Acta Math., 202 (2009), 241–283
DOI: 10.1007/s11511-009-0037-8

On the regularity of solutions of optimal transportation problems

by

Grégoire Loeper

Université Claude Bernard Lyon 1
Lyon, France

Contents

1. Introduction

Given two topological spaces A and B, a cost function $c: A\times B\to\mathbb{R}$ and two probability measures μ_0 and μ_1, respectively on A and B, Monge's problem of optimal transportation consists in finding, among all measurable maps $T: A\to B$ that push forward μ_0 onto μ_1 (hereafter $T_\#\mu_0=\mu_1$), in the sense that

$$\mu_1(E)=\mu_0(T^{-1}(E)) \quad \text{for all } E\subset B \text{ Borel}, \tag{1}$$

a map that realizes

$$\min\left\{\int_A c(x,T(x))\,d\mu_0(x): T_\#\mu_0=\mu_1\right\}. \tag{2}$$

Optimal transportation has undergone a rapid and important development since the pioneering work of Brenier, who discovered that when $A=B=\mathbb{R}^n$ and the cost is the distance squared, optimal maps for the problem (2) are gradients of convex functions [1] (see also [21] where the connection with gradients was first proved). Following this result and its subsequent extensions, the theory of optimal transportation has flourished, with generalizations to other cost functions [8], [16], more general spaces such as Riemannian manifolds [25], applications in many other areas of mathematics such as geometric analysis, functional inequalities, fluid mechanics, dynamical systems, and other more concrete applications such as irrigation and cosmology.

When A and B are domains of the Euclidean space $\mathbb{R}^n$, or of a Riemannian manifold, a common feature to all optimal transportation problems is that optimal maps derive from a (cost-convex) potential, which, assuming some smoothness, is in turn a solution to a fully non-linear elliptic PDE: the Monge–Ampère equation. In all cases, the Monge–Ampère equation arising from an optimal transportation problem reads in local coordinates

$$\det(D^2\phi-\mathcal{A}(x,\nabla\phi))=f(x,\nabla\phi), \tag{3}$$

where $(x,p)\mapsto\mathcal{A}(x,p)$ is a symmetric matrix-valued function, that depends on the cost function $c(x,y)$ through the formula

$$\mathcal{A}(x,p)=-D^2_{xx}c(x,y) \quad \text{for } y \text{ such that } -\nabla_x c(x,y)=p. \tag{4}$$

That there is indeed a unique y such that $-\nabla_x c(x,y)=p$ will be guaranteed by condition (A1) given hereafter. The optimal map will then be

$$x\longmapsto y \quad \text{such that } -\nabla_x c(x,y)=\nabla\phi(x).$$

In the case $\mathcal{A}=0$, equation (3) was well known and studied before optimal transportation since it appears in Minkowski's problem: find a convex hypersurface with prescribed Gauss curvature. In the case of optimal transportation, the boundary condition consists in prescribing that the image of the optimal map equals a certain domain. It is known as the second boundary value problem.

Until recently, except in the particular case of the so-called reflector antenna, treated by Wang [37] (see also [11] for C^1 regularity), the regularity of optimal maps was only known in the case where the cost function is the (Euclidean) squared distance $c(x,y)=|x-y|^2$, which is the cost considered by Brenier in [1], for which the matrix $\mathcal{A}$ in (3) is the identity (which is trivially equivalent to the case $\mathcal{A}=0$). Those results have involved several authors, among them Caffarelli, Urbas and Delanoë. An important step was made recently by Ma, Trudinger and Wang [24], and Trudinger and Wang [30], who introduced a condition (named (A3) and (A3w) in their papers) on the cost function under which they could show existence of smooth solutions to (3). Let us give right away this condition that will play a central role in the present paper. Let $A=\Omega$ and $B=\Omega'$ be bounded domains of $\mathbb{R}^n$ on which the initial and final measures will be supported. Assume that c belongs to $C^4(\Omega\times\Omega')$. For $(x,y)\in\Omega\times\Omega'$ and $(\xi,\nu)\in\mathbb{R}^n\times\mathbb{R}^n$, we define

$$\mathfrak{S}_{\mathfrak{c}}(x,y)(\xi,\nu):=D^2_{p_kp_l}\mathcal{A}_{ij}\,\xi_i\xi_j\,\nu_k\nu_l\,(x,p),\quad p=-\nabla_x c(x,y). \tag{5}$$

Whenever ξ and ν are orthogonal unit vectors, we will say that $\mathfrak{S}_{\mathfrak{c}}(x,y)(\xi,\nu)$ defines the *cost-sectional curvature from x to y in the directions* (ξ,ν). As we will see in Definition 2.13, this definition is intrinsic. Note that this map is in general not symmetric, and that it depends on two points x and y. The reason why we use the word sectional curvature will be clear in a few lines. We will say that the cost function c has *non-negative cost-sectional curvature* on $\Omega\times\Omega'$ if

$$\mathfrak{S}_{\mathfrak{c}}(x,y)(\xi,\nu)\geqslant 0\quad\text{for all } (x,y)\in\Omega\times\Omega' \text{ and all } (\xi,\nu)\in\mathbb{R}^n\times\mathbb{R}^n \text{ such that } \xi\perp\nu. \tag{6}$$

A cost function satisfies condition (Aw) on $\Omega\times\Omega'$ if and only if it has non-negative cost-sectional curvature on $\Omega\times\Omega'$, i.e. if it satisfies (6).

Under condition (Aw) and natural requirements on the domains Ω and Ω', Trudinger and Wang [30] showed that the solution to (3) is globally smooth for smooth positive measures μ_0 and μ_1. They showed that (Aw) is satisfied by a large class of cost functions, that we will give as examples later on. Note that the quadratic cost satisfies assumption (Aw). This result is achieved by the so-called continuity method, for which a key ingredient is to obtain a priori estimates on the second derivatives of the solution. At this stage, condition (Aw) was used in a crucial way. However, even if it was known that not all cost functions can lead to smooth optimal maps, it was unclear whether the condition (Aw) was necessary, or just a technical condition for the a priori estimates to go through.

In this paper we show that the condition (Aw) is indeed the *necessary and sufficient condition for regularity:* one cannot expect regularity without this condition, and more precisely, if $\mathfrak{S}_{\mathfrak{c}}(x,y)(\xi,\nu)<0$ for $(x,y)\in\Omega\times\Omega'$ and $\xi\perp\nu\in\mathbb{R}^n$, one can immediately build a pair of C^∞ strictly positive measures, supported on sets that satisfy the usual smoothness and convexity assumptions, so that the optimal potential is not even C^1, and the optimal map is therefore discontinuous. This result is obtained by analyzing the geometric nature of condition (6). Let us first recall that the solution ϕ of the Monge–Ampère equation is a priori known to be cost-convex (in short c-convex), meaning that at each point $x\in\Omega$, there exist $y\in\Omega'$ and a value $\phi^c(y)$ such that

$$\begin{aligned}
-\phi^c(y)-c(x,y)&=\phi(x),\\
-\phi^c(y)-c(x',y)&\leqslant\phi(x')\quad\text{for all } x'\in\Omega.
\end{aligned}$$

The function $-\phi^c(y)-c(x,y)$ is called a *supporting* function, and the function $y\mapsto\phi^c(y)$ is called the *cost-transform* (in short the *c-transform*) of ϕ, also defined by

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

(These notions will be recalled in greater details hereafter.) We prove that condition (Aw) can be reformulated as a property of cost-convex functions, which we call *connectedness of the contact set*:

$$\text{For all } x\in\Omega,\text{ the contact set } G_\phi(x):=\{y:\phi^c(y)=-\phi(x)-c(x,y)\}\text{ is connected.}\tag{7}$$

Assuming a natural condition on Ω' (namely its c-convexity, see Definition 2.9) this condition involves only the cost function, since it must hold for any ϕ^c defined through a c-transform.

A case of special interest for applications is the generalization of Brenier's cost $\frac{1}{2}|x-y|^2$ to Riemannian manifolds, namely $c(x,y)=\frac{1}{2}d^2(x,y)$. Existence and uniqueness of optimal maps in that case was established by McCann [25], and further examined by several authors, with many interesting applications in geometric and functional analysis (for example [12] and [26]). The optimal map takes the form $x\mapsto\exp_x(\nabla\phi(x))$ for a c-convex potential ϕ and is called a *gradient map*. Then, a natural question is the interpretation of condition (Aw) and of the cost-sectional curvature in this context. We show that for some universal constant K,

$$\text{cost-sectional curvature from } x \text{ to } x=K\cdot\text{Riemannian sectional curvature at } x.$$

(We mean there that the equality holds for every 2-plane and actually $K=\frac{2}{3}$.) As a direct consequence of the previous result, *the optimal* (*gradient*) *map will not be continuous for arbitrary smooth positive data if the manifold does not have non-negative sectional curvature everywhere.* Although the techniques are totally different, it is interesting to notice that in recent works, Lott and Villani [23], and Sturm [27] have recovered the Ricci curvature through a property of optimal transport maps (namely through the displacement convexity of some functionals). Here, we somehow recover the sectional curvature through the continuity of optimal maps.

We next investigate the continuity of optimal maps under the stronger condition of uniformly positive cost-sectional curvature, or condition (As):

$$\begin{gathered}\text{there exists } C_0>0 \text{ such that}\\ \mathfrak{S}_{\mathfrak{c}}(x,y,\xi,\nu)\geqslant C_0|\xi|^2|\nu|^2 \text{ for all } (x,y)\in\Omega\times\Omega' \text{ and all } (\xi,\nu)\in\mathbb{R}^n\times\mathbb{R}^n,\ \xi\perp\nu.\end{gathered}\tag{8}$$

We obtain that the (weak) solution of (3) is C^1 or $C^{1,\alpha}$ under quite mild assumptions on the measures. Namely, for $B_r(x)$ the ball of radius r and center x, and μ_1 being bounded away from 0, we need $\mu_0(B_r(x))=o(r^{n-1})$ to show that the solution of (3) is C^1, and $\mu_0(B_r(x))=O(r^{n-p})$, $p<1$, to show that it is $C^{1,\alpha}$, for $\alpha=\alpha(n,p)\in(0,1)$. Those conditions allow μ_0 and μ_1 to be singular with respect to the Lebesgue measure and μ_0 to vanish.

This result can be seen as analogous to Caffarelli's $C^{1,\alpha}$ estimate [5] for a large class of cost functions and related Monge–Ampère equations. It also shows that the partial regularity results are better under (As) than under (Aw), since Caffarelli's $C^{1,\alpha}$ regularity result required μ_0 and μ_1 to have densities bounded away from 0 and infinity, and it is known to be close to optimal [35].

In a forthcoming work [22] we shall prove that the quadratic cost on the sphere has uniformly positive cost-sectional curvature, i.e. satisfies (As). We obtain therefore regularity of optimal (gradient) maps under adequate conditions.

The rest of the paper is organized as follows: in §2 we gather all definitions and results that we will need throughout the paper. In §3 we state our results. Then each of the subsequent sections is devoted to the proof of a theorem. The reader knowledgeable about the subject might skip directly to §3.

Acknowledgments. At this point I wish to express my gratitude to Neil Trudinger and Xu-Jia Wang, for many discussions, and for sharing results in progress while we were all working on this subject. I thank Cédric Villani for fruitful discussions, and Philippe Delanoë with whom we started to think about the problem of regularity for optimal transportation on the sphere. I wish to thank Robert McCann, who first raised to me the issue of the connectedness of the contact set, in 2003. I gratefully acknowledge the support of a French Australian exchange grant PHC FAST EGIDE No. 12739WA. I am also grateful for the hospitality of the Center for Mathematics and its Applications at the University of Canberra.

2. Preliminaries

2.1. Notation

Hereafter dVol denotes the Lebesgue measure of $\mathbb{R}^n$ and $B_r(x)$ denotes the ball of radius r centered at x. For $\delta>0$, we set classically $\Omega_\delta=\{x\in\Omega: d(x,\partial\Omega)>\delta\}$. When we say that a function (resp. a measure) is smooth without stating the degree of smoothness, we assume that it is C^∞-smooth (resp. has a C^∞-smooth density with respect to the Lebesgue measure).

2.2. Kantorovitch duality and c-convex potentials

In this section, we recall how to obtain the optimal map from a c-convex potential in the general case. This allows us to introduce definitions that we will be using throughout the paper. References concerning the existence of optimal map by Monge–Kantorovitch duality are [1] for the cost $|x-y|^2$, [16] and [8] for general costs, [25] for the Riemannian case, otherwise the book [33] offers a rather complete reference on the topic.

Monge's problem (2) is first relaxed to become a problem of linear programming; one seeks now

$$\mathcal{I}=\inf\left\{\int_{\mathbb{R}^n\times\mathbb{R}^n} c(x,y)\,d\pi(x,y) : \pi\in\Pi(\mu_0,\mu_1)\right\}, \tag{9}$$

where $\Pi(\mu_0,\mu_1)$ is the set of positive measures on $\mathbb{R}^n\times\mathbb{R}^n$ whose marginals are respectively μ_0 and μ_1. Note that the (Kantorovitch) infimum (9) is smaller than the (Monge) infimum of the cost (2), since whenever a map T pushes forward μ_0 onto μ_1, the measure $\pi_T(x):=\mu_0(x)\otimes\delta_{T(x)}(y)$ belongs to $\Pi(\mu_1,\mu_1)$.

Then, the dual Monge–Kantorovitch problem is to find an optimal pair of potentials (ϕ,ψ) that realizes

$$\mathcal{J}=\sup\left\{-\int\phi(x)\,d\mu_0(x)-\int\psi(y)\,d\mu_1(y) : \phi(x)+\psi(y)\geqslant -c(x,y)\right\}. \tag{10}$$

The constraint on ϕ and ψ leads to the definition of c- and c*-transforms.

Definition 2.1. Given a lower semi-continuous function $\phi:\Omega\subset\mathbb{R}^n\to\mathbb{R}\cup\{+\infty\}$, we define its *c-transform* at $y\in\Omega'$ by

$$\phi^c(y)=\sup_{x\in\Omega}(-c(x,y)-\phi(x)). \tag{11}$$

For $\psi:\Omega'\subset\mathbb{R}^n\to\mathbb{R}$ also lower semi-continuous, define its c^**-transform* at $x\in\Omega$ by

$$\psi^{c*}(x)=\sup_{y\in\Omega'}(-c(x,y)-\psi(y)). \tag{12}$$

A function is *cost-convex*, or, in short, *c-convex*, if it is the c*-transform of another function $\psi:\Omega'\to\mathbb{R}$, i.e. for $x\in\Omega$, $\phi(x)=\sup_{y\in\Omega'}(-c(x,y)-\psi(y))$ for some lower semi-continuous $\psi:\Omega'\to\mathbb{R}$. Moreover in this case $(\phi^c)^{c*}=\phi$ on Ω (see [33]).

Our first assumption on c will be:

(A0) The cost-function c belongs to $C^4(\bar{\Omega}\times\bar{\Omega}')$.

We will also always assume that Ω and Ω' are bounded. These assumptions are not the weakest possible for the existence/uniqueness theory.

Proposition 2.2. *If c is Lipschitz and semi-concave with respect to x, locally uniformly with respect to y, and Ω' is bounded, then ϕ^c is locally semi-convex and Lipschitz. In particular, this holds under assumption* (A0). *The symmetric statement holds for ψ^{c*}.*

By Fenchel–Rockafellar's duality theorem, we have $\mathcal{I}=\mathcal{J}$. One can then easily show that the supremum (10) and the infimum (9) are achieved. Since $\phi(x)+\psi(y)\geqslant -c(x,y)$ implies $\psi\geqslant\phi^c$, we can assume that for the optimal pair in $\mathcal{J}$ we have $\psi=\phi^c$ and $\phi=(\phi^c)^{c*}$. Writing the equality of the integrals in (9) and (10) for any optimal γ and any optimal pair (ϕ,ϕ^c), we obtain that γ is supported in $\{(x,y):\phi(x)+\phi^c(y)+c(x,y)=0\}$. This leads us to the following definition.

Definition 2.3. (Gradient mapping) Let ϕ be a c-convex function. We define the set-valued mapping G_ϕ by

$$G_\phi(x)=\{y\in\Omega':\phi(x)+\phi^c(y)=-c(x,y)\}.$$

For all $x\in\Omega$, $G_\phi(x)$ is the contact set between ϕ^c and its supporting function

$$-\phi(x)-c(x,\cdot).$$

Noticing that for all $y\in G_\phi(x)$, the function $\phi(\cdot)+c(\cdot,y)$ has a global minimum at x, we introduce/recall the following definitions:

Definition 2.4. (Subdifferential) For a semi-convex function ϕ, the *subdifferential* of ϕ at x, that we denote by $\partial\phi(x)$, is the set

$$\partial\phi(x)=\{p\in\mathbb{R}^n:\phi(y)\geqslant\phi(x)+p\cdot(y-x)+o(|x-y|)\}.$$

The subdifferential is always a convex set, and is always non-empty for a semi-convex function.

Definition 2.5. (c-subdifferential) If ϕ is c-convex, the *c-subdifferential* of ϕ at x, that we denote by $\partial^c\phi(x)$, is the set

$$\partial^c\phi(x)=\{-\nabla_x c(x,y):y\in G_\phi(x)\}.$$

The inclusion $\varnothing\neq\partial^c\phi(x)\subset\partial\phi(x)$ always holds.

We now introduce two assumptions on the cost-function, which are the usual assumptions made in order to obtain an optimal map. For $x=(x_1,...,x_n)$ and $y=(y_1,...,y_n)$, let us first introduce the notation

$$D^2_{xy}c(x,y)=[\partial_{x_i}\partial_{y_j}c(x,y)]^n_{i,j=1}.$$

(A1) For all $x\in\bar\Omega$, the mapping $y\mapsto-\nabla_x c(x,y)$ is injective on $\bar\Omega'$.

(A2) The cost function c satisfies $\det D^2_{xy}c\neq0$ for all $(x,y)\in\bar\Omega\times\bar\Omega'$.

This leads us to the definition of the c-exponential map.

Definition 2.6. Under assumption (A1), for $x\in\Omega$ we define the *c-exponential map* at x, which we denote by $\mathfrak{T}_x$, such that

$$\mathfrak{T}_x(-\nabla_x c(x,y))=y\quad\text{for all }(x,y)\in\Omega\times\Omega'.$$

Moreover, under assumptions (A0), (A1) and (A2), and assuming that Ω' is connected, there exists a constant $C_\mathfrak{T}>0$ that depends on c, Ω and Ω', such that for all $x\in\Omega$ and all $p_1,p_2\in-\nabla_x c(x,\Omega')$,

$$\frac{1}{C_\mathfrak{T}}\leqslant\frac{|\mathfrak{T}_x(p_2)-\mathfrak{T}_x(p_1)|}{|p_2-p_1|}\leqslant C_\mathfrak{T}. \tag{13}$$

Remark 1. The definition of c-exponential map is again motivated by the case where cost = distance squared, where the c-exponential map is the exponential map. Moreover, notice the important identity

$$[D^2_{xy}c]^{-1}=-D_p\mathfrak{T}_x|_{x,p=-\nabla_x c(x,y)}. \tag{14}$$

Remark 2. Anticipating the extension to Riemannian manifolds, we mention at this point that this definition is intrinsic, i.e. it defines the map $\mathfrak{T}$ as a map going from $M\times TM$ to M in a coordinate independent way. In this setting, the gradients should be computed with respect to the metric g of the manifold.

Under assumptions (A1) and (A2), the mapping G_ϕ is single-valued outside of a set of Hausdorff dimension less than or equal to $n-1$, hence, if μ_0 does not give mass to sets of Hausdorff dimension less than $n-1$, G_ϕ will be the optimal map for Monge's problem while the optimal measure in (9) will be $\pi=\mu_0\otimes\delta_{G_\phi(x)}$. So, after having relaxed the constraint that the optimal π should be supported on the graph of a map, one still obtains a minimizer that satisfies this constraint.

Notice that Monge's historical cost was equal to the distance itself: $c(x,y)=|x-y|$. One sees immediately that for this cost function, there is not a unique y such that $-\nabla_x c(x,y)=\nabla\phi(x)$, hence assumption (A1) is not satisfied and, indeed, there is in general no uniqueness of the optimal map.

We now state a general existence theorem, under assumptions that are clearly not minimal, but that will suffice for the scope of this paper, where we deal with regularity issues.

THEOREM 2.7. *Let Ω and Ω' be two bounded domains of $\mathbb{R}^n$. Let $c\in C^4(\bar\Omega\times\bar\Omega')$ satisfy assumptions* (A0)–(A2). *Let μ_0 and μ_1 be two probability measures on Ω and Ω'. Assume that μ_0 does not give mass to sets of Hausdorff dimension less than or equal to $n-1$. Then there exists a μ_0-a.e. unique minimizer T of Monge's optimal transportation problem* (2). *Moreover, there is a c-convex ϕ on Ω such that $T=G_\phi$* (*see Definition* 2.3). *Finally, if ψ is c-convex and satisfies $(G_\psi)_\#\mu_0=\mu_1$, then*

$$\nabla\psi=\nabla\phi \quad \mu_0\text{-}a.e.$$

2.3. Notion of c-convexity for sets

Following [24] and [30], we here introduce the notions that extend naturally the notions of convexity/strict convexity for a set.

Definition 2.8. (c-segment) Let $p\mapsto\mathfrak{T}_x(p)$ be the mapping given in Definition 2.6. The point x being held fixed, a *c-segment* with respect to x is the image by $\mathfrak{T}_x$ of a segment of $\mathbb{R}^n$.

If for $v_0,v_1\in\mathbb{R}^n$ we have $\mathfrak{T}_x(v_i)=y_i$, $i=0,1$, the c-segment with respect to x joining y_0 to y_1 will be $\{y_\theta:\theta\in[0,1]\}$, where $y_\theta=\mathfrak{T}_x(\theta v_1+(1-\theta)v_0)$. It will be denoted $[y_0,y_1]_x$.

Definition 2.9. (c-convex sets) Let $\Omega,\Omega'\subset\mathbb{R}^n$. We say that Ω' is *c-convex* with respect to Ω if for all $y_0,y_1\in\Omega'$ and $x\in\Omega$, the c-segment $[y_0,y_1]_x$ is contained in Ω'.

Remark. Note that this can be said in the following way: for all $x\in\Omega$, the set $-\nabla_x c(x,\Omega')$ is convex.

Definition 2.10. (Uniform strict c-convexity of sets) Given two subsets Ω and Ω' of $\mathbb{R}^n$, we say that Ω' is *uniformly strictly c-convex* with respect to Ω if the sets $\{-\nabla_x c(x,\Omega')\}_{x\in\Omega}$ are uniformly strictly convex, uniformly with respect to x. We say that Ω is uniformly strictly c*-convex with respect to Ω' if the dual assertion holds true.

Remark 1. In local coordinates, Ω is uniformly strictly c*-convex with respect to Ω' reads

$$[D_i\gamma_j(x)-D_{p_k}\mathcal{A}_{ij}(x,p)\gamma_k]\tau_i\tau_j \geqslant \varepsilon_0 > 0 \quad \text{for some } \varepsilon_0>0, \tag{15}$$

for all $x\in\partial\Omega$, $p\in-\nabla_x c(x,\Omega')$, unit tangent vector τ and outer unit normal γ.

Remark 2. When $\mathcal{A}$ does not depend on p, one recovers the usual convexity.

Remarks on the sub-differential and c-sub-differential. The question is to know if we have, for all c-convex ϕ on Ω and all $x\in\Omega$, $\partial\phi(x)=\partial^c\phi(x)$. Clearly, when ϕ is c-convex and differentiable at x, the equality holds. For an extremal point p of $\partial\phi(x)$, there will be a sequence x_n converging to x such that ϕ is differentiable at x_n and $\lim_{n\to\infty}\nabla\phi(x_n)=p$. Hence, extremal points of $\partial\phi(x)$ belong to $\partial^c\phi(x)$. Then it is not hard to show the following result.

PROPOSITION 2.11. *Assume that Ω' is c-convex with respect to Ω. Then the following assertions are equivalent*:

(1) *For all c-convex ϕ on Ω and all $x\in\Omega$, $\partial^c\phi(x)=\partial\phi(x)$.*
(2) *For all c-convex ϕ on Ω and all $x\in\Omega$, $\partial^c\phi(x)$ is convex.*
(3) *For all c-convex ϕ on Ω and all $x\in\Omega$, $G_\phi(x)$ is c-convex with respect to x.*
(4) *For all c-convex ϕ on Ω and all $x\in\Omega$, $G_\phi(x)$ is connected.*

Proof. We prove only that (4) implies (2). First, the connectedness of $G_\phi(x)$ implies the connectedness of $\partial^c\phi(x)$, since $\nabla_x c$ is continuous. Then, for $x_0\in\Omega$ and $y_0,y_1\in\Omega'$, assume that y_0 and y_1 both belong to $G_\phi(x_0)$. Letting

$$h(x)=\max\{-c(x,y_0)+c(x_0,y_0)+\phi(x_0),-c(x,y_1)+c(x_0,y_1)+\phi(x_0)\},$$

one has $\phi(x)\geqslant h(x)$ on Ω, with equality at $x=x_0$. Hence $\partial^c h(x_0)\subset\partial^c\phi(x_0)$. Since (4) holds, $\partial^c h(x_0)$ is connected, and as it is included in $\partial h(x_0)$ which is a segment, it is equal to the segment $[-\nabla_x c(x_0,y_0),-\nabla_x c(x_0,y_1)]$. This shows that $\partial^c\phi(x_0)$ is convex. $\square$

2.4. The Monge–Ampère equation

In all cases, for a C^2-smooth c-convex potential ϕ such that $(G_\phi)_\#\mu_0=\mu_1$, the conservation of mass is expressed in local coordinates by the following Monge–Ampère equation:

$$\det(D^2_{xx}c(x,G_\phi(x))+D^2\phi)=|\det D^2_{xy}c|\frac{\varrho_0(x)}{\varrho_1(G_\phi(x))}, \tag{16}$$

where $\varrho_i=d\mu_i/\mathrm{dVol}$ denotes the density of μ_i with respect to the Lebesgue measure. (See [24] for a derivation of this equation, or [12] or [14].) Hence, the equation fits into the general form (3).

2.5. Generalized solutions

Definition 2.12. (Generalized solutions) Let $\phi:\Omega\to\mathbb{R}$ be a c-convex function. Then

- ϕ is a *weak Alexandrov solution* to (16) if

$$\mu_0(B)=\mu_1(G_\phi(B)) \quad \text{for all } B\subset\Omega. \tag{17}$$

This will be denoted by

$$\mu_0=(G_\phi)^\#\mu_1.$$

- ϕ is a *weak Brenier solution* to (16) if

$$\mu_1(B')=\mu_0(G_\phi^{-1}(B')) \quad \text{for all } B'\subset\Omega'. \tag{18}$$

This is equivalent to

$$\mu_1=(G_\phi)_\#\mu_0.$$

Alexandrov and Brenier solutions. First notice that in the definition (18), μ_1 is deduced from μ_0, while it is the contrary in (17). As we have seen, the Kantorovitch procedure (10) yields an optimal transport map whenever μ_0 does not give mass to sets of Hausdorff dimension less than $n-1$. Moreover, the map G_ϕ will satisfy (18) by construction, and hence will be a weak Brenier solution to (16). Taking advantage of the c-convexity of ϕ, one can show that whenever μ_1 is absolutely continuous with respect to the Lebesgue measure, $(G_\phi)^\#\mu_1$ is countably additive, and hence is a Radon measure (see [24, Lemma 3.4]); then a Brenier solution is an Alexandrov solution. Note that if one considers $\mu_0=(G_\phi)^\#\mathrm{dVol}$, this will be the Monge–Ampère measure of ϕ. Most importantly, for μ_0 supported in Ω, $(G_\phi)_\#\mu_0=\mathbf{1}_{\Omega'}\mathrm{dVol}$ does not imply $(G_\phi)^\#\mathrm{dVol}=\mu_0$, except if Ω' is c-convex with respect to Ω (see [24]).

2.6. Cost-sectional curvature and conditions (Aw) and (As)

A central notion in the present paper will be the notion of cost-sectional curvature

$$\mathfrak{S}_{\mathfrak{c}}(x,y).$$

Definition 2.13. Under assumptions (A0)–(A2), one can define on $T_x\Omega\times T_x\Omega$ the real-valued map

$$\mathfrak{S}_{\mathfrak{c}}(x_0,y_0)(\xi,\nu)=D^4_{p_\nu p_\nu x_\xi x_\xi}[(x,p)\mapsto -c(x,\mathfrak{T}_{x_0}(p))]|_{x_0,p_0=-\nabla_x c(x_0,y_0)}. \tag{19}$$

When ξ and ν are unit orthogonal vectors, $\mathfrak{S}_{\mathfrak{c}}(x_0,y_0)(\xi,\nu)$ defines the cost-sectional curvature from x_0 to y_0 in directions (ξ,ν). Definition (19) is equivalent to the following:

$$\mathfrak{S}_{\mathfrak{c}}(x_0,y_0)(\xi,\nu)=D^2_{tt}D^2_{ss}[(s,t)\mapsto -c(\exp_{x_0}(t\xi),\mathfrak{T}_{x_0}(p_0+s\nu))]|_{t,s=0}. \tag{20}$$

The fact that (19) and (20) are equivalent follows from the following observation.

PROPOSITION 2.14. *The definition of $\mathfrak{S}_{\mathfrak{c}}(x_0,y_0)(\xi,\nu)$ is intrinsic, i.e. depends only on $(x_0,y_0)\in\Omega\times\Omega'$ and on $(\xi,\nu)\in T_{x_0}(\Omega)\times T_{x_0}(\Omega)$, but not on the choice of local coordinates around x_0 or y_0. Moreover, it is symmetric: letting $c^*(y,x)=c(x,y)$ and $\mathfrak{T}^*$ be the c^*-exponential map, the identity*

$$\mathfrak{S}_{\mathfrak{c}}(x_0,y_0)(\xi,\nu)=\mathfrak{S}_{\mathfrak{c}*}(y_0,x_0)(\tilde{\nu},\tilde{\xi}) \tag{21}$$

holds with

$$\tilde{\nu}=D_p\mathfrak{T}_{x_0}(p_0)\cdot\nu \quad \textit{and} \quad \tilde{\xi}=[D_q\mathfrak{T}^*_{y_0}(q_0)]^{-1}\cdot\xi,$$

with p_0 as above and $q_0=-\nabla_y c(x_0,y_0)$. Notice that whenever $\xi\perp\nu$, one has $\tilde{\xi}\perp\tilde{\nu}$.

The proof is deferred to the appendix.

Remark. The intrinsic nature of the cost-sectional curvature tensor has been observed independently in [20].

We are now ready to introduce the conditions:

(As) The cost-sectional curvature is uniformly positive, i.e. there exists $C_0>0$ such that for all $(x,y)\in\Omega\times\Omega'$ and all $(\nu,\xi)\in\mathbb{R}^n\times\mathbb{R}^n$ with $\xi\perp\nu$,

$$\mathfrak{S}_{\mathfrak{c}}(x,y)(\xi,\nu)\geqslant C_0|\xi|^2|\nu|^2.$$

(Aw) The cost-sectional curvature is non-negative: (As) is satisfied with $C_0=0$.

Remark on the symmetry of the conditions on c. Let $c^*(y,x):=c(x,y)$. From Proposition 2.14, one checks that if c satisfies (Aw) (resp. (As)) then c^* satisfies (Aw) (resp. (As) with a different constant). Conditions (A0) and (A2) are also clearly satisfied by c^* if satisfied by c.

2.7. The Riemannian case

The construction of optimal maps has been extended in a natural way to smooth compact Riemannian manifolds by McCann in [25] for Lipschitz semi-concave costs. All the above definitions can be translated unambiguously to the Riemannian setting. In particular, the notions of c-exponential map and c-convexity are intrinsic notions (see Remark 2 after Definition 2.6). The definition of cost-sectional curvature (2.13) extends also naturally to the Riemannian setting. Since it has been proved in Proposition 2.14 that the value of the cost-sectional curvature is coordinate-independent, this gives sense to conditions (Aw) and (As) on a Riemannian manifold. However, one needs to restrict to the set of pairs (x, y) such that c is smooth in a neighborhood of (x, y), and this becomes an issue for costs that are functions of the distance: Indeed, on a compact manifold, the distance cannot be smooth on the whole of $M\times M$ (due to the cut-locus). Hence the Riemannian case requires to weaken somehow assumption (A0). For x in M, we let Dom_x be the set of y such that $c(x, y)$ is smooth at (x, y). As developed by the author in [22], and with P. Delanoë in [15], but also by Y. Kim and R. McCann [20], or by C. Villani in [34], the relevant geometric condition on M that replaces (A0) is the following:

$$\text{For all } x\in M,\ \mathfrak{T}_x^{-1}(\mathrm{Dom}_x) = -\nabla_x c(x, \mathrm{Dom}_x) \text{ is convex.}$$

A case of interest is when $c(\cdot\,,\cdot)=\frac{1}{2}d^2(\cdot\,,\cdot)$ with $d(\cdot\,,\cdot)$ being the distance function (quadratic cost). Then, the c-exponential map is the exponential map, the mapping G_ϕ is $x\mapsto\exp_x(\nabla_g\phi)$, the gradient $\nabla_g\phi$ being relative to the Riemannian metric g. (We recall that gradient mappings were first introduced by X. Cabré [2], to generalize the Alexandrov–Bakelman–Pucci estimate on Riemannian manifolds.) Then, for x in M, we have $\mathrm{Dom}_x=M\setminus\text{cut-locus}(x)$. In [22], we address the problem of the quadratic cost on the sphere, as well as the cost $c(x, y)=-\log|x-y|$, that appears in the design of optimal reflector antennas. To establish our regularity results, we need to show a priori that $T(x)$ remains uniformly far from the boundary of Dom_x. This is precisely the object of [15].

2.8. Previous regularity results for optimal maps

The regularity of optimal maps follows from the regularity of the c-convex potential solution of the Monge–Ampère equation (16), the former being as smooth as the gradient of the latter. It thus falls into the theory of viscosity solutions of fully non-linear elliptic equations [10]; however, the Monge–Ampère equation is degenerate elliptic. A very complete reference concerning the regularity theory for the quadratic case are the lecture notes by J. Urbas [32]. Two types of regularity results are usually sought for this type of equations.

Classical regularity. Show that the equation has classical C^2 solutions, provided the measures are smooth enough, and assuming some boundary conditions. Due to the log-concavity of the Monge–Ampère operator, and using classical elliptic theory (see for instance [18]), C^∞ regularity of the solution of (16) follows from C^2 a priori estimates.

Partial regularity. Show that a weak solution of (16) is C^1 or $C^{1,\alpha}$ under suitable conditions. We mention also that $W^{2,p}$ regularity results can be obtained.

The Euclidean Monge–Ampère equation and the quadratic cost. This corresponds to the case where the cost function is the Euclidean distance squared

$$c(x,y)=|x-y|^2$$

(or equivalently $c(x,y)=-x\cdot y$), for which c-convexity means convexity in the usual sense, $G_\phi(x)=\nabla\phi(x)$ and equation (16) takes the following form:

$$\det D^2\phi=\frac{\varrho_0(x)}{\varrho_1(\nabla\phi(x))}. \tag{22}$$

Here again, we have $\varrho_i=d\mu_0/\mathrm{dVol}$, $i=0,1$. Classical regularity has been established by Caffarelli [3], [6], [7], [9], Delanoë [13] and Urbas [31]. The optimal classical regularity result, found in [3] and [9], is that for C^α-smooth positive densities, and uniformly strictly convex domains, the solution of (22) is $C^{2,\alpha}(\bar\Omega)$. Partial regularity results have been obtained by Caffarelli [4], [5], [6], [7], where it is shown that for μ_0 and μ_1 having densities bounded away from 0 and infinity, the solution of (22) is $C^{1,\alpha}$. Due to counterexamples by Wang [35], those results are close to optimal.

The reflector antenna. The design of reflector antennas can be formulated as a problem of optimal transportation on the unit sphere with cost equal to $-\log|x-y|$. The potential (height function) $\phi:\mathbb{S}^{n-1}\to\mathbb{R}^+$ parameterizes the antenna A as follows: $A=\{x\phi(x):x\in\mathbb{S}^{n-1}\}$. Then the antenna is admissible if and only if ϕ is c-convex on $\mathbb{S}^{n-1}$ for $c(x,y)=-\log|x-y|$, and $G_\phi(x)$ yields the direction in which the ray coming in the direction x is reflected. This is the first non-quadratic cost for which regularity of solutions has been established. Wang [36], [37] (see also Guan and Wang [19] where the results are extended to higher dimension) has shown classical C^2 (and hence C^∞) regularity of solutions of the associated Monge–Ampère equation when the densities are smooth. In a recent work, with totally different techniques, Caffarelli, Huang and Gutiérrez [11] have shown C^1 regularity for the solution (i.e. continuity of the optimal map) under the condition that the measures μ_0 and μ_1 have densities bounded away from 0 and infinity. This application will also be addressed by our forthcoming paper [22].

General costs and the conditions (As) and (Aw). Recently an important step was achieved in two papers by Ma, Trudinger and Wang . They gave in the first paper [24] a sufficient condition ((As), called (A3) in their paper) for C^2 (and subsequently C^∞) interior regularity. In the second paper [30], they could lower this condition down to (Aw) (condition (A3w) in their paper) to obtain a sufficient condition for global C^2 (and subsequently C^∞) regularity, assuming uniform strict c-convexity and smoothness of the domains. Note that the result under (Aw) recovers the results of Urbas and Delanoë for the quadratic cost. We mention that the results obtained in [24] and [30] have been exposed by Trudinger in [28].

THEOREM 2.15. ([28], [30]) *Let Ω and Ω' be two bounded domains of $\mathbb{R}^n$. Assume that Ω and Ω' are uniformly strictly c- and c^*-convex with respect to each other. Let c and c^* satisfy* (A0)–(A2) *and* (Aw) *on $\Omega\times\Omega'$. Let μ_0 and μ_1 be two probability measures on Ω and Ω' having densities ϱ_0 and ϱ_1. Assume that $\varrho_0\in C^2(\bar\Omega)$ and $\varrho_1\in C^2(\bar\Omega')$ are bounded away from* 0. *Then, for a c-convex ϕ on Ω such that $(G_\phi)_\#\mu_0=\mu_1$, we have*

$$\phi\in C^3(\Omega)\cap C^2(\bar\Omega).$$

We also mention the continuity result obtained in [17] concerning optimal transportation between boundaries of uniformly convex domains, that might have some connections with the present work.

3. Results

We present some answers to the following four questions:

(1) Is there a sharp necessary and sufficient condition on the cost function which would guarantee that when both measures have C^∞-smooth densities, and their supports satisfy usual convexity assumptions, the solution of (16) (and hence the optimal map) is C^∞-smooth?

(2) Is there a necessary and sufficient condition on the cost function and on the data under which optimal maps are continuous?

(3) What are the cost-functions for which connectedness of the contact set (7) holds?

(4) When the cost is set to be the squared distance of a Riemannian manifold, what is the meaning of conditions (Aw) and (As) in terms of the Riemannian metric?

3.1. Condition (Aw), connectedness of the contact set and regularity issues

Answer to questions (1) *and* (3): (Aw) is necessary and sufficient for regularity of optimal maps. Moreover, (Aw) is equivalent to the connectedness of the contact set.

In the following theorem, "smooth" means C^∞-smooth. This is for simplicity, and one can lower the smoothness assumptions on the domains and the measures, see [30].

THEOREM 3.1. *Let Ω and Ω' be two bounded domains of $\mathbb{R}^n$. Let c be a cost function that satisfies* (A0)–(A2) *on $\Omega\times\Omega'$. Assume that Ω and Ω' are smooth, uniformly strictly c- and c^*-convex with respect to each other. Then, the following assertions are equivalent*:

(1) *the cost function c satisfies* (Aw) *in $\Omega\times\Omega'$*;

(2) *for smooth strictly positive probability measures μ_0 and μ_1 in $\bar\Omega$ and $\bar\Omega'$, there exists a c-convex potential $\phi\in C^1(\Omega)$ such that*

$$(G_\phi)_\# \mu_0 = \mu_1;$$

(3) *for smooth strictly positive probability measures μ_0 and μ_1 in $\bar\Omega$ and $\bar\Omega'$, there exists a c-convex potential $\phi\in C^\infty(\bar\Omega)$ such that*

$$(G_\phi)_\# \mu_0 = \mu_1;$$

(4) *for all c-convex ϕ in Ω and all $x\in\Omega$,*

$$\partial^c\phi(x) = \partial\phi(x);$$

(5) *for all c-convex ϕ in Ω and all $x\in\Omega$, the set*

$$\{y : \phi(x) + \phi^c(y) = -c(x,y)\}$$

is c-convex with respect to x;

(6) *continuously differentiable c-convex potentials are dense among c-convex potentials for the topology of local uniform convergence.*

Hence, if condition (Aw) *is violated at some points $(x_0,y_0)\in\Omega\times\Omega'$, there exist smooth positive measures μ_0 and μ_1 on Ω and Ω' such that there exists no C^1 c-convex potential satisfying $(G_\phi)_\#\mu_0 = \mu_1$.*

Remark. Setting $c^*(y,x)=c(x,y)$, we have seen that $\mathfrak{S}_{\mathfrak{c}}\geqslant 0$ implies $\mathfrak{S}_{\mathfrak{c}^*}\geqslant 0$. Hence, all of those assertions are equivalent to their dual counterpart.

We can add the following equivalent condition for (Aw).

THEOREM 3.2. *Under the assumptions of Theorem* 3.1, *condition* (Aw) *holds if and only if for any $x_0\in\Omega$ and any $(y_0,y_1)\in\Omega'$, letting $\bar\phi$ be defined by*

$$\bar\phi(x) = \max\{-c(x,y_0)+c(x_0,y_0), -c(x,y_1)+c(x_0,y_1)\},$$

we have that, for any $y_\theta\in[y_0,y_1]_{x_0}$ (see Definition 2.8),

$$\bar\phi(x) \geqslant -c(x,y_\theta)+c(x_0,y_\theta)$$

holds in Ω.

In other words, $f_\theta(x)=-c(x,y_\theta)+c(x_0,y_\theta)$, which is the supporting function that interpolates at x_0 (non-linearly) between

$$f_0(x) = -c(x,y_0)+c(x_0,y_0) \quad \text{and} \quad f_1(x) = -c(x,y_1)+c(x_0,y_1),$$

has to remain below $\max\{f_0, f_1\}$.

Remark 1. The function $\bar{\phi}$ furnishes the counterexample to regularity when (Aw) is not satisfied, since for a suitable choice of x_0, y_0 and y_1, $\bar{\phi}$ cannot be approximated by C^1 c-convex potentials.

Remark 2. As shown by Propositions 5.1 and 5.12, a quantitative version of Theorem 3.2 holds to express condition (As).

Remark 3. The implications (1) ⇒ (2), (1) ⇒ (3) and (1) ⇒ (6) belong to Trudinger and Wang in [30]. We show here that condition (Aw) is necessary: if it is violated at some point, one can always build a counterexample where the solution to (16) is not C^1 even with C^∞-smooth positive measures and good boundary conditions (hence the optimal map is not continuous). Moreover, condition (Aw) is equivalent to a very natural geometric property of c-convex functions.

3.2. Improved partial regularity under condition (As)

Partial answer to question (2): There is partial (i.e. C^1 and $C^{1,\alpha}$) regularity under (As), requiring much lower assumptions on the measures than what is needed in the quadratic case. There cannot be C^1 regularity without (Aw). When only (Aw) is satisfied, the question of C^1 regularity remains open, except for the case $c(x,y)=|x-y|^2$ treated by Caffarelli [7].

Let us begin by giving the two integrability conditions that will be used in this result. The first one reads:

$$\begin{gathered}\text{For some } p \in \,]n,+\infty] \text{ and } C_{\mu_0} > 0, \\ \mu_0(B_\varepsilon(x)) \leqslant C_{\mu_0}\varepsilon^{n(1-1/p)} \text{ for all } \varepsilon \geqslant 0 \text{ and } x \in \Omega.\end{gathered} \tag{23}$$

The second condition reads:

$$\begin{gathered}\text{For some } f: \mathbb{R}^+ \to \mathbb{R}^+ \text{ with } \lim_{\varepsilon\to 0} f(\varepsilon) = 0, \\ \mu_0(B_\varepsilon(x)) \leqslant f(\varepsilon)\varepsilon^{n(1-1/n)} \text{ for all } \varepsilon \geqslant 0 \text{ and } x \in \Omega.\end{gathered} \tag{24}$$

In order to appreciate the forthcoming theorem, let us mention a few facts on these integrability conditions (the proof of this proposition is given at the end of the paper).

Proposition 3.3. *Let μ_0 be a probability measure on $\mathbb{R}^n$.*

(1) *If μ_0 satisfies* (23) *for some $p>n$, then μ_0 satisfies* (24).

(2) *If $\mu_0 \in L^p(\Omega)$ for some $p>n$, then μ_0 satisfies* (23) *with the same p.*

(3) *If $\mu_0 \in L^n(\Omega)$, then μ_0 satisfies* (24).

(4) *If μ_0 satisfies* (24), *then μ_0 does not give mass to sets of Hausdorff dimension less than or equal to $n-1$, hence* (24) *guarantees the existence of an optimal map.*

(5) *There are probability measures on Ω that satisfy* (23) (*and hence* (24)) *and which are not absolutely continuous with respect to the Lebesgue measure.*

Then our result is the following.

Theorem 3.4. *Let c be a cost function that satisfies assumptions* (A0)–(A2) *and* (As) *on $\Omega \times \Omega'$, Ω and Ω' being bounded domains uniformly strictly c- and c*-convex with respect to each other. Let μ_0 and μ_1 be probability measures respectively on Ω and $\omega' \subset \Omega'$, with ω' c-convex with respect to Ω. Let ϕ be a c-convex potential on Ω such that $(G_\phi)_\# \mu_0 = \mu_1$. Assume that $\mu_1 \geqslant m \, \mathrm{dVol}$ on ω' for some $m>0$.*

(1) *Assume that μ_0 satisfies* (23) *for some $p>n$. Let*

$$\alpha = 1 - \frac{n}{p} \quad \textit{and} \quad \beta = \frac{\alpha}{4n-2+\alpha}.$$

Then, for any $\delta>0$, we have

$$\|\phi\|_{C^{1,\beta}(\Omega_\delta)} \leqslant \mathcal{C},$$

and $\mathcal{C}$ depends only on $\delta>0$, on the constant C_{μ_0} in (23), *on m, on the constants in conditions* (A0)–(A2) *and* (As), *and on $C_{\mathfrak{T}}$ in* (13).

(2) *If μ_0 satisfies* (24), *then ϕ belongs to $C^1(\Omega_\delta)$ and the modulus of continuity of $\nabla \phi$ is controlled by f in* (24).

As an easy corollary of Theorem 3.4, we can extend the C^1 estimates to the boundary if the support of the measure μ_0 is compactly contained in Ω.

Theorem 3.5. *Assume, in addition to the assumptions of Theorem* 3.4, *that μ_0 is supported in $\bar{\omega}$, with ω compactly contained in Ω. Then, if μ_0 satisfies* (24), *$\phi \in C^1(\bar{\omega})$, and if μ_0 satisfies* (23), *$\phi \in C^{1,\beta}(\bar{\omega})$, with β as in Theorem* 3.4.

Remark on the conditions on Ω and Ω'. Our result holds true for μ_0 supported in any subset ω of Ω (hence not necessarily c*-convex), and μ_1 supported in any subset ω' of Ω' c-convex (but not necessarily strictly) with respect to Ω. Hence, what we need is the existence of supersets Ω and Ω' uniformly c- and c*-convex with respect to each other, in order to use the results of [30]. The only point where we need this condition is during the proof of Proposition 5.6, where we rely on Theorem 3.1 to assert that $\partial\phi = \partial^c \phi$. However, in [24], Ma, Trudinger and Wang proved the following theorem.

Theorem 3.6. ([24]) *Let c satisfy* (A0)–(A2) *and* (As) *on* $\Omega\times\Omega'$, Ω' *being c-convex with respect to* Ω. *Then, for* C^2*-smooth positive probability measures* μ_0 *and* μ_1 *on* Ω *and* Ω', *the c-convex potential* ϕ *such that* $(G_\phi)_\#\mu_0=\mu_1$ *is* C^2*-smooth inside* Ω.

Using this result, Proposition 4.4 yields that for all c-convex potentials ϕ on Ω, $\partial^c\phi=\partial\phi$. Hence, we could have relaxed the assumptions of Theorem 3.4 on Ω and Ω', only requiring Ω' to be c-convex with respect to Ω (i.e. no c*-convexity on Ω and no strict c-convexity of Ω'). Note that the proof of Theorem 3.6 has been completed later on by Trudinger and Wang in [29], relying in part on our Proposition 5.1 (which is an independent result). Thus, we can now state the following result.

Theorem 3.7. *The results of Theorem* 3.4 *hold assuming only for* Ω *and* Ω' *that* Ω' *is c-convex with respect to* Ω.

We mention that the results of Kim and McCann [20], obtained simultaneously with those of [29] but using different techniques, also allow for completion of the proof of Theorem 3.6, under the assumption that Ω' and Ω are c- and c*-convex with respect to each other. This allows one to drop the strict convexity assumption in Theorem 3.4.

Remark on the integrability conditions. The integrability conditions on μ_0 and μ_1 are really mild: we only ask that μ_1 be bounded from below, and that $\mu_0(B_r)\leqslant r^{n-p}$ for $p\geqslant 1$ ($p>1$ yields $C^{1,\alpha}$ regularity) (see conditions (23) and (24) and the subsequent discussion). The continuity of the optimal map is also asserted in the case $\mu_0\in L^n$ (that implies (24)), which is somehow surprising: indeed $D^2\phi\in L^n$ does not imply $\phi\in C^1$, but here $\det(D^2\phi-\mathcal{A}(x,\nabla\phi))\in L^n$ implies $\phi\in C^1$. In a forthcoming work, we shall show that our result adapts to the reflector antenna, hence improving the result obtained independently by Caffarelli, Gutiérrez and Huang [11] on reflector antennas. Moreover, our techniques yield quantitative $C^{1,\alpha}$ estimates: the exponent α can be explicitly computed. Finally, our continuity estimates extends up to the boundary (Theorem 3.5). This is achieved through a geometric formulation of condition (As).

A fully satisfactory answer would include a general result of partial regularity under condition (Aw). This result is expected in view of the Euclidean case (since the quadratic cost is really the limit case for condition (Aw)). Note that, in view of counterexamples given in [35], the results under (Aw) cannot be as good as under (As), and cannot be much better than Caffarelli's results [7] that require densities bounded away from 0 and infinity.

3.3. Conditions (Aw) and (As) for the quadratic cost of a Riemannian manifold

We refer the reader to Remark 2 after the definition of the cost-sectional curvature (19), where the intrinsic meaning of (19) on a manifold is discussed.

Partial answer to question (4): When the cost is the Riemannian distance squared, the cost-sectional curvature at $y=x$ equals (up to a multiplicative constant) the Riemannian sectional curvature.

Theorem 3.8. *Let M be a C^4 Riemannian manifold. Let $c(x,y)=\frac{1}{2}d^2(x,y)$ for all $(x,y)\in M\times M$. Let $\mathfrak{S}_{\mathfrak{c}}$ be given by* (19). *Then, for all $\xi,\nu\in T_xM$,*

$$\frac{\mathfrak{S}_{\mathfrak{c}}(x,x)(\nu,\xi)}{|\xi|_g^2|\nu|_g^2-(\xi\cdot\nu)_g^2}=\frac{2}{3}\cdot \textit{sectional curvature of } M \textit{ at } x \textit{ in the 2-plane } (\xi,\nu).$$

Hence, if (Aw) (*resp.* (As)) *is satisfied at (x,x), the sectional curvature of M at x is non-negative* (*resp. strictly positive*).

Corollary 3.9. *Let M be a compact Riemannian manifold. If the sectional curvature of M is not everywhere non-negative, there are smooth positive measures on M such that the optimal map* (*for the cost function $c(x,y)=\frac{1}{2}d^2(x,y)$*) *is not continuous.*

At the end of the proof of Theorem 3.8, we give a counterexample to regularity for a 2-dimensional manifold with negative sectional curvature.

This observation closes (with a negative answer) the open problem of the regularity of optimal gradient maps when the manifold does not have non-negative sectional curvature everywhere. There is a partial converse assertion in the special case of constant sectional curvature: *The quadratic cost on the round sphere $\mathbb{S}^{n-1}$ satisfies* (As). This will be the object of a forthcoming work [22]. Hence our previous result can be adapted to this Riemannian case.

3.4. Examples of costs that satisfy (As) or (Aw)

We repeat the collection of costs that was given in [24] and [30].

- $c(x,y)=\sqrt{1+|x-y|^2}$ satisfies (As).
- $c(x,y)=\sqrt{1-|x-y|^2}$ satisfies (As).
- $c(x,y)=(1+|x-y|^2)^{p/2}$ satisfies (As) for $1\leqslant p<2$ and $|x-y|^2<1/(p-1)$.
- $c(x,y)=|x-y|^2+|f(x)-g(y)|^2$, $f,g\in C^4(\mathbb{R}^n;\mathbb{R})$ convex (resp. strictly convex) with $|\nabla f|,|\nabla g|<1$, satisfies (Aw) (resp. (As)).
- $c(x,y)=\pm|x-y|^p/p$, $p\neq 0$, satisfies (Aw) for $p=\pm 2$ and (As) for $-2<p<1$ (for the cost $-|x-y|^p/p$ only).

- $c(x,y)=-\log|x-y|$ satisfies (As) on

$$(\mathbb{R}^n\times\mathbb{R}^n)\setminus\{(x,x):x\in\mathbb{R}^n\}.$$

- The reflector antenna problem ([36]) corresponds to the case $c(x,y)=-\log|x-y|$ restricted to $\mathbb{S}^n$. As pointed out in [30], this cost satisfies (As) on

$$(\mathbb{S}^{n-1}\times\mathbb{S}^{n-1})\setminus\{(x,x):x\in\mathbb{S}^{n-1}\}.$$

- As shown in the forthcoming paper [22], the squared Riemannian distance on the sphere satisfies (As) on the set

$$(\mathbb{S}^{n-1}\times\mathbb{S}^{n-1})\setminus\{(x,-x):x\in\mathbb{S}^{n-1}\}.$$

Note that it is the restriction to $\mathbb{S}^{n-1}$ of the cost $c(x,y)=\theta^2(x,y)$, where θ is the angle formed by x and y. (For those two cases, see §2.7 where the meaning of conditions (Aw) and (As) on a Riemannian manifold is discussed).

4. Proof of Theorem 3.1

We begin with the following uniqueness result of independent interest.

PROPOSITION 4.1. *Let μ and ν be two probability measures on Ω and Ω', with Ω and Ω' being connected domains of $\mathbb{R}^n$. Assume that either μ or ν is positive Lebesgue almost everywhere in Ω (resp. in Ω'). Then, among all pairs of functions (ϕ,ψ) such that ϕ is c-convex and ψ is c^*-convex, the problem* (10) *has at most one minimizer, up to an additive constant.*

The proof of this proposition is deferred to the end of the paper.

4.1. Condition (Aw) implies connectedness of the contact set

We will begin with the following lemma.

LEMMA 4.2. *Let ϕ be c-convex. Let $\{\phi_\varepsilon\}_{\varepsilon>0}$ be a sequence of c-convex potentials that converges uniformly to ϕ on compact sets of Ω. Then, if $p=-\nabla_x c(x_0,y)\in\partial\phi(x_0)$, $(x_0,y)\in\Omega\times\Omega'$, there is a sequence $\{x_\varepsilon\}_{\varepsilon>0}$ that converges to x_0 and a sequence $\{y_\varepsilon\}_{\varepsilon>0}$ that converges to y such that $p_\varepsilon=-\nabla_x c(x_\varepsilon,y_\varepsilon)\in\partial\phi_\varepsilon(x_\varepsilon)$. Finally, p_ε converges to p.*

Proof. Let $y=\mathfrak{T}_{x_0}(p)$, i.e. $p=-\nabla_x c(x_0,y)$. Since ϕ and ϕ_ε are c-convex and c is semi-concave, there exist $K,r>0$ such that

$$\tilde{\phi}(x):=\phi(x)+\tfrac{1}{2}K|x-x_0|^2+c(x,y) \quad \text{and} \quad \tilde{\phi}_\varepsilon(x):=\phi_\varepsilon(x)+\tfrac{1}{2}K|x-x_0|^2+c(x,y)$$

are convex on $B_r(x_0)$ compactly contained in Ω. One can also assume, by subtracting a constant, that $\tilde{\phi}(x_0)=0$, and that $\tilde{\phi}(x)\geqslant 0$ on Ω. Finally, one can assume (by relabeling the sequence) that on $B_r(x_0)$ we have $|\phi_\varepsilon-\phi|\leqslant\varepsilon$.

Consider then

$$\tilde{\phi}_\varepsilon^\delta=\tilde{\phi}_\varepsilon+\tfrac{1}{2}\delta|x-x_0|^2-\varepsilon.$$

We have $\tilde{\phi}_\varepsilon^\delta(x_0)\leqslant 0$, and on $\partial B_\mu(x_0)$, with $\mu\leqslant r$,

$$\tilde{\phi}_\varepsilon^\delta(z)\geqslant\tilde{\phi}(z)+\tfrac{1}{2}\delta\mu^2-2\varepsilon\geqslant\tfrac{1}{2}\delta\mu^2-2\varepsilon.$$

By taking $\mu=\varepsilon^{1/3}$ and $\delta=4\varepsilon^{1/3}$, we get that $\tilde{\phi}_\varepsilon^\delta$ has a local minimum in $B_\mu(x_0)$, hence at some point $x_\varepsilon\in B_\mu(x_0)$ we have

$$\partial\phi_\varepsilon(x_\varepsilon)\ni-\nabla_x c(x_\varepsilon,y)-K(x_\varepsilon-x_0)-\delta(x_\varepsilon-x_0).$$

Then we have $|(K+\delta)(x_\varepsilon-x_0)|$ small, and, due to (A1) and (A2), there exists y_ε close to y such that $\nabla_x c(x_\varepsilon,y_\varepsilon)=\nabla_x c(x_\varepsilon,y)+K(x_\varepsilon-x_0)+\delta(x_\varepsilon-x_0)$. Thus, $-\phi_\varepsilon(x)-c(x,y_\varepsilon)$ has a critical point at x_ε. This implies that $p_\varepsilon=-\nabla_x c(x_\varepsilon,y_\varepsilon)\in\partial\phi_\varepsilon(x_\varepsilon)$. Finally, since $x_\varepsilon\to x$ and $y_\varepsilon\to y$, we conclude that $p_\varepsilon\to p$. □

Now we prove that $\partial^c\phi=\partial\phi$. In order to do this, we must show that if ϕ is c-convex and $-\phi(\cdot)-c(\cdot,y)$ has a critical point at x_0, then this is a global maximum.

We first have the following observation.

Lemma 4.3. *Let ϕ be c-convex. Assume that $-\phi-c(\cdot,y)$ has a critical point at x_0 (i.e. $0\in\partial\phi(x_0)+\nabla_x c(x_0,y)$) and that it is not a global maximum. Then ϕ is not differentiable at x_0.*

Proof. Indeed, $-\phi(\cdot)-c(\cdot,y)$ has a critical point at x_0, but $\phi(x_0)+\phi^c(y)=-c(x_0,y)$ does not hold. However, there is a point y' such that $\phi(x_0)+\phi^c(y')=-c(x_0,y')$. Hence, $-\nabla_x c(x_0,y),-\nabla_x c(x_0,y')\in\partial\phi(x_0)$, and we have $\nabla_x c(x_0,y)\neq\nabla_x c(x_0,y')$ by (A1). □

We now consider the following assumption:

(D) C^1 c-convex functions are dense in the set $\{\phi \text{ c-convex on } \Omega : G_\phi(\Omega)\subset\Omega'\}$ for the topology of uniform convergence on compact sets of Ω.

Proposition 4.4. *Assume that assumption* (D) *holds. Let $p=-\nabla_x c(x_0,y)\in\partial\phi(x_0)$ with ϕ c-convex. Then $-\phi(\cdot)-c(\cdot,y)$ reaches a global maximum at x_0.*

Proof. Assume the contrary, i.e. $-\phi(x_1)-c(x_1,y)>-\phi(x_0)-c(x_0,y)$ for some $x_1\in\Omega$. By (D), there exists a sequence of C^1 c-convex potentials $\{\phi_\varepsilon\}_{\varepsilon>0}$ that converges to ϕ. By Lemma 4.2, there is a sequence $\{x_\varepsilon\}_{\varepsilon>0}$ such that $x_\varepsilon\to x_0$ and $\nabla\phi_\varepsilon(x_\varepsilon)\to-\nabla_x c(x_0,y)$. Let y_ε be such that $\nabla\phi_\varepsilon(x_\varepsilon)=-\nabla_x c(x_\varepsilon,y_\varepsilon)$. Then $y_\varepsilon\to y$. Since ϕ_ε is C^1, by Lemma 4.3, x_ε, the critical point of $-\phi_\varepsilon(\cdot)-c(\cdot,y_\varepsilon)$, is necessarily a global maximum. Finally, since ϕ_ε converges locally uniformly to ϕ, we see that $-\phi(\cdot)-c(\cdot,y)$ reaches a global maximum at x_0.

LEMMA 4.5. *Assume that Ω and Ω' are bounded, uniformly strictly c- and c^*-convex with respect to each other. Assume that c satisfies* (A0)–(A2) *and* (Aw) *on $\Omega\times\Omega'$. Then* (D) *holds.*

Proof. As we will see, this result is implied immediately by the result of [30] combined with Proposition 4.1. Let ϕ be c-convex and set $\mu_1=(G_\phi)_\#\mathbf{1}_\Omega\mathrm{dVol}$. Note that, by Proposition 4.1, ϕ is the unique, up to a constant, c-convex potential such that $(G_\phi)_\#\mathbf{1}_\Omega\mathrm{dVol}=\mu_1$. Consider a sequence of smooth positive densities $\{\mu_1^\varepsilon\}_{\varepsilon>0}$ in Ω' such that $\mu_1^\varepsilon\mathrm{dVol}$ converges weakly-$*$ to μ_1, and has the same total mass as μ_1. Consider ϕ_ε such that $(G_{\phi_\varepsilon})_\#\mathbf{1}_\Omega\mathrm{dVol}=\mu_1^\varepsilon\mathrm{dVol}$. From [30], ϕ_ε is C^2-smooth inside Ω. Then, by Proposition 4.1, up to a normalizing constant, ϕ_ε is converging to ϕ and $\nabla\phi_\varepsilon$ is converging to $\nabla\phi$ on the points where ϕ is differentiable. □

Hence, under the assumptions of Lemma 4.5, we have $\partial\phi(x)=\partial^c\phi(x)$. In view of Proposition 2.11, the equality $\partial\phi(x)=\partial^c\phi(x)$, for all ϕ and x, is equivalent to the c-convexity of the set

$$G_\phi(x)=\{y:\phi(x)+\phi^c(y)=-c(x,y)\}.$$

This shows that condition (Aw) is sufficient. □

4.2. Condition (Aw) is necessary for smoothness and connectedness of the contact set

We now show that if (Aw) is violated somewhere in $\Omega\times\Omega'$, then there exists a c-convex potential for which we do not have $\partial\phi=\partial^c\phi$. Assuming this, in view of Lemma 4.5 and Proposition 4.4, this implies that this potential cannot be a limit of C^1-smooth c-convex potentials. Hence, considering the sequence $\{\phi_\varepsilon\}_{\varepsilon>0}$ used in the proof of Lemma 4.5, this sequence is not C^1 for ε smaller than some ε_0. This implies in turn that there exists smooth positive densities μ_0 and μ_1 in Ω and Ω' such that the c-convex potential ϕ satisfying $(G_\phi)_\#\mu_0=\mu_1$ is not C^1-smooth.

Assume that for some $x_0\in\Omega$, $y\in\Omega'$, $p=-\nabla_x c(x_0,y)$, and for some unit vectors ξ and ν in $\mathbb{R}^n$ with $\xi\perp\nu$, one has

$$D^2_{p_\nu p_\nu}[p\mapsto D^2_{x_\xi x_\xi} c(x,\mathfrak{T}_x(p))] \geqslant N_0>0. \tag{25}$$

Let $y_0=\mathfrak{T}_{x_0}(p-\varepsilon\nu)$ and $y_1=T_{x_0}(p+\varepsilon\nu)$, with ε small, and recall that $y=\mathfrak{T}_{x_0}(p)$. Hence y is the 'middle' of the c-segment $[y_0,y_1]_x$. Let us define

$$\bar\phi(x)=\max\{-c(x,y_0)+c(x_0,y_0),-c(x,y_1)+c(x_0,y_1)\}. \tag{26}$$

(This function will be used often in the geometric interpretation of (As) and (Aw). It is the "second simplest" c-convex function, as the supremum of two supporting functions. It plays the role of $(x_1,...,x_n)\mapsto|x_1|$ in the Euclidean case.)

Note first that $\xi\perp\nu$ implies that $\xi\perp\nabla_x c(x_0,y_1)-\nabla_x c(x_0,y_0)$. Consider near x_0 a smooth curve $\gamma(t)$ such that $\gamma(0)=x_0$ and $\dot\gamma(0)=\xi$, and such that for $t\in[-\delta,\delta]$ one has

$$f_0(\gamma(t)):=-c(\gamma(t),y_0)+c(x_0,y_0)=-c(\gamma(t),y_1)+c(x_0,y_1)=:f_1(\gamma(t)).$$

Such a curve exists by the implicit function theorem, and it is C^2-smooth. On γ, we have

$$\bar\phi=\tfrac12(f_0+f_1),$$

since $f_0=f_1$ on γ. Then we compare $\frac12(f_0+f_1)$ with $-c(x,y)+c(x_0,y)$. By (25), we have

$$\tfrac12[D^2_{x_\xi x_\xi}c(x_0,y_0)+D^2_{x_\xi x_\xi}c(x_0,y_1)]\geqslant D^2_{x_\xi x_\xi}c(x_0,y)+c(\varepsilon,N_0),$$

where $c(\varepsilon,N_0)$ is positive for ε small enough. Then, of course,

$$\nabla_x c(x_0,y)=\tfrac12[\nabla_x c(x_0,y_0)+\nabla_x c(x_0,y_1)].$$

Hence we have, for ε small enough,

$$\begin{aligned}
&[-c(\gamma(t),y)+c(x_0,y)]-\bar\phi(\gamma(t))\\
&\quad=[-c(\gamma(t),y)+c(x_0,y)]-\tfrac12(f_0+f_1)(\gamma(t))\\
&\quad=\tfrac12\big[\tfrac12[D^2_{xx}c(x_0,y_0)+D^2_{xx}c(x_0,y_1)]-D^2_{xx}c(x_0,y)\big]\cdot(\gamma(t)-x_0,\gamma(t)-x_0)+o(t^2)\\
&\quad=\tfrac12\big[\tfrac12[D^2_{x_\xi x_\xi}c(x_0,y_0)+D^2_{x_\xi x_\xi}c(x_0,y_1)]-D^2_{x_\xi x_\xi}c(x_0,y)\big]t^2+o(t^2)\\
&\quad\geqslant\tfrac12 c(\varepsilon,N_0)t^2+o(t^2).
\end{aligned}$$

This will be strictly positive for $t\in[-\delta,\delta]\setminus\{0\}$ small enough, and of course the difference $-\bar\phi-[c(x,y)-c(x_0,y)]$ vanishes at x_0. Obviously, the function $\bar\phi$ is c-convex, $-\bar\phi(\cdot)-c(\cdot,y)$ has a critical point at x_0, and this is not a global maximum. Hence, from Proposition 4.4, (D) cannot hold true.

The proof of Theorems 3.1 and 3.2 is complete.

5. Proof of Theorem 3.4

5.1. Sketch of the proof

The key argument of the proof is the geometrical translation of condition (As): assume that a c-convex ϕ is not differentiable at $x=0$, hence, for some pair y_0, y_1, $-\phi(\cdot)-c(\cdot, y_0)$ and $-\phi(\cdot)-c(\cdot, y_1)$ both reach a maximum at $x=0$. (From Theorem 3.1, under (As), all critical points of $-\phi(\cdot)-c(\cdot, y)$ are global maxima.) Consider y_θ in the c-segment with respect to $x=0$ joining y_0 to y_1. As we will see in Proposition 5.1, the function $-\phi(\cdot)-c(\cdot, y_\theta)$ will reach a maximum at $x=0$, and condition (As) implies moreover that for $\theta\in[\varepsilon, 1-\varepsilon]$ this maximum will be strict in the following sense: we will have

$$-\phi(x)-c(x, y_\theta) \leqslant -\phi(0)-c(0, y_\theta)-\delta|x|^2+o(|x|^2),$$

with $\delta>0$ depending on $|y_1-y_0|$ and $C_0>0$ in condition (As), and bounded from below for θ away from 0 and 1.

Then, by estimating all supporting functions to ϕ on a small ball $B_\eta(0)$ centered at 0, we will find that for y in a $C\eta$ neighborhood of $\{y_\theta\}_{\theta\in[1/4,3/4]}$, $C>0$ depending on δ above, $-\phi(\cdot)-c(\cdot, y)$ will reach a local maximum in $B_\eta(0)$. Hence $G_\phi(B_\eta(0))$ contains a $C\eta$ neighborhood of $\{y_\theta\}_{\theta\in[1/4,3/4]}$. This is Proposition 5.6. Once this is shown, we can contradict the bound on the Jacobian determinant of G_ϕ.

We now enter into the rigorous proof of Theorem 3.4, which is articulated in three parts.

5.2. Part I. Geometric interpretation of condition (As)

This proposition is the geometrical translation of assumption (As). Actually, as we will see in Proposition 5.12, the result of Proposition 5.1 is equivalent to assumption (As) for a smooth cost function.

PROPOSITION 5.1. *Let c be a cost function that satisfies* (A0)–(A2) *and* (As) *on* $\Omega\times\Omega'$. *For $x_0\in\Omega$ and $y_0, y_1\in\Omega'$, let $\{y_\theta\}_{\theta\in[0,1]}$ be the c-segment with respect to x_0 joining y_0 to y_1, in the sense of Definition* 2.8, *and assume that Ω' is c-convex with respect to x_0. Let*

$$\bar\phi(x)=\max\{-c(x, y_0)+c(x_0, y_0), -c(x, y_1)+c(x_0, y_1)\}.$$

There exist constants $\delta_0, C>0$ and γ such that for all $\varepsilon\in\left]0, \frac12\right[$, all $\theta\in[\varepsilon, 1-\varepsilon]$ and all $x\in\Omega$ such that $|x-x_0|\leqslant C\varepsilon$,

$$\bar\phi(x) \geqslant -c(x, y_\theta)+c(x_0, y_\theta)+\delta_0\theta(1-\theta)|y_1-y_0|^2|x-x_0|^2-\gamma|x-x_0|^3,$$

with lower bounds on δ_0 *and* C *and an upper bound on* γ *that depend on the bounds in assumptions* (A0)–(A2) *and* (As), *on an upper bound on* $|y_1-y_0|$, *and on* $C_{\mathfrak{T}}$ *in* (13).

Shifting and rotating the coordinates, we can assume that $x_0=0$ and that

$$\nabla_x c(0,y_0)-\nabla_x c(0,y_1)$$

is parallel to e_1. Then, we observe the following fact.

Proposition 5.2. *Subtracting from* c *a smooth function* $x\mapsto\lambda(x)$ *that depends only on* x *does not change the map solution of the optimal transportation problem, and the new cost* $c(x,y)-\lambda(x)$ *will still satisfy assumptions* (A0)–(A2) *and* (Aw). *The optimal potential will be changed according to the rule* $\phi\mapsto\phi+\lambda$. *If moreover the function* λ *is affine, this modification does not change the bounds in assumptions* (A2) *and* (As).

Using Proposition 5.2, we can subtract from c the affine function given by

$$\lambda(x)=\nabla_x c(0,y_0)\cdot(x-x^1e_1),$$

so that the new cost c will satisfy

$$\nabla_x c(0,y_0)=-ae_1 \quad\text{and}\quad \nabla_x c(0,y_1)=-be_1 \tag{27}$$

for some $a\neq b$ from assumption (A1) (we will assume hereafter that $b>a$). Note that (27) is equivalent to

$$y_0=\mathfrak{T}_{x=0}(ae_1) \quad\text{and}\quad y_1=\mathfrak{T}_{x=0}(be_1).$$

We then have, for all $\theta\in[0,1]$,

$$-c(x,y_\theta)+c(0,y_\theta)=[\theta b+(1-\theta)a]x^1-\tfrac{1}{2}D^2_{xx}c(0,y_\theta)\cdot(x,x)+o(|x|^2). \tag{28}$$

We now have the following lemma, which is the point where we use assumption (As).

Lemma 5.3. *Under the assumptions and with the notation of Proposition* 5.1, *in particular assuming* (As), *for all* $x\in\mathbb{R}^n$ *and all* $\theta\in[0,1]$, *letting* a *and* b *be defined through* (27), *one has*

$$-D^2_{xx}c(0,y_\theta)\cdot(x,x)\leqslant-[(1-\theta)D^2_{xx}c(0,y_0)+\theta D^2_{xx}c(0,y_1)]\cdot(x,x)-\delta|x|^2+\Delta|x_1|^2,$$

where

$$\delta=\tfrac{1}{4}C_0|b-a|^2\theta(1-\theta) \quad\textit{and}\quad \Delta=\frac{\Delta_0^2}{C_0}|b-a|^2\theta(1-\theta),$$

with C_0 *given in assumption* (As) *and* Δ_0 *depending on*

$$\|c(\cdot,\cdot)\|_{C^4(\Omega\times\Omega')} \quad\textit{and}\quad \|[D_{xy}c]^{-1}\|_{L^\infty(\Omega\times\Omega')}.$$

Note in particular that under (A0) *and* (As), C_0 *is bounded away from* 0 *and infinity.*

We will also need the following elementary estimates, that we state without proof.

LEMMA 5.4. *Under the assumptions and with the notation of Proposition* 5.1, *for all* $x\in\mathbb{R}^n$ *and all* $\theta,\theta'\in[0,1]$,

$$\tfrac{1}{2}|D^2_{xx}c(0,y_\theta)\cdot(x,x)-D^2_{xx}c(0,y_{\theta'})\cdot(x,x)|\leqslant C_1|\theta-\theta'||x|^2, \tag{29}$$

where C_1 *depends on* $|b-a|$, $\|[D_{xy}c]^{-1}\|_{L^\infty(\Omega\times\Omega')}$ *and* $\|c(\cdot,\cdot)\|_{C^3(\bar\Omega\times\bar\Omega')}$.

LEMMA 5.5. *Let* $[t_0,t_1]\subset\mathbb{R}$ *and* f *belong to* $C^2([t_0,t_1],\mathbb{R})$.

(1) *If* $f''\geqslant\alpha$, *we have, for all* $t_0,t_1\in\mathbb{R}$,

$$\theta f(t_0)+(1-\theta)f(t_1)\geqslant f(\theta t_0+(1-\theta)t_1)+\tfrac{1}{2}\alpha\theta(1-\theta)|t_1-t_0|^2.$$

(2) *In all cases, we have*

$$|\theta f(t_0)+(1-\theta)f(t_1)-f(\theta t_0+(1-\theta)t_1)|\leqslant\tfrac{1}{2}\|f\|_{C^2(t_0,t_1)}\theta(1-\theta)|t_1-t_0|^2.$$

Proof of Lemma 5.3. We apply the first part of Lemma 5.5 to the function

$$f: t\longmapsto -D^2_{xx}c(0,\mathfrak{T}_{x=0}(te_1))\cdot(x',x'),$$

where x' is equal to $(0,x^2,...,x^n)$, and hence $x'\perp e_1$. From assumption (As), f satisfies $f''\geqslant C_0|x'|^2$. Then, by choosing $t_0=a$ and $t_1=b$ (note that $y_\theta=\mathfrak{T}_{x=0}((\theta b+(1-\theta)a)e_1))$, we obtain that

$$\begin{aligned}-D^2_{xx}c(0,y_\theta)\cdot(x',x')\leqslant&-[(1-\theta)D^2_{xx}c(0,y_0)+\theta D^2_{xx}c(0,y_1)]\cdot(x',x')\\&-\tfrac{1}{2}C_0|x'|^2\theta(1-\theta)|b-a|^2.\end{aligned}$$

To conclude the lemma, we have to control the terms where x^1 appears. For this, we apply the second part of Lemma 5.5 to

$$g: t\longmapsto D^2_{xx}c(x,\mathfrak{T}_x(te_1))\cdot(x,x)-D^2_{xx}c(x,\mathfrak{T}_x(te_1))\cdot(x',x'),$$

for which we have $|g''|\leqslant 2\Delta_1|x^1||x|$, where Δ_1 depends on $\|c(\cdot,\cdot)\|_{C^4}$ and $\|[D_{xy}c]^{-1}\|_{L^\infty}$. This yields

$$\begin{aligned}-D^2_{xx}c(0,y_\theta)\cdot(x,x)\leqslant&-[(1-\theta)D^2_{xx}c(0,y_0)+\theta D^2_{xx}c(0,y_1)]\cdot(x,x)\\&+\theta(1-\theta)|b-a|^2\big(-\tfrac{1}{2}C_0|x'|^2+\Delta_1|x_1||x|\big)\\\leqslant&-[(1-\theta)D^2_{xx}c(0,y_0)+\theta D^2_{xx}c(0,y_1)]\cdot(x,x)\\&+\theta(1-\theta)|b-a|^2\big(-\tfrac{1}{2}C_0|x|^2+(\Delta_1+C_0)|x_1||x|\big).\end{aligned}$$

We set $\Delta_0 = \Delta_1 + C_0$. Using a standard argument we have

$$\Delta_0|x||x^1| \leqslant \frac{1}{4}C_0|x|^2 + \frac{\Delta_0^2}{C_0}|x^1|^2,$$

and we obtain

$$-\frac{1}{2}C_0|x|^2 + \Delta_0|x||x_1| \leqslant -\frac{1}{4}C_0|x|^2 + \frac{\Delta_0^2}{C_0}|x^1|^2.$$

This concludes the proof of Lemma 5.3. □

Proof of Proposition 5.1. Using the general fact that for $f_0, f_1 \in \mathbb{R}$ and $0 \leqslant \theta \leqslant 1$,

$$\max\{f_0, f_1\} \geqslant \theta f_1 + (1-\theta) f_0,$$

we have, using (28),

$$\bar{\phi}(x) \geqslant (\theta b + (1-\theta)a)x^1 - \tfrac{1}{2}[\theta D_{xx}^2 c(0, y_1) + (1-\theta) D_{xx}^2 c(0, y_0)] \cdot (x, x) + o(|x|^2).$$

We now use assumption (As) through Lemma 5.3 to handle the second term of the right-hand side. This yields the intermediate inequality

$$\bar{\phi}(x) \geqslant (\theta b + (1-\theta)a)x_1 - \tfrac{1}{2} D_{xx}^2 c(0, y_\theta) \cdot (x, x) + \delta|x|^2 - \Delta|x^1|^2 + o(|x|^2), \tag{30}$$

with δ and Δ given in Lemma 5.3. In order to eliminate the term $-\Delta|x^1|^2$ in the right-hand side, we proceed as follows: we first write (30) for some $\theta' \in [0, 1]$, and then change it into

$$\begin{aligned}\bar{\phi}(x) \geqslant{}& (\theta b + (1-\theta)a)x_1 - \tfrac{1}{2} D_{xx}^2 c(0, y_\theta) \cdot (x, x) + \delta|x|^2 + \tfrac{1}{2}[D_{xx}^2 c(0, y_\theta) - D_{xx}^2 c(0, y_{\theta'})] \cdot (x, x) \\ &+ ((b-a)(\theta'-\theta) - \Delta x^1)x^1 + (\delta'-\delta)|x|^2 + (\Delta - \Delta')|x^1|^2 + o(|x|^2),\end{aligned}$$

where $\delta' = \delta(\theta')$ and $\Delta' = \Delta(\theta')$ as in Lemma 5.3. We now have to control the terms

$$\begin{aligned} T_1 &= ((b-a)(\theta'-\theta) - \Delta x^1)x^1, \\ T_2 &= \tfrac{1}{2}[D_{xx}^2 c(0, y_\theta) - D_{xx}^2 c(0, y_{\theta'})] \cdot (x, x), \\ T_3 &= (\delta'-\delta)|x|^2 + (\Delta - \Delta')|x^1|^2. \end{aligned}$$

The term T_1 can be cancelled through an appropriate choice of θ'. We first choose $\varepsilon > 0$ small (but fixed). Taking $\theta \in [\varepsilon, 1-\varepsilon]$, we choose θ' such that

$$\theta' = \theta + \frac{x^1 \Delta}{b-a} = \theta + \frac{x^1}{C}, \tag{31}$$

with

$$C=\frac{b-a}{\Delta}=\frac{C_0}{\theta(1-\theta)|b-a|\Delta_0^2}. \tag{32}$$

Since $\theta\in[\varepsilon,1-\varepsilon]$, this choice of θ' is possible if we restrict to $|x^1|\leqslant C\varepsilon$. Note that C_0 is bounded away from 0, hence C is bounded away from 0 for $|b-a|$ bounded (we do not need C to be bounded from above). Note also that θ' will depend on x^1 but θ has been fixed before.

The second term T_2 is controlled using Lemma 5.4: we have

$$|T_2|\leqslant C_1|\theta'-\theta||x|^2.$$

For the third term T_3, we note, from the definition of δ and Δ in Lemma 5.3, that

$$|\Delta-\Delta'|\leqslant|b-a|^2\frac{\Delta_0^2}{C_0}|\theta'-\theta| \quad\text{and}\quad |\delta-\delta'|\leqslant|C_0||\theta'-\theta|.$$

Hence, using (31),

$$|T_2+T_3|\leqslant C_2|x|^3,$$

where C_2 depends on the bounds in assumptions (A0), (A2) and (As), and on $|b-a|$. We conclude that, for a suitable choice of θ',

$$|T_1+T_2+T_3|\leqslant C_2|x|^3. \tag{33}$$

We now have, for all $\theta\in[\varepsilon,1-\varepsilon]$ and all $x\in\Omega$ with $|x|<C\varepsilon$,

$$\bar\phi(x)\geqslant(\theta b+(1-\theta)a)x_1-\tfrac{1}{2}D^2_{xx}c(0,y_\theta)\cdot(x,x)+\delta|x|^2-C_2|x|^3+o(|x|^2).$$

Using (28), this leads to

$$\bar\phi(x)\geqslant -c(x,y_\theta)+c(0,y_\theta)+\delta|x|^2-C_2|x|^3+o(|x|^2).$$

We now notice that all the terms in $o(|x|^2)$ are error terms in the second order Taylor expansion (28). Under assumption (A0), c belongs to $C^3(\bar\Omega\times\bar\Omega')$, hence there exists γ such that the above inequality still holds true when replacing $-C_2|x|^3+o(|x|^2)$ by $-\gamma|x|^3$. The constant γ will depend on the bounds in (A0), (A2) and (As), and on $|b-a|$. From Lemma 5.3, we have $\delta=\frac{1}{4}C_0\theta(1-\theta)|b-a|^2$. Using now (13), we have

$$\frac{1}{C_{\mathfrak{T}}}|y_1-y_0|\leqslant|b-a|,$$

and letting $\delta_0=C_0/4C_{\mathfrak{T}}^2$, we conclude the proof of Proposition 5.1. □

5.3. Part II. Construction of supporting functions

We let $\mathcal{N}_\mu(B)$ denote the μ-neighborhood of a set B, and we use Proposition 5.1 to prove the following result.

PROPOSITION 5.6. *Let ϕ be c-convex. Let c, Ω and Ω' satisfy the assumptions of Theorem* 3.4. *Let $x_0, x_1 \in \Omega$, $y_0 \in G_\phi(x_0)$ and $y_1 \in G_\phi(x_1)$. Then, there exist constants $C, C', C'' > 0$ and $x_m \in [x_0, x_1]$, such that, if $\mathcal{N}_\eta([x_0, x_1]) \subset \Omega$ and*

$$|y_1 - y_0| \geqslant \max\{|x_1 - x_0|, C|x_1 - x_0|^{1/5}\} > 0, \tag{34}$$

then

$$\mathcal{N}_\mu\big(\big\{y_\theta, \theta \in \big[\tfrac{1}{4}, \tfrac{3}{4}\big]\big\}\big) \cap \Omega' \subset G_\phi(B_\eta(x_m)),$$

where

$$\eta = C' \left(\frac{|x_1 - x_0|}{|y_1 - y_0|} \right)^{1/2}, \tag{35}$$

$$\mu = C'' \eta |y_1 - y_0|^2. \tag{36}$$

Here $\{y_\theta\}_{\theta \in [0,1]} = [y_0, y_1]_{x_m}$ denotes the c-segment from y_0 to y_1 with respect to x_m. Under assumptions (A0)–(A2) *and* (As), *the constants C and C' are bounded away from infinity and C'' is bounded away from* 0.

Remark. If x_0, x_1, y_0 and y_1 satisfying (34) cannot be found, then ϕ is Hölder continuous with exponent $\frac{1}{5}$.

Without loss of generality, we will assume that $\phi(x_0) = \phi(x_1)$: indeed, as remarked in Proposition 5.2, by subtracting from the cost function c an affine function λ that depends only on x, we will not modify the map solution of the optimal transportation problem, and the optimal potential ϕ will be changed into $\phi + \lambda$. Hence one can subtract a suitable affine function from c so that $\phi(x_0) = \phi(x_1)$. Notice that, as λ is chosen affine, the gradient of the "new" potentials are deduced from the "old" ones just by adding the constant vector $\nabla_x \lambda$. Hence this does not change the continuity properties of $\nabla \phi$, neither does it change the derivatives of c of order greater than 1.

As $y_0 \in G_\phi(x_0)$ and $y_1 \in G_\phi(x_1)$ we have, using (11), for all $x \in \Omega$,

$$-c(x, y_0) + c(x_0, y_0) + \phi(x_0) \leqslant \phi(x),$$
$$-c(x, y_1) + c(x_1, y_1) + \phi(x_1) \leqslant \phi(x),$$

with equality at $x = x_0$ in the first line, at $x = x_1$ in the second line. Since $\phi(x_0) = \phi(x_1)$, the difference between the supporting functions

$$x \longmapsto -c(x, y_0) + c(x_0, y_0) + \phi(x_0) \quad \text{and} \quad x \longmapsto -c(x, y_1) + c(x_1, y_1) + \phi(x_1)$$

will vanish at some point x_m in the segment $[x_0,x_1]$. Without loss of generality, we can add a constant to ϕ so that at this point both supporting functions are equal to 0. Hence

$$-c(x_m,y_0)+c(x_0,y_0)+\phi(x_0)=0, \tag{37}$$

$$-c(x_m,y_1)+c(x_1,y_1)+\phi(x_1)=0. \tag{38}$$

LEMMA 5.7. *Under the assumptions made above, and assuming moreover that*

$$|y_1-y_0|\geqslant|x_1-x_0|,$$

we have, for all x in the segment $[x_0,x_1]$,

$$\phi(x)\leqslant C_3|x_1-x_0||y_1-y_0|,$$

where C_3 depends only on $\|c(\,\cdot\,,\cdot)\|_{C^2(\Omega\times\Omega')}$.

Proof. Using (37) and (38), we have

$$H=\phi(x_0)\leqslant-\nabla_x c(x_m,y_0)\cdot(x_0-x_m)+\tfrac{1}{2}\|c\|_{C^2}|x_0-x_m|^2, \tag{39}$$

$$H=\phi(x_1)\leqslant-\nabla_x c(x_m,y_1)\cdot(x_1-x_m)+\tfrac{1}{2}\|c\|_{C^2}|x_1-x_m|^2. \tag{40}$$

By Proposition 2.2, the potential ϕ is semi-convex, with $D^2\phi\geqslant-\|D^2_{xx}c\|_{L^\infty(\Omega\times\Omega')}I$. Applying the first part of Lemma 5.5 to the function $f\colon t\mapsto\phi(x_0+t(x_1-x_0))$ on $[0,1]$, for which $f''\geqslant-D^2\phi\cdot(x_1-x_0,x_1-x_0)$, we find that

$$\phi(x)\leqslant H+C|x_1-x_0|^2\quad\text{for all } x\in[x_0,x_1], \tag{41}$$

where $C=C(\|c\|_{C^2(\Omega\times\Omega')})$. Then we consider two cases.

The first case is when $-\nabla_x c(x_m,y_0)\cdot(x_0-x_m)$ and $-\nabla_x c(x_m,y_1)\cdot(x_1-x_m)$ are not both positive: let us assume for example that $-\nabla_x c(x_m,y_0)\cdot(x_0-x_m)$ is negative. Then we have, using (39), $H\leqslant\frac{1}{2}\|c\|_{C^2}|x_0-x_m|^2$, and using (41), we get that

$$\phi(x)\leqslant\big(C+\tfrac{1}{2}\|c\|_{C^2(\Omega\times\Omega')}\big)|x_1-x_0|^2\quad\text{for all } x\in[x_0,x_1].$$

Thus we can conclude using $|x_1-x_0|\leqslant|y_1-y_0|$.

The second case is when $-\nabla_x c(x_m,y_0)\cdot(x_0-x_m)$ and $-\nabla_x c(x_m,y_1)\cdot(x_1-x_m)$ are both positive. This implies that

$$-\nabla_x c(x_m,y_0)\cdot(x_0-x_m)\leqslant-\nabla_x c(x_m,y_0)\cdot(x_0-x_1),$$
$$-\nabla_x c(x_m,y_1)\cdot(x_1-x_m)\leqslant-\nabla_x c(x_m,y_1)\cdot(x_1-x_0).$$

Combining with (39) and (40), we have

$$\begin{aligned}2H &\leqslant -\nabla_x c(x_m,y_0)\cdot(x_0-x_1)-\nabla_x c(x_m,y_1)\cdot(x_1-x_0)+\|c\|_{C^2}|x_0-x_1|^2\\ &\leqslant |\nabla_x c(x_m,y_0)-\nabla_x c(x_m,y_1)||x_0-x_1|+\|c\|_{C^2}|x_0-x_1|^2\\ &\leqslant \|c\|_{C^2}(|x_1-x_0||y_1-y_0|+|x_0-x_1|^2).\end{aligned}$$

Using $|x_1-x_0|\leqslant|y_1-y_0|$, and then (41), we conclude. □

We now assume the following: letting Γ be defined by

$$\Gamma=\left(\frac{\gamma^2}{\delta_0^3}2^{12}C_3\right)^{1/5}, \tag{42}$$

with $C_3(\|c\|_{C^2(\Omega\times\Omega')})$ defined in Lemma 5.7, x_0, x_1, y_0 and y_1 satisfy

$$|y_1-y_0|\geqslant\max\{\Gamma|x_1-x_0|^{1/5},|x_1-x_0|\}. \tag{43}$$

Hence, the constant C in Proposition 5.6 will be equal to Γ.

We next state the following result, from which the proof of Proposition 5.6 will follow easily.

Lemma 5.8. *Let x_m be defined as above. For $y\in\Omega'$, consider the function*

$$f_y(x)=-c(x,y)+c(x_m,y)+\phi(x_m).$$

Under the assumptions made above, there exist η and μ as in Proposition 5.6, *such that for all $y\in\mathcal{N}_\mu\big(\{y_\theta,\theta\in\big[\frac{1}{4},\frac{3}{4}\big]\}\big)\cap\Omega'$,*

$$\phi-f_y\geqslant 0\quad \text{on } \partial B_\eta(x_m)\cap\Omega. \tag{44}$$

Before proving this lemma, we first show how it leads to Proposition 5.6.

Proof of Proposition 5.6. By construction, $f_y(x_m)=\phi(x_m)$. Hence, if we have $\phi\geqslant f_y$ on $\partial B_\eta(x_m)$, then $\phi-f_y$ will have a local minimum inside $B_\eta(x_m)$, and for some point $x\in B_\eta(x_m)$, we will have $-\nabla_x c(x,y)\in\partial\phi(x)$. Using Theorem 3.1, we have $\partial\phi(x)=\partial^c\phi(x)$, and this implies that $y\in G_\phi(x)\subset G_\phi(B_\eta(x_m))$. □

We now prove the main lemma.

Proof of Lemma 5.8. Using (37) and (38), and then Proposition 5.1 centered at x_m, we obtain

$$\begin{aligned}\phi(x)&\geqslant\max\{-c(x,y_0)+c(x_m,y_0),-c(x,y_1)+c(x_m,y_1)\}\\ &\geqslant -c(x,y_\theta)+c(x_m,y_\theta)+\delta_0\theta(1-\theta)|y_0-y_1|^2|x-x_m|^2-\gamma|x-x_m|^3=\Phi(x)\end{aligned} \tag{45}$$

for $\varepsilon>0$, for all $\theta\in[\varepsilon,1-\varepsilon]$, $|x-x_m|\leqslant C\varepsilon$, and with $\{y_\theta\}_{\theta\in[0,1]}$ being the c-segment with respect to x_m joining y_0 to y_1. Then, for $y\in\Omega'$, we have

$$\begin{aligned}-c(x,y)+c(x_m,y)&=-c(x,y_\theta)+c(x_m,y_\theta)\\&\quad+\int_0^1[\nabla_y c(x_m,y_\theta+s(y-y_\theta))-\nabla_y c(x,y_\theta+s(y-y_\theta))]\cdot(y-y_\theta)\,ds\\&\leqslant -c(x,y_\theta)+c(x_m,y_\theta)+C_4|y-y_\theta||x-x_m|,\end{aligned}$$

where $C_4=\|D^2_{xy}c\|_{L^\infty(\Omega\times\Omega')}$. Combining this with Lemma 5.7 to estimate $\phi(x_m)$, we get

$$\begin{aligned}f_y(x)&=-c(x,y)+c(x_m,y)+\phi(x_m)\\&\leqslant -c(x,y_\theta)+c(x_m,y_\theta)+C_4|y-y_\theta||x-x_m|+C_3|x_1-x_0||y_1-y_0|=F_y(x).\end{aligned}\tag{46}$$

Inequality (44) will be satisfied if we have, for F_y and Φ defined in (45) and (46),

$$F_y(x)\leqslant\Phi(x)\tag{47}$$

on the set $\{x:|x-x_m|=\eta\}$, for some $\eta>0$. First, we restrict θ to $\left[\frac14,\frac34\right]$, (i.e. we take $\varepsilon=\frac14$ in (45)). Then (47) reads

$$\tfrac{3}{16}\delta_0|y_0-y_1|^2\eta^2-\gamma\eta^3\geqslant C_4|y-y_\theta|\eta+C_3|x_1-x_0||y_1-y_0|.\tag{48}$$

Inequality (48) will be satisfied if the three following inequalities are satisfied:

$$\begin{aligned}&\tfrac{1}{16}\delta_0|y_0-y_1|^2\eta^2\geqslant C_3|x_1-x_0||y_1-y_0|,\\&\tfrac{1}{16}\delta_0|y_0-y_1|^2\eta^2\geqslant C_4|y-y_\theta|\eta,\\&\tfrac{1}{16}\delta_0|y_0-y_1|^2\eta^2\geqslant\gamma\eta^3.\end{aligned}$$

In order to satisfy the first inequality, we define η by

$$\eta^2=\frac{16C_3}{\delta_0}\frac{|x_1-x_0|}{|y_1-y_0|}.$$

In order to satisfy the second inequality, we define μ by

$$\mu=C_5\eta|y_1-y_0|^2,$$

where $C_5=\delta_0/16C_4$ (note that C_5 is bounded away from 0), and consider $y\in\Omega'$ such that $|y-y_\theta|\leqslant\mu$. The third inequality will then be implied by

$$\gamma\eta\leqslant\tfrac{1}{16}\delta_0|y_0-y_1|^2,$$

which is equivalent to

$$\frac{\gamma^2}{\delta_0^3}16^3C_3|x_1-x_0|\leqslant|y_1-y_0|^5,$$

and we recognize here assumption (43). The constants C, C' and C'' in Proposition 5.6 are defined by $C=\Gamma$ from assumption (43), $C'=(16C_3/\delta_0)^{1/2}$ and $C''=C_5$. Then, for all $y\in\mathcal{N}_\mu\left\{y_\theta,\theta\in\left[\frac14,\frac34\right]\right\}\cap\Omega'$, the function $f_y(x)=-c(x,y)+c(x_m,y)+\phi(x_m)$ will satisfy $f_y\leqslant\phi$ on the boundary of the ball $B_\eta(x_m)$. This proves Lemma 5.8. □

5.4. Part III. Continuity estimates

PROPOSITION 5.9. *Let ϕ be c-convex with $G_\phi(\Omega)\subset\Omega'$. Let c, Ω and Ω' satisfy the assumptions of Theorem* 3.4, *and let $(G_\phi)^{\#}$ be as in Definition* 2.12.

- *If $(G_\phi)^{\#}$*dVol *satisfies* (23) *for some $p>n$, then $\phi\in C^{1,\beta}_{\mathrm{loc}}(\Omega)$, with $\beta(n,p)$ as in Theorem* 3.4.
- *If $(G_\phi)^{\#}$*dVol *satisfies* (24), *then $\phi\in C^1_{\mathrm{loc}}(\Omega)$.*

We first state the following general result, whose proof is deferred to the appendix.

LEMMA 5.10. *Let Ω' be c-convex with respect to $x_m\in\Omega$ and let $y_0,y_1\in\Omega'$. There exist $C,\mu_0>0$ depending on c, Ω and Ω' such that, for all $\mu\in(0,\mu_0)$,*

$$\mathrm{Vol}(\mathcal{N}_\mu([y_0,y_1]_{x_m})\cap\Omega')\geqslant C\,\mathrm{Vol}(\mathcal{N}_\mu([y_0,y_1]_{x_m})).$$

Proof of Proposition 5.9. Consider

$$\Omega_\delta=\{x\in\Omega:d(x,\partial\Omega)>\delta\}.$$

In order to have $\mathcal{N}_\eta([x_0,x_1])\subset\Omega$, it is enough to have

(1) $x_0,x_1\in\Omega_\delta$,

(2) $|x_0-x_1|<\frac{1}{2}\delta$,

(3) $\eta<\frac{1}{2}\delta$.

If $y_0\in G_\phi(x_0)$ and $y_1\in G_\phi(x_1)$ satisfy (34) in Proposition 5.6, with $|y_1-y_0|\geqslant C|x_1-x_0|^{1/5}$, then $\eta\leqslant E|x_1-x_0|^{2/5}$, with η defined in Proposition 5.6, and E being a constant depending only on C' and C'' in Proposition 5.6. Hence, for $|x_1-x_0|^{2/5}\leqslant\delta/2E$, it follows that $\mathcal{N}_\eta([x_0,x_1])\subset\Omega$, and Proposition 5.6 applies. We now set

$$R_\delta=\inf\left\{\frac{\delta}{2},\left(\frac{\delta}{2E}\right)^{5/2}\right\},\tag{49}$$

and in the remainder of the proof, we chose $x_1,x_0\in\Omega_\delta$ such that $|x_1-x_0|\leqslant R_\delta$. From Proposition 5.6, we will have

$$N_\mu\left\{y_\theta,\theta\in\left[\tfrac{1}{4},\tfrac{3}{4}\right]\right\}\cap\Omega'\subset G_\phi(B_\eta(x_m)).\tag{50}$$

From Lemma 5.10, and the definition of μ in (36), there exist $C,C'>0$ such that

$$\mathrm{Vol}\big(N_\mu\left\{y_\theta,\theta\in\left[\tfrac{1}{4},\tfrac{3}{4}\right]\right\}\cap\Omega'\big)\geqslant C|y_1-y_0|\mu^{n-1}=C'|y_1-y_0|\eta^{n-1}|y_1-y_0|^{2(n-1)}.\tag{51}$$

$C^{1,\beta}$ estimates for data with bounded density. If the Jacobian determinant of the mapping G_ϕ is bounded (in other words, if $(G_\phi)^{\#}$dVol has a density bounded in L^∞ with respect to the Lebesgue measure), then, for some C and C',

$$\mathrm{Vol}(G_\phi(B_\eta(x_m)))\leqslant C\,\mathrm{Vol}(B_\eta(x_m))=C'\eta^n.\tag{52}$$

Using (50) with (51) and (52), we find for some C and C' that

$$|y_1-y_0|^{2n-1} \leqslant C\eta = C'\left(\frac{|x_1-x_0|}{|y_1-y_0|}\right)^{1/2},$$

which yields finally, for another constant $C_6>0$,

$$|y_1-y_0| \leqslant C_6|x_1-x_0|^{1/(4n-1)}.$$

From this, we readily deduce that G_ϕ is single valued, moreover $G_\phi \in C^{1/(4n-1)}_{\mathrm{loc}}(\Omega)$. Since $-\nabla_x c(x,y_i)=\nabla\phi(x_i)$, $i=0,1$, and $\nabla_x c$ is Lipschitz, this also yields $\phi\in C^{1,1/(4n-1)}_{\mathrm{loc}}(\Omega)$.

$C^{1,\beta}$ estimates for data satisfying (23). We can refine the argument: Let again $\nu=(G_\phi)^\#\,\mathrm{dVol}$ and F be defined by

$$\begin{aligned} F(V) &= \sup\{\mathrm{Vol}(G_\phi(B)) : B\subset\Omega \text{ is a ball of volume } V\} \\ &= \sup\{\nu(B) : B\subset\Omega \text{ is a ball of volume } V\}. \end{aligned} \tag{53}$$

Then, by Proposition 5.6, we have $F(\mathrm{Vol}(B_\eta(x_m)))\geqslant \mathrm{Vol}\big(N_\mu\{y_\theta, \theta\in[\frac14,\frac34]\}\cap\Omega'\big)$, which yields, using (51) and the definition of η in (35),

$$F\left(\omega_n C_5^n \frac{|x_1-x_0|^{n/2}}{|y_1-y_0|^{n/2}}\right) \geqslant C_7|x_1-x_0|^{(n-1)/2}|y_1-y_0|^{(3n-1)/2} \tag{54}$$

for some C_7 bounded away from 0, with ω_n being the volume of the n-dimensional unit ball. Assume that $F(V)\leqslant CV^{\varkappa}$ for some $\varkappa\in\mathbb{R}$. Note that $\nu\in L^p$ implies the (stronger) bound $F(V)=o(V^{1-1/p})$, hence it is natural to write $\varkappa=1-1/p$ for some $p\in]1,+\infty]$, and the condition

$$F(V)\leqslant CV^{1-1/p} \tag{55}$$

is then equivalent to condition (23) for ν. We obtain from (54) and (55) that

$$|y_1-y_0|^{2n-1+(1-n/p)/2} \leqslant C_8|x_1-x_0|^{(1-n/p)/2}.$$

We see first that we need $p>n$, and, setting $\alpha=1-n/p$, we obtain

$$|y_1-y_0| \leqslant C_9|x_1-x_0|^{\alpha/(4n-2+\alpha)}.$$

This yields Hölder continuity for G_ϕ. Then we use that $\nabla\phi(x)=-\nabla_x c(x,G_\phi(x))$ and the smoothness of c to obtain a similar Hölder estimate for $\nabla\phi$.

C^1 estimates for data satisfying (24). We only assume condition (24) for

$$\nu=(G_\phi)^\#\,\mathrm{dVol},$$

which we can rewrite under the following form:

$$F(V) \leqslant [f(V^{2/n})]^{2n-1} V^{1-1/n}, \tag{56}$$

for some increasing $f:[0,1]\to\mathbb{R}^+$, with $\lim_{V\to 0} f(V)=0$, F being defined in (53). Consistently with (43), we can assume that, as x_1 goes to x_0, $|x_1-x_0|/|y_1-y_0|$ also goes to 0. Using (56) in (54), we get, for some C_{10} and C_{11} bounded away from 0 and infinity,

$$f^{2n-1}\left(C_{10}\frac{|x_1-x_0|}{|y_1-y_0|}\right) \geqslant (C_{11}|y_1-y_0|)^{2n-1}.$$

Hence we get that $|y_1-y_0|$ goes to 0 when $|x_1-x_0|$ goes to 0. Then, the modulus of continuity g of G_ϕ in Ω_δ satisfies the following:

For all $u\leqslant R_\delta$, either $g(u)\leqslant \max\{u, \Gamma u^{1/5}\}$ or

$$f\left(C_{10}\frac{u}{g(u)}\right) \geqslant C_{11} g(u),$$

which is equivalent to

$$u \geqslant f^{-1}(C_{11} g(u))\frac{g(u)}{C_{10}},$$

which in turn is equivalent to

$$g(u) \leqslant \omega(u),$$

where ω is the inverse of $z\mapsto f^{-1}(C_{11}z)z/C_{10}$. It is easily checked that $\lim_{r\to 0^+}\omega(r)=0$. This shows the continuity of G_ϕ. Finally, we have $\nabla\phi(x)=-\nabla_x c(x, G_\phi(x))$, and the continuity of $\nabla\phi$ is asserted. □

Remark. The power $\beta=\alpha/(4n-2+\alpha)$ is not optimal for example if $n=1$ and $p=+\infty$, for which the $C^{1,1}$ regularity is trivial, but note that in order to obtain this bound, we had to assume (43). Hence, the conclusion should be: either ϕ is $C^{1,1}$, or ϕ is $C^{1,1/5}$, or ϕ is $C^{1,\beta}$. Note that $\beta\leqslant\frac{1}{7}$ for $n\geqslant 2$.

In Proposition 5.9, we use a bound on $(G_\phi)^\#\mathrm{dVol}$. However, in Theorem 3.4, we only have $(G_\phi)_\#\mu_1=\mu_0$, and as we we do not want to assume that $\mu_1\in L^1(\mathbb{R}^n)$, this does not imply necessarily that $(G_\phi)^\#\mu_1=\mu_0$ (see Definition 2.12 and the subsequent discussion). Hence we need the following proposition to finish the proof of Theorem 3.4.

Proposition 5.11. *Let ϕ be c-convex on Ω, with $G_\phi(\Omega)\subset\Omega'$. Assume that*

$$(G_\phi)_\#\mu_0 = \mu_1.$$

Assume that $\mu_1\geqslant m\mathrm{dVol}$ on Ω'. Then, for all $\omega\subset\Omega$, we have

$$\mu_0(\omega) \geqslant m\,\mathrm{Vol}(G_\phi(\omega)),$$

and hence

$$(G_\phi)^\#\mathrm{dVol} \leqslant \frac{1}{m}\mu_0.$$

Proof. In Ω' we consider

$$N=\{y\in\Omega' : \text{there exist } x_1\neq x_2\in\Omega \text{ such that } G_\phi(x_1)=G_\phi(x_2)=y\}.$$

Then

$$N=\{y\in\Omega' : \phi^c \text{ is not differentiable at } y\}.$$

Hence, $\mathrm{Vol}(N)=0$ and $\mathrm{Vol}(G_\phi(\omega)\setminus N)=\mathrm{Vol}(G_\phi(\omega))$. Moreover, on $G_\phi(\omega)\setminus N$, G_ϕ^{-1} is single valued. Then $G_\phi^{-1}(G_\phi(\omega)\setminus N)\subset\omega$. Hence,

$$\mu_0(\omega)\geqslant\mu_0(G_\phi^{-1}(G_\phi(\omega)\setminus N))=\mu_1(G_\phi(\omega)\setminus N)\geqslant m\,\mathrm{Vol}(G_\phi(\omega)\setminus N)=m\,\mathrm{Vol}(G_\phi(\omega)).\quad\square$$

Proof of the boundary regularity. This part is easy: under the assumptions of Theorem 3.5, the density μ_0 satisfies (23) with $p>n$ (resp. satisfies (24)). Hence, Theorem 3.4 applies and $\phi\in C^{1,\beta}_{\mathrm{loc}}(\Omega)$ (resp. $\phi\in C^1_{\mathrm{loc}}(\Omega)$). Since Ω_2 is compactly contained in Ω, we conclude the boundary regularity on Ω_2. This proves Theorem 3.5.

Remark. This proof of the boundary regularity is very simple because we have interior regularity even when μ_0 vanishes. This is not the case for the classical Monge–Ampère equation, where the boundary regularity requires that both Ω and Ω' are convex, and is more complicated to establish (see [6]).

We now show that there is indeed equivalence between assumption (As) at a point x and the conclusion of Proposition 5.1. This is a quantitative version of Theorem 3.2.

PROPOSITION 5.12. *Assume that at a point x_0, for all y_0 and y_1, we have*

$$\bar\phi(x)\geqslant -c(x,y_{1/2})+c(x_0,y_{1/2})+\delta_0|y_0-y_1|^2|x-x_0|^2+O(|x-x_0|^3)$$

with $\bar\phi$ as above, where $y_{1/2}$ is the 'middle' point of $[y_0,y_1]_{x_0}$. Then, the cost function satisfies assumption (As) *at x_0 with $C_0=C\delta_0$, for some constant $C>0$ that depends on the bound in* (A2).

Proof. The proof follows the same lines as the proof of Theorem 3.1, and is omitted here. $\square$

6. Proof of Theorem 3.8

We consider condition (Aw) at $(x_0, y=x_0)$. We recall that

$$\mathfrak{S}_{\mathfrak{c}}(x_0,x_0)(\xi,\nu)=-D^2_{p_\nu p_\nu}D^2_{x_\xi x_\xi}[(x,p)\mapsto c(x,\mathfrak{T}_{x_0}(p))]$$

for any ν and ξ in $T_{x_0}M$. Let us first take a normal system of coordinates at x_0, so that we will compute

$$Q=-D_{tt}^2 D_{ss}^2[(x,p)\mapsto c(\mathfrak{T}_{x_0}(t\xi),\mathfrak{T}_{x_0}(s\nu))].$$

Let us write a finite difference version of this operator. We first introduce

$$y_-=\mathfrak{T}_{x_0}(-h\nu),\quad y_+=\mathfrak{T}_{x_0}(h\nu),\quad x_-=\mathfrak{T}_{x_0}(-h\xi)\quad\text{and}\quad x_+=\mathfrak{T}_{x_0}(h\xi).$$

We use the usual second-order difference quotient, for example

$$D_{x_\xi,x_\xi}^2 c(x,\mathfrak{T}_{x_0}(p))=\lim_{h\to 0}\frac{c(x_+,x_0)-2c(x_0,x_0)+c(x_-,x_0)}{h^2}.$$

(Of course we have $c(x_0,x_0)=0$.) We will have, as h goes to 0,

$$\lim_{h\to 0}\frac{1}{h^4}\biggl(\sum_{i,j=+,-}c(x_i,y_j)-2\sum_{i=+,-}(c(x_i,x_0)+c(y_i,x_0))\biggr)=-Q.$$

Rearranging the terms, we find that the bracket in the left-hand side is equal to

$$\sum_{i,j=+,-}[c(x_i,y_j)-c(x_i,x_0)-c(y_j,x_0)].$$

Each of the terms inside the bracket has a simple geometric interpretation: consider the triangle with vertices (x_0,x_i,y_j) whose sides are geodesics. This is a right-angled triangle. If the metric is flat, by Pythagoras' theorem, the term inside the brackets is 0. In the general case, a standard computation shows that it is equal to $-\frac{1}{6}\varkappa(x_0,\xi,\nu)h^4+o(h^4)$, where $\varkappa(x_0,\xi,\nu)$ is the sectional curvature at x_0 in the 2-plane generated by ξ and ν. Hence, we get that $Q=\frac{2}{3}\varkappa(x,\xi,\nu)$.

Now, to reach the more general formula of Theorem 3.8, we use the following expansion of the distance that Cédric Villani communicated to us.

LEMMA 6.1. *Let M be a smooth Riemannian manifold. Let γ_1 and γ_2 be two unit speed geodesics that leave a point $x_0\in M$. Let θ be the angle between $\dot\gamma_1(0)$ and $\dot\gamma_2(0)$* (*measured with respect to the metric*), *let $\varkappa$ be the sectional curvature of M at x_0 in the* 2-*plane generated by $\dot\gamma_1(0)$ and $\dot\gamma_2(0)$. Then we have*

$$d^2(\gamma_1(t),\gamma_2(t))=2(1-\cos(\theta))\bigl(\bigl(1-\tfrac{1}{6}\varkappa\bigl(\cos^2\bigl(\tfrac{1}{2}\theta\bigr)\bigr)t^2+O(t^4)\bigr)t^2\bigr)^2.$$

Then, we obtain easily, following the same lines as in the case above, that

$$\mathfrak{S}_{\mathfrak{c}}(x_0,x_0)(\xi,\nu)=\tfrac{2}{3}\varkappa(x_0,\xi,\nu)(|\xi|_g^2|\nu|_g^2-(\xi,\nu)_g^2),$$

where $(\cdot,\cdot)_g$ and $|\cdot|_g$ denote the scalar product and the norm with respect to g, respectively. This proves Theorem 3.8.

6.1. Counterexample to regularity for a manifold with negative curvature

Consider the 2-dimensional surface $H=\{(x,y,z):z=x^2-y^2\}\subset\mathbb{R}^3$, endowed with the Riemannian metric inherited from the canonical metric of $\mathbb{R}^3$. Then H has negative sectional curvature around 0. For r sufficiently small, $\Omega=H\cap B_r(0)$ is c-convex with respect to itself. Consider the function

$$\bar\phi(x)=\max\bigl\{-\tfrac{1}{2}d^2(X,X_0),-\tfrac{1}{2}d^2(X,X_1)\bigr\},$$

where $X_0=(0,a,-a^2)$ and $X_1=(0,-a,-a^2)$. Then, as shown by our proof of Theorem 3.1, for a small enough, no sequence of C^1 c-convex potentials can converge uniformly to $\bar\phi$ on Ω. Let μ_0 be the Lebesgue measure of Ω, and $\mu_1=\frac{1}{2}(\delta_{X_0}+\delta_{X_1})$. We have $(G_{\bar\phi})_{\#}\mu_0=\mu_1$. Let $\mu_1^\varepsilon\in C^\infty(\bar\Omega)$ be a positive mollification of μ_1 such that its total mass remains equal to 1, and which preserves the symmetries with respect to $x=0$ and $y=0$. Let ϕ_n be such that $(G_{\phi_n})_{\#}\mu_0=\mu_n$. Then, for n large enough, ϕ_n is not differentiable at the origin. Indeed, for symmetry reasons, 0 belongs to the subdifferential of ϕ_n at 0; on the other hand, ϕ_n converges uniformly to $\bar\phi$, and we know, from the fact that (Aw) is violated at 0, that $-\bar\phi-c(\cdot,0)$ does not reach its global maximum on Ω at 0.

7. Appendix

Proof of Proposition 2.14. We first prove the "intrinsic" part. In order to show this, we consider a C^2 curve γ in Ω defined in a neighborhood of 0, such that

$$\gamma(0)=x_0, \tag{57}$$

$$\dot\gamma(0)=\xi. \tag{58}$$

We then consider the quantity

$$Q_\gamma=D^2_{tt}D^2_{ss}[(s,t)\mapsto c(\gamma(t),\mathfrak{T}_{x_0}(p_0+s\nu))]|_{t,s=0}.$$

We show that this quantity is independent of the choice of γ. We have

$$\begin{aligned}Q_\gamma&=D^2_{ss}[s\mapsto D^2_{\xi\xi}c(x_0,\mathfrak{T}_{x_0}(p_0+s\nu))+D_xc(x_0,\mathfrak{T}_{x_0}(p_0+s\nu))\cdot\ddot\gamma(0)]\\&=D^2_{ss}[s\mapsto D^2_{\xi\xi}c(x_0,\mathfrak{T}_{x_0}(p_0+s\nu))-(p_0+s\nu)\cdot\ddot\gamma(0)]\\&=D^2_{ss}[s\mapsto D^2_{\xi\xi}c(x_0,\mathfrak{T}_{x_0}(p_0+s\nu))],\end{aligned}$$

where the second equality follows from the very definition of the c-exponential map. Hence, the value of the curvature is independent of $\ddot\gamma(0)$, and therefore of the choice of γ

as long as it satisfies (57) and (58). One can now choose a system of geodesic coordinates around x_0, which yields the equivalence of the definitions (19) and (20). Then, the second part of Proposition 2.14 follows by taking the c-geodesics with respect to y_0 as new coordinates around x_0, which yields

$$\mathfrak{S}_{\mathfrak{c}}(x_0,y_0)(\xi,\nu)=D^2_{p_\nu p_\nu q_{\tilde\xi} q_{\tilde\xi}}[c(\mathfrak{T}_{y_0}(q),\mathfrak{T}_{x_0}(p))]|_{q_0=-\nabla_y c(x_0,y_0),p_0=-\nabla_x c(x_0,y_0)}, \qquad (59)$$

where $\tilde\xi$ is chosen such that

$$D_q\mathfrak{T}^*_{y_0}(q_0)\cdot\tilde\xi=\xi.$$

The condition $\xi\perp\nu$ nows reads $D_q\mathfrak{T}^*_{y_0}(q_0)\cdot\tilde\xi\perp\nu$, or equivalently $[D_{x,y}c]^{-1}\cdot(\nu,\tilde\xi)=0$. Then, identity (21) follows by a symmetric argument. □

Proof of Proposition 3.3. We prove only the last point, the other points being elementary. Consider on $\mathbb{R}^n$ a measure locally equal to $\mu_0=\mathcal{L}^{n-1}\otimes\mu$, where $\mathcal{L}^{n-1}$ is the $(n-1)$-dimensional Lebesgue measure, and μ is a probability measure on $[0,1]$ equal to the derivative of the Devil's staircase. Then, $\mu\notin L^1$. On the other hand, for all $[a,b]\subset[0,1]$, $\mu([a,b])\leqslant|b-a|^\alpha$ for some $\alpha\in(0,1]$. Then, for $x=(x_1,...,x_n)$,

$$\mu_0(B_r(x))\leqslant Cr^{n-1}\mu([x_n-r,x_n+r])\leqslant Cr^{n-1+\alpha}=Cr^{n(1-1/p)}$$

for some $p>n$. Hence $\mu_0\notin L^1_{\mathrm{loc}}$ and μ_0 satisfies (23) for some $p>n$. □

Proof of Proposition 4.1. We know (see [33, Chapter 2]) that there exists a probability measure π on $\mathbb{R}^n\times\mathbb{R}^n$, with marginals μ and ν, such that

$$\int_{\mathbb{R}^n}\phi(x)\,d\mu(x)+\psi(x)\,d\nu(x)=-\int c(x,y)\,d\pi(x,y),$$

and moreover, there exists a c-convex potential $\bar\phi$ such that

$$\operatorname{supp}(\pi)\subset\{(x,G_{\bar\phi}(x)):x\in\mathbb{R}^n\}.$$

Let us decompose π as $\pi=\mu\otimes\gamma_x$, where for μ-a.e. $x\in\mathbb{R}^n$, γ_x is a probability measure on $\mathbb{R}^n$ and γ_x is supported in $G_{\bar\phi}(x)$. Hence, we have

$$\int_{\mathbb{R}^n}\int_{\mathbb{R}^n}(\phi(x)+\psi(y)-c(x,y))\,d\gamma_x(y)\,d\mu(x)=0.$$

This implies that $y\in G_\phi(x)$ for μ-a.e. x and γ_x-a.e. y. Since $y\in G_{\bar\phi}(x)$ γ_x-a.e. for μ-a.e. x, we deduce that

$$G_{\bar\phi}(x)\cap G_\phi(x)\neq\varnothing$$

for μ-a.e. x (and hence for Lebesgue-a.e. x, since $\mu>0$ a.e.). This implies that $\nabla\phi=\nabla\bar\phi$ Lebesgue-a.e., and that $\phi-\bar\phi$ is constant. This shows that ϕ is uniquely defined up to a constant. Now, the pair (ψ^{c*},ψ) can only improve the infimum (10) compared to (ϕ,ψ), hence it is also optimal. Hence, ψ^{c*} is also uniquely defined up to a constant. If ψ is c*-convex, then $(\psi^{c*})^c=\psi$, and ψ is thus uniquely defined. □

Proof of Lemma 5.10. From (A1) and (A2), for all $x_m \in \Omega$, $\psi: y \mapsto -\nabla_x c(x_m, y)$ is a diffeomorphism from Ω' to $-\nabla_x c(x_m, \Omega')$. Then,

$$\psi(\mathcal{N}_\eta([y_0, y_1]_{x_m}) \cap \Omega') = \psi(\mathcal{N}_\eta([y_0, y_1]_{x_m})) \cap \psi(\Omega').$$

Letting $p_i = -\nabla_x c(x_m, y_i)$, $i=0,1$, using (A1) and (A2), there exists $C>0$ such that

$$\mathcal{N}_{C\eta}([p_0, p_1]) \subset \psi(\mathcal{N}_\eta([y_0, y_1]_{x_m})).$$

Moreover, as Ω' is c-convex with respect to x_m, $\psi(\Omega')$ is a convex set.

Then we claim the following: for $U \subset \mathbb{R}^n$ convex and for $u, v \in U$, the function

$$r \longmapsto \frac{\operatorname{Vol}(\mathcal{N}_r([u,v]) \cap U)}{\operatorname{Vol}(\mathcal{N}_r([u,v]))}$$

is non-increasing. Indeed, by the convexity of U, for $w \in [u,v]$, if $w + w' \in B_r(w) \cap U$, then $w + \theta w' \in B_{\theta r}(w) \cap U$ for $\theta \in [0,1]$. From this the claim easily follows.

Hence, we have

$$\begin{aligned} \operatorname{Vol}(\psi(\mathcal{N}_\eta([y_0, y_1]_{x_m}) \cap \Omega')) &\geqslant \operatorname{Vol}(\mathcal{N}_{C\eta}([p_0, p_1]) \cap \psi(\Omega')) \\ &\geqslant \frac{\operatorname{Vol}(\mathcal{N}_{C\eta}([p_0, p_1])) \operatorname{Vol}(\mathcal{N}_1([p_0, p_1]) \cap \psi(\Omega'))}{\operatorname{Vol}(\mathcal{N}_1([p_0, p_1]))}, \end{aligned}$$

whenever η is small enough so that $C\eta \leqslant 1$. By compactness, one has

$$\frac{\operatorname{Vol}(\mathcal{N}_1([p_0, p_1]) \cap \psi(\Omega'))}{\operatorname{Vol}(\mathcal{N}_1([p_0, p_1]))} \geqslant C(\Omega').$$

Moreover, for $C>0$, there exists a constant $C'>0$ such that

$$\operatorname{Vol}(\mathcal{N}_{C\eta}([p_0, p_1])) \geqslant C' \operatorname{Vol}(\mathcal{N}_\eta([p_0, p_1]))$$

for all $\eta>0$. Then, as ψ is a smooth diffeomorphism, one has

$$\frac{\operatorname{Vol}(\mathcal{N}_\eta([y_0, y_1]_{x_m}) \cap \Omega')}{\operatorname{Vol}(\mathcal{N}_\eta([y_0, y_1]_{x_m}))} \geqslant C(c, \Omega, \Omega') \frac{\operatorname{Vol}(\psi(\mathcal{N}_\eta([y_0, y_1]_{x_m}) \cap \Omega'))}{\operatorname{Vol}(\psi(\mathcal{N}_\eta([y_0, y_1]_{x_m})))}. \qquad \square$$

References

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

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

[3] Caffarelli, L. A., Interior $W^{2,p}$ estimates for solutions of the Monge–Ampère equation. *Ann. of Math.*, 131 (1990), 135–150.

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

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

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

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

[8] — Allocation maps with general cost functions, in *Partial Differential Equations and Applications*, Lecture Notes in Pure and Appl. Math., 177, pp. 29–35. Dekker, New York, 1996.

[9] — Boundary regularity of maps with convex potentials. II. *Ann. of Math.*, 144 (1996), 453–496.

[10] Caffarelli, L. A. & Cabré, X., *Fully Nonlinear Elliptic Equations*. American Mathematical Society Colloquium Publications, 43. Amer. Math. Soc., Providence, RI, 1995.

[11] Caffarelli, L. A., Gutiérrez, C. E. & Huang, Q., On the regularity of reflector antennas. *Ann. of Math.*, 167 (2008), 299–323.

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

[13] Delanoë, P., Classical solvability in dimension two of the second boundary-value problem associated with the Monge–Ampère operator. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, 8 (1991), 443–457.

[14] — Gradient rearrangement for diffeomorphisms of a compact manifold. *Differential Geom. Appl.*, 20 (2004), 145–165.

[15] Delanoë, P. & Loeper, G., Gradient estimates for potentials of invertible gradient-mappings on the sphere. *Calc. Var. Partial Differential Equations*, 26 (2006), 297–311.

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

[17] — Shape recognition via Wasserstein distance. *Quart. Appl. Math.*, 58 (2000), 705–737.

[18] Gilbarg, D. & Trudinger, N. S., *Elliptic Partial Differential Equations of Second Order*. Grundlehren der Mathematischen Wissenschaften, 224. Springer, Berlin–Heidelberg, 1983.

[19] Guan, P. & Wang, X.-J., On a Monge–Ampère equation arising in geometric optics. *J. Differential Geom.*, 48 (1998), 205–223.

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

[21] Knott, M. & Smith, C. S., On the optimal mapping of distributions. *J. Optim. Theory Appl.*, 43 (1984), 39–49.

[22] Loeper, G., On the regularity of solutions of optimal transportation problems: the sphere and the reflector antenna. To appear in *Arch. Ration. Mech. Anal.*

[23] Lott, J. & Villani, C., Ricci curvature for metric-measure spaces via optimal transport. *Ann. of Math.*, 169 (2009), 903–991.

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

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

[26] Otto, F. & Villani, C., Generalization of an inequality by Talagrand and links with the logarithmic Sobolev inequality. *J. Funct. Anal.*, 173 (2000), 361–400.

[27] Sturm, K.-T., On the geometry of metric measure spaces. I, II. *Acta Math.*, 196 (2006), 65–131, 133–177.

[28] Trudinger, N. S., Recent developments in elliptic partial differential equations of Monge–Ampère type, in *International Congress of Mathematicians.* Vol. III, pp. 291–301. Eur. Math. Soc., Zürich, 2006.

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

[30] — On the second boundary value problem for Monge–Ampère type equations and optimal transportation. *Ann. Scuola Norm. Sup. Pisa Cl. Sci.*, 8 (2009), 143–174.

[31] Urbas, J., On the second boundary value problem for equations of Monge–Ampère type. *J. Reine Angew. Math.*, 487 (1997), 115–124.

[32] — Mass transfer problems. Lecture Notes, University of Bonn, 1998.

[33] Villani, C., *Topics in Optimal Transportation.* Graduate Studies in Mathematics, 58. Amer. Math. Soc., Providence, RI, 2003.

[34] — *Optimal Transport. Old and New.* Grundlehren der Mathematischen Wissenschaften, 338. Springer, Berlin–Heidelberg, 2009.

[35] Wang, X.-J., Some counterexamples to the regularity of Monge–Ampère equations. *Proc. Amer. Math. Soc.*, 123 (1995), 841–845.

[36] — On the design of a reflector antenna. *Inverse Problems*, 12 (1996), 351–375.

[37] — On the design of a reflector antenna. II. *Calc. Var. Partial Differential Equations*, 20 (2004), 329–341.

Grégoire Loeper
Institut Camille Jordan
Université Claude Bernard Lyon 1
43 boulevard du 11 novembre 1918
FR-69622 Villeurbanne cedex
France
loeper@math.univ-lyon1.fr

Received July, 2, 2007