Optimal transportation between hypersurfaces bounding some strictly convex domains

Emmanuel Humbert, Luc Molinet

OPTIMAL TRANSPORTATION BETWEEN HYPERSURFACES BOUNDING SOME STRICTLY CONVEX DOMAINS

E. HUMBERT AND L. MOLINET

ABSTRACT. Let M, N be two smooth compact hypersurfaces of $\mathbb{R}^n$ which bound strictly convex domains equipped with two absolutely continuous measures μ and ν (with respect to the volume measures of M and N). We consider the optimal transportation from μ to ν for the quadratic cost. Let $(\phi : m \to \mathbb{R}, \psi : N \to \mathbb{R})$ be some functions which achieve the supremum in the Kantorovich formulation of the problem and which satisfy

$$\psi(y) = \inf_{z \in M} \Big(\frac{1}{2}|y-z|^2 - \varphi(z) \Big); \varphi(x) = \inf_{z \in N} \Big(\frac{1}{2}|x-z|^2 - \psi(z) \Big).$$

Define for $y \in N$,

$$\varphi^{\square}(y) = \sup_{z \in M} \Big(\frac{1}{2}|y-z|^2 - \varphi(z) \Big).$$

In this short paper, we exhibit a relationship between the regularity of $\varphi^{\square}$ and the existence of a solution to the Monge problem.

Let M and N be two smooth compact hypersurfaces of $\mathbb{R}^n$, $n \geq 2$, which are the boundary of some strictly convex domains. In the present paper, we study the existence of a solution of Monge Problem when considering the optimal transport with quadratic cost between two measures μ and ν respectively supported in M and N. This situations has been already studied: see [3]. In the whole paper, we assume that μ and ν have the form $\mu = f(x)dv(x)$ and $\nu = g(y)dv(y)$ where f, g are some non-zero nonnegative continuous functions on M and N and where dv stands for the volume measures on M and N. The quadratic cost is defined for all $x, y \in \mathbb{R}^n$ by $c_2(x,y) := \frac{1}{2}|x-y|^2$. Here, $|\cdot|$ denotes the standard norm associated to the canonical scalar product of $\mathbb{R}^n$. For all $x, y \in \mathbb{R}^n$, the scalar product of x and y will be denoted by $x \cdot y$.

The standard formulation of the optimal transport from μ to ν for the quadratic cost is

$$T_0 := \inf_{\pi \in \Pi'(\mu,\nu)} I_0'(\pi)$$

where $\Pi'(\mu', \nu')$ is the set of probabilily measures $\pi(\cdot,\cdot)$ on $\mathbb{R}^n \times \mathbb{R}^n$ such that

$$\pi(\mathbb{R}^n, \cdot) = \nu \text{ and } \pi(\cdot, \mathbb{R}^n) = \mu$$

and for all $\pi \in \Pi'(\mu', \nu')$,

$$I_0'(\pi) := \int_{\mathbb{R}^n \times \mathbb{R}^n} c_2(x,y) d\pi(x,y).$$

As easily checked, this is an equivalent formulation to write

$$T_0 := \inf_{\pi \in \Pi(\mu,\nu)} I_0(\pi) \tag{0.1}$$

where $\Pi(\mu, \nu)$ is the set of probabilily measures $\pi(\cdot,\cdot)$ on $M \times N$ such that

$$\pi(M, \cdot) = \nu \text{ and } \pi(\cdot, N) = \mu$$

and for all $\pi \in \Pi(\mu,\nu)$,

$$I_0(\pi) := \int_{M\times N} c_2(x,y)\, d\pi(x,y).$$

By the Monge-Kantorovich duality (see for instance, [5], [6]), one has

$$T_0 = \sup_{(\varphi,\psi)\in\Omega} J_0(\varphi,\psi) \tag{0.2}$$

where Ω is the set of couples of functions $(\varphi,\psi) \in C^0(M)\times C^0(N)$ such that

$$\varphi(x) + \psi(y) \leq c_2(x,y)$$

for all $(x,y) \in M\times N$ and where

$$J_0(g,h) = \int_M g(x)d\mu(x) + \int_N h(y)d\nu(y).$$

for all $(g,h) \in C^0(M)\times C^0(N)$.
Actually, in many situations, one can also show the uniqueness of π (see [1]).
It is standard to prove that (see for instance [1] for references):

(1) the infimum in (0.1) is attained by some probability measure π, which is called a transference plan;
(2) the supremum in (0.2) is attained by some couple of functions (φ,ψ) which satisfy for all $x\in M$, $y\in N$:

$$\psi(y) = \inf_{z\in M}\Big(c_2(z,y) - \varphi(z)\Big); \varphi(x) = \inf_{z\in N}\Big(c_2(x,z) - \psi(z)\Big) \tag{0.3}$$

In the whole paper, if $\eta : M \to \mathbb{R}$, we will note for all $y\in N$

$$\eta^*(y) = \inf_{x\in M}\Big(c_2(x,y) - \eta(x)\Big). \tag{0.4}$$

In the same way, if $\eta : N \to \mathbb{R}$, we will note for all $x\in M$

$$\eta^*(x) = \inf_{y\in N}\Big(c_2(x,y) - \eta(y)\Big), \tag{0.5}$$

so that $\varphi^* = \psi$ and $\psi^* = \varphi$.

An important question about (0.1) is the following : are the transference plans associated with (0.1) supported in a graph ? Indeed, a positive answer to this question would ensure the existence of a solution to the famous Monge problem (see, for instance, [6] for some explanations). Unfortunately, the answer is no in its full generality. Gangbo and McCann [3] could construct some counter examples. Even worse: numerical computations indicate that this is not true either in the simplest situation when $M = N = S^1$ (for $n \geq 1$, S^n denotes the unit sphere of $\mathbb{R}^{n+1}$) and when μ and ν have smooth positive densities (see [2]).

In this paper we try to give some simple criteria that would imply a positive answer. Before, stating our result, we need to introduce some definitions.

For any fonction $\Theta : M \to \mathbb{R}$, we set for all $y\in N$

$$\Theta^{\square}(y) = \sup_{x\in M}\Big(c_2(x,y) - \Theta(x)\Big). \tag{0.6}$$

and in the same way, if $\Theta : N \to \mathbb{R}$, we set for all $x \in M$

$$\Theta^{\square}(x) = \sup_{y \in N} \Big(c_2(x,y) - \Theta(x) \Big). \tag{0.7}$$

While, as proven is [3], the function φ is always C^1, the function $\varphi^{\square}$ has no reason to be C^1 in general but surprisingly, its regularity is directly related to the question above. More precisely, our main result is

Theorem 0.1. *For all $y \in N$, define $\Theta_y : M \to \mathbb{R}$ by*

$$\Theta_y(x) = c_2(x,y) - \varphi(x)$$

so that, for $y \in N$, $\varphi^*(y) = \inf_M \Theta_y$ and $\varphi^{\square}(y) = \sup_M \Theta_y$. Then, the following assertions are equivalent:

(1) *$\varphi^{\square}$ is C^1;*
(2) *for all $y \in N$, the function Θ_y has exactly two critical points: its minimum and it maximum.*

If one of the assertions above is true then $\varphi^{\square\square} = \varphi^{**} = \varphi$. Moreover, the support of π is contained in a graph.

The idea of the proof is as follows. Let Γ be the set of points $x \in M$ which are the maximum of a function Θ_y for some $y \in N$. There is then two crucial observations:

(1) If (x, y) belongs to the support of π and if $x \in \Gamma$, then y is unique. This implies that if $\Gamma = M$ then the support of π is contained in a graph.
(2) Let $x \in \Gamma$, $y \in N$ such that x is the maximum of Θ_y. Then, y is unique. This allows to define a map $T : \Gamma \to N$ such that $T(x) = y$. The main argument of the proof is to show that under the assumptions of Theorem 0.1, T is actually a homeomorphism, which implies that $\Gamma = M$ and allows to conclude.

Even if assumptions 1) or 2) are not easy to check, we think that this theorem gives a new point of view, that we hope useful, to this Monge problem. For convenience of the reader, we stated all the results which seem of particular interest to us in Propositions 1.3, 1.4, 1.6. Theorem 0.1 is a direct consequence of these propositions.

1. Proof of Theorem 0.1

1.1. **Notations and Preliminaries.** We keep the notations of the introduction: (φ, ψ) is a couple of functions maximizing the problem (0.2). By Gangbo and McCann [3], these functions are C^1. Indeed, in their paper Section 3, they show that the convex functions they study are tangentially differentiable and that these tangent differentials are continuous on M and N. Here, the function we consider are the same functions restricted to M and N and are hence C^1. Notice that the proof of this fact is far to be obvious. In addition, the functions φ, ψ satisfy $\varphi^* = \psi$ and $\psi^* = \varphi$. We recall that for all $y \in N$, we defined $\Theta_y : M \to \mathbb{R}$ by

$$\Theta_y(x) = c_2(x,y) - \varphi(x).$$

In the same way, if $x \in M$, we define $\Theta_x : N \to \mathbb{R}$ by

$$\Theta_x(x) = c_2(x,y) - \psi(x).$$

For $x \in M$, $y \in N$, we introduce the sets:

$$\Omega_x := \{ z \in N \,/ \theta_x(z) = \inf_{z' \in N} \theta_x(z') = \varphi^*(z) = \psi(z) \}$$

and

$$\Omega_y := \{z \in M \,/ \theta_y(z) = \inf_{z' \in M} \theta_y(z') = \psi^*(z) = \varphi(z)\}.$$

Note that, by compacity of M and N, these sets are non empty. Note also that

$$\Omega_x := \{z \in N \,/ \varphi(x) + \psi(z) = c_2(x, z)\} \text{ and } \Omega_y := \{z \in M \,/ \varphi(z) + \psi(y) = c_2(z, y)\}$$

which has the immediate consequence that

$$y \in \Omega_x \Leftrightarrow x \in \Omega_y \Leftrightarrow \varphi(x) + \psi(y) = c_2(x, y). \tag{1.1}$$

Now, proving that the support of π is contained in a graph of a continuous preserving map $\alpha : M \to N$ (resp. $\alpha : N \to M$) is reduced to proving that for all $x \in M$ (resp. $y \in N$), the set Ω_x (resp. Ω_y) contains exactly one point. Indeed, by (0.1) and (0.2), one has

$$\int_{M \times N} c_2(x, y) d\pi(x, y) = \int_M \varphi(x) d\mu(x) + \int_N \psi(y) d\nu(y)$$

which can be rewritten, since the marginals of π are μ and ν, by

$$\int_{M \times N} (c_2(x, y) - \varphi(x) - \psi(y))\, d\pi(x, y) = 0.$$

Since $c_2(x, y) - \varphi(x) - \psi(y) \geq 0$, one has identically on the support of π:

$$c_2(x, y) = \varphi(x) + \psi(y)$$

and hence $x \in \Omega_y$ or $y \in \Omega_x$.

For any $y \in N$ we denote by $n_N(y)$ the unitary normal outside vector to N at y and we define the line D_y by

$$D_y = y - \nabla\psi(y) + \text{span}(n_N(y)) \; .$$

Similarly, for $x \in M$ we define the line D_x by

$$D_x = x - \nabla\varphi(x) + \text{span}(n_M(x)) \; .$$

An easy computation of the derivative of Θ_y and Θ_x shows that if $x \in M, y \in N$, then

$$y \in D_x \Leftrightarrow \nabla\Theta_y(x) = 0 \text{ and } x \in D_y \Leftrightarrow \nabla\Theta_x(y) = 0. \tag{1.2}$$

Finally, we will use several times the following Lemma:

Lemma 1.1. *For all $x \in M$, $y \in N$, the functions Θ_x and Θ_y are never constant.*

Proof. Assume for instance that Θ_y is constant. Then for any $x \in M$, $x \in \Omega_y$ and hence, by (1.1), $y \in \Omega_x$. By (1.2), $x \in D_y$ which implies that M is contained in the right line D_x. This is impossible and Lemma 1.1 follows. □

1.2. **Properties of $\varphi^\square$.** In the proof on Theorem 0.1, we will use some basic properties of $\varphi^\square$. Many of them are very standard. We first recall its definition:

$$\varphi^\square(y) = \sup_{x\in M}\Big(c_2(x,y) - \varphi(x)\Big).$$

For convenience, for any $y \in N$, we will denote by $\Omega_y^\square$ the set of points of M achieving the maximum in the definition of $\varphi^\square$. We collect the properties we will need in the following Proposition:

Proposition 1.2. (1) *For all $x \in M$, $ty \in N$, one has $\varphi(x)+\varphi^\square(y) \geq c_2(x,y)$ with equality if and only if $x \in \Omega_y^\square$;*
(2) *$\varphi^{\square\square} \leq \varphi$;*
(3) *$\varphi^\square$ is Lipschitz;*
(4) *$\varphi^{\square\square\square} = \varphi^\square$;*
(5) *Let $y \in M$. Assume that there is only one point such x_y such that*

$$\varphi^\square(x) = c_2(x_y, y) - \varphi(x_y)$$

i.e. the supremum in the definition of $\varphi^\square$ is attained at only one point, then $\varphi^\square$ is differentiable at y.

Proof. We start by proving 1). Let $x \in M$ and $y \in N$. It holds that

$$\begin{aligned}\varphi(x) + \varphi^\square(y) &= \varphi(x) + \sup_{z\in M}\Big(c_2(z,y) - \varphi(z)\Big)\\ &\geq \varphi(x) + (c_2(x,y) - \varphi(x)) = c_2(x,y).\end{aligned}$$

The inequality above becomes an equality if and only if $x \in \Omega_y^\square$. This proves 1).

Let us now deal with 2). Let $x \in M$. By definition:

$$\varphi^{\square\square}(x) = \sup_{y\in N}(c_2(x,y) - \varphi^\square(y)).$$

By compacity of N, there exists $y_x \in M$ such that

$$\varphi^{\square\square}(x) = c_2(x, y_x) - \varphi^\square(y_x). \tag{1.3}$$

By definition, one also have

$$\varphi^\square(y_x) = \sup_{z\in M}(c_2(z,y_x) - \varphi(z)).$$

and, setting $z = x$, one has

$$\varphi^\square(y_x) \geq c_2(x, y_x) - \varphi(x).$$

Together with (1.3), this gives 2).

Let us prove 3). Let $y, z \in M$, $x_y \in \Omega_y^\square$ and $x_z \in \Omega_z^\square$. We prove that

$$\begin{aligned}\frac{1}{2}(z+y-2x_y)\cdot(z-y) &\geq \varphi^\square(z) - \varphi^\square(y)\\ &\geq \frac{1}{2}(z+y-2x_z)\cdot(z-y).\end{aligned} \tag{1.4}$$

The definition of $\Omega_z^\square$ implies that

$$\varphi^\square(z) = c_2(x_z, z) - \varphi(x_z).$$

The construction of $\varphi^\square$ implies that

$$\varphi^\square(y) \geq c_2(x_z, y) - \varphi(x_z).$$

Observing that

$$\frac{1}{2}(z + y - 2x_z) \cdot (z - y) = c_2(x_z, z) - c_2(x_z, y),$$

this provides the right inequality of (1.4). The left inequality is proven in the same way. Now, observe that since M, N are compact, there exists a constant $C > 0$ independant of y, z such that

$$\frac{1}{2}|z + y - 2x_z| \leq C \ \text{ and } \frac{1}{2}|z + y - 2x_y| \leq C.$$

Using that $|z-y|$ is less than the geodesic distance on N, we immediatly deduce that $\varphi^\square$ is lipschitz. Note that this implies that $\varphi^\square$ is continuous and by Rademacher's Theorem, is differentiable almost everywhere.

We now prove 4). By Point 2)

$$\varphi^{\square\square} \leq \varphi.$$

In particular, for all $y \in N$,

$$\varphi^{\square\square\square}(y) = \sup_{x\in M} (c_2(x, y) - \varphi^{\square\square}(x)) \geq \sup_{x\in M} (c_2(x, y) - \varphi(x)) = \varphi^\square(y).$$

For all $x \in M$, $y \in N$, we also have as in Point 1)

$$\varphi^{\square\square}(x) + \varphi^\square(y) \geq c_2(x, y).$$

Then as in Point 2), $\varphi^{\square\square\square} = (\varphi^\square)^{\square\square} \leq \varphi^\square$. This shows 4).

Let us finish by proving 5). Let $y \in M$ and assume that $\Omega^\square_y$ is reduced to one point x. Let (z_k) be a sequence of points of N tending to y. For all k, choose $x_k \in \Omega^\square_{z_k}$. By compacity of M, one can assume that x_k converges to some $x' \in M$. The definition of $\Omega^\square_{z_k}$ and Point 1) implies that

$$\varphi(x_k) + \varphi^\square(z_k) = c_2(x_k, z_k).$$

By continuity of φ and $\varphi^\square$, we obtain as z_k tends to y,

$$\varphi(x') + \varphi^\square(y) = c_2(x', y)$$

which proves that $x' \in \Omega^\square_y$ and hence $x' = x$ Using (1.4), we have

$$\frac{1}{2}(z_k - y) \cdot (z_k + y - 2x_k) \geq \varphi^\square(z_k) - \varphi^\square(y) \geq \frac{1}{2}(z_k - y) \cdot (z_k + y - 2x). \tag{1.5}$$

Until the end of the proof, the notation o_k will stand for a term which is $o(|z_k - y|)$. Since x_k tends to x, we have

$$x_k \cdot (z_k - y) = x \cdot (z_k - y) + o_k.$$

When z_k is close to y,

$$z_k - y = P_y(z_k - y) + o_k,$$

where P_y denotes the orthogonal projection onto the tangent space T_yN. Coming back to (1.5), we obtain that

$$\frac{1}{2}P_y(z_k - y) \cdot (z_k + y - 2x) + o_k \geq \varphi^\square(z_k) - \varphi^\square(y) \geq \frac{1}{2}P_y(z_k - y) \cdot (z_k + y - 2x) + o_k.$$

Since P_y is self-adjoint, this yields

$$\frac{1}{2}(z_k-y)\cdot P_y(z_k+y-2x)+o_k \geq \varphi^{\square}(z_k)-\varphi^{\square}(y) \geq \frac{1}{2}(z_k-y)\cdot P_y(z_k+y-2x)+o_k.$$

Noticing that

$$\lim_k P_y(z_k+y-2x) = 2P_y(y-x)$$

and setting $v := P_y(y-x)$, it follows that

$$(z_k-y)\cdot v + o_k \geq \varphi^{\square}(z_k)-\varphi^{\square}(y) \geq (z_k-y)\cdot v + o_k.$$

This ensures that for any sequence z_k tending to y, one can extract a subsequence such that

$$\varphi^{\square}(z_k)-\varphi^{\square}(y) - v\cdot(z_k-y) = o_k.$$

Since when z_k tends to y, (z_k-y) is equivalent to the geodesic distance from y to z_k in N, this proves that $\varphi^{\square}$ is differentiable and that $\nabla\varphi^{\square}(y)=v$ which completes the proof of Proposition 1.2. □

1.3. **Proof of Theorem 0.1.** We define

$$\Gamma := \{x\in M \,/\, \exists y \in N\,,\, \theta_y(x) = \sup_{x\in M}\theta_y(x) = \varphi^{\square}(y)\}$$

The first observation is the following:

Proposition 1.3. (Properties of the set Γ)

(1) *The set Γ is closed;*
(2) *If $x\in\Gamma$ then $\#\Omega_x = 1$. In particular, if $\Gamma = M$, the support of π is contained in a graph.*
(3) *For all $x\in\Gamma$, one has $\varphi^{\square\square}(x) = \varphi(x) = \varphi^{**}(x)$.*

Proof. Let us first show that Γ is closed: let $(x_n)\subset\Gamma$ be such that $x_n\to x$ in M. There exists $(y_n)\subset N$ such that for all $n\in\mathbb{N}$, $\Theta_{y_n}(x_n) = \max_M \Theta_{y_n}$. Now, let (y_{n_k}) be a subsequence of (y_n) that converges to some $y\in N$. Such subsequence exists by compactness of N. On the one hand, by the continuity of the map $z\mapsto \max_M\Theta_z$, we obtain that $\Theta_{y_{n_k}}(x_{n_k}) = \max_M\theta_{y_n}\to\max_M\Theta_y$. On the other hand, we have $\Theta_{y_{n_k}}(x_{n_k})\to\Theta_y(x)$. Therefore $\Theta_y(x) = \max_M\Theta_y$ and thus $x\in\Gamma$. This proves that Γ is closed.

Les us come to the proof of the second part of the statement. Let $x\in\Gamma$. By definition of Γ, there exists $y_1\in N$ such that x is a maximum for Θ_{y_1}. Assume that $\#\Omega_x\geq 2$ and let also $y_2,y_3\in\Omega_x$, $y_2\neq y_3$. Then, $x\in\Omega_{y_i}$ for $i=2,3$ and then x is a minimum of Θ_{y_i} $i=2,3$. By Equation (1.2), $y_1,y_2,y_3\in D_x$. Since N is the boundary of a strictly convex domain, D_x intersects N at at most two points. Since $y_2\neq y_3$, we must have $y_1=y_2$ or $y_1=y_3$. Let us assume for instance that $y_1=y_2$. This means that x is a minimum as well as a maximum of Θ_{y_1} which forces Θ_{y_1} to be constant on M. By Lemma 1.1, this cannot occur.

Let us prove now the third part of the statement. For all $x\in M$, it holds that $\varphi^{**}(x) = \psi^*(x) = \varphi(x)$. By Proposition 1.2, $\varphi^{\square\square}\leq\varphi$. It thus remains to prove that if $x \in \Gamma$ then $\varphi^{\square\square}(x) \geq \varphi(x)$. For such x, there exists $y \in N$ such that

$$\begin{aligned}\frac{|x-y|^2}{2} - \varphi(x) &= \varphi^{\square}(y) = \varphi^{\square\square\square}(y) \\ &= \sup_{z\in M}\Big(\frac{|z-y|^2}{2} - \varphi^{\square\square}(z)\Big) \\ &\geq \frac{|x-y|^2}{2} - \varphi^{\square\square}(x) \\ &\geq \frac{|x-y|^2}{2} - \varphi(x)\end{aligned}$$

Here, we used the fact that $\varphi^{\square\square\square} = \varphi^{\square}$, which is proven in Proposition 1.2. We then must have equality in all the inequalities above which implies $\varphi^{\square\square}(x) \geq \varphi(x)$. □

We are now in position to define

$$\begin{aligned} T\ :\ \Gamma &\rightarrow N \\ x &\mapsto T(x)\end{aligned}$$

such that $\Theta_{T(x)}(x) = \sup_M \Theta_{T(x)}$. Then,

Proposition 1.4. (Properties of the mapping T) *T is a well defined continuous map which is surjective. Moreover, for all $x \in \Gamma$, the outer unit normal vector to M at x and to N at $T(x)$ satisfy:*

$$n_M(x) \cdot n_N(T(x)) < 0.$$

Proof. To show that T is well defined, we have to show that for all $x \in \Gamma$, there exists one and only one $y \in N$ such that $\Theta_y(x) = \sup_M \Theta_y$. The existence of such y is ensured by the fact that $x \in \Gamma$. Now, assume that y_1 and y_2 satisfy this relation. Then, y_1 and y_2 must belong to the right line D_x (see Relation (1.2)). Moreover, since Ω_x is never empty, let $y_3 \in \Omega_x$ then again $y_3 \in D_x$. Notice that y_3 is distinct from y_1 and y_2 otherwise Θ_{y_1} is constant which is prohibited by Lemma 1.1. Since D_x intersects N at at most two points, y_2 and y_3 must be equal. This prove that T is well defined.

The fact that T is surjective is obvious: if $y \in N$, we choose $x \in M$, which is compact, such that x is a maximum of Θ_y. The definition of T implies that $T(x) = y$. Let us show the continuity of T. Let $(x_n) \subset \Gamma$ such that $x_n \to x$ in Γ. By construction we have $\theta_{T(x_n)}(x_n) = \max_M \theta_{T(x_n)}$. Now, let (x_{n_k}) be a subsequence of (x_n) such that $T(x_{n_k})$ converges to some $y \in N$. Obviously, proceeding as in the proof of Proposition 1.3, $\Theta_y(x) = \max_M \Theta_y$ and thus $y = T(x)$. Therefore $T(x)$ is the unique adherence point of the sequence $(T(x_n))$ which ensures that $T(x_n) \to T(x)$ and proves the continuity of T.

Let us prove the last part of the statement of Proposition 1.4. Let $x \in \Gamma$ and set $y = T(x)$. Since x a maximum for Θ_y, by (1.2), $y \in D_x$. Let also y' be a minimum for Θ_x i.e. $y' \in \Omega_x$. By (1.1), x is also a minimum for $\Theta_{y'}$ which implies, by (1.2), that $y' \in D_x$. Moreover, $y \neq y'$ since otherwise this would imply that Θ_y is constant which would contradict Lemma 1.1. Since N bounds a strictly convex domain, D_x intersects N at at most two points which forces to have

$$D_x \cap N = \{y, y'\}.$$

Note that $n_M(x) \cdot n_N(y) \neq 0$ since otherwise this would imply that D_x is tangent to N and thus intersects N at only one point. Note also that among $n_M(x) \cdot n_N(y)$, $n_M(x) \cdot n_N(y')$, one is positive and the other one is negative. It then suffices to prove that $n_M(x) \cdot n_N(y') > 0$. This fact is proven in [3]: with their notations, $t_+(x) = y'$ satisfies the desired relation. Since the proof is easy we repeat it here for sake of completeness. Let $x' \in \Omega_y$. Then, $x' \neq x$ otherwise we would get that Θ_x is constant. The *monotonicity property* asserts that

$$(x - x') \cdot (y - y') \leq 0. \tag{1.6}$$

Indeed, since $x \in \Omega_{y'}$ and $x' \in \Omega_y$, we have

$$\varphi(x) + \psi(y') = c_2(x, y') \ \text{ and } \varphi(x') + \psi(y) = c_2(x', y).$$

Moreover, by definition of (φ, ψ),

$$\varphi(x) + \psi(y) \leq c_2(x, y) \ \text{ and } \varphi(x') + \psi(y') \leq c_2(x', y').$$

These relations imply that

$$c_2(x', y) + c_2(x, y') \leq c_2(x, y) + c_2(x', y').$$

Coming back to the definition of c_2, we obtain Relation (1.6). Note that since $y, y' \in D_x$, the definition of D_x tells us that $\overrightarrow{yy'} = \lambda n_M(x)$ for some $\lambda \neq 0$. If we assume that $n_M(x) \cdot n_N(y') < 0$ and $n_M(x) \cdot n_N(y) > 0$ then the fact that N bounds a strictly convex domain forces $\lambda < 0$. Therefore, Relation (1.6) becomes $(x - x') \cdot n_M(x) \leq 0$ which is impossible since $x \neq x'$ and since M bounds a strictly convex domain. This ends the proof of Proposition 1.4. □

Note that the preceding proof shows that

Lemma 1.5. *For all $x \in \Gamma$, $D_x \cap N$ has exactly two distinct points y, y' such that x is a maximum for Θ_y and a minimum for $\Theta_{y'}$. Moreover*

$$n_M(x) \cdot n_N(y) < 0 \ \text{ and } \ n_M(x) \cdot n_N(y') > 0.$$

Proposition 1.6. *The following assertions are equivalent:*

1. *$\varphi^\square$ is C^1;*
2. *T is injective;*
3. *for all $y \in M$, Θ_y has exactly two critical points.*

If one of these assertions is true, then $\Gamma = M$.

Proof. Let us show that 1) implies 2). Assume $\varphi^\square$ is C^1. For all $y \in N$, define the right line

$$D^\square_y = y - \nabla \varphi^\square(y) + \text{span}(n_N(y)) \ .$$

An straightforward computation shows that for $x \in M$,

$$x \in D^\square_y \Longleftrightarrow y \text{ is a critical point of } \Theta^\square(\cdot) := c_2(x, \cdot) - \varphi^\square(\cdot). \tag{1.7}$$

Let $x \in \Gamma$. Then x is a maximum of the function Θ_y i.e. $\varphi^\square(y) = \Theta_y(x)$ for some $y \in N$. Now, using the fact that $\varphi^{\square\square\square} = \varphi^\square$ (see Proposition 1.2), one also has

$$\sup_{z \in M} (c_2(z, y) - \varphi^{\square\square}(z)) = \varphi^{\square\square\square}(y) = \varphi^\square(y) = \Theta_y(x). \tag{1.8}$$

On the other hand, by Proposition 1.3, $\varphi(x) = \varphi^{\square\square}(x)$ and hence

$$\Theta_y(x) = c_2(x, y) - \varphi^{\square\square}(x).$$

Together with (1.8), we get that x is a maximum for $z \to c_2(z,y) - \varphi^{\square\square}(z)$. Obviously, mimicking what was done to get (1.1), we also have that y is a maximum for the function of N $z \to c_2(x,z) - \varphi^{\square}(z)$. Relation (1.7) then leads to $x \in D^{\square}_y$. Assume now that $T(x) = T(x')$. We then obtain that $x, x' \in D^{\square}_{T(x)}$. Moreover, Lemma 1.5 also establishes that $n_M(x) \cdot n_N(T(x)) < 0$ and $n_M(x') \cdot n_N(T(x)) < 0$ which forces x and x' to be equal since M bounds a strictly convex domain. This proves that T is injective.

Let us prove that 2) implies 3). At first, we show that under assumption 2), $\Gamma = M$. From Propositions 1.3 and 1.4, $T : \Gamma \to M$ is now bijective, continuous. Since Γ is compact, it sends closed sets on closed sets and thus T is actually a homeomorphism. This ensures that $\Gamma = M$. Indeed, M and N bound some convex domains in $\mathbb{R}^n$ and are then diffeomorphic to S^n. We just proved that Γ is a closed set of M homeomorphic to N and hence to S^n. To prove that $\Gamma = M$, it suffices to notice that it is open in M and to conclude by the fact that M is connected. This follows from the Jordan-Brouwer separation theorem (see for instance [4], Corollay (18.9) Page 110).

We are now in position to prove 3). A consequence of Lemma 1.5 and the fact that $\Gamma = M$ is that for all $(x,y) \in M \times N$ such that $x \in \Omega_y$ then

$$n_M(x) \cdot n_N(y) < 0. \tag{1.9}$$

We already noticed that each Ω_x $(x \in M)$ is reduced to a point (this comes from Proposition 1.3 and the fact that $\Gamma = M$) but this is also true for Ω_y for any $y \in N$. Indeed, if x, z are some minima for Θ_y, they must belong to the right line D_y and they must satisfy (1.9) which is only possible if $x = z$. So, let $y \in N$ and let x be a minimum of Θ_y and x' be a maximum of Θ_y. Assume that Θ_y has some other critical point x''. Then, $y \in D_{x''}$. By Lemma 1.5, x'' must be a maximum or a minimum of Θ_y. The argument above tells us that x'' cannot be a minimum. But x'' cannot be either a maximum: it would imply $T(x) = T(x'')$ which is impossible since we assumed T to be injective. This proves that the only critical points of Θ_y are x, x'.

Finally, we prove that 3) implies 1). Assume that Θ_y has only two critical points for any $y \in N$. Then, for all y, $\Omega^{\square}_y$ is reduced to one point (otherwise, Θ_y has at least two maxima and one minimum). From Point 5) of Proposition 1.2, we obtain that $\varphi^{\square}$ is differentiable on N. It remains to prove that its differential is continuous. Let (y_k) be a sequence of points in N tending to some y. Let $x_k \in \Omega^{\square}_{y_k}$ and $x \in \Omega^{\square}_y$. Let x' be an adherence point of (x_k). Since x_k is a maximum of $z \to c_2(z, y_k) - \varphi(z)$, passing to the limit, x' is a maximum of $z \to c_2(z,y) - \varphi(z)$. This implies that $x' \in \Omega^{\square}_y$, which is reduced to one point. Hence $x' = x$ which shows that x_k tends to x. In particular, the sequence of right lines $(D^{\square}_{y_k})$ (which are orthogonal to the tangent spaces $T_{y_k}N$ and which are such that $x_k \in D^{\square}_{y_k}$) has a limit position which is $D^{\square}_y$. The definition of these right lines gives the continuity at y of the differential of $\varphi^{\square}$. This ends the proof of Proposition 1.6. □

Theorem 0.1 is a direct consequence of Propositions 1.3, 1.4 and 1.6.

References

[1] N. Ahmad, H. K. Kim and R.J. McCann, *Optimal transportation, topology and uniqueness*, Bull. Math. Sci. **1** (2011) 13-32

[2] P.A. Chiappori, R.J. McCann and L.P. Nesheim, *Hedonic price equilibria, stable matching, and opitmal transport: equivalence, topology and uniqueness*, Econom. Theory **42** (2010), 317-354.

[3] W. Gangbo and R.J. McCann, *Shape recognition via Wasserstein distance*, Quart. Appl. Math. **58**, (2000) 705-737.

[4] M.J. Greenberg and J.R. Harper, *Algebraic Topology, a first course*, Mathematic Lecture Note Series, Benjamin/Cummings Publishing Co., Inc., Advanced Book Program, Reading, Mass., 1981

[5] L. Kantorovich, *On the translocations of masses*, C.R. (Doklady) Acad. Sci. URSS (N.S) **37** (1942), 199-201

[6] C. Villani, *Topics on transportations*, AMS, Graduate Studies in Mathematics **58**, 2003.