Optimal transportation for a quadratic cost with convex constraints and applications

C. Jimenez [a], F. Santambrogio [b,*]

[a] *Laboratoire de Mathématiques de Brest, Université de Bretagne Occidentale, 6, Avenue Victor Le Gorgeu, 29200 Brest, France*
[b] *Laboratoire de Mathématiques d'Orsay, Université Paris-Sud, 91405 Orsay cedex, France*

Abstract

We prove existence of an optimal transport map in the Monge–Kantorovich problem associated to a cost $c(x, y)$ which is not finite everywhere, but coincides with $|x - y|^2$ if the displacement $y - x$ belongs to a given convex set C and it is $+\infty$ otherwise. The result is proven for C satisfying some technical assumptions allowing any convex body in $\mathbb{R}^2$ and any convex polyhedron in $\mathbb{R}^d$, $d > 2$. The tools are inspired by the recent Champion–DePascale–Juutinen technique. Their idea, based on density points and avoiding disintegrations and dual formulations, allowed to deal with L^∞ problems and, later on, with the Monge problem for arbitrary norms.
© 2012 Elsevier Masson SAS. All rights reserved.

Résumé

Dans cet article on démontre l'existence d'un transport optimal pour le problème de Monge–Kantorovich associé à un coût $c(x, y)$ qui n'est pas fini partout, mais coïncide avec $|x - y|^2$ si le déplacement $y - x$ appartient à un corps convexe C et est $+\infty$ sinon. Le résultat est démontré sous la condition que C satisfasse des hypothèses techniques qui sont vérifiées par tout convexe de $\mathbb{R}^2$ et par tout polyhèdre de $\mathbb{R}^d$, $d > 2$. Les outils s'inspirent à une technique récente due à Champion–DePascale–Juutinen. Leur idée, qui est basée sur des points de densité et qui n'utilise pas de désintegrations et de formulations duales, a déjà permis de traiter des problèmes L^∞ et, plus tard, le problème de Monge pour des normes arbitraires.
© 2012 Elsevier Masson SAS. All rights reserved.

Keywords: Optimal transportation; Existence of optimal maps; Convex bodies; c-monotonicity

1. Introduction

The optimal transport problem, introduced by G. Monge in [14] at the end of the 18th century, has become very famous in the past 20 years, when the relaxed formulation given by L.V. Kantorovich in the '40s, coupled with the most recent advances, has finally allowed to give a complete solution to many useful issues of the problem. Given
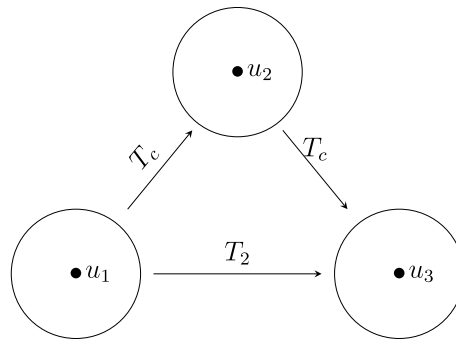
Fig. 1. The situation in Example 1.1.

$f_0 \ll dx$ and f_1 two probability measures on $\mathbb{R}^d$, the issue is to push forward f_0 to f_1 with a map $T : \mathbb{R}^d \to \mathbb{R}^d$ called optimal transport map which solves the following optimization problem:

$$(M) \quad \inf\left\{ \int_{\mathbb{R}^d} c\big(x, T(x)\big) \, df_0(x) \colon T_\sharp f_0 = f_1 \right\},$$

where $c : \mathbb{R}^d \times \mathbb{R}^d \to [0, +\infty]$ is a fixed cost function. The original problem by G. Monge concerned the case where c is the Euclidian distance, but later on many other cost functions have been investigated, and in particular the quadratic one $c(x, y) = |x - y|^2$, which has been popularized by [5]. We will not enter now into the details of the formulation that L.V. Kantorovich gave to overcome the difficulties of the nonlinear behavior of problem (M) since we want to present as soon as possible the particular cost function that we want to consider. In this paper we will actually focus on a quadratic cost with a closed convex constraint $C \subset \mathbb{R}^d$:

$$c(x, y) = \begin{cases} |y - x|^2 & \text{if } y - x \in C, \\ +\infty & \text{otherwise,} \end{cases} \tag{1.1}$$

where we denote by $| \cdot |$ the Euclidian norm.

Solving (M) for this cost seems a natural question, both for its modelization meaning (imagine that C is the unit ball in $\mathbb{R}^d$: in such a case the problem aims at finding the optimal displacement minimizing the total effort, subject to a pointwise constraint $|T(x) - x| \leqslant 1$) and for mathematical applications. Actually, it appears in [7] that a strategy to solve more general transport problems passes through the solutions of some constrained cases, where it is imposed that $T(x) - x$ belongs to a given set, which is often the face of a convex body. The L^∞ problems studied in [10] can also be considered constrained problems (if the minimal L^∞ norm $\|T(x) - x\|_{L^\infty}$ is L, this means that we could consider maps constrained to satisfy $T(x) - x \in \overline{B(0, L)}$). In many of these cases, once the problem is transformed into a convexly constrained one, then it is only necessary to find an arbitrary transport map satisfying such a constraint. The minimization of the quadratic cost among them is just an additional criterion to select a special one, and it is chosen since it usually guarantees better properties. We refer anyway to [7] for a more detailed discussion about convex costs in general and quadratic costs with convex constraints.

To illustrate the difference between our setting and the usual quadratic case $c(x, y) = |x - y|^2$, we give an example.

Example 1.1. Let u_i, with $i = 1, 2, 3$ three points of the plan with the following distances: $|u_1 - u_2| = |u_2 - u_3| = 1$ and $|u_1 - u_3| = 5/4$. Take $B(u_i, \varepsilon)$ with $\varepsilon > 0$ and $i = 1, 2, 3$ three tiny balls and set f_0 and f_1 absolutely continuous with respect to the Lebesgue measure with densities $1_{B(u_1, \varepsilon)} + 1_{B(u_2, \varepsilon)}$ and $1_{B(u_2, \varepsilon)} + 1_{B(u_3, \varepsilon)}$ respectively. The unique optimal transport plan for the usual quadratic cost is $T_2(x) = x + (u_3 - u_1)$ if $x \in B(u_1, \varepsilon)$ and the identity elsewhere, but this map is not admissible for the cost c defined by (1.1) when C is the unit ball. The unique optimal transport plan for this cost is hence $T_c(x) = x + (u_{i+1} - u_i)$ if $x \in B(u_i, \varepsilon)$ for $i = 1, 2$. (See Fig. 1.)

As we said, the quadratic cost has often better properties than other general cost functions, and it is not astonishing that the first existence result for (M) was actually proven for $c(x, y) = |y - x|^2$. Actually, Y. Brenier proved in 1991, existence and uniqueness of an optimal map T, and the crucial point in the proof is the existence of a dual formulation

$$\sup_{(u,v)\in\mathcal{C}(\mathbb{R}^d)\times\mathcal{C}(\mathbb{R}^d)}\left\{\int_{\mathbb{R}^d} v(y)\,df_1(y) - \int_{\mathbb{R}^d} u(x)\,df_0(x)\colon v(y) - u(x) \leqslant c(x,y)\right\},$$

where the sup is actually achieved. This allowed Brenier to use the primal–dual optimality condition and prove the optimality of the map

$$T(x) = x - \frac{1}{2}\nabla u(x).$$

This scheme of proof was successfully extended to the cases $c(x,y) = |x-y|^p$ with $p > 1$ (or, more generally, to any strictly convex function of $x - y$) by W. Gangbo and R.J. McCann [13], and by L.A. Caffarelli [6]. The original problem by Monge ($p = 1$) happens to be more difficult and needed other arguments and techniques. It was solved later (see L.C. Evans and W. Gangbo [11], N.S. Trudinger and X.J. Wang [16], L. Ambrosio and A. Pratelli [3], M. Feldman R.J. McCann [12]). The existence for the dual problem was also crucial in these proofs.

When we add a constraint, the existence of an optimal pair of functions for the dual formulation (D) is not guaranteed: the usual proofs of this result require continuity or finiteness of the cost (see the books of C. Villani [17] and [18]). Some very irregular pairs of solutions (not even functions) were found for non-finite Borel costs by M. Beiglböck, C. Léonard and W. Schachermayer [4]. Let us mention also that under the additional restrictive assumption $\int c(x,y)\,df_0(x)\,df_1(y) < +\infty$, W. Schachermayer and J. Teichmann proved the existence of an optimal pair of Borel functions for (D) [15]. Hence, though very interesting, these results cannot be used in our setting.

Nevertheless, more recently, T. Champion, L. De Pascale and P. Juutinen developed a completely different idea and manage to solve the "L^∞ case" (see [10]). It is the case where, instead of minimizing the integral $\int |T(x) - x|^p\,df_0(x)$, i.e. the L^p norm, they considered the minimization of the L^∞ norm (as we mentioned before). This frame is very peculiar because, due to its non-integral formulation, there is no dual formulation for the Monge problem. Hence, they developed a completely different idea, which was powerful enough to be extended later on to some cases where the cost is still integral, but for a function $c(x,y)$ which is not strictly convex in $x - y$. In particular, this gave the proof of the existence of a solution T to Monge's problem for $c(x,y) = \|x - y\|$ for a general norm ([8] and [9]), which was unsolved before.

In this paper, we show that we can solve (M) with c defined by (1.1) using the same tools as T. Champion, L. De Pascale and P. Juutinen.

In Section 2, we state the main results. In Section 3, we make some comments and give a proof in dimension 1. The general proof is given in Section 4.2 after proving some preliminary results in Section 4.1. Finally, in the last section, we give an application of our result to the case of crystalline norms and to some L^∞ problems.

2. Main result

In all the following, the cost c is defined by (1.1), f_0 and f_1 are two probability measures on $\mathbb{R}^d$ and f_0 is absolutely continuous with respect to the Lebesgue measure. Let us introduce the relaxed formulation of the Monge problem introduced by Kantorovich. This new problem, (MK), satisfies $\inf(MK) = \inf(M)$ (see [17] and [18]):

$$(MK) \quad \inf\left\{\int_{\mathbb{R}^d\times\mathbb{R}^d} c(x,y)\,d\pi(x,y)\colon \pi \in \Pi(f_0, f_1)\right\},$$

where $\Pi(f_0, f_1)$ is the set of probability measure on $\mathbb{R}^d \times \mathbb{R}^d$ with fixed marginals f_0 and f_1. The following classical result holds:

Proposition 2.1. *If there exists $\pi \in \Pi(x, y)$ such that $\int_{\mathbb{R}^d\times\mathbb{R}^d} c(x,y)\,d\pi(x,y)$ is finite, then (MK) has a solution.*

The classical strategy to solve (M) is to show that at least a solution γ of (MK) is concentrated on the graph of a map T i.e.:

$$\int \varphi(x,y)\,d\gamma(x,y) = \int \varphi\big(x, T(x)\big)\,df_0(x) \quad \forall \varphi \in \mathcal{C}_c(\mathbb{R}^d).$$

Indeed, the map T is admissible for (M) if γ is admissible for (MK) and it is optimal if γ is optimal. Conversely for any T admissible for (M), $\gamma = (\mathrm{Id} \times T)_\sharp f_0$ is admissible for (MK), and if T is optimal, γ is optimal too.

In order to state the main theorem of this paper, we need to give some definitions on the geometry of convex sets.

Definition 2.2. Consider a convex set $C \subset \mathbb{R}^d$.

1. If x_0 is any point of C, we call *dimension* of C the dimension of $\mathrm{Span}(C - x_0)$, i.e. the dimension of the smallest affine space containing C. This number is denoted $\dim(C)$ and does not depend on the choice of x_0. The interior and boundary of C in the canonical topology associated to the Euclidian distance in such an affine space are called relative interior and *relative boundary* and denoted by $\mathrm{ri}\,C$ and $\mathrm{r}\partial C$.
2. A convex subset $F \subset C$ containing more than one point is called a *straight part* of C if it is contained in relative boundary $\mathrm{r}\partial C$.
3. A subset F of a convex set C is called a *maximal flat part* of C if it is maximal for the inclusion among straight parts of C.
4. A subset F of a convex set C is called a *flat part* of C if it there is a finite chain of convex sets $F = F_1 \subset F_2 \subset \cdots \subset F_{k-1} \subset F_k = C$ ($k \geqslant 2$) such that, for each $i = 1, \ldots, k-1$, F_i is a maximal flat part of F_{i+1}. Notice that in this case one has $\dim(F_1) < \dim(F_2) < \cdots < \dim(F_k)$. In particular, for every flat part F of C we have $1 \leqslant \dim(F) \leqslant (d-1)$.
5. A set $C \subset \mathbb{R}^d$ is said to be *strictly convex* if it has no flat parts.

The aim of this paper is to prove the following result:

Theorem 2.3. *Assume that*:

(i) *there exists $\pi \in \Pi(f_0, f_1)$ such that $\int_{\mathbb{R}^d \times \mathbb{R}^d} c(x, y)\, d\pi(x, y) < +\infty$,*
(ii) *C has at most a countable number of flat parts.*

Then, if f_0 is absolutely continuous with respect to the Lebesgue measure on $\mathbb{R}^d$, any solution γ of (MK) is concentrated on the graph of a map T which is optimal for (M).

Remark 2.4. Assumption (ii) is obviously satisfied if C is strictly convex. Moreover, if C satisfies (ii), then any flat part of C is a convex subset of an affine space $\mathbb{R}^k$ for $k < d$ which satisfies the same assumption. Finally, this assumption allows every bounded convex set in $\mathbb{R}^2$ (because flat parts should be subsets of positive length of its boundary, which has finite perimeter, and hence they cannot be more than countably many) and all compact polytopes in any dimension.

Corollary 2.5. *Under the assumption of Theorem 2.3, the solutions of (MK) and (M) are unique.*

Proof. The idea is classical: let γ_1 and γ_2 be two solutions of (MK), let T_1 and T_2 the associated transport maps (given by Theorem 2.3). The probability measure $\frac{\gamma_1 + \gamma_2}{2}$ is also associated to a transport map T_3. Then for any $\varphi \in \mathcal{C}_c(\mathbb{R}^d \times \mathbb{R}^d)$:

$$1/2\left(\int_{\mathbb{R}^d} \varphi\big(x, T_1(x)\big)\, df_0(x) + \int_{\mathbb{R}^d} \varphi\big(x, T_2(x)\big)\, df_0(x) \right) = \int_{\mathbb{R}^d} \varphi\big(x, T_3(x)\big)\, df_0(x).$$

This implies $f_0(\{x: T_1(x) \neq T_2(x)\}) = 0$, that is $\gamma_1 = \gamma_2$ and $T_1 = T_2$ almost everywhere. □

3. Proof and comments in dimension 1

The one-dimensional case is a very special one. Actually, for a very wide class of transport costs, the optimal transport is always characterized in the same way and is always the same, namely the unique nondecreasing transport map sending f_0 to f_1. This is summarized in the following theorem (the proof is written here in its full generality, even if it is essentially taken from [1]):

Theorem 3.1. *Let $\phi : \mathbb{R} \to \mathbb{R} \cup \{+\infty\}$ be an l.s.c. strictly convex function (which actually means strictly convex on its domain $\{\phi < +\infty\}$), and f_0 and f_1 two probability measures on $\mathbb{R}$, with f_0 non-atomic. Then the optimal transport problem*

$$\inf\left\{ \int_{\mathbb{R}^2} \phi(y - x)\, d\pi(x, y) \colon \pi \in \Pi(f_0, f_1) \right\}$$

has a unique solution, which is given by $\gamma = (\mathrm{Id} \times T)_\# f_0$, where T is the nondecreasing transport map from f_0 to f_1, which is well defined f_0-a.e.

Moreover, if the strict convexity assumption is dropped and ϕ is only convex, then the same T is actually an optimal transport map, but no uniqueness is guaranteed anymore.

Proof. We will use the usual strategy based on c-cyclical monotonicity, i.e. the fact that any optimal γ is concentrated on a cyclical monotone set Γ, where we can also assume $\phi(y - x) < +\infty$. This means that $(x, y), (x', y') \in \Gamma$ implies

$$\phi(y - x) + \phi(y' - x') \leqslant \phi(y' - x) + \phi(y - x'). \tag{3.1}$$

We only need to show that this implies (in the strictly convex case) a monotone behavior: we will actually deduce from (3.1) that $x < x'$ implies $y \leqslant y'$. This allows to say that Γ is included in the graph of a monotone (nondecreasing) multifunction. Hence, up to a countable set, for every x there is unique image $y = T(x)$, and the map T is nondecreasing.

To prove $y \leqslant y'$ suppose by contradiction $y > y'$ and denote $a = y - x$, $b = y' - x'$ and $c = x' - x$. Condition (3.1) reads as $\phi(a) + \phi(b) \leqslant \phi(b + c) + \phi(a - c)$. Moreover, we have

$$b + c = (1 - t)b + ta, \qquad a - c = tb + (1 - t)a, \quad \text{for } t = \frac{c}{a - b}.$$

The assumption $y' < y$ reads as $b + c < a$, which gives $t \in \,]0, 1[$ (since it implies $b < a$, and hence $t > 0$, and $c < a - b$, and hence $t < 1$). Thus, convexity yields

$$\phi(a) + \phi(b) \leqslant \phi(b + c) + \phi(a - c)$$
$$\leqslant (1 - t)\phi(b) + t\phi(a) + t\phi(b) + (1 - t)\phi(a) = \phi(a) + \phi(b).$$

Since we assumed $\phi(a), \phi(b) < +\infty$, we also deduce $\phi(b + c), \phi(a - c) < +\infty$ and the strict convexity implies a strict inequality, thus getting to a contradiction.

This shows that a strictly convex cost always admits a unique optimal transport plan, which is given by the monotone nondecreasing transport map. The statement when ϕ is only convex is obtained by approximation through $\phi(y - x) + \varepsilon |y - x|^2$ (or through any other strictly convex approximation). $\square$

It is important to notice that in this one-dimensional convex framework the optimal transport does not really depend on the cost. This is very useful for approximation procedures (if, for instance, one considers costs of the form $|y - x|^p$ and lets $p \to \infty$ it is possible to deduce the optimality of the same nondecreasing transport for the L^∞ transportation problem, see again [10]). Not only, in our case when $\phi(z) = |z|^2$ for $z \in C$ and $\phi = +\infty$ outside C, this shows that the optimal transport is not really affected by the constraint (differently from the example in $\mathbb{R}^2$ that we gave in the introduction). What is affected by the constraint is the possibility of transporting f_0 onto f_1 *at finite cost*. Depending on the two measures, it is possible that the nondecreasing map T satisfies $T(x) - x \in C$ or not. If yes, then it is optimal; if not, then the problem has no solution with a finite cost.

4. Proof of Theorem 2.3

4.1. Preliminary results

Let γ be an optimal transport map for (MK). We recall that when the cost is l.s.c. (which is the case of our cost function c), then any optimal γ is concentrated on a c-cyclically monotone set Γ. This means in particular

$$(x_0, y_0), (x_1, y_1) \in \Gamma \quad \Rightarrow \quad c(x_0, y_0) + c(x_1, y_1) \leqslant c(x_1, y_0) + c(x_0, y_1)$$

(we refer again to [1] for the actual definition of c-cyclically monotone sets, which imposes a similar condition for families of k points in Γ and arbitrary permutations of their second coordinates; here we only need the condition for $k = 2$). In our case, thanks to the definition of c, if we use the equivalence

$$|x_0 - y_0|^2 + |x_1 - y_1|^2 \leqslant |x_0 - y_1|^2 + |x_1 - y_0|^2 \quad \Leftrightarrow \quad \langle x_1 - x_0, y_1 - y_0 \rangle \geqslant 0$$

then we get

$$\left. \begin{array}{l} \text{If } (x_0, y_0), \ (x_1, y_1) \in \Gamma \text{ are such that:} \\ y_0 - x_1 \in C \text{ and } y_1 - x_0 \in C \end{array} \right\} \quad \Rightarrow \quad \langle y_1 - y_0, x_1 - x_0 \rangle \geqslant 0. \tag{4.1}$$

Without loss of generality, we may assume that Γ is σ-compact (which is useful for measurability reasons) and that any $(x, y) \in \Gamma$ is such that $y - x \in C$.

Arguing as in [10], we introduce the following object:

Definition 4.1. For any $r > 0$ and any $y \in \mathbb{R}^d$, we set:

$$\Gamma^{-1}\big(\overline{B(y, r)}\big) := \pi_x\big(\Gamma \cap \big(\mathbb{R}^d \times \overline{B(y, r)}\big)\big)$$

where $B(y, r)$ denotes the open ball of radius $r > 0$ centered at y. The assumption of σ-compactness on Γ guarantees that the set that we just defined is measurable.

We recall the following results extracted from [10] and [8]:

Lemma 4.2. *The measure γ is concentrated on a σ-compact set $R(\Gamma)$ such that for all (x_0, y_0) in $R(\Gamma)$ and for all $r > 0$, the point x_0 is a Lebesgue point of $\Gamma^{-1}(B(y_0, r))$, i.e.:*

$$\lim_{\varepsilon \to 0} \frac{|\Gamma^{-1}(\overline{B(y_0, r)}) \cap B(x_0, \varepsilon)|}{|B(x_0, \varepsilon)|} = 1.$$

Proposition 4.3. *Let $(x_0, y_0) \in R(\Gamma)$, let $r > 0$, $\delta \in \,]0, 1[$ and ξ a unit vector in $\mathbb{R}^d$, then for any $\varepsilon > 0$ the following set has a positive Lebesgue measure:*

$$\Gamma^{-1}\big(\overline{B(y_0, r)}\big) \cap B(x_0, \varepsilon) \cap C(x_0, \xi, \delta),$$

where the set $C(x_0, \xi, \delta)$ is the following convex cone:

$$\big\{ x \colon \langle x - x_0, \xi \rangle > (1 - \delta)|x - x_0| \big\}$$

(notice that it is an open cone, with $x_0 \notin C(x_0, \xi, \delta)$).

Lemma 4.4. *Let $x \in C$, ξ a unit vector $\eta > 0$ such that $x + \eta\xi \in \mathrm{Int}(C)$. Let $0 < \eta' < \eta$ be fixed. Then there exist $r > 0$ and $\delta \in \,]0, 1[$ such that:*

$$\forall y \in B(x, r) \cap C, \quad \big(C(y, \xi, \delta) \cap \overline{B(y, \eta')}\big) \subset \mathrm{Int}(C).$$

The direction ξ is called an *entering direction* at the point x.

Proof. First of all notice that, if $x + \eta\xi \in \mathrm{Int}(C)$, then all the interior points of the segment connecting x to $x + \eta\xi$ must belong to $\mathrm{Int}(C)$.

We prove the existence of $r > 0$ and $\delta \in \,]0, 1[$ such that:

$$\forall \, y \in B(x, r) \cap C, \quad \big(C(y, \xi, \delta) \cap \big[\overline{B(y, \eta')} \setminus B(y, \eta'/2)\big]\big) \subset \mathrm{Int}(C). \tag{4.2}$$

If we prove (4.2), the result then follows by passing to the convex envelope. The thesis (4.2) may be proven by contradiction. Suppose that it is false, and find a sequence $y_n \in C$ such that $y_n \to x$, together with a sequence $z_n \notin \mathrm{Int}(C)$ and $z_n \in (C(y_n, \xi, \delta_n) \cap [\overline{B(y_n, \eta')} \setminus B(y_n, \eta'/2)])$, for $\delta_n \to 0$. Up to subsequence we may suppose $z_n \to z$. This limit point z must satisfy $z \notin \mathrm{Int}(C)$ (since $\mathrm{Int}(C)$ is open) and $|z - x| \in [\eta'/2, \eta']$. Moreover, from $\langle z_n - y_n, \xi \rangle \geqslant (1 - \delta_n)|z_n - y_n|$, we get $\langle z - x, \xi \rangle \geqslant |z - x|$, i.e. $z - x = \lambda\xi$ for $\lambda \geqslant 0$. Then, we get $\lambda \in [\eta'/2, \eta']$ and finally we obtain a contradiction since z belongs to the interior of the segment connecting x to $x + \eta\xi$, but not to $\mathrm{Int}(C)$. $\quad\square$

4.2. The proof in dimension $d > 1$

We start proving a fundamental lemma:

Lemma 4.5. *Let (x_0, y_0) and (x_0, y_0') two points of $R(\Gamma)$. We assume*

$$\left\{(1 - t)(y_0 - x_0) + t(y_0' - x_0): t \in [0, 1]\right\} \cap \mathrm{Int}(C) \neq \emptyset.$$

Then $y_0 = y_0'$.

Proof. To show this lemma, we proceed by contradiction and adapt the method developed in [10] to our case. Let (x_0, y_0) and (x_0, y_0') two points of $R(\Gamma)$ satisfying the lemma's assumption with $y_0 \neq y_0'$.

We set $\xi := \frac{y_0' - y_0}{|y_0' - y_0|}$. The direction ξ is an entering direction at $y_0 - x_0$ and $-\xi$ is an entering direction at $y_0' - x_0$. Using Lemma 4.4, we take $r > 0$, $\rho \in {]0, 1[}$ and $\varepsilon > 0$, such that:

$$\left(C(y_0' - x_0, -\xi, \rho) \cap \overline{B(y_0' - x_0, \varepsilon)}\right) \subset \mathrm{Int}(C), \tag{4.3}$$

$$\forall z \in B(y_0 - x_0, r + \varepsilon), \quad \left(C(z, \xi, \rho) \cap \overline{B(z, \varepsilon)}\right) \subset \mathrm{Int}(C). \tag{4.4}$$

Moreover, if ρ and r are small enough, we also have

$$\forall x \in C(x_0, \xi, \rho), \ \forall y \in \overline{B(y_0, r)}: \quad \langle x - x_0, y - y_0'\rangle < 0. \tag{4.5}$$

Indeed the direction of $x - x_0$ is almost that of ξ (up to an error ρ) and that of $y - y_0'$ almost $-\xi$ (up to an error of order $r/|y_0 - y_0'|$). The two vectors being opposite, it is clear that if both δ and r are small then the scalar product stays negative. By use of Proposition 4.3, we get the existence of a couple $(x, y) \in \Gamma$ such that $x \in \overline{B(x_0, \varepsilon)} \cap C(x_0, \xi, \rho)$ and $y \in \overline{B(y_0, r)}$.

Then the couple $(x, y) \in \Gamma$ satisfies:

- $y_0' - x = y_0' - x_0 + (x_0 - x) \in C$, as a consequence of (4.3),
- $\langle x - x_0, y - y_0'\rangle < 0$, by (4.5).

To get a contradiction with (4.1), it remains only to show $y - x_0 \in C$. This comes from (4.4). Indeed $y - x = (y_0 - x_0) + (x_0 - x) + (y - y_0)$ is in $B(y_0 - x_0, r + \varepsilon)$ and $x - x_0 \in (C(0, \xi, \rho) \cap \overline{B(0, \varepsilon)})$ so that

$$y - x_0 = (y - x) + (x - x_0) \in \mathrm{Int}\, C. \qquad \square$$

We easily deduce

Corollary 4.6. *If C is strictly convex, then any solution of (MK) is supported on a graph.*

Moreover, if C is not strictly convex, any solution of (MK) is supported on a set $R(\Gamma)$ satisfying the following property:

If (x_0, y_0) and (x_0, y_0') are in $R(\Gamma)$, then $(y_0 - x_0)$ and $(y_0' - x_0)$ belong to a same flat part of C.

To prove Theorem 2.3 itself, we proceed by induction on the dimension d of the space. The theorem in dimension $d = 1$ has already been proved. Let us show that if the theorem holds in any dimension lower or equal to $d - 1$, it also holds in dimension d.

Denote by C_0 the set of points of C which do not belong to any flat part of C, and by $(C_i)_{i \in I}$ the family of the maximal flat parts of C. Let γ an optimal transport plan and $R(\Gamma)$ defined as previously so that γ is concentrated on $R(\Gamma)$. We set:

$$\Gamma_i := \left\{(x, y): y - x \in C_i\right\} \cap R(\Gamma), \quad \forall i \in I \cup \{0\}.$$

By Lemma 4.5, the measure $\gamma \llcorner \Gamma_0$ is concentrated on a graph, and, for any $i \in I$, we have the following property:

$$(x, y) \in \Gamma_i \quad \Rightarrow \quad \left(\left\{(x, y'): (x, y') \in R(\Gamma)\right\} \subset \Gamma_i\right).$$

Then it is enough to show that for any $i \in I$, $\gamma_i := \gamma \llcorner \Gamma_i$ is concentrated on a graph and to glue the results together. Each measure γ_i is an optimal transport plan for the same cost c between its marginals. The assumption that I is at most countable allows to decompose γ into a sum of these measures, thus avoiding the need for disintegration results. In particular, it is clear that the key assumption that f_0 is absolutely continuous stays true for the first marginal of γ_i (which would not be easy in case of a more-than-countable disintegration).

Let us fix $i \in I$, let $n \leqslant d - 1$ be the dimension of C_i and z a fixed vector in C_i. Up to a change of variables, we may assume

$$\mathrm{Span}(C_i - z) = \mathrm{Span}(e^1, \ldots, e^n)$$

where $(e^1, \ldots, e^d)$ is the canonic base of $\mathbb{R}^d$.

Notations. We denote:

$$\mathbf{x}_1 := (x^1, \ldots, x^n), \qquad \mathbf{x}_2 := (x^{n+1}, \ldots, x^d), \qquad \mathbf{z}_2 = (z^{n+1}, \ldots, z^d),$$
$$\hat{\gamma} := \gamma \llcorner \Gamma_i, \qquad \hat{f}_0 := \pi_x \sharp \hat{\gamma}, \qquad \hat{f}_1 := \pi_y \hat{\gamma}.$$

Note that $y - x \in C_i$ implies $\mathbf{y}_2 - \mathbf{x}_2 = \mathbf{z}_2$ so $\hat{\gamma}$ is concentrated on

$$\{(x, (\mathbf{y}_1, \mathbf{x}_2 + \mathbf{z}_2)) \colon x = (\mathbf{x}_1, \mathbf{x}_2) \in \mathbb{R}^d, \ \mathbf{y}_1 \in \mathbb{R}^d\}.$$

Making a disintegration of $\hat{\gamma}$, $\hat{f}_0$, $\hat{f}_1$ we write:

$$\hat{\gamma}(x, (\mathbf{y}_1, \mathbf{x}_2 + \mathbf{z}_2)) = \hat{\mu}^{\mathbf{x}_2}(\mathbf{x}_1, \mathbf{y}_1) \otimes \zeta(\mathbf{x}_2, \mathbf{y}_2), \quad \text{with } \zeta := \pi_{\mathbf{x}_2, \mathbf{y}_2} \sharp \hat{\gamma},$$
$$\hat{f}_0 = \hat{f}_0^{\mathbf{x}_2}(\mathbf{x}_1) \otimes (\pi_{\mathbf{x}_2} \sharp \hat{f}_0)(\mathbf{x}_2), \qquad \hat{f}_1 = \hat{f}_1^{\mathbf{y}_2}(\mathbf{y}_1) \otimes (\pi_{\mathbf{y}_2} \sharp \hat{f}_1)(\mathbf{y}_2).$$

Notice moreover that the measure ζ is concentrated on the pairs of the form $(\mathbf{x}_2, \mathbf{x}_2 + \mathbf{z}_2)$. Then we have:

$$\int c(x, y) \, d\hat{\gamma}(x, y) = \int |\mathbf{x}_1 - \mathbf{y}_1|^2 \, d\hat{\gamma}(x, y) + |\mathbf{z}_2|^2 \gamma(\Gamma_i)$$
$$= \int |\mathbf{x}_1 - \mathbf{y}_1|^2 \, d\hat{\mu}^{\mathbf{x}_2}(\mathbf{x}_1, \mathbf{y}_1) \, d\zeta(\mathbf{x}_2, \mathbf{x}_2 + \mathbf{z}_2) + |\mathbf{z}_2|^2 \gamma(\Gamma_i).$$

We set:

$$\hat{c}(\mathbf{x}_1, \mathbf{y}_1) = \begin{cases} |\mathbf{x}_1 - \mathbf{y}_1|^2 & \text{if } (\mathbf{y}_1 - \mathbf{x}_1, \mathbf{z}_2) \in C_i, \\ +\infty & \text{otherwise.} \end{cases}$$

Lemma 4.7. *ζ-almost every $(\mathbf{x}_2, \mathbf{x}_2 + \mathbf{z}_2)$, we have*:

(i) *$\hat{f}_0^{\mathbf{x}_2} = \pi_{\mathbf{x}_1} \sharp \hat{\mu}^{\mathbf{x}_2}$, and $\hat{f}_1^{\mathbf{y}_2} = \pi_{\mathbf{y}_1} \sharp \hat{\mu}^{\mathbf{x}_2}$, moreover $\hat{f}_0^{\mathbf{x}_2}$ is absolutely continuous with respect to the Lebesgue measure,*
(ii) *the measure $\hat{\mu}^{\mathbf{x}_2}$ on $\mathbb{R}^d \times \mathbb{R}^d$ is a solution of the following transport problem*:

$$(P^{\mathbf{x}_2}) \quad \min \left\{ \int \hat{c}(\mathbf{x}_1, \mathbf{y}_1) \, d\pi(\mathbf{x}_1, \mathbf{y}_1) \colon \pi \in \Pi(\hat{f}_0^{\mathbf{x}_2}, \hat{f}_1^{\mathbf{y}_2}) \right\}.$$

Then using Theorem 2.3 in dimension $n \leqslant d - 1$, we get that, for ζ-almost every $(\mathbf{x}_2, \mathbf{x}_2 + \mathbf{z}_2)$ the measure $\hat{\mu}^{\mathbf{x}_2}$ is concentrated on a graph. Then, recalling that ζ is concentrated on $\{(\mathbf{x}_2, \mathbf{x}_2 + \mathbf{z}_2)\}$, we get the desired result.

Proof of Lemma 4.7. (i) Let us show $\hat{f}_0^{\mathbf{x}_2} = \pi_{\mathbf{x}_1} \sharp \hat{\mu}^{\mathbf{x}_2}$. On the one hand, we have:

$$\hat{f}_0 = \pi_{\mathbf{x}_1, \mathbf{x}_2} \sharp \hat{\gamma} = \pi_{\mathbf{x}_1, \mathbf{x}_2} \sharp (\hat{\mu}^{\mathbf{x}_2}(\mathbf{x}_1, \mathbf{y}_1) \otimes \zeta(\mathbf{x}_2, \mathbf{y}_2)) = (\pi_{\mathbf{x}_1} \sharp \hat{\mu}^{\mathbf{x}_2}) \otimes (\pi_{\mathbf{x}_2} \sharp \zeta(\mathbf{x}_2, \mathbf{y}_2)).$$

On the other hand:

$$\pi_{\mathbf{x}_2} \sharp \zeta(\mathbf{x}_2, \mathbf{y}_2) = \pi_{\mathbf{x}_2} \sharp \pi_{\mathbf{x}_2, \mathbf{y}_2} \sharp \hat{\gamma}(x, y) = \pi_{\mathbf{x}_2} \sharp \hat{f}_0.$$

Then, by uniqueness of the disintegration $\hat{f}_0^{\mathbf{x}_2}$ of $\hat{f}_0$ outside a $\pi_{\mathbf{x}_2} \sharp f_0$ negligible set, we get the result.

In the same way, we can show $\hat{f}_1^{\mathbf{y}_2} = \pi_{\mathbf{y}_1} \sharp \hat{\mu}^{\mathbf{x}_2}$. Moreover as f_0 is absolutely continuous with respect to Lebesgue, also is $\hat{f}_0^{\mathbf{x}_2}$ (with respect to the n-dimensional Lebesgue measure, instead of the d-dimensional one).

(ii) We proceed by contradiction. Assume there exists $E \subset \{(\mathbf{x}_2, \mathbf{x}_2 + \mathbf{z}_2)\}$ not negligible such that, for all $(\mathbf{x}_2, \mathbf{x}_2 + \mathbf{z}_2) \in E$, the measure $\hat{\mu}^{\mathbf{x}_2}$ is not optimal for $(P^{\mathbf{x}_2})$. For any $(\mathbf{x}_2, \mathbf{x}_2 + \mathbf{z}_2) \in E$, let us pick another measure $\bar{\mu}^{\mathbf{x}_2}$ chosen so as to be optimal for $(P^{\mathbf{x}_2})$. We set:

$$\bar{\gamma} = \bar{\mu}^{\mathbf{x}_2} \otimes \zeta \mathbin{\llcorner} E + \hat{\mu}^{\mathbf{x}_2} \otimes \zeta \mathbin{\llcorner} E^c$$

where E^c is the complementary of E in $\tilde{C}_i$. Then $\bar{\gamma} \in \Pi(\hat{f}_0, \hat{f}_1)$, indeed:

$$\begin{aligned}
\pi_x \sharp \bar{\gamma} &= \pi_x \sharp \big(\bar{\mu}^{\mathbf{x}_2} \otimes \zeta \mathbin{\llcorner} E\big) + \pi_x \sharp \big(\hat{\mu}^{\mathbf{x}_2} \otimes \zeta \mathbin{\llcorner} E^c\big) \\
&= \pi_{\mathbf{x}_1} \sharp \bar{\mu}^{\mathbf{x}_2} \otimes \pi_{\mathbf{x}_2} \sharp \zeta \mathbin{\llcorner} E + \pi_{\mathbf{x}_1} \sharp \hat{\mu}^{\mathbf{x}_2} \otimes \pi_{\mathbf{x}_2} \sharp \zeta \mathbin{\llcorner} E^c \\
&= \hat{f}_0^{\mathbf{x}_2} \otimes \big(\pi_{\mathbf{x}_2} \zeta \mathbin{\llcorner} E + \pi_{\mathbf{x}_2} \sharp \zeta \mathbin{\llcorner} E^c\big) \\
&= \hat{f}_0^{\mathbf{x}_2} \otimes \pi_{\mathbf{x}_2} \zeta = \hat{f}_0,
\end{aligned}$$

the same holds for the y marginal. Moreover:

$$\begin{aligned}
\int c(x, y) \, d\bar{\gamma}(x, y) &= \iint_E |\mathbf{x}_1 - \mathbf{y}_1|^2 \, d\bar{\mu}^{\mathbf{x}_2}(\mathbf{x}_1) \, d\zeta(\mathbf{x}_2, \mathbf{x}_2 + \mathbf{z}_2) \\
&\quad + \iint_{E^c} |\mathbf{x}_1 - \mathbf{y}_1|^2 \, d\hat{\mu}^{\mathbf{x}_2}(\mathbf{x}_1) \, d\zeta(\mathbf{x}_2, \mathbf{x}_2 + \mathbf{z}_2) + |\mathbf{z}_2|^2 \gamma(\hat{C}_i) \\
&< \iint_E |\mathbf{x}_1 - \mathbf{y}_1|^2 \, d\hat{\mu}^{\mathbf{x}_2}(\mathbf{x}_1) \, d\zeta(\mathbf{x}_2, \mathbf{x}_2 + \mathbf{z}_2) \\
&\quad + \iint_{E^c} |\mathbf{x}_1 - \mathbf{y}_1|^2 \, d\hat{\mu}^{\mathbf{x}_2}(\mathbf{x}_1) \, d\zeta(\mathbf{x}_2, \mathbf{x}_2 + \mathbf{z}_2) + |\mathbf{z}_2|^2 \gamma(\hat{C}_i) \\
&= \int c(x, y) \, d\hat{\gamma}(x, y).
\end{aligned}$$

This is impossible because $\hat{\gamma}$ is optimal for the transport problem with cost c between its marginals $\hat{f}_0$ and $\hat{f}_1$. $\square$

5. An application to crystalline norms and L^∞ problems

We present in this section how the results of this paper may be applied, following the strategy presented in [7], to the problem of the existence of optimal transport maps when the cost is a crystalline norm $\|x - y\|$. We recall that a norm is called crystalline if its unit ball is a convex polytope containing 0 in its interior with a finite number of faces; In particular, this means that it has the form $\|z\| = \max_i z \cdot v_i$ for a certain finite set of vectors $(v_i)_{i=1,\dots,k}$. This case has been first studied in [2] and was one of the first steps towards the extension of the Monge problem to general norms. The existence of an optimal map for norm costs has finally been established in full generality in [8].

The strategy proposed in [7] concerns the minimization of a transport cost $\int c(x - y) \, d\pi$ for a convex, but not necessarily strictly convex, function c. In the same spirit of the proof of Theorem 2.3, a decomposition may be performed on an optimal transport plan γ, according to the "faces" of the cost. In this case, since the crystalline norm has a finite number of faces F_i (which are the cones on the maximal flat parts of its unit ball), one can write $\gamma = \sum_i \gamma_i$ where the measures γ^i are obtained in the following way. If $f_0 \ll \mathcal{L}^d$, duality implies (see [7]) that for a.e. x there is an index $i = i(x)$ such that $(x, y) \in \mathrm{spt}(\gamma)$ implies $x - y \in F_i$ (in case of non-uniqueness of this index just pick one at random, for instance by fixing an order relation on the faces). Then define $\gamma_i = \gamma \mathbin{\llcorner} \{(x, y): i(x) = i\}$.

This means that it is enough to prove that every plan γ^i may be actually induced by a transport map (more precisely, that there exists a new plan, with the same marginals and an improved transport cost, which is induced by a map). This is quite easy since the cost, when restricted to a face, is actually affine, and hence every transport plan with the same marginals and supported on $\{(x, y): x - y \in F_i\}$ gives the same cost. The issue is thus reduced to finding an arbitrary

admissible transport map T, with prescribed marginals f_0^i and f_1^i (defined as the marginals of the original γ^i), satisfying $x - T(x) \in F_i$.

Thanks to the results of the present paper this is possible by considering, for instance, the minimization problem (MK) with convex set $C = F_i$. This selects a particular transport map (the one minimizing the quadratic cost among admissible ones), but gives in particular the existence of one such transport map. The only point to be noticed is the fact that a transport plan with finite cost is actually known to exist, and it is γ^i.

As a consequence, we have a new proof of

Theorem 5.1. *If $f_0 \ll dx$ and $\| \cdot \|$ is a crystalline norm, the transport problem*

$$\min\left\{ \int \|x - y\| \, d\pi \colon \pi \in \Pi(f_0, f_1) \right\}$$

admits at least a solution induced by a transport map T.

Analogously, the arguments used in this paper may also be applied to L^∞ problems like those studied in [10]. We already pointed out that the strategy of the quadratic perturbation is useful when one only wants to prove the existence of at least a transport map T sending f_0 onto f_1 and satisfying $T(x) - x \in C$. Suppose now that we want to solve

$$(M_\infty) \quad \min\{\pi - \operatorname{ess\,sup} \|y - x\| \colon \pi \in \Pi(f_0, f_1)\},$$

where $\| \cdot \|$ is a (possibly asymmetric) norm whose unit ball is C. Notice that we have the equality $\pi - \operatorname{ess\,sup} \|y - x\| = \max\{\|y - x\| \colon (x, y) \in \operatorname{spt}(\pi)\}$ and that this quantity measures the maximal displacement in terms of the norm $\| \cdot \|$. Then it is sufficient to denote by L the minimal value of such an ess sup and to notice that π is optimal if and only if it is concentrated on $\{(x, y) \colon y - x \in LC\}$. In order to find a solution induced by a transport map T it is enough, for instance, to take a map T solving (M) for the cost c defined with the dilated body LC instead of C. Hence we have the following theorem, generalizing the results in [10] (and using, by the way, a different selection principle, since the plan which was selected in [10] was the infinitely cyclically monotone one, and not that minimizing the quadratic cost).

Theorem 5.2. *Suppose that C satisfies the assumptions of Theorem 2.3, that $0 \in \operatorname{int}(C)$ and that $f_0 \ll dx$. Let $\| \cdot \|$ be defined through the formula $\|z\| := \inf\{\lambda > 0 \colon z/\lambda \in C\}$. Then, the L^∞ transport problem (M_∞) admits at least a solution induced by a transport map T.*

Acknowledgements

References

[1] L. Ambrosio, Lecture notes on optimal transport problems, in: Mathematical Aspects of Evolving Interfaces, CIME Summer School in Madeira, in: Lecture Notes in Math., vol. 1812, Springer, 2003.

[2] L. Ambrosio, B. Kirchheim, A. Pratelli, Existence of optimal transport maps for crystalline norms, Duke Math. J. 125 (2) (2004) 207–241.

[3] L. Ambrosio, A. Pratelli, Existence and stability results in the $L^1 L1$ theory of optimal transportation, in: Optimal Transportation and Applications, Martina Franca, 2001, in: Lecture Notes in Math., vol. 1813, Springer, Berlin, 2003, pp. 123–160.

[4] M. Beiglböck, Ch. Léonard, W. Schachermayer, A generalized dual maximizer for the Monge–Kantorovich transport problem, preprint available at http://www.mat.univie.ac.at/~schachermayer/pubs/.

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

[6] L.A. Caffarelli, Allocation maps with general cost function, in: Partial Differential Equations and Applications, in: Lecture Notes in Pure and Appl. Math., vol. 177, Dekker, New York, 1996, pp. 29–35.

[7] G. Carlier, L. De Pascale, F. Santambrogio, A strategy for non-strictly convex transport costs and the example of $\|x - y\|^p$ in $\mathbb{R}^2$, Commun. Math. Sci. 8 (4) (2010) 931–941.

[8] T. Champion, L. De Pascale, The Monge problem in $\mathbb{R}^d$, Duke Math. J. 157 (3) (2011) 551–572.

[9] T. Champion, L. De Pascale, The Monge problem for strictly convex norms in $\mathbb{R}^d$, J. Eur. Math. Soc. (JEMS) 12 (6) (2010) 1355–1369.

[10] T. Champion, L. De Pascale, P. Juutinen, The ∞-Wasserstein distance: local solutions and existence of optimal transport maps, SIAM J. Math. Anal. 40 (1) (2008) 1–20.

[11] L.C. Evans, W. Gangbo, Differential equations methods for the Monge–Kantorovich mass transfer problem, Mem. Amer. Math. Soc. 137 (653) (1999), viii+66 pp.

[12] M. Feldman, R.J. McCann, Uniqueness and transport density in Monge's mass transportation problem, Calc. Var. Partial Differential Equations 15 (1) (2002) 81–113.

[13] W. Gangbo, R.J. McCann, The geometry of optimal transportation, Acta Math. 177 (2) (1996) 113–161.

[14] G. Monge, Mémoire sur la théorie des déblais et des remblais, Hist. de l'Acad. des Sci. de Paris, 1781.

[15] W. Schachermayer, J. Teichmann, Characterization of optimal transport plans for the Monge–Kantorovich problem, Proc. Amer. Math. Soc. 137 (2) (2009) 519–529.

[16] N.S. Trudinger, X.J. Wang, On the Monge mass transfer problem, Calc. Var. Partial Differential Equations 13 (1) (2001) 19–31.

[17] C. Villani, Topics in Optimal Transportation, Grad. Stud. Math., vol. 58, American Mathematical Society, Providence, RI, 2003.

[18] C. Villani, Optimal Transport. Old and New, Grundlehren Math. Wiss. (Fundamental Principles of Mathematical Sciences), vol. 338, Springer-Verlag, Berlin, 2009, xxii+973 pp.