# Existence and uniqueness of optimal maps on Alexandrov spaces

Jérôme Bertrand

**Abstract**

The purpose of this paper is to show that in a finite dimensional metric space with Alexandrov's curvature bounded below, Monge's transport problem for the quadratic cost admits a unique solution.

## 1 Introduction

In this paper, we provide a solution to the Monge-Kantorovich problem in Alexandrov space when the cost function is the square of the metric. We begin by explaining what Monge's problem is, and briefly describing Alexandrov spaces. The mass transportation problem, raised by Monge in 1781 [14], is to move one distribution of mass onto another as efficiently as possible, where the criterion of efficiency is to minimise a certain cost. In the original formulation, the cost was the average distance covered by the mass. In other words, given a set $X$ endowed with a cost function, and given two probability measures $\mu_0, \mu_1$ on $X$ as constraints, the problem is to determine, if it exists, a minimizer of

$$\inf \int_X c(x, s(x))\, d\mu_0(x)$$

among all measurable maps $s$ sending the initial measure $\mu_0$ onto $\mu_1$ (we denote this condition by $\mu_1 = s_\sharp \mu_0$).

Monge's original problem has only been solved quite recently (Sudakov gave a proof first [20] but later, part of his proof appeared to be incorrect, we refer to [3] for details). In the meantime, the importance of the mass transport problem for other cost functions was recognized and much work has been devoted to its study. One of the most interesting cases is that of the quadratic cost (the square of the metric distance). In Euclidean space, Monge's problem for the quadratic cost was solved in the eighties by Brenier [6] (and Knott-Smith [17] independently) who was motivated by fluid mechanics problems. Brenier showed the existence and the uniqueness of the optimal map, under the hypothesis that the initial measure is absolutely continous with respect to the Lebesgue measure. The optimal map is the gradient of a convex function on $\mathbb{R}^n$. Subsequently, this result was generalised to the setting of Riemannian manifolds by McCann [13], taking advantage of

a notion of generalised convex functions (so-called $c$-concave functions, see Definition 3.1 below).

A difficulty in solving Monge's problem is that the problem may be ill-posed; a measure preserving map between two distributions of mass may not exist. To avoid this problem, Kantorovich introduced a relaxed version of Monge's problem, which is to minimise the quantity

$$\int_{X\times X} c(x,y)\ d\Pi(x,y)$$

among all plans, namely all probability measures $\Pi$ whose marginals are $\mu_0$ and $\mu_1$ (in this case, the mass is allowed to split). This problem admits solutions in very general settings (see Section 3). A solution $s$ of Monge's problem for $\mu_0$ and $\mu_1$ induces a solution of Kantorovich's problem. More precisely, the plan defined by $(Id, s)_\sharp\mu_0$ is a solution. Classically, in the quadratic cost case, existence and uniqueness of the optimal map follow from the converse property: any optimal plan is actually induced by a map. In this paper, we implement this strategy for a class of metric spaces called Alexandrov spaces.

Alexandrov spaces are the natural generalization of Riemannian manifolds whose sectional curvature is bounded below (a precise definition is given in Section 2; we do note however that an Alexandrov space is not necessarily a Gromov-Hausdorff limit of Riemannian manifolds). In their seminal paper [8], Burago, Gromov, and Perelman showed that such (finite dimensional) metric spaces possess a certain kind of "Riemannian structure". This viewpoint was developed further by Otsu and Shioya [15]. We use their results to extend McCann's theorem to this setting. Our main result is the following

**Theorem 1.1** *Let $(X,d)$ be a finite dimensional Alexandrov space and $\mu_H$ be the corresponding Hausdorff measure. Let $\mu_0, \mu_1$ be probability measures on $X$ with compact supports such that $\mu_0$ is absolutly continuous with respect to $\mu_H$.*

*Under these assumptions, Kantorovich's problem 4.2 admits a solution, and any optimal plan is supported in the graph of a Borel function $F$. This map $F$ is also a minimizer of Monge's problem and satisfies for $\mu$ almost every $x \in X$,*

$$F(x) = exp(-\nabla\phi(x)),$$

*where $\phi$ is a $d^2$-concave function (see Definition 3.1).*

*Moreover, up to modifications on negligeable sets, the map $\nabla\phi$ is unique, and hence so is the optimal map $F$.*

**Remark 1.2** *Observe that in the above theorem, the lower bound on the (Alexandrov) curvature does not appear explicitly in the statement. Consequently, our result also applies to any compact Riemannian manifold and allows us to give another proof of McCann's theorem [13]. The original proof, relying strongly on the regularity of the Riemannian exponential map, cannot be adapted to this case.*

To put our result in perspective, we end this section by citing known results about Monge's problem in a singular setting (all but one for $\mathbb{R}^n$, the other result is for a compact manifold). First, Gangbo and McCann [10] extended Brenier's result to the case of strictly convex cost functions (including the case of the square of an arbitrary norm). Ambrosio and Rigot treated the case of Heisenberg group [4]. Ambrosio, Kirchheim and Pratelli proved the existence of optimal maps for Monge's original problem for crystalline norms [2]. Very recently, Bernard and Buffoni proved the existence of optimal maps on Finsler manifolds with strictly convex norms [5]. We refer to the books [22, 16, 21] for more on optimal mass transport.

Let us mention another geometric motivation which led us to study the problem above. Recently, the notion of having bounded below Ricci curvature has been extended to the setting of metric measure spaces independently by Lott-Villani, and by Sturm [11, 12, 18, 19]. This equivalent definition uses optimal mass transport theory. For a Riemannian manifold, the proof of the equivalence between both definitions basically relies on McCann's theorem [13] and on a non-smooth change of variables formula [9]. A natural question (mentioned in [11]) is whether or not an Alexandrov space has Ricci curvature bounded below in this generalised sense. Our result may be useful in investigating this.

The rest of this paper is organised as follows. In the next section, we present properties of Alexandrov spaces, refering to the book [7] for a background of the theory. We then use these results in Section 3 to prove our main theorem. A sketch of our proof is given at the beginning of this section. In the last section, we indicate how to adapt our result to other strictly convex costs of the distance function and how to relax the compactness assumption on the supports of the given measures.

## 2 Properties of Alexandrov spaces

In this section, we summarize known properties of Alexandrov spaces used in the rest of the paper. These results are taken from a paper by Burago, Gromov, and Perelman [8] and a paper by Otsu and Shioya [15]. Most of the results of [8] and proofs can also be found in the book [7].

**Definition 2.1 (Alexandrov space)** *Let $S_k^2$ be the 2-dimensional space form of curvature $k$ and $\delta_k$ be the metric induced by the Riemannian metric. A finite dimensional Alexandrov space $X$ of curvature bounded below by $k$ is a complete, connected, locally compact, geodesic space such that for all geodesics $\gamma$ in $X$ and $\overline{\gamma}$ in $S_k^2$ such that their lengths are equal , $d(p,\gamma(0)) = \delta_k(\bar{p},\overline{\gamma}(0))$ and $d(p,\gamma(1)) = \delta_k(\bar{p},\overline{\gamma}(1))$, the inequality below is satisfied for all $t \in [0,1]$:*

$$d(p,\gamma(t)) \geq \delta_k(\bar{p},\overline{\gamma}(t)).$$

The lower bound on the curvature of an Alexandrov space allows one to define the angle between two geodesics starting at the same point. Using this property, we can prove the first variation formula for distance functions.

**Lemma 2.2 (First variation formula)** *Let $a$ and $x$ be two distinct points of an Alexandrov space $X$ and $\gamma$ be a unitary geodesic starting from $x$. Then, the equality*

$$d(a, \gamma(t)) = d(a, x) - t \cos \angle_{min} + o(t),$$

*holds for all nonnegative $t$ where $\angle_{min}$ is the smallest angle between $\gamma$ and a geodesic between $x$ and $a$.*

For a proof, we refer to [7, Corollary 4.5.7 and Remark 4.5.12] or [15, Theorem 3.5].

To prove our main result, we will use the fact that an Alexandrov space is not far from being a (Riemannian) manifold. We explain this point of view, starting with a definition.

**Definition 2.3** *Let $X$ be an $n$-dimensional Alexandrov space. A point $p \in X$ is said to be regular if the tangent cone at $p$ is isometric to Euclidean space and singular otherwise. Throughout the rest of this paper, we denote by $Reg(X)$ (respectively $Sing(X)$) the set of regular (respectively singular) points of $X$.*

**Remark 2.4** *The curvature bound implies that the tangent cone at each point is unique, and is isometric to the Euclidean cone over the space of directions at $x$. The Hausdorff dimension of any tangent cone is equal to the dimension of $X$. (See [7] for a proof.)*

We gather together the main properties of the regular set of an Alexandrov space in three theorems. The first one states that, in a certain sense, the regular set covers almost all of $X$. The second theorem shows there is a certain kind of differential structure on $Reg(X)$. The last establishes the existence of a Riemannan structure on $Reg(X)$ and the compatibility of this structure with the Alexandrov metric. The order in which these theorems are given below is not chronological; furthermore, we have mixed results proved by various authors. We hope that this non-standard presentation helps the readers' understanding.

**Theorem 2.5** *The subset $Reg(X)$ is a dense measurable set (because it is the intersection of countably many dense open sets) of full measure in $X$. More precisely, the Hausdorff dimension of the singular set satisfies $dim_H(Sing(X)) \leq n - 1$.*

**Proof :** The first statement was proved by Burago-Gromov-Perelman in [8], see also [7, Chapter 10]. The Hausdorff measure property was obtained by Burago-Gromov-Perelman [8] and Otsu-Shioya [15, Theorem A] independently. ■

**Theorem 2.6 (Charts on $Reg(X)$)** *There exists an atlas $(\phi, U_\phi)_{\phi\in\Phi}$ on $Reg(X)$. In other words, $\forall\phi \in \Phi$, $\phi : U_\phi \mapsto \mathbb{R}^n$ is a bilipschitz map defined on an open set $U_\phi$ of $X$. Moreover, for any $\phi \in \Phi$, there exists $V_\phi \subset U_\phi$, which is dense and of full measure such that*

$$\bigcup_{\phi\in\Phi} V_\phi \supset Reg(X).$$

*Now, if $\phi, \psi \in \Phi$ are such that $U_\phi \cap U_\psi \neq \emptyset$, $\psi \circ \phi^{-1}$ is continuously differentiable on $\phi(V_\phi \cap V_\psi \cap Reg(X))$ (see Definition 2.12 for a precise definition).*

**Proof :** This statement is proved in [15, Theorem 4.2]. ■

**Remark 2.7** *Notice however that $Sing(X)$ can be a dense subset of $X$; an example is provided in [15].*

**Theorem 2.8 (Riemannian structure on $Reg(X)$)**
*a) There exists a continuous Riemannian metric on $Reg(X)$, i.e. a family $(g_\phi)_{\phi\in\Phi}$ of maps such that*

$$g_\phi : U_\phi \longrightarrow Sym^+(\mathbb{R}^n)$$

*is a continuous map ($Sym^+(\mathbb{R}^n)$ denotes the set of symmetric positive definite matrices in $\mathbb{R}^n$), and these maps satisfy the usual formula*

$$g_\phi =^t (d(\phi \circ \psi^{-1}))\, g_\psi\, d(\phi \circ \psi^{-1}).$$

*b) The Riemannian structure is compatible with the Alexandrov metric, in the following sense:*
*i) For any $x \in Reg(X)$ any chart $\phi \in \Phi$ such that $x \in U_\phi$ and any $\delta > 0$, there exists a neighbourhood of $x$ in $U_\phi$ such that the map $\phi$ is a bilipschitz homeomorphism with Lipschitz constants smaller than $1 + \delta$ on this neighbourhood.*
*ii) The tangent cone based at a point $x \in Reg(X)$, endowed with the induced metric, is isometric to $(\mathbb{R}^n, g_\phi(x))$ (assuming that $x \in U_\phi$).*
*iii) The metric induced by the Riemannian metric coincides with the original metric.*

**Proof :** Statement b)i) was obtained by Burago, Gromov, and Perelman in [8], see also [7, Chapter 10] for a proof. The others were proved by Otsu and Shioya. ■

**Remark 2.9** *Otsu and Shioya also showed that, up to some modifications (taking averages of suitable distance functions), the natural maps can be made $C^1$ on the whole image of $U_\phi \cap U_\psi \cap Reg(X)$ (see [15, section 5] for a statement). However, for our purpose, the previous version is more convenient.*

In order to give some idea of the proof of the theorem of Otsu and Shioya, let us give some details of these charts.

**Definition 2.10 (strained points and natural charts)** *Let $(X, d)$ be an $n$-dimensional Alexandrov space. A point $p$ is called a $(n, \varepsilon)$ (or simply $\varepsilon$ if the dimension is implicit) strained point if there are $n$ pairs of points $(x_i, y_i)_{i\in\{1,\cdots,n\}}$ in $X$ such that $\forall i, j\ (i \neq j)\ \in \{1, \cdots, n\}$,*

$$\tilde{\angle} x_i p y_i > \pi - \varepsilon,$$

$$\tilde{\angle} x_i p x_j > \frac{\pi}{2} - 10\,\varepsilon,$$

$$\tilde{\angle} x_i p y_j > \frac{\pi}{2} - 10\,\varepsilon,$$

$$\tilde{\angle} y_i p y_j > \frac{\pi}{2} - 10\,\varepsilon,$$

*where $\tilde{\angle}$ denotes the comparison angle. The collection $(x_i, y_i)$ itself is called a $(n, \varepsilon)$-strainer for $p$.*

*We denote by $\phi_{x_1,\cdots,x_n}$ (or simply $\phi$ if there is no ambiguity) the following map :*

$$\begin{array}{rcl} \phi : U & \longrightarrow & \mathbb{R}^n \\ x & \longmapsto & (d(x_1, x), \cdots, d(x_n, x)) \end{array}$$

**Remark 2.11** *Strained points were introduced in [8]. In particular, any regular point $x$ is an $\varepsilon$-strained point for arbitrary positive $\varepsilon$ (actually, a point is regular if and only if it is $\varepsilon$-strained for arbitrary $\varepsilon > 0$). It follows that we can consider the above map $\phi$ for any strainer at $x$, and it can be shown that this map induces a chart (in the sense of Theorem 2.6) for $\varepsilon$ sufficiently small. Throughout the rest of this text, we refer to such a map as a "natural map".*

**Definition 2.12** *Let $p$ be a point of $X$. We denote by $V_p$ the set of points $q$ such that there exists a unique geodesic between $p$ and $q$.*
*Let $\phi_{x_1,\cdots,x_n}$ be a natural map. We set*

$$V_\phi = \bigcap_{i=1}^{n} V_{x_i}.$$

One of the main ingredients used to prove Theorem 2.6 is the following lemma on regularity of distance functions on Alexandrov space, which is of independent interest.

**Lemma 2.13 ([15, Lemma 4.1] )** *Let $\phi = \phi_{x_1,\cdots,x_n}$ be a natural map in a neighbourhood of a regular point $p$ and $q$ be an arbitary point in $X$. The function $d_q \circ \phi^{-1}$ is continuously differentiable on $\phi(V_\phi \cap V_q)$.*

An important consequence of the above theorems is that Rademacher's theorem on Lipschitz maps holds in this setting.

**Corollary 2.14 (Rademacher's theorem)** *On an $n$-dimensional Alexandrov space of curvature bounded below, the usual notion of differentiability (as in the Riemannian setting), (gradient) vector field and first order expansion are well-defined on $Reg(X)$ hence almost everywhere. Moreover, any Lipschitz function is differentiable almost everywhere with respect to the $n$-dimensional Hausdorff measure.*

**Proof :** More precisely, we say that a map $f$ is differentiable at $x$ if there exists a natural chart $\phi$ such that $x \in V_\phi$ and $f \circ \phi^{-1}$ is differentiable at $\phi(x)$. With this definition, the differentiability statement follows easily from the above results. Let us prove Rademacher's theorem. $Reg(X)$ is a subset of a separable metric space, hence separable. As a consequence, there exists a countable subset $\Phi_{\mathbb{N}}$ of $\Phi$ such that

$$\bigcup_{\phi \in \Phi_{\mathbb{N}}} V_\phi \supset Reg(X).$$

Thanks to this property and the fact that $Reg(X)$ is of full measure in $X$, it is sufficient to prove that any Lipschitz map $f$ is differentiable almost everywhere on any open set $U_\phi$. Now, let $\phi$ be a natural chart defined on $U_\phi$. The map $f \circ \phi^{-1}$ is a Lipschitz map (Theorem 2.8 b)1), hence differentiable almost everywhere thanks to the usual Rademacher theorem. We conclude the proof by noticing that $\phi$ is a Lipschitz map (Theorem 2.8) and $V_\phi$ is a subset of full measure of $U_\phi$ (Theorem 2.6). ■

# 3 Optimal map on an Alexandrov space

The goal of this section is to prove Theorem 1.1. The proof is divided in several steps. First, we use the Kantorovich duality. This result (see Theorem 3.3) gives us the existence of an optimal plan and characterizes the support of such a plan. More precisely, the support of an optimal plan is related to special maps called $c$-concave functions. This part holds in the general setting of Polish spaces. The second step, which is the core of the proof, is a proof of the fact that any optimal plan is supported in the graph of a map $F$ (up to a negligeable set). In the penultimate step, we prove that the map $F$ sends the initial measure onto the final one. In the last part, we establish the uniqueness of such a map (up to a negligeable set).

## 3.1 Kantorovich duality

Throughout this paragraph, we refer to $\mu_0, \mu_1$ as Borel probability measures on a complete separable metric space $X$ without any further assumption on their support.

The dual Kantorovich problem is the problem of maximizing the following quantity

$$J(\phi, \psi) = \int_X \phi(x) d\mu_0(x) + \int_X \psi(y) d\mu_1(y)$$

where $\phi$ and $\psi$ are elements of the space $C^b(X)$ of continuous bounded functions on $X$ such that $\forall x, y \in X \times X$,

$$\phi(x) + \psi(y) \leq c(x, y).$$

Under quite general assumptions, it is possible to show that extrema of both problems coincide, and, moreover, any minimizer of the

Kantorovich problem is associated to a pair of maximizers for the dual problem. Before we state this duality theorem, we recall some definitions which will enable us to describe properties of a maximizing pair.

**Definition 3.1** *A cost function on $X$ is a lower semicontinuous function $c : U \times V \mapsto \mathbb{R}^+$, where $U, V$ are Borel subsets of $X$. Let $\phi : U \mapsto \mathbb{R} \cup \{+\infty\}$ be a proper (i.e. $\neq +\infty$) measurable function. We define $\phi^c : V \mapsto \mathbb{R} \cup \{-\infty\}$, the c-transform of $\phi$ by the formula*

$$\phi^c(y) = \inf_{x \in U} c(x, y) - \phi(x).$$

*We define in a similar manner the c-transform of a function defined on $V$ (keeping $c(x, y)$ unchanged because of the possible asymmetry of $c$) and we call $\phi$ as above, a c-concave function if $(\phi^c)^c = \phi$ (in the rest of the paper, we will write $\phi^{cc}$).*

**Remark 3.2** *Assume that a cost function is locally Lipschitz (e.g. $c(x, y) = d^2/2(x, y)$) and $U, V$ are compact sets, then any c-concave function is Lipschitz also (see [13] for a proof).*

We refer to [22] or [16] for a more detailed analysis of $c$-concavity. There exist numerous versions of the Kantorovich duality. We give here a version borrowed from the book [1, Theorem 6.1.5].

**Theorem 3.3 (Kantorovich duality)** *Let $(X, d)$ be a complete separable metric space, $\mu_0, \mu_1$ be Borel probability measures on $X$ and $c$ be a cost function such that*

$$\int_{X \times X} c \, d\mu_0 d\mu_1 < +\infty.$$

*Then*

$$\sup_{(\psi,\phi) \in C^b(X) \times C^b(X)} J(\psi, \phi) = \min_{\pi \in S(\mu_0,\mu_1)} \int_{X \times X} c(x, y) d\pi(x, y) \qquad (1)$$

*where $S(\mu_0, \mu_1)$ denotes the set of probability measures whose marginals are $\mu_0, \mu_1$. In addition, there exists a maximizing pair $(\phi, \phi^c) \in L^1(\mu_0) \times L^1(\mu_1)$ for the dual problem where $\phi$ is a $c-$concave function. Moreover there exists an optimal plan $\gamma$ such that for the above pair*

$$\phi(x) + \phi^c(y) = c(x, y) \quad \gamma - a.e. \text{ in } X \times X. \qquad (2)$$

*Conversely, if there exists $\phi \in L^1(\mu_0)$ such that (2) holds, then $\gamma$ is optimal.*

## 3.2 The support of an optimal plan is a graph

In this section, we prove the main technical result of this paper, namely that the support of an optimal plan is concentrated on the graph of a function. First, we state a lemma on the differentiability of a function along geodesics.

**Lemma 3.4 (Differentiability along geodesics)** *Let $f$ be a function on $X$, differentiable at a point $x \in Reg(X)$ and $\gamma$ be a unitary geodesic defined on $[0,T]$. Then, the Taylor expansion given below holds:*

$$f(\gamma(t)) = f(x) + t\langle \nabla f(x), \gamma'(0)\rangle + o(t).$$

**Proof :** Let $\phi = \phi_{p_1,\cdots,p_n}$ be a natural map around $x$ such that $x \in V_\phi$. By definition, $f \circ \phi^{-1}$ is then differentiable in the usual sense. It remains to prove that $\phi \circ \gamma$ is differentiable at 0. This follows from the first variation formula (Lemma 2.2) using the fact that $x \in V_\phi$. ■

**Lemma 3.5** *Let $\psi$ be a $d^2/2$-concave function on an Alexandrov space $(X,d)$. Let us assume that the supports of $\psi$ and $\psi^c$ are compact sets. Then*

$$\frac{1}{2}d^2(x,y) \geq \psi(x) + \psi^c(y) \tag{3}$$

*for all $x, y \in X$. If $x \in Reg(X)$ is a point where $\psi$ is differentiable, then equality holds in (3) if and only if $y = \exp_x(-\nabla\psi(x))$.*

**Remark 3.6** *$exp_x(-\nabla\psi(x))$ denotes the endpoint of the geodesic whose direction is $\frac{\nabla\psi}{|\nabla\psi|}$ and is parametrized on $[0, |\nabla\psi(x)|]$. In particular, $d(x, \exp_x(-\nabla\psi(x)) = |\nabla\psi(x)|$.*

**Proof :** The inequality follows from the definition of $\psi^c$. The compactness of the support of $\psi^c$ implies the existence of a pair which satisfies the equality. So, let us assume that the equality holds in (3) for a pair $(x,y)$. We set $\theta(z) = d^2(y,z)/2$. Let $\gamma(t)$ be a unitary geodesic which starts at $x$ and is parametrized by $[0,T]$. By assumption on $x$ and $y$, we have

$$\begin{aligned}
\theta(\gamma(t)) - \theta(x) &= d^2(\gamma(t),y)/2 - d^2(x,y)/2 \\
&\geq \psi(\gamma(t)) - \psi(x) \\
&\geq t\langle \nabla\psi(x), \gamma'(0)\rangle + o(t)
\end{aligned}$$

where we get the first inequality from (3) and the second from Lemma 3.4. The function $\theta$ is said to be subdifferentiable at $x$ with subgradient $\nabla\psi(x)$. Now, we are going to prove that $\theta$ is also superdifferentiable at $x$ (actually everywhere). As a consequence, we will get that $\theta$ is differentiable at $x$ and the subgradient and the supergradient coincide, yielding the equality at the end of the proof.

Applying the first variation formula (Lemma 2.2) to $\theta$ and $\gamma$ yields

$$d^2(\gamma(t),y) \leq d^2(x,y) + -2td(x,y)\cos\angle(\sigma'(0),\gamma'(0)) + o(t)$$

where $\sigma$ is a geodesic between $x$ and $y$, and $\angle(\sigma'(0),\gamma'(0))$ is the angle between $\sigma$ and $\gamma$. Let us remark that the above inequality holds true for any geodesic $\gamma$ starting at $x$ and any other geodesic linking $x$ and $y$ (if any).

Letting $t$ go to 0, we obtain

$$\langle \nabla\psi(x) + d(x,y)\sigma'(0), \gamma'(0)\rangle \leq 0.$$

To prove the reverse inequality, one would like to consider a geodesic which starts at $x$ and whose direction is $-\gamma'(0)$. Unfortunately, we cannot suppose such a geodesic exists; however, applying the above argument to a sequence of geodesics starting at $x$ and whose directions converge to $-\gamma'(0)$ allows to conclude that the function $\theta$ is differentiable at $x$ and

$$\nabla\psi(x) = -d(x,y)\sigma'(0).$$

This gives the characterization of $y$ stated in the lemma.

■

**Remark 3.7** *Note that we also showed that if $x$ is a point such that $\psi$ is differentiable at $x$, there exists a unique geodesic between $x$ and $exp_x(-\nabla\psi(x))$.*

Now, we can deduce from the previous lemma that any optimal plan is actually supported in the graph of a Borel function.

**Proposition 3.8** *Let $\Pi_0$ be a minimizer of the variational problem*

$$\min \int_{X\times X} d^2(x,y)\, d\Pi(x,y)$$

*among all couplings on $X \times X$ whose marginals are two probability measures $\mu_0, \mu_1$ that satisfy the assumptions of Theorem 1.1. Then, there exists a measurable function $F$ such that*

$$\Pi_0 = (Id, F)_\sharp \mu_0.$$

*Moreover, the map $F$ is defined by the formula*

$$F(x) = exp_x(-\nabla\psi(x)) \tag{4}$$

*and $\psi$ is a $d^2/2$-concave function.*

**Remark 3.9** *By abuse of language, we will say that $F$ is a minimizer of the above problem.*

**Proof :** We deduce from Kantorovich duality (Theorem 3.3) the existence of a pair $(\psi, \psi^c)$ of $c$-concave maps which are locally Lipschitz maps (see Remark 3.2) and such that

$$\int_{X\times X} d^2/2(x,y)\, d\Pi_0(x,y) = \int_X \psi(x) d\mu_0(x) + \int_X \psi^c(y) d\mu_1(y)$$

where $(\psi, \psi^c)$ satisfy (3) by definition of the $c$-transform. Now, Rademacher's theorem and Lemma 3.5 imply that the support of $\Pi_0$ is concentrated on the graph of the map

$$F(x) = exp_x(-\nabla\psi(x)).$$

Note that the map $F$ restricted to the subset of $Reg(X)$ of points where $\phi$ is differentiable is a continuous map. Hence, the measurability of $Reg(X)$ (Theorem 2.5) entails the measurability of $F$.

It remains to prove that $F_\sharp \mu_0 = \mu_1$. Let $Z$ be a subset of $Reg(X)$ of full measure, such that $\forall x \in Z, \psi$ is differentiable at $x$. By definition of $\Pi_0$ and by assumption on $\mu_0$, $\Pi_0(Z \times X) = 1$. As a consequence, if $A$ is a Borel set of $X$, then the following equalities hold

$$\begin{aligned} \mu_1(A) &= \int_{X\times A} d\Pi_0 = \int_{Z\times A} d\Pi_0 \\ &= \mu_0(Z \cap F^{-1}(A)) = F_\sharp \mu_0(A). \end{aligned}$$

■

### 3.3 Uniqueness property

It remains to prove the uniqueness of the optimal plan and optimal map. The results of Subsection 3.2 reduce the proof to the case of the optimal map. Therefore, to complete the proof of Theorem 1.1, we just need the following result.

**Proposition 3.10** *Under the assumptions of Theorem 1.1, we set $t(x) = \exp_x(-\nabla\psi(x))$ where $\psi$ is a $d^2/2$-concave function, a solution of Monge's problem. Then, up to modifications on a negligeable subset, the map $\nabla\psi$ is uniquely determined.*

**Proof :** Let $s$ be another solution. Namely, $s$ is a Borel function, mapping $\mu_0$ to $t_\sharp \mu_0$, such that:

$$\int_X \psi \, d\mu_0 + \int_X \psi^c \, dt_\sharp \mu_0 = \int_X \frac{1}{2} d^2(x, s(x)) \, d\mu_0.$$

As $t_\sharp \mu_0 = s_\sharp \mu_0$ by assumption, we get

$$\int_X \psi(x) + \psi^c(s(x)) - \frac{1}{2} d^2(x, s(x)) \, d\mu_0(x) = 0,$$

where the integrand is nonpositive by definition of the $c$-transform, and consequently is equal to 0 $\mu_0$ almost everywhere. Lemma 3.5 and Rademacher's theorem allow us to conclude that $s = t$ almost everywhere. The uniqueness of $\nabla\psi$ follows from Remarks 3.6 and 3.7. ■

## 4 Generalisation to other costs

As in the Riemannian case, our main result can be adapted to other strictly convex costs of the distance function. Compactness of supports can also be relaxed. The proof of the theorem below is similar to the proof in the Riemannian case, so we only sketch it. We refer to [1, Theorem 6.2.4] for a detailed proof in the Euclidean case.

Throughout the section, we consider a cost $c(x, y)$ defined by

$$c(x, y) = h(d(x, y)),$$

where $h : \mathbb{R}^+ \mapsto \mathbb{R}^+$ is a strictly convex and nondecreasing differentiable function.

We also use the approximate differential of a map, which we recall the definition of in this setting.

**Definition 4.1** *We say that $f : X \mapsto \mathbb{R}$ has an approximate differential at $x \in Reg(X)$ if there exists a map $g : X \mapsto \mathbb{R}$ differentiable at $x$ such that the set $\{f \neq g\}$ has density $0$ at $x$.*

**Theorem 4.2** *Let $(X, d)$ be a finite dimensional Alexandrov space and $\mu_H$ be the corresponding Hausdorff measure. Let $\mu_0, \mu_1$ be probability measures on $X$ such that $\mu_0$ is absolutly continuous with respect to $\mu_H$ and*

$$\int_{X\times X} c \, d\mu_0 d\mu_1 < +\infty.$$

*Under these assumptions, the Kantorovitch problem admits a solution, and any optimal plan is supported in the graph of a Borel function $F$. This map $F$ is also a minimizer of Monge's problem and satisfies for $\mu$ almost every $x \in X$,*

$$F(x) = exp(-\frac{(h')^{-1}(|\tilde{\nabla}\phi(x)|)}{|\tilde{\nabla}(\phi(x))|}\tilde{\nabla}\phi(x)),$$

*if $|\tilde{\nabla}\phi(x)| \neq 0$ and $F(x) = x$ otherwise. $\phi$ is a c-concave function (see Definition 3.1) and $\tilde{\nabla}$ denotes the approximate gradient of $\phi$.*

*Moreover, up to modifications on negligeable sets, the map $\tilde{\nabla}\phi$ is unique, and as a consequence so is the optimal map $F$.*

**Proof :** To get the result, we only have to prove an analogue of Lemma 3.5. To this aim, we set $(\phi, \phi^c)$ a maximizing pair of the dual Kantorovich problem. To circumvent difficulties arising from the non-compactness of supports, we use auxilliary maps defined on compact subsets. We fix $o \in X$, $R$ a positive number, and define

$$\phi_R(x) = \inf_{B(o,R)} c(x, y) - \phi^c(y).$$

As $h$ is a convex function, the cost $c$ is a locally Lipschitz map and so is $\phi_R$. By compactness, for any $x \in support(\mu_0)$, there exists $y \in B(o, R)$ such that

$$\phi_R(x) + \phi^c(y) = h(d(x, y)).$$

Now, let us assume that $\phi_R$ is differentiable at $x$ and set $\theta(x) = c(x, y)$. Arguing as in Lemma 3.5, we get that $\theta$ is subdifferentiable at $x$ with subgradient $\nabla\phi_R(x)$. The differentiability of $h$ entails that $\theta$ is also superdifferentiable at $x$ (see Lemma 3.5)

$$\theta(\gamma(t)) \leq \theta(x) - th'(d(x, y)) \cos\angle(\sigma'(0), \gamma'(0)) + o(t) \tag{5}$$

where we use the notations of Lemma 3.5.

Consequently, using the strict convexity of $h$, we get the following equivalence (assuming $\phi_R$ is differentiable at $x$):

$$\phi_R(x) + \phi^c(y) = c(x, y)$$

if and only if

$$y = exp(-\frac{(h')^{-1}(|\nabla\phi_R(x)|)}{|\nabla(\phi_R(x))|}\nabla\phi_R(x))$$

where the equality reads $y = x$ in the case where $|\nabla(\phi_R(x))| = 0$.

Now, let $\Pi$ be an optimal plan, the Kantorovich duality implies

$$\phi(x) + \phi^c(y) = c(x, y) \;\; \Pi - \text{a.e.}$$

Therefore, for $\mu_0$ almost every $x \in X$, there exists $y$ such that the above equality holds. This entails that $\bigcup_{R\in\mathbb{N}^*}\{\phi = \phi_R\}$ is a subset of full measure. We conclude the proof as in [1] (note that Lebesgue's theorem on approximate differentiability holds in our setting thanks to the existence of charts on $Reg(X)$ satisfying Property b)i) of Theorem 2.8 and Theorem 2.5).

■

**Acknowledgments.** This work was done during a one-year stay at the Scuola Normale Superiore di Pisa. I would like to thank warmly Luigi Ambrosio for useful discussions and remarks and the institute for its hospitality and support. Many thanks to Catriona also for careful rereading. I thank the referee for useful remarks on the previous version of this paper.

# References

[1] Luigi Ambrosio, Nicola Gigli, and Giuseppe Savaré. *Gradient flows in metric spaces and in the space of probability measures.* Lectures in Mathematics ETH Zürich. Birkhäuser Verlag, Basel, 2005.

[2] Luigi Ambrosio, Bernd Kirchheim, and Aldo Pratelli. Existence of optimal transport maps for crystalline norms. *Duke Math. J.*, 125(2):207–241, 2004.

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

[4] Luigi Ambrosio and Séverine Rigot. Optimal mass transportation in the Heisenberg group. *J. Funct. Anal.*, 208(2):261–301, 2004.

[5] Patrick Bernard and Boris Buffoni. The Monge problem for supercritical Mañé potentials on compact manifolds. *Adv. Math.*, 207(2):691–706, 2006.

[6] Yann Brenier. Polar factorization and monotone rearrangement of vector-valued functions. *Comm. Pure Appl. Math.*, 44(4):375–417, 1991.

[7] Dmitri Burago, Yuri Burago, and Sergei Ivanov. *A course in metric geometry*, volume 33 of *Graduate Studies in Mathematics.* American Mathematical Society, Providence, RI, 2001.

[8] Yuri Burago, Mikhail Gromov, and Gregory Perel′man. A. D. Aleksandrov spaces with curvatures bounded below. *Uspekhi Mat. Nauk*, 47(2(284)):3–51, 222, 1992.

[9] Dario Cordero-Erausquin, Robert J. McCann, and Michael Schmuckenschläger. A Riemannian interpolation inequality à la Borell, Brascamp and Lieb. *Invent. Math.*, 146(2):219–257, 2001.

[10] Wilfrid Gangbo and Robert J. McCann. The geometry of optimal transportation. *Acta Math.*, 177(2):113–161, 1996.

[11] John Lott and Cédric Villani. Ricci curvature for metric-measure spaces via optimal mass transport. *To appear in Annals of math.*

[12] John Lott and Cédric Villani. Weak curvature conditions and poincaré inequalities. *To appear in Journal of Functional Analysis.*

[13] Robert J. McCann. Polar factorization of maps on Riemannian manifolds. *Geom. Funct. Anal.*, 11(3):589–608, 2001.

[14] Gaspard Monge. *Mémoire sur la théorie des déblais et des remblais, Académie royale des sciences de Paris.* 1781.

[15] Yukio Otsu and Takashi Shioya. The Riemannian structure of Alexandrov spaces. *J. Differential Geom.*, 39(3):629–658, 1994.

[16] Svetlozar T. Rachev and Ludger Rüschendorf. *Mass transportation problems. Vol. I.* Probability and its Applications (New York). Springer-Verlag, New York, 1998. Theory.

[17] Cyril S. Smith and Martin Knott. Note on the optimal transportation of distributions. *J. Optim. Theory Appl.*, 52(2):323–329, 1987.

[18] Karl-Theodor Sturm. On the geometry of metric measure spaces. I. *Acta Math.*, 196(1):65–131, 2006.

[19] Karl-Theodor Sturm. On the geometry of metric measure spaces. II. *Acta Math.*, 196(1):133–177, 2006.

[20] V. N. Sudakov. Geometric problems in the theory of infinite-dimensional probability distributions. *Proc. Steklov Inst. Math.*, (2):i–v, 1–178, 1979. Cover to cover translation of Trudy Mat. Inst. Steklov **141** (1976).

[21] Cédric Villani. *Optimal transport, old and new (to appear).*

[22] Cédric Villani. *Topics in optimal transportation.*, volume 58 of *Graduate Studies in Mathematics.* American Mathematical Society, Providence, RI, 2003.

Jérôme Bertrand
Institut de Mathématiques
Université Paul Sabatier
118 route de Narbonne
F31062 Cedex 4 Toulouse
*email:* bertrand@math.ups-tlse.fr