PARTIAL REGULARITY FOR OPTIMAL TRANSPORT MAPS

by Guido De PHILIPPIS *and* Alessio FIGALLI

ABSTRACT

We prove that, for general cost functions on $\mathbf{R}^n$, or for the cost $d^2/2$ on a Riemannian manifold, optimal transport maps between smooth densities are always smooth outside a closed singular set of measure zero.

1. Introduction

A natural and important issue in optimal transport theory is the regularity of optimal transport maps. Indeed, apart from being a typical PDE/analysis question, knowing whether optimal maps are smooth or not is an important step towards a qualitative understanding of them.

It is by now well known that, for the smoothness of optimal maps, conditions on both the cost function and on the geometry of the supports of the measures are needed.

In the special case $c(x,y) = |x-y|^2/2$ on $\mathbf{R}^n$, Caffarelli [3–6] proved regularity of optimal maps under suitable assumptions on the densities and on the geometry of their support. More precisely, in its simplest form, Caffarelli's result states as follows:

Theorem **1.1.** — *Let f and g be smooth probability densities, respectively bounded away from zero and infinity on two bounded open sets X and Y, and let $T : X \to Y$ denote the unique optimal transport map from f to g for the quadratic cost $|x-y|^2/2$. If Y is convex, then T is smooth inside X. On the other hand, if Y is not convex, then there exist smooth densities f and g (both bounded away from zero and infinity on X and Y, respectively) for which the map T is not continuous.*

A natural question which arises from the previous result is whether one may prove some partial regularity on T when the convexity assumption on Y is removed. In [16, 18] the authors proved the following result:

Theorem **1.2.** — *Let f and g be smooth probability densities, respectively bounded away from zero and infinity on two bounded open sets X and Y, and let $T : X \to Y$ denote the unique optimal transport map from f to g for the quadratic cost $|x-y|^2/2$. Then there exist two open sets $X' \subset X$ and $Y' \subset Y$, with $|X \setminus X'| = |Y \setminus Y'| = 0$, such that $T : X' \to Y'$ is a smooth diffeomorphism.*

In the case of general cost functions on $\mathbf{R}^n$, or when $c(x,y) = d(x,y)^2/2$ on a Riemannian manifold M ($d(x,y)$ being the Riemannian distance), the situation is much more complicated. Indeed, as shown by Ma, Trudinger, and Wang [33], and Loeper [31], in addition to suitable convexity assumptions on the support of the target density (or on the cut locus of the manifold when $\mathrm{supp}(g) = M$ [24]), a very strong structural condition on

DOI 10.1007/s10240-014-0064-7

the cost function, the so-called *MTW condition*, is needed to ensure the smoothness of the map.

More precisely, if the MTW condition holds (together with some suitable convexity assumptions on the target domain), then the optimal map is smooth [19, 21, 30, 35, 36]. On the other hand, if the MTW condition fails at one point, then one can construct smooth densities (both supported on domains which satisfy the needed convexity assumptions) for which the optimal transport map is not continuous [31] (see also [15]).

In the case of Riemannian manifolds, the MTW condition for $c = d^2/2$ is very restrictive: indeed, as shown by Loeper [31], it implies that M has non-negative sectional curvature, and actually it is much stronger than the latter [23, 28]. In particular, if M has negative sectional curvature, then the MTW condition fails at *every* point. Let us also mention that, up to now, the MTW condition is known to be satisfied only for very special classes of Riemannian manifolds, such as spheres, their products, their quotients and submersions, and their perturbations [10, 11, 20, 22, 25, 29, 32], and for instance it is known to fail on sufficiently flat ellipsoids [23].

The goal of the present paper is to show that, even without any condition on the cost function or on the supports of the densities, optimal transport maps are always smooth outside a closed singular set of measure zero. In order to state our results, we first have to introduce some basic assumptions on the cost functions which are needed to ensure existence and uniqueness of optimal maps. As before, X and Y denote two open subsets of $\mathbf{R}^n$.

(C0) The cost function $c : \mathrm{X} \times \mathrm{Y} \to \mathbf{R}$ is of class C^2 with $\|c\|_{\mathrm{C}^2(\mathrm{X}\times\mathrm{Y})} < \infty$.
(C1) For any $x \in \mathrm{X}$, the map $\mathrm{Y} \ni y \mapsto -\mathrm{D}_x c(x, y) \in \mathbf{R}^n$ is injective.
(C2) For any $y \in \mathrm{Y}$, the map $\mathrm{X} \ni x \mapsto -\mathrm{D}_y c(x, y) \in \mathbf{R}^n$ is injective.
(C3) $\det(\mathrm{D}_{xy} c)(x, y) \neq 0$ for all $(x, y) \in \mathrm{X} \times \mathrm{Y}$.

Here are our main results:

Theorem **1.3.** — *Let* $\mathrm{X}, \mathrm{Y} \subset \mathbf{R}^n$ *be two bounded open sets, and let* $f : \mathrm{X} \to \mathbf{R}^+$ *and* $g : \mathrm{Y} \to \mathbf{R}^+$ *be two continuous probability densities, respectively bounded away from zero and infinity on* X *and* Y. *Assume that the cost* $c : \mathrm{X} \times \mathrm{Y} \to \mathbf{R}$ *satisfies* (C0)–(C3), *and denote by* $\mathrm{T} : \mathrm{X} \to \mathrm{Y}$ *the unique optimal transport map sending* f *onto* g. *Then there exist two relatively closed sets* $\Sigma_\mathrm{X} \subset \mathrm{X}$, $\Sigma_\mathrm{Y} \subset \mathrm{Y}$ *of measure zero such that* $\mathrm{T} : \mathrm{X} \setminus \Sigma_\mathrm{X} \to \mathrm{Y} \setminus \Sigma_\mathrm{Y}$ *is a homeomorphism of class* $\mathrm{C}^{0,\beta}_{\mathrm{loc}}$ *for any* $\beta < 1$. *In addition, if* $c \in \mathrm{C}^{k+2,\alpha}_{\mathrm{loc}}(\mathrm{X} \times \mathrm{Y})$, $f \in \mathrm{C}^{k,\alpha}_{\mathrm{loc}}(\mathrm{X})$, *and* $g \in \mathrm{C}^{k,\alpha}_{\mathrm{loc}}(\mathrm{Y})$ *for some* $k \geq 0$ *and* $\alpha \in (0, 1)$, *then* $\mathrm{T} : \mathrm{X} \setminus \Sigma_\mathrm{X} \to \mathrm{Y} \setminus \Sigma_\mathrm{Y}$ *is a diffeomorphism of class* $\mathrm{C}^{k+1,\alpha}_{\mathrm{loc}}$.

Theorem **1.4.** — *Let* M *be a smooth Riemannian manifold, and let* $f, g : \mathrm{M} \to \mathbf{R}^+$ *be two continuous probability densities, locally bounded away from zero and infinity on* M. *Let* $\mathrm{T} : \mathrm{M} \to \mathrm{M}$ *denote the optimal transport map for the cost* $c = d^2/2$ *sending* f *onto* g. *Then there exist two closed sets* $\Sigma_\mathrm{X}, \Sigma_\mathrm{Y} \subset \mathrm{M}$ *of measure zero such that* $\mathrm{T} : \mathrm{M} \setminus \Sigma_\mathrm{X} \to \mathrm{M} \setminus \Sigma_\mathrm{Y}$ *is a homeomorphism of class* $\mathrm{C}^{0,\beta}_{\mathrm{loc}}$ *for any* $\beta < 1$. *In addition, if both* f *and* g *are of class* $\mathrm{C}^{k,\alpha}$, *then* $\mathrm{T} : \mathrm{M} \setminus \Sigma_\mathrm{X} \to \mathrm{M} \setminus \Sigma_\mathrm{Y}$ *is a diffeomorphism of class* $\mathrm{C}^{k+1,\alpha}_{\mathrm{loc}}$.

The paper is structured as follows: in the next section we introduce some notation and preliminary results. Then, in Section 3, we show how both Theorem 1.3 and Theorem 1.4 are a direct consequence of some local regularity results around differentiability points of T, see Theorems 4.3 and 5.3. Finally, Sections 4 and 5 are devoted to the proof of these local results.

2. Notation and preliminary results

Through a well established procedure, maps that solve optimal transport problems derive from a c-convex potential, itself solution to a Monge-Ampère type equation.

More precisely, given a cost function $c : X \times Y \to \mathbf{R}$, a function $u : X \to \mathbf{R}$ is said *c-convex* if it can be written as

$$u(x) = \sup_{y \in Y} \{ -c(x,y) + \lambda_y \}, \tag{2.1}$$

for some constants $\lambda_y \in \mathbf{R} \cup \{-\infty\}$.

Similarly to the subdifferential for convex function, for c-convex functions one can talk about their c-subdifferential: if $u : X \to \mathbf{R}$ is a c-convex function as above, the *c-subdifferential* of u at x is the (nonempty) set

$$\partial_c u(x) := \{ y \in \overline{Y} : \ u(z) \geq -c(z,y) + c(x,y) + u(x) \ \forall z \in X \}. \tag{2.2}$$

If $x_0 \in X$ and $y_0 \in \partial_c u(x_0)$, we will say that the function

$$C_{x_0,y_0}(\cdot) := -c(\cdot, y_0) + c(x_0, y_0) + u(x_0) \tag{2.3}$$

is a *c-support* for u at x_0. We also define the *Frechet subdifferential* of u at x as

$$\partial^- u(x) := \{ p \in \mathbf{R}^n : \ u(z) \geq u(x) + p \cdot (z - x) + o(|z - x|) \}.$$

We will use the following notation: if $E \subset X$ then

$$\partial_c u(E) := \bigcup_{x \in E} \partial_c u(x), \qquad \partial^- u(E) := \bigcup_{x \in E} \partial^- u(x).$$

It is easy to check that, if c is of class C^1, then the following inclusion holds:

$$y \in \partial_c u(x) \quad \Longrightarrow \quad -D_x c(x,y) \in \partial^- u(x). \tag{2.4}$$

In addition, if c satisfies (C0)–(C2), then we can define the *c-exponential map*:

$$\text{for any } x \in X, y \in Y, p \in \mathbf{R}^n, \quad \begin{cases} \text{c-exp}_x(p) = y & \Leftrightarrow \quad p = -D_x c(x,y) \\ \text{c}^*\text{-exp}_y(p) = x & \Leftrightarrow \quad p = -D_y c(x,y) \end{cases} \tag{2.5}$$

Using (**2.5**), we can rewrite (**2.4**) as

$$\partial_c u(x) \subset \text{c-exp}_x\big(\partial^- u(x)\big). \tag{2.6}$$

Notice that, if $c \in C^1$ and Y is bounded, it follows immediately from (**2.1**) that c-convex functions are Lipschitz, so in particular they are differentiable a.e.

The following notation will be convenient: given a c-convex function $u : X \to \mathbf{R}$, we define (at almost every point) the map $T_u : X \to Y$ as

$$T_u(x) := \text{c-exp}_x\big(\nabla u(x)\big). \tag{2.7}$$

(Of course T_u depends also on c, but to keep the notation lighter we prefer not to make this dependence explicit. The reader should keep in mind that, whenever we write T_u, the cost c is always the one for which u is c-convex.)

Finally, let us observe that if c satisfies (C0) and Y is bounded, then it follows from (**2.1**) that u is semiconvex (i.e., there exists a constant $C > 0$ such that $u + C|x|^2/2$ is convex, see for instance [13]). In particular, by Alexandrov's Theorem, c-convex functions are twice differentiable a.e. (see [37, Theorem 14.25] for a list of different equivalent definitions of this notion).

The following is a basic result in optimal transport theory (see for instance [37, Chapter 10]):

Theorem **2.1.** — *Let* $c : X \times Y \to \mathbf{R}$ *satisfy* (C0)–(C1). *Given two probability densities* f *and* g *supported on* X *and* Y *respectively, there exists a c-convex function* $u : X \to \mathbf{R}$ *such that* $T_u : X \to Y$ *is the unique optimal transport map sending* f *onto* g.

In the particular case $c(x, y) = -x \cdot y$ (which is equivalent to the quadratic cost $|x - y|^2/2$), c-convex functions are convex and the above result takes the following simple form [2]:

Theorem **2.2.** — *Let* $c(x, y) = -x \cdot y$. *Given two probability densities* f *and* g *supported on* X *and* Y *respectively, there exists a convex function* $v : X \to \mathbf{R}$ *such that* $T_v = \nabla v : X \to Y$ *is the unique optimal transport map sending* f *onto* g.

Although on Riemannian manifolds the cost function $c = d^2/2$ is not smooth everywhere, one can still prove existence of optimal maps [13, 17, 34] (let us remark that, in this case, the c-exponential map coincides with the classical exponential map in Riemannian geometry):

Theorem **2.3.** — *Let* M *be a smooth Riemannian manifold, and* $c = d^2/2$. *Given two probability densities* f *and* g *supported on* M, *there exists a c-convex function* $u : M \to \mathbf{R} \cup \{+\infty\}$ *such that* u *is differentiable* f*-a.e., and* $T_u(x) = \exp_x(\nabla u(x))$ *is the unique optimal transport map sending* f *onto* g.

We conclude this section by recalling that c-convex functions arising in optimal transport problems solve a Monge-Ampère type equation almost everywhere, referring to [1, Section 6.2], [37, Chapters 11 and 12], and [15] for more details.

Whenever c satisfies (C0)–(C3), then the transport condition $(T_u)_\sharp f = g$ gives

$$|\det(DT_u(x))| = \frac{f(x)}{g(T_u(x))} \quad \text{a.e.} \tag{2.8}$$

In addition, the c-convexity of u implies that, at every point x where u is twice differentiable,

$$D^2u(x) + D_{xx}c(x, \text{c-exp}_x(\nabla u(x))) \geq 0. \tag{2.9}$$

Hence, writing (**2.7**) as

$$-D_x c(x, T_u(x)) = \nabla u(x),$$

differentiating the above relation with respect to x, and using (**2.8**) and (**2.9**), we obtain

$$\begin{aligned} &\det(D^2u(x) + D_{xx}c(x, \text{c-exp}_x(\nabla u(x)))) \\ &\quad = |\det(D_{xy}c(x, \text{c-exp}_x(\nabla u(x))))| \frac{f(x)}{g(\text{c-exp}_x(\nabla u(x)))} \end{aligned} \tag{2.10}$$

at every point x where u it is twice differentiable. In particular, when $c(x,y) = -x \cdot y$, the convex function v provided by Theorem 2.2 solves the classical Monge-Ampère equation

$$\det(D^2v(x)) = \frac{f(x)}{g(\nabla v(x))} \quad \text{a.e.}$$

3. The localization argument and proof of the results

The goal of this section is to prove Theorems 1.3 and 1.4 by showing that the assumptions of Theorems 4.3 and 5.3 below are satisfied near almost every point.

The rough idea is the following: if $\bar{x}$ is a point where the semiconvex function u is twice differentiable, then around that point u looks like a parabola. In addition, by looking close enough to $\bar{x}$, the cost function c will be very close to the linear one and the densities will be almost constant there. Hence we can apply Theorem 4.3 to deduce that u is of class $C^{1,\beta}$ in neighborhood of $\bar{x}$ (resp. u is of class $C^{k+2,\alpha}$ by Theorem 5.3, if $c \in C^{k+2,\alpha}_{\text{loc}}$ and $f, g \in C^{k,\alpha}_{\text{loc}}$), which implies in particular that T_u is of class $C^{0,\beta}$ in neighborhood of $\bar{x}$ (resp. T_u is of class $C^{k+1,\alpha}$ by Theorem 5.3, if $c \in C^{k+2,\alpha}_{\text{loc}}$ and $f, g \in C^{k,\alpha}_{\text{loc}}$). Being our assumptions completely symmetric in x and y, we can apply the same argument to the optimal map T^* sending g onto f. Since $T^* = (T_u)^{-1}$ (see the discussion below), it follows that T_u is a global homeomorphism of class $C^{0,\beta}_{\mathrm{loc}}$ (resp. T_u is a global diffeomorphism of class $C^{k+1,\alpha}_{\mathrm{loc}}$) outside a closed set of measure zero.

We now give a detailed proof.

Proof of Theorem 1.3. — Let us introduce the "c-conjugate" of u, that is, the function $u^c : Y \to \mathbf{R}$ defined as

$$u^c(y) := \sup_{x \in X} \big\{ -c(x, y) - u(x) \big\}.$$

Then u^c is c^*-convex, where

$$c^*(y, x) := c(x, y), \quad \text{and} \quad x \in \partial_{c^*} u^c(y) \iff y \in \partial_c u(x) \tag{3.1}$$

(see for instance [37, Chapter 5]).

Being our assumptions completely symmetric in x and y, c^* satisfies the same assumptions as c. In particular, by Theorem 2.1, there exists an optimal map T^* (with respect to c^*) sending g onto f. In addition, it is well-known that T^* is actually equal to

$$T_{u^c}(y) = \mathrm{c}^*\text{-}\exp_y\big(\nabla u^c(y)\big),$$

and that T_u and T_{u^c} are inverse to each other, that is

$$T_{u^c}\big(T_u(x)\big) = x, \quad T_u\big(T_{u^c}(y)\big) = y \quad \text{for a.e. } x \in X, y \in Y \tag{3.2}$$

(see, for instance, [1, Remark 6.2.11]).

Since semiconvex functions are twice differentiable a.e., there exist sets $X_1 \subset X$, $Y_1 \subset Y$ of full measure such that (**3.2**) holds for every $x \in X_1$ and $y \in Y_1$, and in addition u is twice differentiable for every $x \in X_1$ and u^c is twice differentiable for every $y \in Y_1$. Let us define

$$X' := X_1 \cap (T_u)^{-1}(Y_1).$$

Using that T_u transports f on g and that the two densities are bounded away from zero and infinity, we see that X' is of full measure in X.

We fix a point $\bar{x} \in X'$. Since u is differentiable at $\bar{x}$ (being twice differentiable), it follows by (**2.6**) that the set $\partial_c u(\bar{x})$ is a singleton, namely $\partial_c u(\bar{x}) = \{\mathrm{c}\text{-}\exp_{\bar{x}}(\nabla u(\bar{x}))\}$. Set $\bar{y} := \mathrm{c}\text{-}\exp_{\bar{x}}(\nabla u(\bar{x}))$. Since $\bar{y} \in Y_1$ (by definition of X'), u^c is twice differentiable at $\bar{y}$ and $\bar{x} = T_{u^c}(\bar{y})$. Up to a translation in the system of coordinates (both in x and y) we can assume that both $\bar{x}$ and $\bar{y}$ coincide with the origin $\mathbf{0}$.

Let us define

$$\begin{aligned}
\bar{u}(z) &:= u(z) - u(\mathbf{0}) + c(z, \mathbf{0}) - c(\mathbf{0}, \mathbf{0}),\\
\bar{c}(z, w) &:= c(z, w) - c(z, \mathbf{0}) - c(\mathbf{0}, w) + c(\mathbf{0}, \mathbf{0}),\\
\bar{u}^{\bar{c}}(w) &:= u^c(w) - u(\mathbf{0}) + c(\mathbf{0}, w) - c(\mathbf{0}, \mathbf{0}).
\end{aligned}$$

Then $\bar u$ is a $\bar c$-convex function, $\bar u^{\bar c}$ is its $\bar c$-conjugate, $T_{\bar u} = T_u$, and $T_{\bar u^{\bar c}} = T_{u^c}$, so in particular $(T_{\bar u})_\sharp f = g$ and $(T_{\bar u^{\bar c}})_\sharp g = f$. In addition, because by assumption $\mathbf{0} \in X'$, $\bar u$ is twice differentiable at $\mathbf{0}$ and $\bar u^{\bar c}$ is twice differentiable at $\mathbf{0} = T_{\bar u}(\mathbf{0})$. Let us define $P := D^2\bar u(\mathbf{0})$, and $M := D_{xy}\bar c(\mathbf{0}, \mathbf{0})$. Then, since $\bar c(\cdot, \mathbf{0}) = \bar c(\mathbf{0}, \cdot) \equiv 0$ and $\bar c \in C^2$, a Taylor expansion gives

$$\bar u(z) = \frac{1}{2} Pz \cdot z + o\big(|z|^2\big), \qquad \bar c(z, w) = Mz \cdot w + o\big(|z|^2 + |w|^2\big).$$

Let us observe that, since by assumption f and g are bounded away from zero and infinity, by (C3) and (**2.10**) applied to $\bar u$ and $\bar c$ we get that $\det(P), \det(M) \neq 0$. In addition (**2.9**) implies that P is a positive definite symmetric matrix. Hence, we can perform a second change of coordinates: $z \mapsto \tilde z := P^{1/2} z$, $w \mapsto \tilde w := -P^{-1/2}M^* w$ (M^* being the transpose of M), so that, in the new variables,

$$\tilde u(\tilde z) := \bar u(z) = \frac{1}{2}|\tilde z|^2 + o\big(|\tilde z|^2\big), \qquad \tilde c(\tilde z, \tilde w) := \bar c(z, w) = -\tilde z \cdot \tilde w + o\big(|\tilde z|^2 + |\tilde w|^2\big). \tag{3.3}$$

By an easy computation it follows that $(T_{\tilde u})_\sharp \tilde f = \tilde g$, where[1]

$$\tilde f(\tilde z) := \det\big(P^{-1/2}\big) f\big(P^{-1/2}\tilde z\big), \qquad \tilde g(\tilde w) := \big|\det\big(\big(M^*\big)^{-1} P^{1/2}\big)\big|\, g\big(-\big(M^*\big)^{-1} P^{1/2}\tilde w\big). \tag{3.4}$$

Notice that

$$D_{\tilde z\tilde z}\tilde c(\mathbf{0}, \mathbf{0}) = D_{\tilde w\tilde w}\tilde c(\mathbf{0}, \mathbf{0}) = \mathbf{0}_{n\times n}, \qquad -D_{\tilde z\tilde w}\tilde c(\mathbf{0}, \mathbf{0}) = \mathrm{Id}, \qquad D^2\tilde u(\mathbf{0}) = \mathrm{Id}, \tag{3.5}$$

so, using (**2.10**), we deduce that

$$\frac{\tilde f(\mathbf{0})}{\tilde g(\mathbf{0})} = \frac{\det(D^2\tilde u(\mathbf{0}) + D_{\tilde z\tilde z}\tilde c(\mathbf{0}, \mathbf{0}))}{|\det(D_{\tilde z\tilde w}\tilde c(\mathbf{0}, \mathbf{0}))|} = 1. \tag{3.6}$$

To ensure that we can apply Theorems 4.3 and 5.3, we now perform the following dilation: for $\rho > 0$ we define

$$u_\rho(\tilde z) := \frac{1}{\rho^2}\tilde u(\rho\tilde z), \qquad c_\rho(\tilde z, \tilde w) := \frac{1}{\rho^2}\tilde c(\rho\tilde z, \rho\tilde w).$$

We claim that, provided ρ is sufficiently small, u_ρ and c_ρ satisfy the assumptions of Theorems 4.3 and 5.3.

Indeed, it is immediate to check that u_ρ is a c_ρ-convex function. Also, by the same argument as above, from the relation $(T_{\tilde u})_\sharp \tilde f = \tilde g$ we deduce that T_{u_ρ} sends $\tilde f(\rho\tilde z)$

---

[1] An easy way to check this is to observe that the measures $\mu := f(x)dx$ and $\nu := g(y)dy$ are independent of the choice of coordinates, hence (**3.4**) follows from the identities

$$f(x)dx = \tilde f(\tilde x)d\tilde x, \qquad g(y)dy = \tilde g(\tilde y)d\tilde y.$$

onto $\tilde{g}(\rho\tilde{w})$. In addition, since we can freely multiply both densities by a same constant, it actually follows from (**3.6**) that $(\mathrm{T}_{u_\rho})_\sharp f_\rho = g_\rho$, where

$$f_\rho(\tilde{z}) := \frac{\tilde{f}(\rho\tilde{z})}{\tilde{f}(\mathbf{0})}, \qquad g_\rho(\tilde{w}) := \frac{\tilde{g}(\rho\tilde{w})}{\tilde{g}(\mathbf{0})}.$$

In particular, since f and g are continuous, we get

$$|f_\rho - 1| + |g_\rho - 1| \to 0 \quad \text{inside } \mathrm{B}_3 \tag{3.7}$$

as $\rho \to 0$. Also, by (**3.3**) we get that, for any $\tilde{z}, \tilde{w} \in \mathrm{B}_3$,

$$u_\rho(\tilde{z}) = \frac{1}{2}|\tilde{z}|^2 + o(1), \qquad c_\rho(\tilde{z},\tilde{w}) = -\tilde{z}\cdot\tilde{w} + o(1), \tag{3.8}$$

where $o(1) \to 0$ as $\rho \to 0$. In particular, (**4.9**) and (**4.10**) hold with any positive constants δ_0, η_0 provided ρ is small enough.

Furthermore, by the second order differentiability of $\tilde{u}$ at $\mathbf{0}$ it follows that the multivalued map $\tilde{z} \mapsto \partial^- \bar{u}(\tilde{z})$ is differentiable at $\mathbf{0}$ (see [37, Theorem 14.25]) with gradient equal to the identity matrix (see (**3.3**)), hence

$$\partial^- u_\rho(\tilde{z}) \subset \mathrm{B}_{\gamma_\rho}(\tilde{z}) \quad \forall \tilde{z} \in \mathrm{B}_2,$$

where $\gamma_\rho \to 0$ as $\rho \to 0$. Since $\partial_{c_\rho} u_\rho \subset \mathrm{c}_\rho\text{-exp}(\partial^- u_\rho)$ (by (**2.6**)) and $\|\mathrm{c}_\rho\text{-exp} - \mathrm{Id}\|_\infty = o(1)$ (by (**3.8**)), we get

$$\partial_{c_\rho} u_\rho(\tilde{z}) \subset \mathrm{B}_{\delta_\rho}(\tilde{z}) \quad \forall \tilde{z} \in \mathrm{B}_3, \tag{3.9}$$

with $\delta_\rho = o(1)$ as $\rho \to 0$. Moreover, the c_ρ-conjugate of u_ρ is easily seen to be

$$u_\rho^{c_\rho}(\tilde{w}) = \bar{u}^{\bar{c}}\big(\rho\big(\mathrm{M}^*\big)^{-1}\mathrm{P}^{1/2}\tilde{w}\big).$$

Since u^c is twice differentiable at $\mathbf{0}$, so is $u_\rho^{c_\rho}$. In addition, an easy computation[2] shows that $\mathrm{D}^2 u_\rho^{c_\rho}(\mathbf{0}) = \mathrm{Id}$. Hence, arguing as above we obtain that

$$\partial_{c_\rho^*} u_\rho^{c_\rho}(\tilde{w}) \subset \mathrm{B}_{\delta'_\rho}(\tilde{w}) \quad \forall \tilde{w} \in \mathrm{B}_3, \tag{3.10}$$

with $\delta'_\rho = o(1)$ as $\rho \to 0$.

We now define

$$\mathcal{C}_1 := \overline{\mathrm{B}}_1, \qquad \mathcal{C}_2 := \partial_{c_\rho} u_\rho(\mathcal{C}_1).$$

---

[2] For instance, this follows by differentiating both relations

$$\mathrm{D}_{\tilde{z}} c_\rho\big(\tilde{z}, \mathrm{T}_{u_\rho}(\tilde{z})\big) = -\nabla u_\rho(\tilde{z}) \quad \text{and} \quad \mathrm{D}_{\tilde{w}} c_\rho\big(\mathrm{T}_{u_\rho^{c_\rho}}(\tilde{w}), \tilde{w}\big) = -\nabla u_\rho^{c_\rho}(\tilde{w})$$

at $\mathbf{0}$, and using then (**3.5**) and the fact that $\nabla \mathrm{T}_{u_\rho^{c_\rho}}(\mathbf{0}) = [\nabla \mathrm{T}_{u_\rho}(\mathbf{0})]^{-1}$ and $\mathrm{D}^2 u_\rho(\mathbf{0}) = \mathrm{Id}$.

Observe that both $\mathcal{C}_1$ and $\mathcal{C}_2$ are closed (since the c-subdifferential of a compact set is closed). Also, thanks to (**3.9**), by choosing ρ small enough we can ensure that $B_{1/3} \subset \mathcal{C}_2 \subset B_3$. Finally, it follows from (**2.6**) that

$$(T_{u_\rho})^{-1}(\mathcal{C}_2) \setminus \mathcal{C}_1 \subset (T_{u_\rho})^{-1}\big(\{\text{points of non-differentiability of } u_\rho^{c_\rho}\}\big),$$

and since this latter set has measure zero, a simple computation shows that

$$(T_{u_\rho})_\sharp(f_\rho \mathbf{1}_{\mathcal{C}_1}) = g_\rho \mathbf{1}_{\mathcal{C}_2}.$$

Thus, thanks to (**4.8**), we get that for any $\beta < 1$ the assumptions of Theorem 4.3 are satisfied, provided we choose ρ sufficiently small. Moreover, if in addition $c \in C^{k+2,\alpha}_{\mathrm{loc}}(X \times Y)$, $f \in C^{k,\alpha}_{\mathrm{loc}}(X)$, and $g \in C^{k,\alpha}_{\mathrm{loc}}(Y)$, then also the assumptions of Theorem 5.3 are satisfied.

Hence, by applying Theorem 4.3 (resp. Theorem 5.3) we deduce that $u_\rho \in C^{1,\beta}(B_{1/7})$ (resp. $u_\rho \in C^{k+2,\alpha}(B_{1/9})$), so going back to the original variables we get the existence of a neighborhood $\mathcal{U}_{\bar{x}}$ of $\bar{x}$ such that $u \in C^{1,\beta}(\mathcal{U}_{\bar{x}})$ (resp. $u \in C^{k+2,\alpha}(\mathcal{U}_{\bar{x}})$). This implies in particular that $T_u \in C^{0,\beta}(\mathcal{U}_{\bar{x}})$ (resp. $T_u \in C^{k+1,\alpha}(\mathcal{U}_{\bar{x}})$). Moreover, it follows by Corollary 4.6 that $T_u(\mathcal{U}_{\bar{x}})$ contains a neighborhood of $\bar{y}$.

We now observe that, by symmetry, we can also apply Theorem 4.3 (resp. Theorem 5.3) to $u_\rho^{c_\rho}$. Hence, there exists a neighborhood $\mathcal{V}_{\bar{y}}$ of $\bar{y}$ such that $T_{u^c} \in C^{0,\beta}(\mathcal{V}_{\bar{y}})$. Since T_u and T_{u^c} are inverse to each other (see (**3.2**)) we deduce that, possibly reducing the size of $\mathcal{U}_{\bar{x}}$, T_u is a homeomorphism (resp. diffeomorphism) between $\mathcal{U}_{\bar{x}}$ and $T_u(\mathcal{U}_{\bar{x}})$. Let us consider the open sets

$$X'' := \bigcup_{\bar{x} \in X'} \mathcal{U}_{\bar{x}}, \qquad Y'' := \bigcup_{\bar{x} \in X'} T_u(\mathcal{U}_{\bar{x}}),$$

and define the (relatively) closed $\Sigma_X := X \setminus X''$, $\Sigma_Y := Y \setminus Y''$. Since $X'' \supset X'$, X'' is a set of full measure, so $|\Sigma_X| = 0$. In addition, since $\Sigma_Y = Y \setminus Y'' \subset Y \setminus T_u(X')$ and $T_u(X')$ has full measure in Y, we also get that $|\Sigma_Y| = 0$.

Finally, since $T_u : X \setminus \Sigma_X \to Y \setminus \Sigma_Y$ is a local homeomorphism (resp. diffeomorphism), by (**3.2**) it follows that $T_u : X \setminus \Sigma_X \to Y \setminus \Sigma_Y$ is a global homeomorphism (resp. diffeomorphism), which concludes the proof. □

Proof of Theorem 1.4. — The only difference with respect to the situation in Theorem 1.3 is that now the cost function $c = d^2/2$ is not smooth on the whole $M \times M$. However, even if $d^2/2$ is not everywhere smooth and M is not necessarily compact, it is still true that the c-convex function u provided by Theorem 2.3 is locally semiconvex (i.e., it is locally semiconvex when seen in any chart) [13, 17]. In addition, as shown in [9, Proposition 4.1] (see also [14, Section 3]), if u is twice differentiable at x, then the point $T_u(x)$ is not in the cut-locus of x. Since the cut-locus is closed and $d^2/2$ is smooth outside the cut-locus, we deduce the existence of a set X of full measure such that, if $x_0 \in X$, then: (1) u is twice differentiable at x_0; (2) there exists a neighborhood $\mathcal{U}_{x_0} \times \mathcal{V}_{T_u(x_0)} \subset M \times M$ of $(x_0, T_u(x_0))$ such that $c \in C^\infty(\mathcal{U}_{x_0} \times \mathcal{V}_{T_u(x_0)})$. Hence, by taking a local chart around $(x_0, T_u(x_0))$, the same proof as the one of Theorem 1.3 shows that T_u is a local homeomorphism (resp. diffeomorphism) around almost every point. Using as before that $T_u : M \to M$ is invertible a.e., it follows that T_u is a global homeomorphism (resp. diffeomorphism) outside a closed singular set of measure zero. We leave the details to the interested reader. □

4. $C^{1,\beta}$ regularity and strict *c*-convexity

In this and the next section we prove that, if in some open set a c-convex function u is sufficiently close to a parabola and the cost function is close to the linear one, then u is smooth in some smaller set.

The idea of the proof (which is reminiscent of the argument introduced by Caffarelli in [6] to show $W^{2,p}$ and $C^{2,\alpha}$ estimates for the classical Monge-Ampère equation, though several additional complications arise in our case) is the following: since the cost function is close to the linear one and both densities are almost constant, u is close to a convex function v solving an optimal transport problem with linear cost and constant densities (Lemma 4.1). In addition, since u is close to a parabola, so is v. Hence, by [18] and Caffarelli's regularity theory, v is smooth, and we can use this information to deduce that u is even closer to a second parabola (given by the second order Taylor expansion of v at the origin) inside a small neighborhood around of origin. By rescaling back this neighborhood at scale 1 and iterating this construction, we obtain that u is $C^{1,\beta}$ at the origin for some $\beta \in (0, 1)$. Since this argument can be applied at every point in a neighborhood of the origin, we deduce that u is $C^{1,\beta}$ there, see Theorem 4.3. (A similar strategy has also been used in [7] to show regularity optimal transport maps for the cost $|x - y|^p$, either when p is close to 2 or when X and Y are sufficiently far from each other.)

Once this result is proved, we know that $\partial^- u$ is a singleton at every point, so it follows from (**2.6**) that

$$\partial_c u(x) = \text{c-exp}_x\big(\partial^- u(x)\big),$$

see Remark 4.4 below. (The above identity is exactly what in general may fail for general c-convex functions, unless the MTW condition holds [31].) Thanks to this fact, we obtain that u enjoys a comparison principle (Proposition 5.2), and this allows us to use a second approximation argument with solutions of the classical Monge-Ampère equation (in the spirit of [6, 27]) to conclude that u is $C^{2,\sigma'}$ in a smaller neighborhood, for some $\sigma' > 0$. Then higher regularity follows from standard elliptic estimates, see Theorem 5.3.

Lemma **4.1.** — *Let $\mathcal{C}_1$ and $\mathcal{C}_2$ be two closed sets such that*

$$B_{1/K} \subset \mathcal{C}_1, \qquad \mathcal{C}_2 \subset B_K \tag{4.1}$$

for some $K \geq 1$, f *and* g *two densities supported respectively in* $\mathcal{C}_1$ *and* $\mathcal{C}_2$, *and* $u : \mathcal{C}_1 \to \mathbf{R}$ *a c-convex function such that* $\partial_c u(\mathcal{C}_1) \subset B_K$ *and* $(T_u)_\sharp f = g$. *Let* $\rho > 0$ *be such that* $|\mathcal{C}_1| = |\rho\, \mathcal{C}_2|$ *(where* $\rho\, \mathcal{C}_2$ *denotes the dilation of* $\mathcal{C}_2$ *with respect to the origin), and let* v *be a convex function such that* $\nabla v_\sharp \mathbf{1}_{\mathcal{C}_1} = \mathbf{1}_{\rho \mathcal{C}_2}$ *and* $v(\mathbf{0}) = u(\mathbf{0})$. *Then there exists an increasing function* $\omega : \mathbf{R}^+ \to \mathbf{R}^+$, *depending only* K, *and satisfying* $\omega(\delta) \geq \delta$ *and* $\omega(0^+) = 0$, *such that, if*

$$\|f - \mathbf{1}_{\mathcal{C}_1}\|_\infty + \|g - \mathbf{1}_{\mathcal{C}_2}\|_\infty \leq \delta \tag{4.2}$$

and

$$\left\| c(x,y) + x \cdot y \right\|_{C^2(B_K \times B_K)} \leq \delta, \tag{4.3}$$

then

$$\|u - v\|_{C^0(B_{1/K})} \leq \omega(\delta).$$

Proof. — Assume the lemma is false. Then there exists $\varepsilon_0 > 0$, a sequence of closed sets $\mathcal{C}_1^h$, $\mathcal{C}_2^h$ satisfying (**4.1**), functions f_h, g_h satisfying (**4.2**) with $\delta = 1/h$, and costs c_h converging in C^2 to $-x \cdot y$, such that

$$u_h(\mathbf{0}) = v_h(\mathbf{0}) = 0 \quad \text{and} \quad \|u_h - v_h\|_{C^0(B_{1/K})} \geq \varepsilon_0,$$

where u_h and v_h are as in the statement. First, we extend u_h an v_h to B_K as

$$u_h(x) := \sup_{z \in \mathcal{C}_1^h,\, y \in \partial_{c_h} u_h(z)} \left\{ u_h(z) - c_h(x,y) + c_h(z,y) \right\},$$

$$v_h(x) := \sup_{z \in \mathcal{C}_1^h,\, p \in \partial^- v_h(z)} \left\{ v_h(z) + p \cdot (x - z) \right\}.$$

Notice that, since by assumption $\partial_{c_h} u_h(\mathcal{C}_1^h) \subset B_K$, we have $\partial_{c_h} u_h(B_K) \subset B_K$. Also, $(T_{u_h})_\sharp f_h = g_h$ gives that $\int f_h = \int g_h$, so it follows from (**4.2**) that

$$\rho_h = \left(|\mathcal{C}_1^h| / |\mathcal{C}_2^h| \right)^{1/n} \to 1 \quad \text{as } h \to \infty,$$

which implies that $\partial^- v_h(B_K) \subset B_{\rho_h K} \subset B_{2K}$ for h large. Thus, since the C^1-norm of c_h is uniformly bounded, we deduce that both u_h and v_h are uniformly Lipschitz. Recalling that $u_h(\mathbf{0}) = v_h(\mathbf{0}) = 0$, we get that, up to a subsequence, u_h and v_h uniformly converge inside B_K to u_∞ and v_∞ respectively, where

$$u_\infty(\mathbf{0}) = v_\infty(\mathbf{0}) = 0 \quad \text{and} \quad \|u_\infty - v_\infty\|_{C^0(B_{1/K})} \geq \varepsilon_0. \tag{4.4}$$

In addition f_h (resp. g_h) weak-$*$ converge in L^∞ to some density f_∞ (resp. g_∞) supported in $\overline{B}_K$. Also, since $\rho_h \to 1$, using (**4.2**) we get that $\mathbf{1}_{\mathcal{C}_1^h}$ (resp. $\mathbf{1}_{\rho_h \mathcal{C}_2^h}$) weak-$*$ converges in L^∞

to f_∞ (resp. g_∞). Finally we remark that, because of (**4.2**) and the fact that $\mathcal{C}_1^h \supset B_{1/K}$, we also have

$$f_\infty \geq \mathbf{1}_{B_{1/K}}.$$

In order to get a contradiction we have to show that $u_\infty = v_\infty$ in $B_{1/K}$. To see this, we apply [37, Theorem 5.20] to deduce that both ∇u_∞ and ∇v_∞ are optimal transport maps for the linear cost $-x \cdot y$ sending f_∞ onto g_∞. By uniqueness of the optimal map (see Theorem 2.2) we deduce that $\nabla v_\infty = \nabla u_\infty$ almost everywhere inside $B_{1/K} \subset \operatorname{spt} f_\infty$, hence $u_\infty = v_\infty$ in $B_{1/K}$ (since $u_\infty(\mathbf{0}) = v_\infty(\mathbf{0}) = 0$), contradicting (**4.4**). □

Here and in the sequel, we use $\mathcal{N}_r(E)$ to denote the r-neighborhood of a set E.

Lemma **4.2.** — *Let u and v be, respectively, c-convex and convex, let $D \in \mathbf{R}^{n\times n}$ be a symmetric matrix satisfying*

$$\mathrm{Id}/K \leq D \leq K\,\mathrm{Id} \tag{4.5}$$

for some $K \geq 1$, and define the ellipsoid

$$E(x_0, h) := \big\{x:\ D(x - x_0)\cdot(x - x_0) \leq h\big\}, \qquad h > 0.$$

Assume that there exist small positive constants ε, δ such that

$$\|v - u\|_{C^0(E(x_0,h))} \leq \varepsilon, \qquad \|c + x\cdot y\|_{C^2(E(x_0,h)\times\partial_c u(E(x_0,h)))} \leq \delta. \tag{4.6}$$

Then

$$\partial_c u\big(E(x_0, h - \sqrt{\varepsilon})\big) \subset \mathcal{N}_{K'(\delta+\sqrt{h\varepsilon})}\big(\partial v\big(E(x_0, h)\big)\big) \quad \forall 0 < \varepsilon < h^2 \leq 1, \tag{4.7}$$

where K' depends only on K.

Proof. — Up to a change of coordinates we can assume that $x_0 = \mathbf{0}$, and to simplify notation we set $E_h := E(x_0, h)$. Let us define

$$\bar v(x) := v(x) + \varepsilon + 2\sqrt{\varepsilon}(Dx\cdot x - h),$$

so that $\bar v \geq u$ outside E_h, and $\bar v \leq u$ inside $E_{h-\sqrt{\varepsilon}}$. Then, taking a c-support to u in $E_{h-\sqrt{\varepsilon}}$ (i.e., a function $C_{x,y}$ as in (**2.3**), with $x \in E_{h-\sqrt{\varepsilon}}$ and $y \in \partial_c u(x)$), moving it down and then lifting it up until it touches $\bar v$ from below, we see that it has to touch the graph of $\bar v$ at some point $\bar x \in E_h$: in other words[3]

$$\partial_c u(E_{h-\sqrt{\varepsilon}}) \subset \partial_c \bar v(E_h).$$

[3] Even if $\bar v$ is not c-convex, it still makes sense to consider his c-subdifferential (notice that the c-subdifferential of $\bar v$ may be empty at some points). In particular, the inclusion $\partial_c \bar v(x) \subset \text{c-exp}_x(\partial^- \bar v(x))$ still holds.

By (**4.5**) we see that $\operatorname{diam} E_h \le 2\sqrt{Kh}$, so by a simple computation (using again (**4.5**)) we get

$$\partial^- \bar{v}(E_h) \subset \mathcal{N}_{4K\sqrt{Kh\varepsilon}}\big(\partial^- v(E_h)\big).$$

Thus, since $\partial_c \bar{v}(E_h) \subset \text{c-exp}(\partial^- \bar{v}(E_h))$ (by (**2.6**)) and $\|\text{c-exp} - \mathrm{Id}\|_{C^0} \le \delta$ (by (**4.6**)), we easily deduce that

$$\partial_c u(E_{h-\sqrt{\varepsilon}}) \subset \mathcal{N}_{K'(\delta+\sqrt{h\varepsilon})}\big(\partial^- v(E_h)\big),$$

proving (**4.7**). □

Theorem **4.3.** — *Let $\mathcal{C}_1$ and $\mathcal{C}_2$ be two closed sets satisfying*

$$B_{1/3} \subset \mathcal{C}_1, \qquad \mathcal{C}_2 \subset B_3,$$

let f, g be two densities supported in $\mathcal{C}_1$ and $\mathcal{C}_2$ respectively, and let $u : \mathcal{C}_1 \to \mathbf{R}$ be a c-convex function such that $\partial_c u(\mathcal{C}_1) \subset B_3$ and $(T_u)_\sharp f = g$. Then, for every $\beta \in (0,1)$ there exist constants $\delta_0, \eta_0 > 0$ such that the following holds: if

$$\|f - \mathbf{1}_{\mathcal{C}_1}\|_\infty + \|g - \mathbf{1}_{\mathcal{C}_2}\|_\infty \le \delta_0, \tag{4.8}$$

$$\big\|c(x,y) + x\cdot y\big\|_{C^2(B_3\times B_3)} \le \delta_0, \tag{4.9}$$

and

$$\left\|u - \frac{1}{2}|x|^2\right\|_{C^0(B_3)} \le \eta_0, \tag{4.10}$$

then $u \in C^{1,\beta}(B_{1/7})$.

Proof. — We divide the proof into several steps.

- *Step 1: u is close to a strictly convex solution of the Monge Ampère equation.*

Let $v : \mathbf{R}^n \to \mathbf{R}$ be a convex function such that $\nabla v_\sharp \mathbf{1}_{\mathcal{C}_1} = \mathbf{1}_{\rho\mathcal{C}_2}$ with $\rho = (|\mathcal{C}_1|/|\mathcal{C}_2|)^{1/n}$ (see Theorem 2.2). Up to adding a constant to v, without loss of generality we can assume that $v(\mathbf{0}) = u(\mathbf{0})$. Hence, we can apply Lemma 4.1 to obtain

$$\|v - u\|_{C^0(B_{1/3})} \le \omega(\delta_0) \tag{4.11}$$

for some (universal) modulus of continuity $\omega : \mathbf{R}^+ \to \mathbf{R}^+$, which combined with (**4.10**) gives

$$\left\|v - \frac{1}{2}|x|^2\right\|_{C^0(B_{1/3})} \le \eta_0 + \omega(\delta_0).$$

Also, since $\int_{\mathcal{C}_1} f = \int_{\mathcal{C}_2} g$, it follows easily from (**4.8**) that $|\rho - 1| \le 3\delta_0$. By these two facts we get that $\partial^- v(\mathrm{B}_{1/4}) \subset \mathrm{B}_{7/24} \subset \rho\mathcal{C}_2$ provided δ_0 and η_0 are small enough (recall that v is convex and that $\mathrm{B}_{1/3} \subset \mathcal{C}_2$), so we can apply [18, Proposition 3.4] to deduce that v is a strictly convex Alexandrov solution to the Monge-Ampère equation

$$\det \mathrm{D}^2 v = 1 \quad \text{in } \mathrm{B}_{1/4}. \tag{4.12}$$

In addition, by a simple compactness argument, we see that the modulus of strict convexity of v inside $\mathrm{B}_{1/4}$ is universal. So, by classical Pogorelov and Schauder estimates, we obtain the existence of a universal constant $\mathrm{K}_0 \ge 1$ such that

$$\|v\|_{\mathrm{C}^3(\mathrm{B}_{1/5})} \le \mathrm{K}_0, \qquad \mathrm{Id}/\mathrm{K}_0 \le \mathrm{D}^2 v \le \mathrm{K}_0\,\mathrm{Id} \quad \text{in } \mathrm{B}_{1/5}. \tag{4.13}$$

In particular, there exists a universal value $\bar{h} > 0$ such that, for all $x \in \mathrm{B}_{1/7}$,

$$\mathrm{Q}(x, v, h) := \left\{z:\ v(z) \le v(x) + \nabla v(x)\cdot(z - x) + h\right\} \Subset \mathrm{B}_{1/6} \quad \forall h \le \bar{h}.$$

- *Step 2: Sections of u are close to sections of v.*

Given $x \in \mathrm{B}_{1/7}$ and $y \in \partial_c u(x)$, we define

$$\mathrm{S}(x, y, u, h) := \left\{z:\ u(z) \le u(x) - c(z, y) + c(x, y) + h\right\}.$$

We claim that, if δ_0 is small enough, then for all $x \in \mathrm{B}_{1/7}$, $y \in \partial_c u(x)$, and $h \le \bar{h}/2$, it holds

$$\mathrm{Q}\big(x, v, h - \mathrm{K}_1\sqrt{\omega(\delta_0)}\big) \subset \mathrm{S}(x, y, u, h) \subset \mathrm{Q}\big(x, v, h + \mathrm{K}_1\sqrt{\omega(\delta_0)}\big), \tag{4.14}$$

where $\mathrm{K}_1 > 0$ is a universal constant.

Let us show the first inclusion. For this, take $x \in \mathrm{B}_{1/7}$, $y \in \partial_c u(x)$, and define

$$p_x := -\mathrm{D}_x c(x, y) \in \partial^- u(x).$$

Since v has universal C^2 bounds (see (**4.13**)) and u is semi-convex (with a universal bound), a simple interpolation argument gives

$$\big|p_x - \nabla v(x)\big| \le \mathrm{K}'\sqrt{\|u - v\|_{\mathrm{C}^0(\mathrm{B}_{1/5})}} \le \mathrm{K}'\sqrt{\omega(\delta_0)} \quad \forall x \in \mathrm{B}_{1/7}. \tag{4.15}$$

In addition, by (**4.9**),

$$|y - p_x| \le \|\mathrm{D}_x c + \mathrm{Id}\|_{\mathrm{C}^0(\mathrm{B}_3\times\mathrm{B}_3)} \le \delta_0, \tag{4.16}$$

hence

$$\big|z\cdot p_x + c(z, y)\big| \le |z\cdot p_x - z\cdot y| + \big|z\cdot y + c(z, y)\big| \le 2\delta_0 \quad \forall x, z \in \mathrm{B}_{1/7}. \tag{4.17}$$

Thus, if $z \in Q(x, v, h - K_1\sqrt{\omega(\delta_0)})$, by (**4.11**), (**4.15**), and (**4.17**) we get

$$\begin{aligned} u(z) &\le v(z) + \omega(\delta_0) \le v(x) + \nabla v(x)\cdot(z-x) + h - K_1\sqrt{\omega(\delta_0)} + \omega(\delta_0) \\ &\le u(x) + p_x \cdot z - p_x \cdot x + h - K_1\sqrt{\omega(\delta_0)} + 2\omega(\delta_0) + 2K'\sqrt{\omega(\delta_0)} \\ &\le u(x) - c(z,y) + c(x,y) + h - K_1\sqrt{\omega(\delta_0)} + 2\omega(\delta_0) \\ &\quad + 2K'\sqrt{\omega(\delta_0)} + 4\delta_0 \\ &\le u(x) - c(z,y) + c(x,y) + h, \end{aligned}$$

provided $K_1 > 0$ is sufficiently large. This proves the first inclusion, and the second is analogous.

- *Step 3: Both the sections of u and their images are close to ellipsoids with controlled eccentricity, and u is close to a smooth function near x_0.*

We claim that there exists a universal constant $K_2 \ge 1$ such that the following holds: For every $\eta_0 > 0$ small, there exist small positive constants $h_0 = h_0(\eta_0)$ and $\delta_0 = \delta_0(h_0, \eta_0)$ such that, for all $x_0 \in B_{1/7}$, there is a symmetric matrix A satisfying

$$\mathrm{Id}/K_2 \le A \le K_2\,\mathrm{Id}, \quad \det(A) = 1, \tag{4.18}$$

and such that, for all $y_0 \in \partial_c u(x_0)$,

$$\begin{aligned} &A\big(B_{\sqrt{h_0/8}}(x_0)\big) \subset S(x_0, y_0, u, h_0) \subset A\big(B_{\sqrt{8h_0}}(x_0)\big), \\ &A^{-1}\big(B_{\sqrt{h_0/8}}(y_0)\big) \subset \partial_c u\big(S(x_0, y_0, u, h_0)\big) \subset A^{-1}\big(B_{\sqrt{8h_0}}(y_0)\big). \end{aligned} \tag{4.19}$$

Moreover

$$\left\| u - C_{x_0,y_0} - \frac{1}{2}\left|A^{-1}(x-x_0)\right|^2 \right\|_{C^0(A(B_{\sqrt{8h_0}}(x_0)))} \le \eta_0 h_0, \tag{4.20}$$

where $C_{x_0 y_0}$ is a c-support function for u at x_0, see (**2.3**).

In order to prove the claim, take $h_0 \ll \bar{h}$ small (to be fixed) and $\delta_0 \ll h_0$ such that $K_1\sqrt{\omega(\delta_0)} \le h_0/2$, where K_1 is as in Step 2, so that

$$Q(x_0, v, h_0/2) \subset S(x_0, y_0, u, h_0) \subset Q(x_0, v, 3h_0/2) \Subset B_{1/6}. \tag{4.21}$$

By (**4.13**) and Taylor formula we get

$$v(x) = v(x_0) + \nabla v(x_0)\cdot(x-x_0) + \frac{1}{2}D^2 v(x_0)(x-x_0)\cdot(x-x_0) + O\big(|x-x_0|^3\big), \tag{4.22}$$

so that defining

$$E(x_0, h_0) := \left\{ x : \ \frac{1}{2}D^2 v(x_0)(x-x_0)\cdot(x-x_0) \le h_0 \right\} \tag{4.23}$$

and using (**4.13**), we deduce that, for every h_0 universally small,

$$E(x_0, h_0/2) \subset Q(x_0, v, h_0) \subset E(x_0, 2h_0). \tag{4.24}$$

Moreover, always for h_0 small, thanks to (**4.22**) and the uniform convexity of v

$$\nabla v\big(E(x_0, h_0)\big) \subset E^*\big(\nabla v(x_0), 2h_0\big) \subset \nabla v\big(E(x_0, 3h_0)\big) \tag{4.25}$$

where we have set

$$E^*(\bar y, h_0) := \left\{ y : \frac{1}{2}\big[D^2 v(\bar y)\big]^{-1}(y-\bar y)\cdot(y-\bar y) \le h_0 \right\}.$$

By Lemma 4.2, (**4.24**), and (**4.25**) applied with $3h_0$ in place of h_0, we deduce that for $\delta_0 \ll h_0 \ll \bar h$

$$\partial_c u\big(S(x_0, y_0, u, h_0)\big) \subset \mathcal{N}_{K''\sqrt{\omega(\delta_0)}}\Big(\nabla v\big(E(x_0, 3h_0)\big)\Big) \subset E^*\big(\nabla v(x_0), 7h_0\big). \tag{4.26}$$

Moreover, by (**4.15**), if $y_0 \in \partial_c u(x_0)$ and we set $p_{x_0} := -D_x c(x_0, y_0)$, then

$$\big|y_0 - \nabla v(x_0)\big| \le \big|p_{x_0} - \nabla v(x_0)\big| + \|D_x c + \mathrm{Id}\|_{C^0(B_3\times B_3)} \le K'\sqrt{\omega(\delta_0)} + \delta_0.$$

Thus, choosing δ_0 sufficiently small, it holds

$$E^*\big(\nabla v(x_0), 7h_0\big) \subset E^*(y_0, 8h_0) \quad \forall y_0 \in \partial_c u(x_0). \tag{4.27}$$

We now want to show that

$$E^*(y_0, h_0/8) \subset \partial_c u\big(S(x_0, y_0, u, h_0)\big) \quad \forall y_0 \in \partial_c u(x_0).$$

Observe that, arguing as above, we get

$$E^*(y_0, h_0/8) \subset E^*\big(\nabla v(x_0), h_0/7\big) \quad \forall y_0 \in \partial_c u(x_0) \tag{4.28}$$

provided δ_0 is small enough, so it is enough to prove that

$$E^*\big(\nabla v(x_0), h_0/7\big) \subset \partial_c u\big(S(x_0, y_0, u, h_0)\big).$$

For this, let us define the c^*-convex function $u^c : B_3 \to \mathbf{R}$ and the convex function $v^* : B_3 \to \mathbf{R}$ as

$$u^c(y) := \sup_{x\in B_{1/5}} \big\{-c(x,y) - u(x)\big\}, \qquad v^*(y) := \sup_{x\in B_{1/5}} \big\{x\cdot y - v(x)\big\}$$

(see (**3.1**)). Then it is immediate to check that

$$\big|u^c - v^*\big| \le \omega(\delta_0) + \delta_0 \le 2\omega(\delta_0) \quad \text{on } B_3. \tag{4.29}$$

Also, in view of (**4.13**), v^* is a uniformly convex function of class C^3 on the open set $\nabla v(B_{1/5})$. In addition, since

$$F \subset \partial_c u\big(\partial_{c^*} u^c(F)\big) \quad \text{for any set } F, \tag{4.30}$$

thanks to (**4.21**) and (**4.24**) it is enough to show

$$\partial_{c^*} u^c\big(E^*\big(\nabla v(x_0), h_0/7\big)\big) \subset E(x_0, h_0/4). \tag{4.31}$$

For this, we apply Lemma 4.2 to u^c and v^* to infer

$$\begin{aligned}\partial_{c^*} u^c\big(E^*\big(\nabla v(x_0), h_0/7\big)\big) &\subset \mathcal{N}_{K'''\sqrt{\omega(\delta)}}\big(\nabla v^*\big(E^*\big(\nabla v(x_0), h_0/7\big)\big)\big)\\ &\subset E(x_0, h_0/4),\end{aligned}$$

where we used that

$$\nabla v^* = [\nabla v]^{-1} \quad \text{and} \quad D^2 v^*\big(\nabla v(x_0)\big) = \big[D^2 v(x_0)\big]^{-1}.$$

Thus, recalling (**4.26**), we have proved that there exist h_0 universally small, and δ_0 small depending on h_0, such that

$$E^*\big(\nabla v(x_0), h_0/7\big) \subset \partial_c u\big(S(x_0, y_0, u, h_0)\big) \subset E^*\big(\nabla v(x_0), 7h_0\big) \quad \forall x_0 \in B_{1/7}. \tag{4.32}$$

Using (**4.21**), (**4.24**), (**4.27**), and (**4.28**), this proves (**4.19**) with $A := [D^2 v(x_0)]^{-1/2}$. Also, thanks to (**4.12**) and (**4.13**), (**4.18**) holds.

In order to prove the second part of the claim, we exploit (**4.11**), (**4.9**), (**4.16**), (**4.15**), (**4.22**), and (**4.18**) (recall that C_{x_0,y_0} is defined in (**2.3**) and that $A = [D^2 v(x_0)]^{-1/2}$):

$$\begin{aligned}&\left\| u - C_{x_0,y_0} - \frac{1}{2}\big|A^{-1}(x - x_0)\big|^2 \right\|_{C^0(E(x_0,8h_0))}\\
&\quad = \left\| u - C_{x_0,y_0} - \frac{1}{2} D^2 v(x_0)(x - x_0)\cdot(x - x_0)\right\|_{C^0(E(x_0,8h_0))}\\
&\quad \le 2\|u - v\|_{C^0(E(x_0,8h_0))} + \big\|c(x, y_0) + x\cdot y_0\big\|_{C^0(E(x_0,8h_0))}\\
&\qquad + \big\|c(x_0, y_0) + x_0\cdot y_0\big\|_{C^0(E(x_0,8h_0))} + \big\|(y_0 - p_{x_0})\cdot(x - x_0)\big\|_{C^0(E(x_0,8h_0))}\\
&\qquad + \big\|\big(p_{x_0} - \nabla v(x_0)\big)\cdot(x - x_0)\big\|_{C^0(E(x_0,8h_0))}\\
&\qquad + \Big\| v - v(x_0) - \nabla v(x_0)\cdot(x - x_0)\\
&\qquad - \frac{1}{2} D^2 v(x_0)(x - x_0)\cdot(x - x_0)\Big\|_{C^0(E(x_0,8h_0))}\\
&\quad \le 2\omega(\delta_0) + 3\delta_0 + K'\sqrt{\omega(\delta_0)} + K\big(K_2\sqrt{8h_0}\big)^3 \le \eta_0 h_0,\end{aligned}$$

where the last inequality follows by choosing first h_0 sufficiently small, and then δ_0 much smaller than h_0.

- *Step 4: A first change of variables.*

Fix $x_0 \in B_{1/7}, y_0 \in \partial_c u(x_0)$, define $M := -D_{xy}c(x_0, y_0)$, and consider the change of variables

$$\begin{cases} \bar{x} := x - x_0 \\ \bar{y} := M^{-1}(y - y_0). \end{cases}$$

Notice that, by (**4.9**), it follows that

$$|M - \mathrm{Id}| + \left|M^{-1} - \mathrm{Id}\right| \leq 3\delta_0 \tag{4.33}$$

for δ_0 sufficiently small. We also define

$$\begin{aligned} \bar{c}(\bar{x}, \bar{y}) &:= c(x, y) - c(x, y_0) - c(x_0, y) + c(x_0, y_0), \\ \bar{u}(\bar{x}) &:= u(x) - u(x_0) + c(x, y_0) - c(x_0, y_0), \\ \bar{u}^{\bar{c}}(\bar{y}) &:= u^c(y) - u^c(y_0) + c(x_0, y) - c(x_0, y_0). \end{aligned}$$

Then $\bar{u}$ is $\bar{c}$-convex, $\bar{u}^{\bar{c}}$ is $\bar{c}^*$-convex (where $\bar{c}^*(\bar{y}, \bar{x}) = \bar{c}(\bar{x}, \bar{y})$), and

$$\bar{c}(\cdot, \mathbf{0}) = \bar{c}(\mathbf{0}, \cdot) \equiv 0, \qquad D_{\bar{x}\bar{y}}\bar{c}(\mathbf{0}, \mathbf{0}) = -\mathrm{Id}\,. \tag{4.34}$$

We also notice that

$$\partial_{\bar{c}}\bar{u}(\bar{x}) = M^{-1}\big(\partial_c u(\bar{x} + x_0) - y_0\big). \tag{4.35}$$

Thus, recalling (**4.19**), and using (**4.33**) and (**4.35**), for δ_0 sufficiently small we obtain

$$\begin{aligned} &A(B_{\sqrt{h_0/9}}) \subset S(\mathbf{0}, \mathbf{0}, \bar{u}, h_0) \subset A(B_{\sqrt{9h_0}}), \\ &A^{-1}(B_{\sqrt{h_0/9}}) \subset M^{-1}A^{-1}(B_{\sqrt{h_0/8}}) \subset \partial_{\bar{c}}\bar{u}\big(S(\mathbf{0}, \mathbf{0}, \bar{u}, h_0)\big) \subset M^{-1}A^{-1}(B_{\sqrt{8h_0}}) \\ &\qquad \subset A^{-1}(B_{\sqrt{9h_0}}). \end{aligned} \tag{4.36}$$

Since $(T_u)_\sharp f = g$, it follows that $T_{\bar{u}} = \bar{c}\text{-exp}(\nabla\bar{u})$ satisfies

$$(T_{\bar{u}})_\sharp \bar{f} = \bar{g}, \quad \text{with } \bar{f}(\bar{x}) := f(\bar{x} + x_0),\ \bar{g}(\bar{y}) := \det(M)\, g(M\bar{y} + y_0)$$

(see for instance the footnote in the proof of Theorem 1.3). Notice that, since $|M - \mathrm{Id}| \leq \delta_0$ (by (**4.9**)), we have $|\det(M) - 1| \leq (1 + 2n)\delta_0$ (for δ_0 small), so by (**4.8**) we get

$$\|\bar{f} - \mathbf{1}_{\mathcal{C}_1 - x_0}\|_\infty + \|\bar{g} - \mathbf{1}_{M^{-1}(\mathcal{C}_2 - y_0)}\|_\infty \leq 2(1 + n)\delta_0. \tag{4.37}$$

- *Step 5: A second change of variables and the iteration argument.*

We now perform a second change of variable: we set

$$\begin{cases} \tilde{x} := \frac{1}{\sqrt{h_0}} \mathrm{A}^{-1} \bar{x}, \\ \tilde{y} := \frac{1}{\sqrt{h_0}} \mathrm{A} \bar{y}, \end{cases} \tag{4.38}$$

and define

$$c_1(\tilde{x}, \tilde{y}) := \frac{1}{h_0} \bar{c}\big(\sqrt{h_0}\mathrm{A}\tilde{x}, \sqrt{h_0}\mathrm{A}^{-1}\tilde{y}\big),$$

$$u_1(\tilde{x}) := \frac{1}{h_0} \bar{u}\big(\sqrt{h_0}\mathrm{A}\tilde{x}\big),$$

$$u_1^{c_1}(\tilde{y}) := \frac{1}{h_0} \bar{u}^{\bar{c}}\big(\sqrt{h_0}\mathrm{A}^{-1}\tilde{y}\big).$$

We also define

$$f_1(\tilde{x}) := \bar{f}\big(\sqrt{h_0}\mathrm{A}\tilde{x}\big), \qquad g_1(\tilde{y}) := \bar{g}\big(\sqrt{h_0}\mathrm{A}^{-1}\tilde{y}\big).$$

Since $\det(\mathrm{A}) = 1$ (see (**4.18**)), it is easy to check that $(\mathrm{T}_{u_1})_\sharp f_1 = g_1$ (see the footnote in the proof of Theorem 1.3). Also, since $(\|\mathrm{A}\| + \|\mathrm{A}^{-1}\|)\sqrt{h_0} \ll 1$, it follows from (**4.37**) that

$$|f_1 - 1| + |g_1 - 1| \le 2(1+n)\delta_0 \quad \text{inside } \mathrm{B}_3. \tag{4.39}$$

Moreover, defining

$$\mathcal{C}_1^{(1)} := \mathrm{S}(\mathbf{0}, \mathbf{0}, u_1, 1), \qquad \mathcal{C}_2^{(1)} := \partial_{c_1} u_1\big(\mathrm{S}(\mathbf{0}, \mathbf{0}, u_1, 1)\big),$$

both $\mathcal{C}_1^{(1)}$ and $\mathcal{C}_2^{(1)}$ are closed, and thanks to (**4.36**)

$$\mathrm{B}_{1/3} \subset \mathcal{C}_1^{(1)}, \qquad \mathcal{C}_2^{(1)} \subset \mathrm{B}_3. \tag{4.40}$$

Also, since $(\mathrm{T}_{u_1})_\sharp f_1 = g_1$, arguing as in the proof of Theorem 1.3 we get

$$(\mathrm{T}_{u_1})_\sharp (f_1 \mathbf{1}_{\mathcal{C}_1^{(1)}}) = (g_1 \mathbf{1}_{\mathcal{C}_2^{(1)}}),$$

and by (**4.39**)

$$\|f_1 \mathbf{1}_{\mathcal{C}_1^{(1)}} - \mathbf{1}_{\mathcal{C}_1^{(1)}}\|_\infty + \|g_1 \mathbf{1}_{\mathcal{C}_2^{(1)}} - \mathbf{1}_{\mathcal{C}_2^{(1)}}\|_\infty \le 2(1+n)\delta_0.$$

Finally, by (**4.34**) and (**4.20**), it is easy to check that

$$\big\|c_1(\tilde{x}, \tilde{y}) + \tilde{x} \cdot \tilde{y}\big\|_{\mathrm{C}^2(\mathrm{B}_3 \times \mathrm{B}_3)} \le \delta_0, \qquad \left\| u_1 - \frac{1}{2}|\tilde{x}|^2 \right\|_{\mathrm{C}^0(\mathrm{B}_3)} \le \eta_0.$$

This shows that u_1 satisfies the same assumptions as u with δ_0 replaced by $2(1+n)\delta_0$. Hence, up to take δ_0 slightly smaller, we can apply Step 3 to u_1, and we find a symmetric matrix A_1 satisfying

$$
\begin{aligned}
&\mathrm{Id}/K_2 \le A_1 \le K_2\,\mathrm{Id}, \qquad \det(A_1)=1,\\
&A_1(B_{\sqrt{h_0/8}}) \subset S(\mathbf{0},\mathbf{0},u_1,h_0) \subset A_1(B_{\sqrt{8h_0}}),\\
&A_1^{-1}(B_{\sqrt{h_0/8}}) \subset \partial_{c_1}u_1\big(S(\mathbf{0},\mathbf{0},u_1,h_0)\big) \subset A_1^{-1}(B_{\sqrt{8h_0}}),\\
&\left\| u_1 - \frac{1}{2}\big|A_1^{-1}\tilde{x}\big|^2 \right\|_{C^0(A_1(B(0,\sqrt{8h_0})))} \le \eta_0 h_0.
\end{aligned}
$$

(Here K_2 and h_0 are as in Step 3.)

This allows us to apply to u_1 the very same construction as the one used above to define u_1 from $\bar{u}$: we set

$$
c_2(\tilde{x},\tilde{y}) := \frac{1}{h_0}c_1\big(\sqrt{h_0}A_1\tilde{x}, \sqrt{h_0}A_1^{-1}\tilde{y}\big), \qquad u_2(\tilde{x}) := \frac{1}{h_0}u_1\big(\sqrt{h_0}A_1\tilde{x}\big),
$$

so that $(T_{u_2})_\sharp f_2 = g_2$ with

$$
f_2(\tilde{x}) := f_1\big(\sqrt{h_0}A_1\tilde{x}\big), \qquad g_2(\tilde{y}) := \bar{g}\big(\sqrt{h_0}A_1^{-1}\tilde{y}\big).
$$

Arguing as before, it is easy to check that u_2, c_2, f_2, g_2 satisfy the same assumptions as u_1, c_1, f_1, g_1 with exactly the same constants.

So we can keep iterating this construction, defining for any $k \in \mathbf{N}$

$$
c_{k+1}(\tilde{x},\tilde{y}) := \frac{1}{h_0}c_k\big(\sqrt{h_0}A_k\tilde{x}, \sqrt{h_0}A_k^{-1}\tilde{y}\big), \qquad u_{k+1}(\tilde{x}) := \frac{1}{h_0}u_k\big(\sqrt{h_0}A_k\tilde{x}\big),
$$

where A_k is the matrix constructed in the k-th iteration. In this way, if we set

$$
M_k := A_k \cdot \ldots \cdot A_1, \quad \forall k \ge 1,
$$

we obtain a sequence of symmetric matrices satisfying

$$
\mathrm{Id}/K_2^k \le M_k \le K_2^k\,\mathrm{Id}, \quad \det(M_k)=1, \tag{4.41}
$$

and such that

$$
M_k(B_{(h_0/8)^{k/2}}) \subset S\big(\mathbf{0},\mathbf{0},u_k,h_0^k\big) \subset M_k(B_{(8h_0)^{k/2}}). \tag{4.42}
$$

- *Step 6:* $C^{1,\beta}$ *regularity.*

We now show that, for any $\beta \in (0,1)$, we can choose h_0 and $\delta_0 = \delta_0(h_0)$ small enough so that u_1 is $C^{1,\beta}$ at the origin (here u_1 is the function constructed in the previous step).

This will imply that u is $C^{1,\beta}$ at x_0 with universal bounds, which by the arbitrariness of $x_0 \in B_{1/7}$ gives $u \in C^{1,\beta}(B_{1/7})$.

Fix $\beta \in (0,1)$. Then by (**4.41**) and (**4.42**) we get

$$B_{(\sqrt{h_0}/(\sqrt{8}K_2))^k} \subset S\big(\mathbf{0},\mathbf{0},u_1,h_0^k\big) \subset B_{(K_2\sqrt{8h_0})^k}, \tag{4.43}$$

so defining $r_0 := \sqrt{h_0}/(\sqrt{8}K_2)$ we obtain

$$\|u_1\|_{C^0(B_{r_0^k})} \le h_0^k = \big(\sqrt{8}K_2 r_0\big)^{2k} \le r_0^{(1+\beta)k},$$

provided h_0 (and so r_0) is sufficiently small. This implies the $C^{1,\beta}$ regularity of u_1 at $\mathbf{0}$, concluding the proof. □

Remark **4.4** *(Local to global principle)*. — If u is differentiable at x and c satisfies (C0)–(C1), then every "local support" at x is also a "global c-support" at x, that is, $\partial_c u(x) = \text{c-exp}_x(\partial^- u(x))$. To see this, just notice that

$$\emptyset \ne \partial_c u(x) \subset \text{c-exp}_x\big(\partial^- u(x)\big) = \big\{\text{c-exp}_x\big(\nabla u(x)\big)\big\}$$

(recall (**2.6**)), so necessarily the two sets have to coincide.

Corollary **4.5.** — *Let u be as in Theorem 4.3. Then u is strictly c-convex in $B_{1/7}$. More precisely, for every $\gamma > 2$ there exist $\eta_0, \delta_0 > 0$ depending only on γ such that, if the hypotheses of Theorem 4.3 are satisfied, then, for all $x_0 \in B_{1/7}$, $y_0 \in \partial_c u(x_0)$, and C_{x_0,y_0} as in (**2.3**), we have*

$$\inf_{\partial B_r(x_0)} \{u - C_{x_0,y_0}\} \ge c_0 r^\gamma \quad \forall r \le \operatorname{dist}(x_0, \partial B_{1/7}), \tag{4.44}$$

with $c_0 > 0$ universal.

Proof. — With the same notation as in the proof of Theorem 4.3, it is enough to show that

$$\inf_{\partial B_r} u_1 \ge r^{1/\beta},$$

where u_1 is the function constructed in Step 5 of the proof of Theorem 4.3. Defining $\varrho_0 := K_2\sqrt{8h_0}$, it follows from (**4.43**) that

$$\inf_{\partial B_{\varrho_0^k}} u_1 \ge h_0^k = \big(\varrho_0/\big(\sqrt{8}K_2\big)\big)^{2k} \ge \varrho_0^{\gamma k},$$

provided h_0 is small enough. □

A simple consequence of the above results is the following:

Corollary **4.6.** — *Let u be as in Theorem 4.3, then $T_u(B_{1/7})$ is open.*

Proof. — Since $u \in C^{1,\beta}(B_{1/7})$ we have that $T_u(B_{1/7}) = \partial_c u(B_{1/7})$ (see Remark 4.4). We claim that it is enough to show that if $y_0 \in \partial_c u(B_{1/7})$, then there exists $\varepsilon = \varepsilon(y_0) > 0$ small such that, for all $|y - y_0| < \varepsilon$, the function $u(\cdot) + c(\cdot, y)$ has a local minimum at some point $\bar{x} \in B_{1/7}$. Indeed, if this is the case, then

$$\nabla u(\bar{x}) = -D_x c(\bar{x}, y),$$

and so $y \in \partial_c u(\bar{x})$ (by Remark 4.4), hence $B_\varepsilon(y_0) \subset T_u(B_{1/7})$.

To prove the above fact, fix $r > 0$ such that $B_r(x_0) \subset B_{1/7}$, and pick $\bar{x}$ a point in $\overline{B}_r(x_0)$ where the function $u(\cdot) + c(\cdot, y)$ attains its minimum, i.e.,

$$\bar{x} \in \operatorname*{argmin}_{\overline{B}_r(x_0)} \big\{ u(x) + c(x, y) \big\}.$$

Since, by (**4.44**),

$$\begin{aligned} \min_{x \in \partial B_r(x_0)} \big\{ u(x) + c(x, y) \big\} &\geq \min_{x \in \partial B_r(x_0)} \big\{ u(x) + c(x, y_0) \big\} - \varepsilon \|c\|_{C^1} \\ &\geq u(x_0) + c(x_0, y_0) + c_0 r^\gamma - \varepsilon \|c\|_{C^1}, \end{aligned}$$

while

$$u(x_0) + c(x_0, y) \leq c(x_0, y_0) + u(x_0) + \varepsilon \|c\|_{C^1},$$

choosing $\varepsilon < \frac{c_0}{2\|c\|_{C^1}} r^\gamma$ we obtain that $\bar{x} \in B_r(x_0) \subset B_{1/7}$. This implies that $\bar{x}$ is a local minimum for $u(\cdot) + c(\cdot, y)$, concluding the proof. □

5. Comparison principle and $C^{2,\alpha}$ regularity

We begin this section with a change of variable formula for the c-exponential map.

Lemma **5.1.** — *Let Ω be an open set, $v \in C^2(\Omega)$, and assume that $\nabla v(\Omega) \subset \mathrm{Dom\,c\text{-}exp}$ and that*

$$D^2 v(x) + D_{xx} c\big(x, \mathrm{c\text{-}exp}_x\big(\nabla v(x)\big)\big) \geq 0 \quad \forall x \in \Omega.$$

Then, for every Borel set $A \subset \Omega$,

$$\big|\mathrm{c\text{-}exp}\big(\nabla v(A)\big)\big| \leq \int_A \frac{\det(D^2 v(x) + D_{xx} c(x, \mathrm{c\text{-}exp}_x(\nabla v(x))))}{|\det(D_{xy} c(x, \mathrm{c\text{-}exp}_x(\nabla v(x))))|}\, dx.$$

In addition, if the map $x \mapsto \mathrm{c\text{-}exp}_x(\nabla v(x))$ is injective, then equality holds.

Proof. — The result follows from a direct application of the Area Formula [12, Section 3.3.2, Theorem 1] once one notices that, differentiating the identity

$$\nabla v(x) = -D_x c\big(x, \text{c-exp}_x\big(\nabla v(x)\big)\big)$$

(see (**2.5**)), the Jacobian determinant of the C^1 map $x \mapsto \text{c-exp}_x(\nabla v(x))$ is given precisely by

$$\frac{\det(D^2 v(x) + D_{xx}c(x, \text{c-exp}_x(\nabla v(x))))}{|\det(D_{xy}c(x, \text{c-exp}_x(\nabla v(x))))|}. \qquad \square$$

In the next proposition we show a comparison principle between C^1 c-convex functions and smooth solutions to the Monge-Ampère equation.[4] As already mentioned at the beginning of Section 4 (see also Remark 4.4), the C^1 regularity of u is crucial to ensure that the c-subdifferential coincides with its local counterpart $\text{c-exp}(\partial^- u)$.

Here and in the sequel, we use co[E] to denote the convex hull of a set E. Also, recall that $\mathcal{N}_r(E)$ denotes the r-neighborhood of E.

Proposition **5.2** *(Comparison principle). — Let u be a c-convex function of class* C^1 *inside the set* $S := \{u < 1\}$, *and assume that* $u(\mathbf{0}) = 0$, $B_{1/K} \subset S \subset B_K$, *and that* $\nabla u(S) \Subset \text{Dom-exp}$. *Let f, g be two densities such that*

$$\|f/\lambda_1 - 1\|_{C^0(S)} + \|g/\lambda_2 - 1\|_{C^0(T_u(S))} \leq \varepsilon \tag{5.1}$$

for some constants $\lambda_1, \lambda_2 \in (1/2, 2)$ *and* $\varepsilon \in (0, 1/4)$, *and assume that* $(T_u)_\sharp f = g$. *Furthermore, suppose that*

$$\|c + x \cdot y\|_{C^2(B_K \times B_K)} \leq \delta. \tag{5.2}$$

Then there exist a universal constant $\gamma \in (0, 1)$, *and* $\delta_1 = \delta_1(K) > 0$ *small, such that the following holds: Let v be the solution of*

$$\begin{cases} \det(D^2 v) = \lambda_1/\lambda_2 & \text{in } \mathcal{N}_{\delta^\gamma}(\text{co}[S]), \\ v = 1 & \text{on } \partial(\mathcal{N}_{\delta^\gamma}(\text{co}[S])). \end{cases}$$

Then

$$\|u - v\|_{C^0(S)} \leq C_K\big(\varepsilon + \delta^{\gamma/n}\big) \quad \textit{provided } \delta \leq \delta_1, \tag{5.3}$$

where C_K *is a constant independent of* λ_1, λ_2, ε, *and* δ *(but which depends on* K*).*

[4] A similar result for the case $c(x,y) = |x-y|^p$ appeared in [7, Theorem 6.2]. Here, however, we have to deal with some additional difficulties due to the fact that the c-exponential map is not necessarily defined on the whole $\mathbf{R}^n$.

Proof. — First of all we observe that, since $u(\mathbf{0})=0$, $u=1$ on ∂S, $S\subset B_K$, and $\|c+x\cdot y\|_{C^2(B_K)}\le\delta\ll 1$, it is easy to check that there exists a universal constant $a_1>0$ such that

$$|D_x c(x,y)|\ge a_1 \quad \forall x\in\partial S,\ y=\text{c-exp}_x\big(\nabla u(x)\big). \tag{5.4}$$

Thanks to (**5.4**) and (**5.2**), it follows from the Implicit Function Theorem that, for each $x\in\partial S$, the boundary of the set

$$E_x:=\Big\{z\in B_K:\ c(z,y)-c(x,y)+u(x)\le 1\Big\}$$

is of class C^2 inside B_K, and its second fundamental form is bounded by $C_K\delta$, where $C_K>0$ depends only on K. Hence, since S can be written as

$$S:=\bigcap_{x\in\partial S}E_x,$$

it follows that

$$S \text{ is a } (C_K\delta)\text{-semiconvex set},$$

that is, for any couple of points $x_0,x_1\in S$ the ball centered at $x_{1/2}:=(x_0+x_1)/2$ of radius $C_K\delta|x_1-x_0|^2$ intersects S. Since $S\subset B_K$, this implies that $\mathrm{co}[S]\subset\mathcal{N}_{C'_K\delta}(S)$ for some positive constant C'_K depending only on K. Thus, for any $\gamma\in(0,1)$ we obtain

$$\mathcal{N}_{\delta^\gamma}\big(\mathrm{co}[S]\big)\subset\mathcal{N}_{(1+C'_K)\delta^\gamma}(S).$$

Since $v=1$ on $\partial(\mathcal{N}_{\delta^\gamma}(\mathrm{co}[S]))$ and $\lambda_1/\lambda_2\in(1/4,4)$, by standard interior estimates for solution of the Monge-Ampère equation with constant right hand side (see for instance [8, Lemma 1.1]), we obtain

$$\operatorname*{osc}_S v\le C''_K \tag{5.5}$$

$$1-C''_K\delta^{\gamma/n}\le v<1 \quad \text{on } \partial S, \tag{5.6}$$

$$D^2v\ge\delta^{\gamma/\tau}\mathrm{Id}/C''_K \quad \text{in } \mathrm{co}[S], \tag{5.7}$$

for some $\tau>0$ universal, and some constant C''_K depending only on K.

Let us define

$$v^+:=\big(1+4\varepsilon+2\sqrt{\delta}\big)v-4\varepsilon-2\sqrt{\delta},$$
$$v^-:=\big(1-4\varepsilon-\sqrt{\delta}/2\big)v+4\varepsilon+\sqrt{\delta}/2+2C''_K\delta^{\gamma/n}.$$

Our goal is to show that we can choose γ universally small so that $v^-\ge u\ge v^+$ on S. Indeed, if we can do so, then by (**5.5**) this will imply (**5.3**), concluding the proof.

First of all notice that, thanks to (**5.6**), $v^- > u > v^+$ on ∂S. Let us show first that $v^+ \le v$.

Assume by contradiction this is not the case. Then, since $u > v^+$ on ∂S,

$$\emptyset \neq Z := \{u < v^+\} \Subset S.$$

Since v^+ is convex, taking any supporting plane to v^+ at $x \in Z$, moving it down and then lifting it up until it touches u from below, we deduce that

$$\nabla v^+(Z) \subset \nabla u(Z) \tag{5.8}$$

(recall that both u and v^+ are of class C^1), thus by Remark 4.4

$$\left|\text{c-exp}\big(\nabla v^+(Z)\big)\right| \le \left|T_u(Z)\right|. \tag{5.9}$$

We show that this is impossible. For this, using (**5.7**) and choosing $\gamma := \tau/4$, for any $x \in Z$ we compute

$$\begin{aligned} &D^2 v^+(x) + D_{xx}c\big(x, \text{c-exp}_x\big(\nabla v^+(x)\big)\big) \\ &\quad \ge \big(1+\sqrt{\delta}+4\varepsilon\big)D^2 v + \sqrt{\delta}D^2 v - \delta\,\mathrm{Id} \\ &\quad \ge \big(1+\sqrt{\delta}+4\varepsilon\big)D^2 v + \big(\delta^{3/4}/C''_K - \delta\big)\,\mathrm{Id} \\ &\quad \ge \big(1+\sqrt{\delta}+4\varepsilon\big)D^2 v, \end{aligned}$$

provided δ is sufficiently small, the smallness depending only on K. Thus, thanks (**5.2**) we have

$$\begin{aligned} \frac{\det(D^2 v^+(x) + D_{xx}c(x, \text{c-exp}_x(\nabla v^+(x))))}{|\det(D_{xy}c(x, \text{c-exp}_x(\nabla v^+(x))))|} &\ge \frac{\det((1+\sqrt{\delta}+4\varepsilon)D^2 v)}{1+\delta} \\ &\ge \big(1+\sqrt{\delta}+4\varepsilon\big)^n(1-2\delta)\frac{\lambda_1}{\lambda_2} \\ &\ge (1+4n\varepsilon)\frac{\lambda_1}{\lambda_2}. \end{aligned}$$

In addition, thanks (**5.7**) and (**5.2**), since $\delta^{\gamma/\tau} = \delta^{1/4} \gg \delta$ we see that

$$D^2 v^+ > \|D_{xx}c\|_{C^0(B_K \times B_K)}\,\mathrm{Id} \quad \text{inside } \mathrm{co}[S].$$

Hence, for any $x, z \in Z$, $x \neq z$ and $y = \text{c-exp}_x(\nabla v^+(x))$ (notice that $\text{c-exp}_x(\nabla v^+(x))$ is well-defined because of (**5.8**) and the assumption $\nabla u(S) \Subset \mathrm{Dom\,c\text{-}exp}$), it follows

$$\begin{aligned} v^+(z) + c(z,y) &\ge v^+(x) + c(x,y) + \frac{1}{2}\int_0^1 \Big(D^2 v^+\big(tz+(1-t)x\big) \\ &\qquad + D_{xx}c\big(tz+(1-t)x, y\big)\Big)[z-x, z-x]\,dt \\ &> v^+(x) + c(x,y), \end{aligned}$$

where we used that $\nabla v^+(x) + D_x c(x,y) = 0$. This means that the supporting function $z \mapsto -c(z,y) + c(x,y) + v^+(x)$ can only touch v^+ from below at x, which implies that the map $Z \ni x \mapsto \text{c-exp}_x(\nabla v^+(x))$ is injective. Thus, by Lemma 5.1 we get

$$\left|\text{c-exp}\big(\nabla v^+(Z)\big)\right| \geq (1 + 4n\varepsilon)\frac{\lambda_1}{\lambda_2}|Z|. \tag{5.10}$$

On the other hand, since u is C^1, it follows from $(T_u)_\sharp f = g$ and (**5.1**) that

$$\left|T_u(Z)\right| = \int_Z \frac{f(x)}{g(T_u(x))}\,dx \leq \frac{\lambda_1(1+\varepsilon)}{\lambda_2(1-\varepsilon)}|Z| \leq (1+3\varepsilon)\frac{\lambda_1}{\lambda_2}|Z|.$$

This estimate combined with (**5.10**) shows that (**5.9**) is impossible unless Z is empty. This proves that $v^+ \leq u$.

The proof of the inequality $v^- \leq u$ follows by the same argument except for a minor modification. More precisely, let us assume by contradiction that $W := \{u > v^-\}$ is nonempty. In order to apply the previous argument we would need to know that $\nabla v^-(W) \subset \text{Dom}\,\text{c-exp}$. However, since the gradient of v can be very large near ∂S, this may be a problem.

To circumvent this issue we argue as follows: since W is nonempty, there exists a positive constant $\bar\mu$ such that u touches $v^- + \bar\mu$ from below inside S. Let E be the contact set, i.e., $E := \{u = v^- + \bar\mu\}$. Since both u and v^- are C^1, $\nabla u = \nabla v^-$ on E. Thus, if $\eta > 0$ is small enough, then the set $W_\eta := \{u > v^- + \bar\mu - \eta\}$ is nonempty and $\nabla v^-(W_\eta)$ is contained in a small neighborhood of $\nabla u(W_\eta)$, which is compactly contained in $\text{Dom}\,\text{c-exp}$. At this point, one argues exactly as in the first part of the proof, with W_η in place of Z, to find a contradiction. □

Theorem **5.3.** — *Let $u, f, g, \eta_0, \delta_0$ be as in Theorem 4.3, and assume in addition that $c \in C^{k,\alpha}(B_3 \times B_3)$ and $f, g \in C^{k,\alpha}(B_{1/3})$ for some $k \geq 0$ and $\alpha \in (0,1)$. There exist small constants $\eta_1 \leq \eta_0$ and $\delta_1 \leq \delta_0$ such that, if*

$$\|f - \mathbf{1}_{\mathcal{C}_1}\|_\infty + \|g - \mathbf{1}_{\mathcal{C}_2}\|_\infty \leq \delta_1, \tag{5.11}$$

$$\left\|c(x,y) + x\cdot y\right\|_{C^2(B_3\times B_3)} \leq \delta_1, \tag{5.12}$$

and

$$\left\|u - \frac{1}{2}|x|^2\right\|_{C^0(B_3)} \leq \eta_1, \tag{5.13}$$

then $u \in C^{k+2,\alpha}(B_{1/9})$.

Proof. — We divide the proof in two steps.

- *Step 1:* $C^{1,1}$ *regularity.*

Fix a point $x_0 \in B_{1/8}$, and set $y_0 := \text{c-exp}_{x_0}(\nabla u(x_0))$. Up to replace u (resp. c) with the function u_1 (resp. c_1) constructed in Steps 4 and 5 in the proof of Theorem 4.3, we can assume that $u \geq 0$, $u(\mathbf{0}) = 0$, that

$$S_h := S(\mathbf{0}, \mathbf{0}, u, h) = \{u \leq h\},$$

and that

$$D_{xy}c(\mathbf{0}, \mathbf{0}) = -\operatorname{Id}. \tag{5.14}$$

Under these assumptions we will show that the sections of u are of "good shape", i.e.,

$$B_{\sqrt{h}/K} \subset S_h \subset B_{K\sqrt{h}} \quad \forall h \leq h_1, \tag{5.15}$$

for some universal h_1 and K. Arguing as in Step 6 of Theorem 4.3, this will give that u is $C^{1,1}$ at the origin, and thus at every point in $B_{1/8}$.

First of all notice that, thanks to (**5.13**), for any $h_1 > 0$ we can choose $\eta_1 = \eta_1(h_1) > 0$ small enough such that (**5.15**) holds for S_{h_1} with $K = 2$. Hence, assuming without loss of generality that $\delta_1 \leq 1$, we see that

$$B_{\sqrt{h_1}/3} \subset \mathcal{N}_{\delta_1^\gamma \sqrt{h_1}}\big(\text{co}[S_{h_1}]\big) \subset B_{3\sqrt{h_1}},$$

where γ is the exponent from Proposition 5.2. Let v_1 solve the Monge-Ampère equation

$$\begin{cases} \det(D^2 v_1) = f(\mathbf{0})/g(\mathbf{0}) & \text{in } \mathcal{N}_{\delta_1^\gamma \sqrt{h_1}}(\text{co}[S_{h_1}]), \\ v_1 = h_1 & \text{on } \partial\mathcal{N}_{\delta_1^\gamma \sqrt{h_1}}(\text{co}[S_{h_1}]). \end{cases}$$

Since $B_{1/3} \subset N_{\delta_1^\gamma \sqrt{h_1}}(\text{co}[S_{h_1}])/\sqrt{h_1} \subset B_3$, by standard Pogorelov estimates applied to the function $v_1(\sqrt{h_1}x)/h_1$ (see for instance [26, Theorem 4.2.1]), it follows that $|D^2 v_1(0)| \leq$ M, with $M > 0$ some large universal constant.

Let $h_k := h_1 2^{-k}$ and define $\bar{K} \geq 3$ to be the largest number such that any solution w of

$$\begin{cases} \det(D^2 w) = f(\mathbf{0})/g(\mathbf{0}) & \text{in } Z, \\ w = 1 & \text{on } \partial Z, \end{cases} \quad \text{with } B_{1/\bar{K}} \subset Z \subset B_{\bar{K}}, \tag{5.16}$$

satisfies $|D^2 w(0)| \leq M + 1$.[5] We prove by induction that (**5.15**) holds with $K = \bar{K}$.

---

[5] The fact that $\bar{K}$ is well defined (i.e., $3 \leq \bar{K} < \infty$) follows by the following facts: first of all, by definition, M is an a-priori bound for $|D^2 w(0)|$ whenever w is a solution of (**5.16**) with $B_{1/3} \subset Z \subset B_3$, so $\bar{K} \geq 3$. On the other hand $\bar{K} \leq \sqrt{2(M+1)}$. Indeed, since $1/2 \leq f(\mathbf{0})/g(\mathbf{0}) \leq 2$ (by (**5.11**)) and $M \geq 1$, the function

$$\bar{w} := (M+1)x_1^2 + \frac{f(\mathbf{0})}{g(\mathbf{0})} \frac{x_2^2}{M+1} + x_3^2 + \cdots + x_n^2$$

is a solution of (**5.16**) such that $B_{1/\sqrt{2(M+1)}} \subset B_{1/\sqrt{M+1}} \subset \{\bar{w} \leq 1\} \subset B_{\sqrt{2(M+1)}}$ and $|D^2 \bar{w}(0)| = 2(M+1)$.

If $h=h_1$ then we already know that (**5.15**) holds with $K=2$ (and so with $K=\bar{K}$).

Assume now that (**5.15**) holds with $h=h_k$ and $K=\bar{K}$, and we want to show that it holds with $h=h_{k+1}$. For this, for any $k\in\mathbf{N}$ we consider u_k the solution of

$$\begin{cases} \det(D^2 v_k) = f(\mathbf{0})/g(\mathbf{0}) & \text{in } \mathcal{N}_{\delta_k^\gamma\sqrt{h_k}}(\mathrm{co}[S_{h_k}]),\\ v_k = h_1 2^{-k} & \text{on } \partial N_{\delta_k^\gamma\sqrt{h_k}}(\mathrm{co}[S_{h_k}]),\end{cases}$$

where

$$\delta_k := \left\| c(x,y) + x\cdot y\right\|_{C^2(S_{h_k}\times T_u(S_{h_k}))} \le \delta_1.$$

Let us consider the rescaled functions

$$\bar{u}_k(x) := u\big(\sqrt{h_k}x\big)/h_k, \qquad \bar{v}_k(x) := v_k\big(\sqrt{h_k}x\big)/h_k.$$

Since by the inductive hypothesis $B_{1/\bar{K}} \subset \bar{S}_k := \{\bar{u}_k \le 1\} \subset B_{\bar{K}}$, we can apply Proposition 5.2 to deduce that

$$\|\bar{u}_k - \bar{v}_k\|_{C^0(\bar{S}_k)} \le C_{\bar{K}}\Big(\operatorname*{osc}_{S_{h_k}} f + \operatorname*{osc}_{T_u(S_{h_k})} g + \delta_k^{\gamma/n}\Big) \le C_{\bar{K}}\big(\delta_1 + \delta_1^{\gamma/n}\big). \tag{5.17}$$

This implies in particular that, if δ_1 is sufficiently small, $B_{1/(2\bar{K})} \subset \{\bar{v}_k \le 1\} \subset B_{2\bar{K}}$. By standard estimates on the sections of solutions to the Monge-Ampère equation, the shapes of $\{\bar{v}_k \le 1\}$ and $\{\bar{v}_k \le 1/2\}$ are comparable, and in addition sections are well included into each other [26, Theorem 3.3.8]: there exists a universal constant $L>1$ such that

$$\begin{aligned} &B_{1/(L\bar{K})} \subset \{\bar{v}_k \le 1/2\} \subset B_{L\bar{K}},\\ &\mathrm{dist}\big(\{\bar{v}_k \le 1/4\}, \partial\{\bar{v}_k \le 1/2\}\big) \ge 1/(LK).\end{aligned}$$

Using again (**5.17**) we deduce that, if δ_1 is sufficiently small,

$$\begin{aligned} &B_{1/(2L\bar{K})} \subset \{\bar{u}_k \le 1/2\} \subset B_{2L\bar{K}},\\ &\mathrm{dist}\big(\{\bar{u}_k \le 1/4\}, \partial\{\bar{u}_k \le 1/2\}\big) \ge 1/(2LK)\end{aligned}$$

so, by scaling back,

$$B_{\sqrt{h_{k+1}}/(2L\bar{K})} \subset S_{h_{k+1}} \subset B_{2L\bar{K}\sqrt{h_{k+1}}}, \qquad \mathrm{dist}(S_{h_{k+2}}, \partial S_{h_{k+1}}) \ge \sqrt{h_k}/(2LK). \tag{5.18}$$

This allows us to apply Proposition 5.2 also to $\bar{u}_{k+1}$ to get

$$\|\bar{u}_{k+1} - \bar{v}_{k+1}\|_{C^0(\bar{S}_{k+1})} \le C_{2L\bar{K}}\Big(\operatorname*{osc}_{S_{h_{k+1}}} f + \operatorname*{osc}_{T_u(S_{h_{k+1}})} g + \delta_{k+1}^{\gamma/n}\Big). \tag{5.19}$$

We now observe that, by (**5.15**) and the $C^{1,\beta}$ regularity of u (see Theorem 4.3), it follows that

$$\mathrm{diam}(S_{h_k}) + \mathrm{diam}\big(T_u(S_{h_k})\big) \le C h_k^{\beta/2},$$

so by the $C^{0,\alpha}$ regularity of f and g, and the $C^{2,\alpha}$ regularity of c, we have (recall that $\gamma < 1$)

$$\operatorname*{osc}_{S_{h_k}} f + \operatorname*{osc}_{T_u(S_{h_k})} g + \delta_k^{\gamma/n} \le C' h_k^{\sigma}, \quad \sigma := \frac{\alpha\beta\gamma}{2n}. \tag{5.20}$$

Hence, by (**5.17**) and (**5.19**),

$$\|\bar{u}_k - \bar{v}_k\|_{C^0(\bar{S}_k)} + \|\bar{u}_{k+1} - \bar{v}_{k+1}\|_{C^0(\bar{S}_{k+1})} \le C(C_{\bar{K}} + C_{2L\bar{K}}) h_k^{\sigma},$$

from which we deduce (recall that $h_k = 2h_{k+1}$)

$$\begin{aligned}\|v_k - v_{k+1}\|_{C^0(S_{h_{k+1}})} &\le \|v_k - u\|_{C^0(S_{h_k})} + \|u - v_{k+1}\|_{C^0(S_{h_{k+1}})} \\ &= h_k \|\bar{u}_k - \bar{v}_k\|_{C^0(S_k)} + h_{k+1}\|\bar{u}_{k+1} - \bar{v}_{k+1}\|_{C^0(S_{k+1})} \\ &\le C(C_{\bar{K}} + C_{2L\bar{K}}) h_k^{1+\sigma}.\end{aligned}$$

Since v_k and v_{k+1} are two strictly convex solutions of the Monge Ampère equation with constant right hand side inside $S_{h_{k+1}}$, and since $S_{h_{k+2}}$ is "well contained" inside $S_{h_{k+1}}$, by classical Pogorelov and Schauder estimates we get

$$\big\|D^2 v_k - D^2 v_{k+1}\big\|_{C^0(S_{h_{k+2}})} \le C'_{\bar{K}} h_k^{\sigma}, \tag{5.21}$$

$$\big\|D^3 v_k - D^3 v_{k+1}\big\|_{C^0(S_{h_{k+2}})} \le C'_{\bar{K}} h_k^{\sigma - 1/2}, \tag{5.22}$$

where $C'_{\bar{K}}$ is some constant depending only on $\bar{K}$. By (**5.21**) applied to v_j for all $j = 1, \dots, k$ (this can be done since, by the inductive assumption, (**5.15**) holds for $h = h_j$ with $j = 1, \dots, k$) we obtain

$$\begin{aligned}\big|D^2 v_{k+1}(0)\big| &\le \big|D^2 v_1(0)\big| + \sum_{j=1}^{k} \big|D^2 v_j(0) - D^2 v_{j+1}(0)\big| \\ &\le M + C'_{\bar{K}} h_1^{\sigma} \sum_{j=0}^{k} 2^{-j\sigma} \\ &\le M + \frac{C'_{\bar{K}}}{1 - 2^{-\sigma}} h_1^{\sigma} \le M + 1,\end{aligned}$$

provided we choose h_1 small enough (recall that $h_k = h_1 2^{-k}$). By the definition of $\bar{K}$ it follows that also $S_{h_{k+1}}$ satisfies (**5.15**), concluding the proof of the inductive step.

- *Step 2: higher regularity.*

Now that we know that $u \in C^{1,1}(B_{1/8})$, Equation (**2.10**) becomes uniformly elliptic. So one may use Evans-Krylov Theorem to obtain that $u \in C^{2,\sigma'}_{\rm loc}(B_{1/9})$ for some $\sigma' > 0$, and then standard Schauder estimates to conclude the proof. However, for the convenience of the reader, we show here how to give a simple direct proof of the $C^{2,\sigma'}$ regularity of u with $\sigma' = 2\sigma$.

As in the previous step, it suffices to show that u is $C^{2,\sigma'}$ at the origin, and for this we have to prove that there exists a sequence of paraboloids P_k such that

$$\sup_{B_{r_0^k/C}} |u - P_k| \le C r_0^{k(2+\sigma')} \tag{5.23}$$

for some $r_0, C > 0$.

Let v_k be as in the previous step, and let P_k be their second order Taylor expansion at $\mathbf{0}$:

$$P_k(x) = v_k(\mathbf{0}) + \nabla v_k(\mathbf{0}) \cdot x + \frac{1}{2} D^2 v_k(\mathbf{0}) x \cdot x.$$

We observe that, thanks to (**5.15**),

$$\|v_k - P_k\|_{C^0(B(0,\sqrt{h_{k+2}}/K))} \le \|v_k - P_k\|_{C^0(S_{h_{k+2}})} \le C \left\| D^3 v_k \right\|_{C^0(S_{h_{k+2}})} h_k^{3/2}. \tag{5.24}$$

In addition, by (**5.22**) applied with $j = 1, \ldots, k$ and recalling that $h_k = h_1 2^{-k}$ and $2\sigma < 1$ (see (**5.20**)), we get

$$\begin{aligned} \left\| D^3 v_k \right\|_{C^0(S_{h_{k+2}})} &\le \left\| D^3 v_1 \right\|_{C^0(S_{h_3})} + \sum_{j=1}^{k} \left\| D^3 v_j - D^3 v_{j+1} \right\|_{C^0(S_{h_{j+2}})} \\ &\le C \left(1 + \sum_{j=1}^{k} h_j^{(\sigma - 1/2)} \right) \le C h_k^{\sigma - 1/2}. \end{aligned} \tag{5.25}$$

Combining (**5.15**), (**5.24**), (**5.25**), and recalling (**5.17**) and (**5.20**), we obtain

$$\|u - P_k\|_{C^0(B_{\sqrt{h_{k+2}}/K})} \le \|v_k - P_k\|_{C^0(S_{h_{k+2}})} + \|v_k - u\|_{C^0(S_{h_{k+2}})} \le C h_k^{1+\sigma},$$

so (**5.23**) follows with $r_0 = 1/\sqrt{2}$ and $\sigma' = 2\sigma$. □

Acknowledgements

We wish to thank Luigi Ambrosio for his careful reading of a preliminary version of this manuscript. AF is partially supported by NSF Grant DMS-0969962. Both authors acknowledge the support of the ERC ADG Grant GeMeThNES.

REFERENCES

1. L. Ambrosio, N. Gigli and G. Savaré, *Gradient Flows in Metric Spaces and in the Space of Probability Measures*, 2nd ed., Lectures in Mathematics ETH Zürich, Birkhäuser, Basel, 2008.
2. Y. Brenier, Polar factorization and monotone rearrangement of vector-valued functions, *Commun. Pure Appl. Math.*, **44** (1991), 375–417.
3. L. A. Caffarelli, A localization property of viscosity solutions to the Monge-Ampère equation and their strict convexity, *Ann. Math. (2)*, **131** (1990), 129–134.
4. L. A. Caffarelli, Some regularity properties of solutions of Monge Ampère equation, *Commun. Pure Appl. Math.*, **44** (1991), 965–969.
5. L. A. Caffarelli, The regularity of mappings with a convex potential, *J. Am. Math. Soc.*, **5** (1992), 99–104.
6. L. A. Caffarelli, Interior $W^{2,p}$ estimates for solutions of the Monge-Ampère equation, *Ann. Math. (2)*, **131** (1990), 135–150.
7. L. A. Caffarelli, M. M. Gonzáles and T. Nguyen, A perturbation argument for a Monge-Ampère type equation arising in optimal transportation. Preprint (2011).
8. L. A. Caffarelli and Y. Y. Li, A Liouville theorem for solutions of the Monge-Ampère equation with periodic data, *Ann. Inst. Henri Poincaré, Anal. Non Linéaire*, **21** (2004), 97–120.
9. D. Cordero-Erausquin, R. J. McCann and M. Schmuckenschläger, A Riemannian interpolation inequality à la Borell, Brascamp and Lieb, *Invent. Math.*, **146** (2001), 219–257.
10. P. Delanoë and Y. Ge, Regularity of optimal transportation maps on compact, locally nearly spherical, manifolds, *J. Reine Angew. Math.*, **646** (2010), 65–115.
11. P. Delanoë and F. Rouvière, Positively curved Riemannian locally symmetric spaces are positively squared distance curved, *Can. J. Math.*, **65** (2013), 757–767.
12. L. C. Evans and R. F. Gariepy, *Measure Theory and Fine Properties of Functions*, Studies in Advanced Mathematics, CRC Press, Boca Raton, 1992.
13. A. Fathi and A. Figalli, Optimal transportation on non-compact manifolds, *Isr. J. Math.*, **175** (2010), 1–59.
14. A. Figalli, Existence, uniqueness, and regularity of optimal transport maps, *SIAM J. Math. Anal.*, **39** (2007), 126–137.
15. A. Figalli, Regularity of optimal transport maps [after Ma-Trudinger-Wang and Loeper]. (English summary) Séminaire Bourbaki. Volume 2008/2009. Exposés 997–1011. *Astérisque No. 332* (2010), Exp. No. 1009, ix, 341–368.
16. A. Figalli, Regularity properties of optimal maps between nonconvex domains in the plane, *Commun. Partial Differ. Equ.*, **35** (2010), 465–479.
17. A. Figalli and N. Gigli, Local semiconvexity of Kantorovich potentials on non-compact manifolds, *ESAIM Control Optim. Calc. Var.*, **17** (2011), 648–653.
18. A. Figalli and Y. H. Kim, Partial regularity of Brenier solutions of the Monge-Ampère equation, *Discrete Contin. Dyn. Syst.*, **28** (2010), 559–565.
19. A. Figalli, Y. H. Kim and R. J. McCann, Hölder continuity and injectivity of optimal maps, *Arch. Ration. Mech. Anal.*, **209** (2013), 747–795.
20. A. Figalli, Y. H. Kim and R. J. McCann, Regularity of optimal transport maps on multiple products of spheres, *J. Eur. Math. Soc. (JEMS)*, **5** (2013), 1131–1166.
21. A. Figalli and G. Loeper, C^1 regularity of solutions of the Monge-Ampère equation for optimal transport in dimension two, *Calc. Var. Partial Differ. Equ.*, **35** (2009), 537–550.
22. A. Figalli and L. Rifford, Continuity of optimal transport maps and convexity of injectivity domains on small deformations of $\mathbf{S}^2$, *Commun. Pure Appl. Math.*, **62** (2009), 1670–1706.
23. A. Figalli, L. Rifford and C. Villani, On the Ma-Trudinger-Wang curvature on surfaces, *Calc. Var. Partial Differ. Equ.*, **39** (2010), 307–332.
24. A. Figalli, L. Rifford and C. Villani, Necessary and sufficient conditions for continuity of optimal transport maps on Riemannian manifolds, *Tohoku Math. J. (2)*, **63** (2011), 855–876.
25. A. Figalli, L. Rifford and C. Villani, Nearly round spheres look convex, *Am. J. Math.*, **134** (2012), 109–139.
26. C. Gutierrez, *The Monge-Ampére Equation*, Progress in Nonlinear Differential Equations and Their Applications, vol. 440, Birkhäuser, Boston, 2001.
27. H.-Y. Jian and X.-J. Wang, Continuity estimates for the Monge-Ampère equation, *SIAM J. Math. Anal.*, **39** (2007), 608–626.
28. Y.-H. Kim, Counterexamples to continuity of optimal transport maps on positively curved Riemannian manifolds, *Int. Math. Res. Not. IMRN* 2008, Art. ID rnn120, 15 pp.

29. Y.-H. Kim and R. J. McCann, Towards the smoothness of optimal maps on Riemannian submersions and Riemannian products (of round spheres in particular), *J. Reine Angew. Math.*, **664** (2012), 1–27.
30. J. Liu, N. S. Trudinger and X.-J. Wang, Interior $C^{2,\alpha}$ regularity for potential functions in optimal transportation, *Commun. Partial Differ. Equ.*, **35** (2010), 165–184.
31. G. Loeper, On the regularity of solutions of optimal transportation problems, *Acta Math.*, **202** (2009), 241–283.
32. G. Loeper, Regularity of optimal maps on the sphere: The quadratic cost and the reflector antenna, *Arch. Ration. Mech. Anal.*, **199** (2011), 269–289.
33. X. N. Ma, N. S. Trudinger and X. J. Wang, Regularity of potential functions of the optimal transportation problem, *Arch. Ration. Mech. Anal.*, **177** (2005), 151–183.
34. R. J. McCann, Polar factorization of maps on Riemannian manifolds, *Geom. Funct. Anal.*, **11** (2001), 589–608.
35. N. S. Trudinger and X.-J. Wang, On the second boundary value problem for Monge-Ampère type equations and optimal transportation, *Ann. Sc. Norm. Super. Pisa Cl. Sci. (5)*, **8** (2009), 143–174.
36. N. S. Trudinger and X.-J. Wang, On strict convexity and continuous differentiability of potential functions in optimal transportation, *Arch. Ration. Mech. Anal.*, **192** (2009), 403–418.
37. C. Villani, *Optimal Transport. Old and New*, Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 338, Springer, Berlin, 2009.

G. P.
Scuola Normale Superiore,
p.za dei Cavalieri 7,
56126 Pisa, Italy
guido.dephilippis@sns.it

A. F.
Department of Mathematics,
The University of Texas at Austin,
1 University Station C1200,
Austin, TX 78712, USA
figalli@math.utexas.edu

Manuscrit reçu le 1 octobre 2012
Manuscrit accepté le 15 juillet 2014
publié en ligne le 29 juillet 2014.