pi^1(C_{\delta,N})=\mathcal{R}_\delta$. We finally set

$$L := Q_{+\delta,r}(0,\tilde{y}) \,\cap\, \bigcap_{K\geq0}\overline{\bigcup_{N\geq K}T(C_{\delta,N})}$$

and we claim that $L\subset T(\Gamma\cap\mathcal{R}_\delta\times\overline{B(\tilde{y},r)})\cap Q_{+\delta,r}(0,\tilde{y})$. Indeed let $x\in L$, then there exists $x'\in\mathcal{R}_\delta$ and $z'\in\overline{B(\tilde{y},r)}$ such that $x\in[x',z']$ and

$$u(z')+\|x'-z'\| \;=\; \inf_{k\geq0}\{u(y_k)+\|x'-y_k\|\} \;=\; \min_{y\in\overline{B(\tilde{y},r)}}\{u(y)+\|x'-y\|\}.$$

Since $x'\in\mathcal{R}_\delta\subset\Gamma^{-1}(\overline{B(\tilde{y},r)})$, we infer that there exists $y'\in\overline{B(\tilde{y},r)}$ such that $(x',y')\in\Gamma$. As a consequence, one has

$$u(x') \;=\; u(y')+\|x'-y'\| \;=\; \min_{y\in\overline{B(\tilde{y},r)}}\{u(y)+\|x'-y\|\}.$$

We thus obtain that $u(x') = u(z') + \|x' - z'\|$ and we conclude from $\mathcal{R}_\delta \subset \mathrm{Diff}(u)$ and Lemma 2.2 that either $z' \in [x', y']$ or $y' \in [x', z']$. Therefore z' belongs to the line passing through x' and y', and then by (5.3) we get that x belongs to $[x', y']$ and thus to $T(\Gamma \cap \mathcal{R}_\delta \times \overline{B(\tilde{y}, r)}) \cap Q_{+\delta,r}(0, \tilde{y})$.

Step 4: a lower bound on $\mathcal{L}^d(T(C_{\delta,N}) \cap Q_{+\delta,r}(0,\tilde{y}))$. Fix $N \geq 0$, and define $k \in \{0, \dots, N\}$ the Borel set

$$D_k := \{x \in \mathcal{R}_\delta \,:\, k = \min\{j \,:\, j \in M_N(x)\}\}\,.$$

For any $k \in \{0, \dots, N\}$ the cone $T(D_k \times \{y_k\})$ with basis D_k and vertex y_k is included in $T(C_{\delta,N})$. We claim that these cones do not overlap:

$$k \neq l \quad \Rightarrow \quad T(D_k \times \{y_k\}) \cap T(D_l \times \{y_l\}) = \emptyset.$$

We argue by contradiction and assume that for some $k < l$, $x_k \in D_k$ and $x_l \in D_l$ there exists $z \in [x_k, y_k] \cap [x_l, y_l]$. Then it follows from the definitions of D_k that

$$u(y_k) + \|x_k - y_k\| \leq u(y_l) + \|x_k - y_l\|$$

and from $k < l$ and the definition of D_l that

$$u(y_l) + \|x_l - y_l\| < u(y_k) + \|x_l - y_k\|.$$

We now compute

$$\begin{aligned}
u(y_k) + \|z - y_k\| &= u(y_k) + \|x_k - y_k\| - \|x_k - z\| \\
&\leq u(y_l) + \|x_k - y_l\| - \|x_k - z\| \\
&\leq u(y_l) + \|z - y_l\| \;=\; u(y_l) + \|x_l - y_l\| - \|x_l - z\| \\
&< u(y_k) + \|x_l - y_k\| - \|x_l - z\| \;\leq\; u(y_k) + \|z - y_k\|
\end{aligned}$$

which is a contradiction and proves the claim.

We infer from the choice of Δ_r t(see Figure 1) hat

$$T(D_k \times \{y_k\}) \cap [0, \delta] \times \mathbb{R}^{d-1} \;\subset\; Q_{+\delta,r}(0, \tilde{y})$$

and then we get from (5.3) the following estimate:

$$\forall k \in \{0, \dots, N\}, \qquad \mathcal{L}^d(T(D_k \times \{y_k\}) \cap Q_{+\delta,r}(0, \tilde{y}))) \;\geq\; \delta\, \frac{\eta}{\eta + 2s}\, \mathcal{H}^{d-1}(D_k).$$

Since the cones $T(D_k \times \{y_k\})$ do not overlap, we obtain from the preceding that

$$\mathcal{L}^d(T(C_{\delta,N}) \cap Q_{+\delta,r}(0, \tilde{y})) \;\geq\; \delta\, \frac{\eta}{\eta + 2s} \sum_{k=0}^{N} \mathcal{H}^{d-1}(D_k) \;=\; \delta\, \frac{\eta}{\eta + 2s}\, \mathcal{H}^{d-1}(\mathcal{R}_\delta)$$

and thus

$$\mathcal{L}^d(T(C_{\delta,N}) \cap Q_{+\delta,r}(0, \tilde{y})) \;\geq\; \frac{\eta}{4(\eta + 2s)}\, r^{d-1}\, \delta^d\, \omega_{d-1}. \tag{5.4}$$

Step 5. We now conclude the proof by noticing that

$$L \;=\; \bigcap_{K \geq 0} \overline{\bigcup_{N \geq K} T(C_{\delta,N}) \cap Q_{+\delta,r}(0, \tilde{y})}$$

so that

$$\mathcal{L}^d(L) \;\geq\; \frac{\eta}{4(\eta + 2s)}\, \frac{1}{\Delta_r^{d-1}}\, \mathcal{L}^d(Q_{+\delta,r}(0, \tilde{y})).$$

We then infer from $L \subset T(\Gamma \cap \mathcal{R}_\delta \times \overline{B(\tilde{y}, r)}) \cap Q_{+\delta,r}(0, \tilde{y})$ that (5.2) holds. □

Remark 5.3. In the above proof, we only use the strict convexity of the norm $\|\cdot\|$ to apply Lemma 2.2.

6. Proof of the main theorem

Now we are in position to prove Theorem 1.1 which is, in fact, a corollary of the following more precise result.

Theorem 6.1. *Assume that the norm $\|\cdot\|$ is strictly convex and $\mu << \mathcal{L}^d$. Then for every $\gamma \in \Pi(\mu,\nu) \cap \mathcal{O}_2(\mu,\nu)$ there exists a map $T_\gamma \in \mathcal{T}(\mu,\nu)$ such that $\gamma = (id \times T_\gamma)_\sharp \mu$. Moreover, the solution $\gamma \in \Pi(\mu,\nu) \cap \mathcal{O}_2(\mu,\nu)$ is unique.*

Proof. By Proposition 2.1 in [1], it is sufficient to prove that γ is concentrated on a Borel graph.

It follows from Proposition 3.2 that γ is concentrated on a σ-compact set Γ satisfying (3.3). We then apply Proposition 5.2 to get that γ is concentrated on a σ-compact subset $R_T(\Gamma)$ of $R(\Gamma)$ satisfying (5.1).

We claim that $R_T(\Gamma)$ is a contained in a graph. To prove this, we show that if (x_0, y_0) and (x_0, y_1) both belong to $R_T(\Gamma)$ then $y_0 = y_1$. We argue by contradiction, and then we assume that $y_1 \neq y_0$. As a consequence, one either has $(y_1 - y_0) \cdot (y_0 - x_0) < 0$ or $(y_0 - y_1) \cdot (y_1 - x_0) < 0$. Without loss of generality, we assume that

$$(y_1 - y_0) \cdot (y_0 - x_0) \ < \ 0.$$

We fix $r > 0$ small enough so that

$$\forall x \in Q_{+r,r}(x_0, y_0),\ \forall y' \in \overline{B(y_0, r)},\ \forall y \in \overline{B(y_1, r)}, \qquad (y - y') \cdot (y' - x) \ < \ 0. \tag{6.1}$$

Since $(x_0, y_1) \in R_T(\Gamma)$, we infer that x_0 is a Lebesgue point for $\Gamma^{-1}(\overline{B(y_1, r)})$. Moreover, we also get from $(x_0, y_0) \in R_T(\Gamma)$ and (5.1) that

$$\liminf_{\delta \to 0^+} \frac{\mathcal{L}^d \left(T \left(\Gamma \cap Q_{-\delta,r}(x_0, y_0) \times \overline{B(y_0, r)} \right) \cap Q_{+\delta,r}(x_0, y_0) \right)}{\mathcal{L}^d(Q_{+\delta,r}(x_0, y_0))} \ > \ 0$$

As a consequence, for $\delta \in]0, r[$ small enough there exists (x', y') and (x, y) in Γ such that

$$x' \in Q_{-\delta,r}(x_0, y_0), \quad y' \in \overline{B(y_0, r)}, \quad x \in [x', y'] \cap Q_{+\delta,r}(x_0, y_0) \quad \textit{and} \quad y \in \overline{B(y_1, r)}.$$

It follows from (3.3) applied to (x', y') and (x, y) that

$$(y - y') \cdot (x - x') \geq 0$$

but since $x \in [x', y']$ one also has $x - x' = \frac{|x - x'|}{|y' - x|}(y' - x)$ and we get a contradiction with (6.1).

The uniqueness of $\gamma \in \Pi(\mu,\nu) \cap \mathcal{O}_2(\mu,\nu)$ is obtained as in Step 5 of the proof of Theorem B in [2]: if γ_1 and γ_2 are two such transport plans, then $\frac{\gamma_1 + \gamma_2}{2}$ also belongs to $\Pi(\mu,\nu) \cap \mathcal{O}_2(\mu,\nu)$. It follows from the preceding that these plans are all induced by transport maps, which then coincide μ almost everywhere.

□

7. Norms which are not strictly convex and further remarks

It is remarkable in the precedind proofs that the strict convexity assumption on the norm $\|\cdot\|$ is only used through Lemma 2.2: as explained in the introduction of [2], the direction of transportation is totally detemined at any point of differentiability of a Kantorovich potential u when the the norm $\|\cdot\|$ is strictly convex, and this information is sufficient to conclude in the proof of 5.2. Without this assumption, the optimality of the transport plan γ is not enough to obtain the density property of Proposition 5.2. This is shown by the following example constructed in [2]:

Theorem 7.1 (Theorem A of [2]). *There exist a Borel set $M \subset [-1,1]^3$ with $|M| = 8$ and two Borel maps $f_i : M \to [-2,2] \times [-2,2]$ for $i = 1,2$ such that the following holds. For $x \in M$ denote by l_x the segment connecting $(f_1(x), -2)$ to $(f_2(x), 2)$ then*

1. *$\{x\} = l_x \cap M$ for all $x \in M$,*
2. *$l_x \cap l_y = \emptyset$ for all $x, y \in M$ different.*

To give a counterexample to Proposition 5.2 without the assumption of strict convexity of $\|\cdot\|$, consider the map

$$T(x) := (f_2(x), 2)$$

and observe that, for the norm $\|(x,y,z)\| := \max\{|x|, |y|, 3|z|\}$, the map T is an optimal transport map for (1.2) betwen $\mu = \mathcal{L}^d\lfloor M$ and $\nu = T_\sharp \mu$. However, the open transport set $T(\mathrm{supp}((id \times T)_\sharp \mu))$ has density 0 at every point of M.

A significant quantity related to the transport set is the so called transport density, i.e. a positive measure σ which solves together with any transport potential the system of PDEs

$$\begin{cases} -div(\sigma Du) = \mu - \nu \\ \|Du\|^* = 1 \qquad\qquad \sigma - a.e.. \end{cases} \tag{7.1}$$

The relationship between the transport density and the Monge-Kantorovich problem is given by the following formula first discovered in [9]. Let γ be an optimal transport plan, and for each Borel set $B \subset \Omega$ consider

$$\sigma_\gamma(B) := \int_{\Omega\times\Omega} \mathcal{H}^1(B \cap [x,y]))d\gamma(x,y),$$

then σ_γ is a solution of (7.1) above. Clearly σ_γ is supported on the transport set $T(\mathrm{supp}(\gamma))$. In practical terms the measure $\sigma_\gamma(D)$ of a set D represents the work done in the set D while transporting μ to ν following the plan γ. A detailed discussion of the properties of such measures is beyond the scope of this paper. The transport density plays a crucial role in the proof of existence given in [18] and good estimates from above are available for σ_γ [1, 14, 13, 15]. Proving some estimate from below for σ_γ could be interesting for the approach of this paper. In fact, assume for example that σ_γ has an L^∞ density a_γ (see for example [14, 18]) and that at a point x one has $0 < a_\gamma(x)$. Then the lower density of the transport set $T(\gamma)$ at x satisfies $\theta_*(T(\mathrm{supp}(\gamma)), x) > 0$ because

$$a_\gamma(x) = \lim_{r\to 0} \frac{1}{\omega_d r^d} \int_{B(x,r)} a_\gamma(y)dy \leq \liminf_{r\to 0} \|a_\gamma\|_\infty \frac{|T(\mathrm{supp}(\gamma)) \cap B(x,r)|}{\omega_d r^d}.$$

Because of the above example, we however can not expect an estimate from below on σ_γ for any solution γ of (1.3), but this may hold for example for an element of $\mathcal{O}_2(\mu,\nu)$.

Acknowledgments

References

[1] Ambrosio, L., Lecture Notes on Optimal Transportation Problems, Mathematical aspects of evolving interfaces (Funchal, 2000), Lecture Notes in Math., **1812**, Springer, Berlin, 2003, 1–52.

[2] Ambrosio, L., Kirchheim, B., Pratelli, A., *Existence of optimal transport maps for crystalline norms*, Duke Math. J., **125** (2004), no. 2, 207–241.

[3] Ambrosio, L., Pratelli, A.,, Existence and stability results in the L^1 theory of optimal transportation, Optimal transportation and applications (Martina Franca, 2001), Lecture Notes in Math., **1813**, Springer, Berlin, 2003, 123–160.

[4] Anzellotti, G., Baldo, S., *Asymptotic development by* Γ*-convergence*, Appl. Math. Optim. **27** (1993), no. 2, 105–123.

[5] Attouch, H., *Viscosity solutions of minimization problems*, SIAM J. Optim. **6** (1996), no. 3, 769–806.

[6] Bianchini, S., *On the Euler-Lagrange equation for a variational problem*, Discrete Contin. Dyn. Syst., **17** (2007), no. 3, 449–480.

[7] Bianchini, S., Caravenna, L. Talk given by L.Caravenna in Pisa, november 2007. See also Bibliographical Notes to chapter 5 of [32]

[8] Bianchini, S., Gloyer, M. *On the Euler-Lagrange equation for a variational problem: The general case II*, Preprint SISSA (2008) available at http://digitallibrary.sissa.it/index.jsp

[9] Bouchitté, G., Buttazzo, G. *Characterization of optimal shapes and masses through Monge–Kantorovich equation*, Journal European Math. Soc., **3** (2001), 139–168.

[10] Benamou, J. D., Brenier, Y., *A computational fluid mechanics solution to the Monge–Kantorovich mass transfer problem*, Numer. Math., **84** (2000), 375–393.

[11] Caffarelli, L.A., Feldman, M., McCann, R.J., *Constructing optimal maps for Monge's transport problem as a limit of strictly convex costs*, J. Amer. Math. Soc. **15** (2002), no. 1, 1–26.

[12] Champion, T. De Pascale, L., Juutinen, P. *The* ∞*-Wasserstein distance: local solutions and existence of optimal transport maps*, SIAM J. of Mathematical Analysis **40** (2008), no. 1, 1-20.

[13] De Pascale,L., Evans, L.C., Pratelli, A., *Integral Estimates for Transport Densities*, Bulletin of the London Mathematical Society, **36** (2004), no.3, 383-396.

[14] De Pascale, L., Pratelli, A., *Regularity properties for Monge transport density and for solutions of some shape optimization problem*, Calc. Var. Partial Differ. Equ., **14** (2002), no. 3, 249–274.

[15] De Pascale, L., Pratelli, A. *Interpolation and sharp summability for Monge Transport density*, ESAIM Control, Optimization and Calculus of Variations, **10** (2004), no. 4, 549-552.

[16] Ekeland, I., Temam, R., Convex Analysis and Variational Problems. North-Holland Publishing Company-Amsterdam (1976).

[17] Evans, L. C., *Partial Differential Equations and Monge–Kantorovich Mass Transfer*, Current developments in mathematics, 1997 (Cambridge, MA), Int. Press, Boston, MA, (1999), 65–126.

[18] Evans, L. C., Gangbo, W., *Differential Equations Methods for the Monge–Kantorovich Mass Transfer Problem*, Mem. Amer. Math. Soc., Vol. **137** (1999).

[19] Gangbo, W., McCann, R. J., *The geometry of optimal transportation*, Acta Math., **177** (1996), 113–161.

[20] Kantorovich, L.V., *On the translocation of masses*, C.R. (Dokl.) Acad. Sci. URSS, **37** (1942), 199-201.

[21] Kantorovich, L.V., *On a problem of Monge (in Russian)*, Uspekhi Mat. Nauk., **3** (1948), 225-226.

[22] Monge, G., *Mémoire sur la théorie des Déblais et des Remblais*, Histoire de l'Académie des Sciences de Paris, 1781.

[23] Pratelli, A., *On the sufficiency of c-cyclical monotonicity for optimality of transport plans*, Math. Z., **258** (2008), no. 3, 677–690

[24] Rachev, S., Rüschendorf, L., Mass transportation problems. Vol. I. Theory. Probability and its Applications, Springer-Verlag, New York (1998).

[25] Rockafellar, R.T., Convex analysis, Princeton University Press, Princeton, N. J. (1970).

[26] Rüschendorf, L., *Optimal solutions of multivariate coupling problems*, Appl. Math. (Warsaw) 23 (1995), no. 3, 325–338.

[27] Rüschendorf, L., *On $c-$optimal random variables*, Statistic & Probability letters, **27** (1996), 267–270.

[28] Schachermayer, W., and Teichmann, J., *Characterization of optimal Transport Plans for the Monge-Kantorovich-Problem*, Proc. Amer. Math. Soc., to appear.

[29] Sudakov, V. N., Geometric problems in the theory of infinite-dimensional probability distributions. Cover to cover translation of Trudy Mat. Inst. Steklov 141 (1976). Proc. Steklov Inst. Math. 1979, no. 2, i–v, 1–178.

[30] Trudinger, N.S., Wang, X.J., *On the Monge mass transfer problem*, Calc. Var. Partial Differential Equations **13** (2001), no. 1, 19–31.

[31] Villani, C., Topics in optimal transportation. Graduate Studies in Mathematics, 58, American Mathematical Society (2003)

[32] Villani, C., Optimal Transport, Old and New. Available at http://www.umpa.ens-lyon.fr/ cvillani/surveys.html#oldnew

T.C. Institut de Mathématiques de Toulon et du Var, U.F.R. des Sciences et Techniques, Université du Sud Toulon-Var, Avenue de l'Université, BP 20132, 83957 La Garde cedex, FRANCE

L.D.P. Dipartimento di Matematica Applicata, Universitá di Pisa, Via Buonarroti 1/c, 56127 Pisa, ITALY