THE ∞-WASSERSTEIN DISTANCE: LOCAL SOLUTIONS AND EXISTENCE OF OPTIMAL TRANSPORT MAPS.

THIERRY CHAMPION, LUIGI DE PASCALE, AND PETRI JUUTINEN

Abstract. We consider the nonlinear optimal transportation problem of minimizing the cost functional $\mathcal{C}_\infty(\lambda) = \lambda\text{-}\operatorname{ess\,sup}_{(x,y)\in\Omega^2} |y-x|$ in the set of probability measures on Ω^2 having prescribed marginals. This corresponds to the question of characterizing the measures that realize the infinite Wasserstein distance. We establish the existence of "local" solutions and characterize this class with the aid of an adequate version of cyclical monotonicity. Moreover, under natural assumptions, we show that local solutions are induced by transport maps.

1. Introduction

In this paper, we consider the nonlinear optimal transportation problem that can be mathematically stated as the problem of minimizing the cost functional

$$\mathcal{C}_\infty(\lambda) := \lambda\text{-} \operatorname*{ess\,sup}_{(x,y)\in\Omega^2} |y-x| \tag{1.1}$$

in the set of probability measures on Ω^2 having prescribed marginals. Here, and throughout the paper, we assume that Ω is a compact subset of $\mathbb{R}^d$, $d \geq 1$, $|\cdot|$ denotes the usual Euclidean norm in $\mathbb{R}^d$, μ, ν are the two (given) Borel probability measures on Ω and $\Pi(\mu,\nu)$ denotes the set of admissible transport plans, i.e., the set of Borel probability measures λ on $\Omega^2 := \Omega\times\Omega$ with first marginal $\pi_{1\#}\lambda = \mu$ and second marginal $\pi_{2\#}\lambda = \nu$. Informally, if λ is induced by a transport map $T : \Omega \to \Omega$, i.e., $\lambda = (id \times T)_\#\mu$, then $\mathcal{C}_\infty(\lambda)$ is simply the maximum of the transport distances $|T(x) - x|$.

The problem formulated above corresponds to the question of characterizing the measures that realize the infinite Wasserstein distance

$$(P_\infty) \qquad W_\infty(\mu,\nu) \;=\; \inf\left\{ \mathcal{C}_\infty(\lambda) \;=\; \lambda\text{-} \operatorname*{ess\,sup}_{(x,y)\in\Omega^2} |y-x| \;:\; \lambda\in\Pi(\mu,\nu)\right\}$$

between μ and ν. Clearly, this is the limiting case, as $p\to\infty$, of the more familiar (see e.g. [1], [2], [32]) p-Wasserstein distance problem

$$(P_p) \qquad W_p(\mu,\nu) \;=\; \inf\left\{ \left(\int_{\Omega^2} |y-x|^p d\lambda(x,y)\right)^{\frac{1}{p}} \;:\; \lambda\in\Pi(\mu,\nu)\right\}, \qquad 1\leq p<\infty,$$

Date: 12/11(nov.)/2007.
2000 *Mathematics Subject Classification.* 49Q20, 49K30.
Key words and phrases. infinite Wasserstein distance, restrictable solutions, infinite cyclical monotonicity.

which is a model example of a Monge-Kantorovich type optimal transport problem. Despite the close relationship, there are fundamental differences between these two problems. Most importantly, while (P_p) is linear in λ (removing the $1/p$-power does not change the solution set), the mapping $\lambda \mapsto \mathcal{C}_\infty(\lambda)$ is not even convex. In particular, the problem (P_∞) is not additive, which implies that, unlike in the case of (P_p), a restriction of an optimal transport plan need not be optimal for its own marginals.

In view of simple examples, it turns out that only imposing this "local optimality" property can lead to a satisfactory class of solutions. Hence we introduce in this paper the notion of *restrictable solutions*; this subclass of minimizers of (1.1) is characterized by the property that every portion of μ is transported onto its target in an optimal way, see Definition 4.1 below. The existence of a restrictable solution is obtained with the aid of approximating (P_∞) by the problems (P_p). The same strategy also provides us with the notion of *infinite cyclical monotonicity*, which is derived from the standard c-cyclical monotonicity by applying it to the sequence of costs $c_p(x, y) = |x - y|^p$ and then taking the limit as $p \to \infty$. A reader familiar with the theory of infinity Laplacian and related problems [5] should recognize the analogy between restrictable solutions and absolute minimizers of supremum functionals.

It is one of the main results of this paper that restrictable and infinitely cyclically monotone solutions coincide; see Theorems 3.4 and 4.4 below. We would like to emphasize that although both of these notions are derived via an approximation argument, the proof for their equivalence is completely independent of the derivation. Moreover, this result holds without any further assumptions on the marginals μ and ν. The second principal question we address in this paper is existence and uniqueness of an optimal transport map. Our main result in this direction, Theorem 5.5, states that if $\mu << \mathcal{L}^d$, then any infinitely cyclically monotone solution γ to (P_∞) is induced by a map $T : \Omega \to \Omega$, i.e., $\gamma = (id \times T)_\# \mu$. Regarding the question of uniqueness, we are able to show that if, in addition to the previous assumptions, the second marginal ν is discrete, then the infinitely cyclically monotone solution to (P_∞) is unique.

A major technical difficulty that we are facing in the proofs is the absence of a useful duality theory, which is due to the non-convexity of the objective functional (see §5.4). As a consequence, we must rely on ad hoc arguments designed for the problem at hand. On the other hand, it is quite clear from the proofs that the machinery we are developing applies to more general problems than just (P_∞). In fact, we could have just as well considered a functional $\lambda \mapsto \lambda\text{-}\operatorname{ess\,sup}_{(x,y)\in\Omega^2} c(x, y)$, were $c(x, y)$ is, say, non-negative and lower semicontinuous to begin with. In this work we concentrate on the model case $c(x, y) = |y - x|$ so as to identify the useful tools and notions without coping with the additional technical difficulties required by a more general cost c, which seems to be the natural next step in this study.

Let us finish this introduction by discussing some applications in which the infinite Wasserstein distance W_∞ appears. First on our list is the optimal design problem

$$\sup\left\{ \frac{W_\infty(\mu,\nu)^{p+d}}{W_p(\mu,\nu)^p \|(\frac{d\mu}{dx})^{-1}\|_{L^\infty(U)}} \;:\; (\mu,\nu) \in \mathcal{P}_{a.c.}(U) \times \mathcal{P}(\overline{U}) \right\} \tag{1.2}$$

that appears in [8] in connection with stability estimates for optimal transport maps; here $\mathcal{P}$ and $\mathcal{P}_{a.c.}$ denote the spaces of probability measures and absolutely continuous probability measures, respectively, and $\frac{d\mu}{dx}$ is the Radon-Nikodym derivative of μ with respect to the Lebesgue measure. In [8], the authors prove that if $U \subset \mathbb{R}^d$ is a bounded Lipschitz domain, then the estimate

$$W_\infty(\mu,\nu)^{p+d} \le C_{p,d}(U) \|(\tfrac{d\mu}{dx})^{-1}\|_{L^\infty(U)} W_p(\mu,\nu)^p \tag{1.3}$$

holds for every $p > 1$. The inequality (1.3) is an intrinsic counterpart of a beautiful uniform estimate for the optimal transport maps proved in [8] under stronger regularity assumptions. The optimal constant in (1.3) is given by the supremum in (1.2) and it is conjectured (based on 1-dimensional examples and remarks on increasing transport maps) that it does not blow up when $p \to 1$.

Secondly, during the last few years, models of branching processes using in a way or another tools from the optimal transportation theory have been proposed by several authors, see for example [7, 9, 23, 34]. Roughly speaking, these models favor joint transportation, which in many real world situations, such as in the design of communication or irrigation networks, is more economical than individualized transportation. In particular, in [9] the authors propose to minimize a certain cost functional (which penalizes diffused measures) on the p-Wasserstein space of probability measures. It is remarked in [29] that this model is somewhat less realistic than the others cited above but it has the advantage of being mathematically simpler. Moreover, as pointed out in Remark 6.2.7, Chapter 4 and Section 0.2 of [29], the use of the infinite Wasserstein distance W_∞ in the model of [9] produces results which are closer to the ones derived from the other models.

Then there are applications to PDEs. The metric structure associated to the ∞-Wasserstein distance is a crucial tool in proving the existence of stable solutions for a compressible fluid model of rotating binary stars in [24]. The same metric was used also to bound the growth of the wetted regions in the porous medium flow [13] and to study the long time asymptotics of nonlinear scalar conservation laws [12]. Moreover the ∞-Wasserstein distance is being used in some $N-$particles approximations of the Vlaslov equations [21, 22].

Finally, in [11] the authors have considered a mathematical model of the optimal pricing policy for the use of a public transportation network. This model assumes that the price of a ticket (for the use of the network) is a function of the distance traveled. This seems reasonable in the case when each citizen is associated with a single journey, but not so realistic if we allow multiple journeys and an inexpensive season ticket is available. In the latter case, the price of a season ticket could be assumed to be a piece-wise constant function of the maximal distance traveled, and hence it might be a good idea to insert a component similar to the functional we have considered into the model. It is also quite easy to imagine that in many other transportation problems a significant portion of the total cost is in a way or another connected with the maximal transportation distance. For example, if we assume that the physical transportation device (airplane, car, etc.) is the same for all distances, then it has to be chosen so that the longest transportation can be handled.

Contents

2. Existence of global solutions

As it is pointed out in the introduction, the objective functional

$$\begin{aligned}\lambda \mapsto \mathcal{C}_\infty(\lambda) &:= \lambda\text{-}\operatorname*{ess\,sup}_{(x,y)\in\Omega^2} |y-x| \\ &= \inf\Big\{t \geq 0 : \lambda(\{(x,y)\in\Omega^2 : |y-x| > t\}) = 0\Big\}\end{aligned}$$

is not linear (and not even convex) in λ, contrary to what is usually the case in classical optimal transport problems. However, it is, quite interestingly, level convex in the sense that if $\lambda_1, \lambda_2 \in \Pi(\mu,\nu)$ then

$$\mathcal{C}_\infty(t\lambda_1 + (1-t)\lambda_2) \leq \max\{\mathcal{C}_\infty(\lambda_1), \mathcal{C}_\infty(\lambda_2)\} \qquad \text{for all } t \in (0,1).$$

Note that this implies that the set of solutions to (P_∞) is convex. Moreover, it should be observed that $\mathcal{C}_\infty(\lambda)$ depends on the measure λ only via its support. More precisely, one has

$$\mathcal{C}_\infty(\lambda) \;=\; \sup\{|y-x| \,:\, (x,y) \in \operatorname{supp}(\lambda)\}. \tag{2.1}$$

Thanks to this last property, we are in position to give the following existence result:

Proposition 2.1. *Assume that Ω is a compact subset of $\mathbb{R}^d$ and μ,ν are two probability measures on Ω. Then the problem*

$$(P_\infty) \qquad W_\infty(\mu,\nu) \;=\; \inf\left\{\mathcal{C}_\infty(\lambda) := \lambda\text{-}\operatorname*{ess\,sup}_{(x,y)\in\Omega^2} |y-x| \,:\, \lambda \in \Pi(\mu,\nu)\right\},$$

admits at least one solution $\lambda \in \Pi(\mu,\nu)$.

The optimal set of (P_∞) may be very large thanks to (2.1):

Example 2.2. Let $\mu := \frac{1}{2}\mathcal{L}^2\lfloor_{[0,1]^2\cup[2,3]^2}$ and $\nu := \frac{1}{2}(\delta_{(2,1)} + \delta_{(1,2)})$. Then it is clear that the value of (P_∞) is $\sqrt{5}$, and that any admissible transport plan $\lambda \in \Pi(\mu,\nu)$ is a solution of (P_∞).

In the proof of Proposition 2.1, we shall use the following Lemma.

Lemma 2.3. *If the sequence $(\lambda_n)_n$ converges weakly to λ in $\Pi(\mu,\nu)$, then for any $(x,y) \in \operatorname{supp}(\lambda)$ there exists a sequence $((x_n,y_n))_{n\in\mathbb{N}}$ such that*

$$(x_n,y_n) \to (x,y) \text{ as } n\to\infty \quad \text{and} \quad (x_n,y_n)\in \operatorname{supp}(\lambda_n) \quad \text{for all } n\in\mathbb{N}. \tag{2.2}$$

Proof. Suppose $(x,y)\in\Omega$ is such that (2.2) does not hold. Then we may assume without loss of generality that there exists $r>0$ such that $B((x,y),r)\cap\operatorname{supp}(\lambda_n)=\emptyset$ for any $n\in\mathbb{N}$. It then obviously follows from the weak convergence that $B((x,y),r)\cap\operatorname{supp}(\lambda)=\emptyset$, which concludes the proof. □

Proof of Proposition 2.1. Since the set Ω is a compact subset of $\mathbb{R}^d$ and the measures μ and ν are probability measures on Ω, the non-empty set $\Pi(\mu,\nu)$ is compact for the weak convergence of measures (cf. [32, p. 49]). To apply the direct method of the Calculus of Variations, it remains to notice that $\lambda\mapsto\mathcal{C}_\infty(\lambda)$ is lower semicontinuous for this topology: this is a direct consequence of (2.1) and Lemma 2.3. □

3. Infinitely cyclically monotone solutions

The proof of the existence of a solution to (P_∞) given in Proposition 2.1 is intrinsic, but one may obtain this result also via an approximation argument involving the family of problems $(P_p)_{p\geq 1}$ given by

$$(P_p) \qquad W_p(\mu,\nu) \;=\; \inf\left\{\mathcal{C}_p(\lambda) := \left(\int_{\Omega^2}|y-x|^p d\lambda(x,y)\right)^{\frac{1}{p}} \;:\; \lambda\in\Pi(\mu,\nu)\right\};$$

that is, the functional $\lambda\mapsto\mathcal{C}_p(\lambda)$ is being minimized over the set $\Pi(\mu,\nu)$.

Alternative proof of Proposition 2.1. Under the assumptions made on Ω, μ and ν, for any $p\geq 1$ the problem (P_p) admits at least one solution $\gamma_p\in\Pi(\mu,\nu)$, see e.g. [32, Theorem 1.3]. Since $\Pi(\mu,\nu)$ is compact, we infer that $(\gamma_p)_{p\geq1}$ converges weakly (up to a subsequence) to some $\gamma_\infty\in\Pi(\mu,\nu)$ as $p\to\infty$. Then, for any $\lambda\in\Pi(\mu,\nu)$, we have by the optimality of γ_p and Hölder's inequality that

$$\mathcal{C}_q(\gamma_p) \;=\; \left(\int_{\Omega^2}|y-x|^q d\gamma_p(x,y)\right)^{\frac{1}{q}} \;\leq\; \mathcal{C}_p(\gamma_p) \;\leq\; \mathcal{C}_p(\lambda)$$

for any $p\geq q\geq 1$. For a fixed $q\geq1$, since the function $(x,y)\mapsto|y-x|^q$ is continuous and bounded on Ω^2 one has $\mathcal{C}_q(\gamma_p)\to\mathcal{C}_q(\gamma_\infty)$ as $p\to\infty$. Therefore, taking the limit in p and then in q in the above inequality we obtain $\mathcal{C}_\infty(\gamma_\infty)\leq\mathcal{C}_\infty(\lambda)$. Since this holds for any $\lambda\in\Pi(\mu,\nu)$, γ_∞ is a minimizer of (P_∞). □

The reason for considering the problems (P_p) in this context is not merely the fact that they provide an alternative route to the existence. Namely, it is known that an element $\gamma_p\in\Pi(\mu,\nu)$ is a solution of (P_p) if and only if its support is p-cyclically monotone, that is

$$\sum_{i=1}^n |y_i-x_i|^p \;\leq\; \sum_{i=1}^n |y_{\sigma(i)}-x_i|^p \tag{3.1}$$

for every $n\geq2$, $(x_1,y_1),\dots,(x_n,y_n)\in\operatorname{supp}(\gamma_p)$ and for every permutation $\sigma\in\mathcal{S}_n$. We refer the reader for example to Theorem 3.2 in [2] (or to the references [20, 28, 33]).

By analogy with the p-cyclical monotonicity when $1 \leq p < \infty$, we introduce the corresponding notion for the case $p = \infty$ obtained by taking the limit in (3.1).

Definition 3.1. A transport plan $\gamma \in \Pi(\mu, \nu)$ is *infinitely cyclically monotone* if

$$\max_{1 \leq i \leq n} |y_i - x_i| \;\leq\; \max_{1 \leq i \leq n} |y_{\sigma(i)} - x_i|$$

for every $n \geq 2$, $(x_1, y_1), \ldots, (x_n, y_n) \in \operatorname{supp}(\gamma)$ and $\sigma \in \mathcal{S}_n$.

Using again the approximation of (P_∞) by the problems (P_p), we obtain the existence of an infinitely cyclically monotone solution to (P_∞).

Theorem 3.2. *For $1 \leq p < \infty$, let $\gamma_p \in \Pi(\mu, \nu)$ be a solution to (P_p). Then any cluster point γ_∞ of $(\gamma_p)_{p \geq 1}$ in $\Pi(\mu, \nu)$ as $p \to \infty$ is an infinitely cyclically monotone solution to (P_∞).*

Proof. For simplicity, let us assume that the entire family $(\gamma_p)_{p \geq 1}$ converges weakly to $\gamma_\infty \in \Pi(\mu, \nu)$. It suffices to show that γ_∞ is infinitely cyclically monotone. To this end, let $n \geq 2$, $(x_1, y_1), \ldots, (x_n, y_n) \in \operatorname{supp}(\gamma)$ and $\sigma \in \mathcal{S}_n$. We apply Lemma 2.3 to each pair (x_i, y_i) to obtain the existence of sequences $(x_1^p, y_1^p), \ldots, (x_n^p, y_n^p)$ such that $(x_i^p, y_i^p) \to (x_i, y_i)$ for any i as $p \to \infty$, and $(x_i^p, y_i^p) \in \operatorname{supp}(\gamma_p)$ for all $1 \leq p < \infty$ and $i = 1, \ldots, n$. Since the support of γ_p is p-cyclically monotone, one has

$$\sum_{i=1}^{n} |y_i^p - x_i^p|^p \;\leq\; \sum_{i=1}^{n} |y_{\sigma(i)}^p - x_i^p|^p \qquad \text{for all } 1 < p < \infty.$$

Taking the $1/p$-power on both sides and letting p go to ∞, one obtains the desired inequality. □

Since for any $1 \leq p < \infty$ an admissible transport plan $\lambda \in \Pi(\mu, \nu)$ is a minimizer of (P_p) if and only if it is p-cyclically monotone, it is natural to ask whether this still holds for $p = \infty$. It is, however, quite clear that a generic minimizer of (P_∞) need not be infinitely cyclically monotone; the following gives a counter-example for this implication.

Example 3.3. As an admissible transport plan for the measures μ and ν in Example 2.2 one may take

$$\lambda := \frac{1}{2}\left(\mathcal{L}^2 \lfloor_{[0,1]^2} \times \delta_{(2,1)} + \mathcal{L}^2 \lfloor_{[2,3]^2} \times \delta_{(1,2)}\right).$$

Then λ is a minimizer of (P_∞) but it is not infinitely cyclically monotone: for example, $((0,1),(2,1))$ and $((3,2),(1,2))$ belong to $\operatorname{supp}(\lambda)$, and

$$\max\{|(0,1) - (1,2)|, |(3,2) - (2,1)|\} \;<\; \max\{|(0,1) - (2,1)|, |(3,2) - (1,2)|\}.$$

Notice that in this case, for any $1 < p < \infty$, problem (P_p) admits a unique solution (up to a μ-negligible set) (see [20]), which in fact does not depend on p.

On the other hand, the following result shows that the reverse implication does hold: infinite cyclical monotonicity is indeed a sufficient condition for an admissible plan to be a minimizer of (P_∞).

Theorem 3.4. *Any infinitely cyclically monotone transport plan $\gamma \in \Pi(\mu, \nu)$ is a solution of the problem (P_∞).*

Proof. We make a proof by contradiction. Let $\gamma \in \Pi(\mu,\nu)$ be infinitely cyclically monotone and assume that

$$\gamma\text{-}\operatorname*{ess\,sup}_{(x,y)\in\Omega^2} |y-x| \;\geq\; 10\,\varepsilon + \tilde\gamma\text{-}\operatorname*{ess\,sup}_{(x,y)\in\Omega^2} |y-x| \tag{3.2}$$

for some $\tilde\gamma \in \Pi(\mu,\nu)$ and $\varepsilon > 0$.

Since Ω is compact, there exists a finite family $(c_i)_{1\leq i\leq k}$ such that $\Omega \subset \bigcup_{i=1}^{k} B(c_i,\varepsilon)$. We shall denote $C := \{c_1,\dots,c_k\}$, $V_1 := B(c_1,\varepsilon)$ and for any $i \in \{2,\dots,k\}$ we set $V_i := B(c_i,\varepsilon) \setminus \bigcup_{j=1}^{i-1} V_j$; without loss of generality, we assume that $V_i \neq \emptyset$ for all $i \in \{1,\dots,k\}$.

Next we define two discrete measures γ^ε and $\tilde\gamma^\varepsilon$ on Ω^2 by

$$\gamma^\varepsilon \;:=\; \sum_{1\leq i,j\leq k} \gamma(V_i\times V_j)\,\delta_{(c_i,c_j)}$$

and

$$\tilde\gamma^\varepsilon \;:=\; \sum_{1\leq i,j\leq k} \tilde\gamma(V_i\times V_j)\,\delta_{(c_i,c_j)}.$$

Notice that since γ and $\tilde\gamma$ have the same marginals, the same holds for γ^ε and $\tilde\gamma^\varepsilon$. In particular, one has

$$(x,y)\in \operatorname{supp}(\gamma^\varepsilon) \quad\Rightarrow\quad \text{there exists } \tilde x \in C \text{ such that } (\tilde x, y) \in \operatorname{supp}(\tilde\gamma^\varepsilon) \tag{3.3}$$

and

$$(\tilde x,\tilde y)\in \operatorname{supp}(\tilde\gamma^\varepsilon) \quad\Rightarrow\quad \text{there exists } y \in C \text{ such that } (\tilde x, y) \in \operatorname{supp}(\gamma^\varepsilon). \tag{3.4}$$

The following properties will also be useful in our argument:

Claim 1. There exists (x_0,y_0) in the support of γ^ε such that

$$|y_0-x_0| \;\geq\; 5\,\varepsilon + \max\big\{|y-x| \;:\; (x,y)\in\operatorname{supp}(\tilde\gamma^\varepsilon)\big\}.$$

Claim 2. For any $n\geq 1$, $(x_1,y_1),\dots,(x_n,y_n)\in\operatorname{supp}(\gamma^\varepsilon)$ and $\sigma\in\mathcal{S}_n$,

$$\max_{1\leq i\leq n} |y_i-x_i| \;\leq\; 4\,\varepsilon + \max_{1\leq i\leq n} |y_{\sigma(i)}-x_i|.$$

Above, the first claim is simply a counterpart of the antithesis (3.2) for the discretized measures, while the second says that γ^ε is "almost" infinitely cyclically monotone. We postpone the verification of these two claims until the end of this proof.

Let $(x_0,y_0)\in\operatorname{supp}(\gamma^\varepsilon)$ be given by Claim 1. Owing to (3.3) and (3.4), we can recursively define two sequences $(D_m)_{m\geq 1}$ and $(E_m)_{m\geq 0}$ of subsets of C by setting $E_0 := \{y_0\}$ and for $m\geq 1$,

$$D_m \;:=\; \{\tilde x \,:\text{there exists } y\in E_{m-1} \text{ such that } (\tilde x,y)\in\operatorname{supp}(\tilde\gamma^\varepsilon)\}$$

and

$$E_m \;:=\; \{y \;:\; \text{there exists } \tilde x\in D_m \text{ such that } (\tilde x,y)\in\operatorname{supp}(\gamma^\varepsilon)\}\,.$$

We then set $D := \bigcup_{m\geq 1} D_m$ and $E := \bigcup_{m\geq 0} E_m$.

There are now two alternatives: either x_0 belongs to D or not.

First case: $x_0 \in D$. In this case, there exists $m \geq 1$ such that $x_0 \in D_m$, and by going backwards from D_m to E_0 it is possible to define two finite families $(x_i)_{0\leq i\leq m}$ and $(y_i)_{0\leq i\leq m-1}$ such that

$$\text{for all } i \in \{0,\dots,m-1\}, \quad (x_i,y_i) \in \operatorname{supp}(\gamma^\varepsilon) \;\text{ and }\; (x_{i+1},y_i) \in \operatorname{supp}(\tilde{\gamma}^\varepsilon),$$

where we have set $x_m := x_0$. Claim 2 then yields

$$\max_{0\leq i\leq m-1} |y_i - x_i| - 4\,\varepsilon \;\leq\; \max_{0\leq i\leq m-1} |y_i - x_{i+1}|.$$

Since $\max_{0\leq i\leq m-1} |y_i - x_i| \geq |y_0 - x_0|$, we infer from Claim 1 and the previous inequality that

$$\max\big\{|y-x| \,:\, (x,y) \in \operatorname{supp}(\tilde{\gamma}^\varepsilon)\big\} + \varepsilon \;\leq\; \max_{0\leq i\leq m-1} |y_i - x_{i+1}|.$$

Since $(x_{i+1},y_i) \in \operatorname{supp}(\tilde{\gamma}^\varepsilon)$ for any $i \in \{0,\dots,m-1\}$, this yields a contradiction.

Second case: $x_0 \notin D$. From the definitions of D and E, we notice the following two facts:

$$x \in D,\ (x,y) \in \operatorname{supp}(\gamma^\varepsilon) \quad \Rightarrow \quad y \in E, \tag{3.5}$$

and

$$\tilde{y} \in E,\ (\tilde{x},\tilde{y}) \in \operatorname{supp}(\tilde{\gamma}^\varepsilon) \quad \Rightarrow \quad \tilde{x} \in D. \tag{3.6}$$

As a consequence of (3.5) and since γ^ε and $\tilde{\gamma}^\varepsilon$ have the same marginals, one has

$$\gamma^\varepsilon(D\times E) \;=\; \gamma^\varepsilon(D\times C) \;=\; \tilde{\gamma}^\varepsilon(D\times C).$$

Similarly, one has

$$\tilde{\gamma}^\varepsilon(D\times E) \;=\; \tilde{\gamma}^\varepsilon(C\times E) \;=\; \gamma^\varepsilon(C\times E).$$

We then obtain

$$\gamma^\varepsilon(D\times E) \;=\; \tilde{\gamma}^\varepsilon(D\times C) \;\geq\; \tilde{\gamma}^\varepsilon(D\times E) \;=\; \gamma^\varepsilon(C\times E).$$

This implies that $\gamma^\varepsilon((C\backslash D)\times E) = 0$, whereas by hypothesis one has $(x_0,y_0) \in (C\backslash D)\times E$ and $\gamma^\varepsilon(\{(x_0,y_0)\}) > 0$ since (x_0,y_0) belongs to the support of the discrete measure γ^ε. This yields a contradiction.

To complete the proof of Theorem 3.4, it remains to prove Claims 1 and 2.

Proof of Claim 1. We infer from (3.2) that

$$\gamma\Big(\{(x,y) \,:\, |y-x| \geq 9\,\varepsilon + \tilde{\gamma}\text{-}\operatorname*{ess\,sup}_{(x,y)\in\Omega^2} |y-x|\}\Big) > 0.$$

As a consequence, there exist $i_1, i_2 \in \{1,\dots,k\}$ such that

$$\gamma\Big(\,(V_{i_1}\times V_{i_2}) \cap \{(x,y) \,:\, |y-x| \geq 9\,\varepsilon + \tilde{\gamma}\text{-}\operatorname*{ess\,sup}_{(x,y)\in\Omega^2} |y-x|\}\,\Big) > 0.$$

Since $V_m \subset B(c_m,\varepsilon)$ for $m = i_1, i_2$, one then has

$$|c_{i_2} - c_{i_1}| \;\geq\; 7\,\varepsilon + \tilde{\gamma}\text{-}\operatorname*{ess\,sup}_{(x,y)\in\Omega^2} |y-x| \tag{3.7}$$

and (c_{i_1}, c_{i_2}) belongs to the support of γ^ε. On the other hand, if (c_{j_1}, c_{j_2}) belongs to the support of $\tilde{\gamma}^\varepsilon$ then $\tilde{\gamma}(V_{j_1} \times V_{j_2}) > 0$ and thus

$$|c_{j_2} - c_{j_1}| \;\le\; 2\,\varepsilon + \tilde{\gamma}\text{-}\operatorname*{ess\,sup}_{(x,y)\in\Omega^2} |y - x|.$$

Since this inequality holds whenever $(c_{j_1}, c_{j_2}) \in \operatorname{supp}(\tilde{\gamma}^\varepsilon)$, one has

$$\max\{|y - x| \,:\, (x, y) \in \operatorname{supp}(\tilde{\gamma}^\varepsilon)\} \;\le\; 2\,\varepsilon + \tilde{\gamma}\text{-}\operatorname*{ess\,sup}_{(x,y)\in\Omega^2} |y - x|.$$

This together with (3.7) shows that $(x_0, y_0) := (c_{i_1}, c_{i_2})$ has the desired property.

Proof of Claim 2. Let $n \ge 1$, $\sigma \in \mathcal{S}_n$ and (x_i, y_i) belong to the support of γ^ε for $i \in \{1, \dots, n\}$. For any $i \in \{1, \dots, n\}$, one has $(x_i, y_i) = (c_{j_1}, c_{j_2})$ for some $j_1, j_2 \in \{1, \dots, k\}$ with $\gamma(V_{j_1} \times V_{j_2}) > 0$, and thus there exists $(x'_i, y'_i) \in (V_{j_1} \times V_{j_2}) \cap \operatorname{supp}(\gamma)$. As a consequence,

$$\big|\,|y'_r - x'_s| - |y_r - x_s|\,\big| \;\le\; 2\,\varepsilon \qquad \text{for all } r, s \in \{1, \dots, n\}.$$

Since γ is infinitely cyclically monotone, we have

$$\max_{1\le i\le n} |y'_{\sigma(i)} - x'_i| \;\ge\; \max_{1\le i\le n} |y'_i - x'_i|.$$

It follows that

$$4\,\varepsilon + \max_{1\le i\le n} |y_{\sigma(i)} - x_i| \;\ge\; \max_{1\le i\le n} |y_i - x_i|$$

which proves the claim. □

Remark 3.5. Observe that in the proof above, we in fact always have $x_0 \in D$, that is, the *First case* always occurs. This is a consequence of the conservation of the masses: all the mass transported to E by $\tilde{\gamma}^\varepsilon$ originates from D, while all the mass in D is transported to E by γ^ε. Since γ^ε and $\tilde{\gamma}^\varepsilon$ have the same marginals, this implies that both plans transport D exactly to E. The fact that the *Second case* never occurs in the above proof also underlies the recursive construction of *Step II* in the proof of Theorem A in [25], even if the arguments are different. That paper deals with the sufficiency of cyclical monotonicity for optimality in the classical case (see also [31]).

Remark 3.6. There is a variation of the self-contained proof given above that relies directly on the fact that the total cost $\mathcal{C}_\infty(\lambda) = \lambda\text{-}\operatorname{ess\,sup}_{(x,y)\in\Omega^2} |y - x|$ depends only on the support of λ and not on its density. In the discrete case this means that, as far as the total cost is concerned, the exact amount of mass transferred from any given point to another is irrelevant: it only matters whether the amount is positive or not. Hence in the proof we are allowed to change the transport plans γ_ε and $\tilde{\gamma}_\varepsilon$, along with their marginals, as long as we do not change their supports and make sure that the marginals of the new transport plans agree with each other. Now assuming that we can change the measures in such a way that all the point masses are of integer size, the problem can be interpreted as a pairing problem in which the infinite cyclical monotonicity is both a necessary and sufficient condition for optimality.

The existence of the required integer transport plans with given supports is non-trivial and follows from Dines' algorithm [17] that provides positive solutions for a system of linear equations. We would like to thank Jouni Parkkonen for bringing this paper to our attention.

4. Restrictable solutions

In the previous section, we derived the notion of infinitely cyclically monotone plans from the approximation of the problem (P_∞) by the family of problems (P_p). Another interesting notion may be derived in the same way: let $1 \le p < \infty$ and $\gamma_p \in \Pi(\mu,\nu)$ be a solution to (P_p). Then it follows from the linearity of the functional

$$\gamma \mapsto \mathcal{C}_p(\gamma)^p \;=\; \int_{\Omega^2} |y-x|^p d\lambda(x,y)$$

that any non-zero measure γ' that is majorized by γ_p, i.e., $\gamma'(B) \le \gamma_p(B)$ for all Borel sets $B \subset \Omega \times \Omega$, is an optimal transport plan for the problem

$$\mathcal{C}_p(\gamma') = \inf\left\{\mathcal{C}_p(\gamma) : \gamma \in \Pi(\mu',\nu')\right\}$$

where $\mu' := \pi_{1\#}\gamma'$ and $\nu' := \pi_{2\#}\gamma'$. In other words, optimality is automatically inherited by restriction, and hence we may say that a solution $\gamma_p \in \Pi(\mu,\nu)$ of (P_p) is a *restrictable solution* of this problem. By analogy, we may define a similar notion of restrictable solutions for problem (P_∞) as follows.

Definition 4.1. A transport plan $\gamma \in \Pi(\mu,\nu)$ is a *restrictable solution* of (P_∞) if any non-zero Borel measure γ' in $\Omega \times \Omega$ that is majorized by γ is a solution to the problem

$$(P'_\infty) \qquad \inf\left\{\lambda\text{-}\operatorname*{ess\,sup}_{(x,y)\in\Omega^2} |y-x| \;:\; \lambda \in \Pi(\mu',\nu')\right\}$$

where $\mu' := \pi_{1\#}\gamma'$ and $\nu' := \pi_{2\#}\gamma'$.

The reader should notice the obvious abuse of notation above as the measures μ' and ν' in (P'_∞) are not in general probability measures. However, $\mu'(\Omega) = \nu'(\Omega) > 0$, and that is really all that is needed.

Example 4.2. It is quite clear that not every solution of (P_∞) is restrictable. Indeed, the optimal plan λ considered in Example 3.3 admits the following restriction:

$$\lambda' := \frac{1}{2}\left(\mathcal{L}^2\lfloor_{S_1}\times\tfrac{1}{2}\delta_{(2,1)} + \mathcal{L}^2\lfloor_{S_2}\times\tfrac{1}{2}\delta_{(1,2)}\right),$$

where $S_1 := [0,\frac{1}{2}] \times [\frac{1}{2},1]$ and $S_2 := [\frac{5}{2},3] \times [2,\frac{5}{2}]$. But λ' is not optimal for its own marginals: a better transport plan is the one that takes all the mass that lies in S_1 to $(1,2)$ and the mass in S_2 to $(2,1)$.

Remark 4.3. The notion of a restrictable solution bears a strong resemblance to that of an *absolute minimizer* used in connection with the L^∞ variational problems. We recall that a locally Lipschitz continuous function $u : D \to \mathbb{R}^m$, $m \ge 1$, is called an absolute minimizer of the functional $S(\varphi,D) := \operatorname{ess\,sup}_{x\in D} H(x,\varphi(x),D\varphi(x))$, if

$$S(u,V) \le S(v,V)$$

for every open $V \subset\subset D$ and $v \in W^{1,\infty}(V) \cap C(\overline{V}))$ such that $v|_{\partial V} = u|_{\partial V}$. It has turned out that this is the proper notion of a solution for this type of minimization problems in the sense that important properties such as uniqueness, regularity and characterization via an Euler-Lagrange equation can be obtained for this class of functions. Absolute minimizers were introduced by Aronsson [3], see also e.g. [6], [5], [30], [14] for further details and background.

It is natural to ask whether any cluster point γ_∞ of $(\gamma_p)_{p\geq 1}$ in $\Pi(\mu,\nu)$ as $p\to\infty$ is a restrictable solution of (P_∞). In view of Theorem 3.2, this can be established by showing that the restrictable solutions coincide with the class of infinitely cyclically monotone solutions.

Theorem 4.4. *A transport plan $\gamma\in\Pi(\mu,\nu)$ is infinitely cyclically monotone if and only if it is a restrictable solution of the problem (P_∞).*

Notice that Theorem 3.2 provides the existence of restrictable solutions to (P_∞).

Proof. If $\gamma\in\Pi(\mu,\nu)$ is infinitely cyclically monotone then the same holds for any restriction $\gamma'\leq\gamma$, and Theorem 3.4 then yields that such a restriction γ' is a solution of the corresponding problem (P'_∞).

We now turn to the proof of the sufficiency. Let $\gamma\in\Pi(\mu,\nu)$ be a restrictable solution to (P_∞), and let us fix points $(x_1,y_1),\dots,(x_m,y_m)\in\operatorname{supp}(\gamma)$, $m\geq 2$, and a permutation σ of $\{1,\dots,m\}$. Without loss of generality, we may assume that $(x_i,y_i)\neq(x_j,y_j)$ whenever $i\neq j$. Then there is $\varepsilon_0>0$ such that for all $0<\varepsilon\leq\varepsilon_0$, the sets $B_i:=B(x_i,\varepsilon)\times B(y_i,\varepsilon)$ are pairwise disjoint and $\gamma(B_i)>0$ for all $i=1,\dots,m$.

We define two measures γ' and γ_σ by setting

$$\gamma':=\sum_{i=1}^m c_i\gamma\lfloor_{B_i},\qquad\text{and}\qquad \gamma_\sigma:=\sum_{i=1}^m c_i T^i_{\#}\gamma\lfloor_{B_i}.$$

Here $T^i_{\#}\gamma\lfloor_{B_i}$ is the push-forward of $\gamma\lfloor_{B_i}$ by the mapping $T^i(x,y):=(x,y+y_{\sigma(i)}-y_i)$, and the positive numbers

$$c_i:=\frac{\min_k\gamma(B_k)}{\gamma(B_i)}$$

are chosen so that $\gamma'(B_i)=\min_k\gamma(B_k)>0$ is independent of i. Observe that the support of $T^i_{\#}\gamma\lfloor_{B_i}$ lies in $B_i^\sigma:=B(x_i,\varepsilon)\times B(y_{\sigma(i)},\varepsilon)$ and $\gamma_\sigma(B_i^\sigma)=\gamma'(B_i)$. Moreover, the first marginals $\mu'=\pi_{1\#}\gamma'$ and $\mu_\sigma=\pi_{1\#}\gamma_\sigma$ are equal.

Since γ' is majorized by the restrictable solution γ, we have

$$\gamma'\text{-}\operatorname*{ess\,sup}_{(x,y)\in\Omega^2}|x-y|=W_\infty(\mu',\nu')=\inf\{\tilde\gamma\text{-}\operatorname*{ess\,sup}_{(x,y)\in\Omega^2}|x-y|\,:\,\tilde\gamma\in\Pi(\mu',\nu')\},$$

where $\nu'=\pi_{2\#}\gamma'$. On the other hand, since the supports of both ν' and $\nu_\sigma=\pi_{2\#}\gamma_\sigma$ are contained in the union of the balls $B(y_i,\varepsilon)$ and $\nu'(B(y_i,\varepsilon))=\nu_\sigma(B(y_i,\varepsilon))$ for every i by the construction of γ' and γ_σ, we can re-arrange ν' to ν_σ by transporting mass only within the balls $B(y_i,\varepsilon)$. Thereby we obtain that $W_\infty(\nu',\nu_\sigma)\leq 2\varepsilon$, and hence

$$\begin{aligned}\gamma'\text{-}\operatorname*{ess\,sup}_{(x,y)\in\Omega^2}|x-y|&=W_\infty(\mu',\nu')\leq W_\infty(\mu_\sigma,\nu_\sigma)+W_\infty(\nu_\sigma,\nu')\\&\leq 2\varepsilon+\gamma_\sigma\text{-}\operatorname*{ess\,sup}_{(x,y)\in\Omega^2}|x-y|.\end{aligned}$$

Now clearly

$$\gamma'\text{- }\operatorname*{ess\,sup}_{(x,y)\in\Omega^2}|x-y| \geq \max_{1\leq i\leq m}|x_i-y_i|$$

and

$$\gamma_\sigma\text{- }\operatorname*{ess\,sup}_{(x,y)\in\Omega^2}|x-y| \leq \max_{1\leq i\leq m}|x_i-y_{\sigma(i)}|+2\varepsilon,$$

and thus the preceding inequality yields

$$\max_{1\leq i\leq m}|x_i-y_i| \leq \max_{1\leq i\leq m}|x_i-y_{\sigma(i)}|+4\varepsilon.$$

Since this holds for all $\varepsilon>0$ small enough we are done.

□

Observe that for any $\gamma\in\Pi(\mu,\nu)$ and any Borel set $B\subset\Omega\times\Omega$ such that $\gamma(B)>0$, the measure $\gamma\lfloor_B$ is majorized by γ. Thus, if γ is a restrictable solution of (P_∞), then each such measure $\gamma\lfloor_B$ is an optimal transport plan for its own marginals. It turns out that in general the converse is false, that is, this family of measures alone does not suffice for characterizing the restrictable solutions, as the following example 4.5 shows. Notice that this is another difference with the case of integral costs functionals (like the costs $\mathcal{C}_p$), since for those functionals the converse would be true: if $\gamma\lfloor_B$ is an optimal transport plan for its own marginals for any $B\subset\Omega\times\Omega$ with $\gamma(B)>0$, then γ is a restrictable solution.

Example 4.5. Take $\Omega=[0,1]$ and let

$$\mu:=\frac{1}{3}\delta_0+\frac{2}{3}\delta_1, \qquad \text{and} \qquad \nu:=\frac{2}{3}\delta_0+\frac{1}{3}\delta_1.$$

Then the plan $\gamma:=\frac{1}{3}\delta_{(0,1)}+\frac{2}{3}\delta_{(1,0)}$ is not a restrictable solution since $\gamma':=\frac{1}{3}\delta_{(0,1)}+\frac{1}{3}\delta_{(1,0)}$ is majorized by γ and it is clearly not an optimal transport plan for its own marginals $\mu'=\frac{1}{3}\delta_0+\frac{1}{3}\delta_1$ and $\nu'=\frac{1}{3}\delta_0+\frac{1}{3}\delta_1$. On the other hand, one can check that $\gamma\lfloor_B$ is an optimal transport plan, for its own marginals, for each Borel set $B\subset\Omega^2$. Notice that the only restrictable solution of (P_∞) in this case (which is also the unique solution to (P_p) when $p\geq 1$) is $\gamma_\infty=\frac{1}{3}\delta_{(0,0)}+\frac{1}{3}\delta_{(1,0)}+\frac{1}{3}\delta_{(1,1)}$.

In view of the above example, and under some regularity assumption on the measures μ and ν, it is possible to obtain the following refined version of Theorem 4.4:

Proposition 4.6. *Let $\gamma\in\Pi(\mu,\nu)$ be an optimal transport plan and assume that neither μ nor ν concentrates on sets of dimension $d-1$. Then the following are equivalent*

1. *γ is infinitely cyclically monotone,*
2. *for each Borel set $B\subset\Omega\times\Omega$, $\gamma\lfloor_B$ is optimal between its projections.*

Proof. We only need to prove that under these assumptions (2) implies (1). Assume by contradiction that γ is not infinitely cyclically monotone. Then there exist a family $\{(x_i,y_i)\}_{i=1,\dots,n}$ in $\operatorname{supp}(\gamma)$ and a permutation $\sigma\in\mathcal{S}_n$ such that

$$\max\{|x_1-y_{\sigma(1)}|,\dots,|x_n-y_{\sigma(n)}|\} < \max\{|x_1-y_1|,\dots,|x_n-y_n|\}.$$

By continuity, the same inequality holds true for any family $\{(x'_i, y'_i)\}_{i=1,\dots,n}$ for which $(x'_i, y'_i) \in B(x_i, \varepsilon) \times B(y_i, \varepsilon)$ for all $i = 1, \dots, n$ and for some small enough $\varepsilon > 0$. Notice that $\gamma(B(x_i, \varepsilon) \times B(y_i, \varepsilon))$ is positive for any i.

For each i we define $g_i : [0, \varepsilon] \to \mathbb{R}^+$ by $g_i(r) := \gamma(B(x_i, r) \times B(y_i, r))$. Since μ and ν do not concentrate on sets of dimension $d - 1$, the function g_i is continuous. Let $\alpha = \min_i \gamma(B(x_i, \varepsilon) \times B(y_i, \varepsilon))$ and choose $0 < \tilde{\varepsilon}_i \le \varepsilon$ so that $g_i(\tilde{\varepsilon}_i) = \alpha$ for all i; this is possible since $\lim_{r \to 0} g_i(r) = 0$.

Following now the proof of sufficiency of the previous theorem we obtain that $\gamma\lfloor_B$ for $B = \cup_{i=1}^n B(x_i, \tilde{\varepsilon}_i) \times B(y_i, \tilde{\varepsilon}_i)$ and for $\varepsilon > 0$ small enough is not an optimal transport between its marginals, which contradicts (2). □

Remark 4.7. It is natural to ask what happens if $|x - y|$ is replaced by a more general (real-valued) cost function $c(x, y)$, that is, we consider the functional

$$\gamma \mapsto \gamma\text{-} \operatorname*{ess\,sup}_{(x,y) \in \Omega^2} c(x, y).$$

As expected, the basic existence results, Proposition 2.1 and Theorem 3.2, remain valid provided that c is non-negative and lower semicontinuous with all the relevant concepts appropriately redefined: in particular, in the definition of *infinite cyclical monotonicity* one should replace the support of γ with some appropriate set on which γ is concentrated. Moreover, the proof of the equivalence of restrictable and infinitely cyclically monotone solutions works in this generality if $c(x, y)$ is (uniformly) continuous.

5. Existence and uniqueness of an optimal transport map

In this section, we prove that under reasonably weak assumptions an infinitely cyclically monotone transport plan is induced by a transport map. Moreover, we start the analysis of the uniqueness of such transport maps, and then comment our method of proof in the light of the duality issue for problem (P_∞).

5.1. **Properties of transport plans.** We begin by considering some generic properties of transport plans. This subsection is largely independent of the cost, and the technique detailed below has applications also in the framework of classical transportation problems involving cost functionals in integral form, see [15].

Definition 5.1. Let $y \in \Omega$, $r > 0$ and let $\gamma \in \Pi(\mu, \nu)$ be a transport plan. We define

$$\gamma^{-1}(B(y, r)) := \pi^1((\Omega \times B(y, r)) \cap \operatorname{supp} \gamma).$$

In other words, $\gamma^{-1}(B(y, r))$ is the set of points whose mass is partially or completely transported to $B(y, r)$ by γ. We recognize the slight abuse of notation, but if γ is thought as a device that transports mass, then this seems justifiable. Notice also that $\gamma^{-1}(B(y, r))$ is a Borel set. In fact, it is a countable union of compact sets as shown by the equation

$$\begin{aligned} \pi^1((\Omega \times B(y, r)) \cap \operatorname{supp} \gamma) &= \pi^1(\bigcup_n (\Omega \times \overline{B(y, r - 1/n)}) \cap \operatorname{supp} \gamma) \\ &= \bigcup_n \pi^1((\Omega \times \overline{B(y, r - 1/n)}) \cap \operatorname{supp} \gamma). \end{aligned}$$

Since this notion is important in the sequel, we recall that when A is $\mathcal{L}^d$-measurable, one has

$$\lim_{r\to 0}\frac{\mathcal{L}^d(A\cap B(x,r))}{\mathcal{L}^d(B(x,r))} \;=\; 1$$

for almost every x in A: we shall call such a point x a Lebesgue point of A, this terminology deriving from the fact that such a point may also be considered as a Lebesgue point of χ_A .

The following Lemma, although quite simple, is the cornerstone of the proof of Theorem 5.5 below.

Lemma 5.2. *Let $\gamma \in \Pi(\mu,\nu)$ and assume that $\mu << \mathcal{L}^d$. Then γ is concentrated on a σ-compact set $R(\gamma)$ such that for all $(x,y)\in R(\gamma)$ the point x is a Lebesgue point of $\gamma^{-1}(B(y,r))$ for all $r>0$.*

Proof. In the following, we shall denote by $\mathrm{Leb}(E)$ the set of points $x\in E$ which are Lebesgue points of E. Let

$$A := \{(x,y)\in \mathrm{supp}(\gamma) \;:\; x\notin \mathrm{Leb}(\gamma^{-1}(B(y,r))) \text{ for some } r>0\};$$

we first intend to show that $\gamma(A)=0$. To this end, for each positive integer n we consider a finite covering $\Omega\subset \displaystyle\bigcup_{i\in I(n)} B(y_i^n,\frac{1}{2n})$ by balls of radius $\frac{1}{2n}$. We notice that if $(x,y)\in\mathrm{supp}(\gamma)$ and x is not a Lebesgue point of $\gamma^{-1}(B(y,r))$ for some $r>0$, then for any $n\geq\frac{1}{r}$ and y_i^n such that $|y_i^n-y|<\frac{1}{2n}$ the point x belongs to $\gamma^{-1}(B(y_i^n,\frac{1}{2n}))$ but is not a Lebesgue point of this set. Then

$$\pi^1(A)\;\subset\;\bigcup_{n\geq 1}\;\bigcup_{i\in I(n)}\left(\gamma^{-1}(B(y_i^n,\frac{1}{2n}))\setminus \mathrm{Leb}(\gamma^{-1}(B_\gamma(y_i^n,\frac{1}{2n})))\right).$$

Notice that the set on the right hand side has Lebesgue measure 0, and thus μ-measure 0. It follows that $\gamma(A)\leq\gamma(\pi^1(A)\times\Omega)=\mu(\pi^1(A))=0$.

Finally, since $\mathcal{L}^d(\pi^1(A))=0$, there exists a sequence $(U_k)_{k\geq 0}$ of open sets such that

$$\text{for all } k\geq 0,\qquad \pi^1(A)\subset U_k,\qquad \text{and}\qquad \lim_{k\to\infty}\mathcal{L}^d(U_k)=0.$$

Then the set $R(\gamma)\;:=\;\mathrm{supp}(\gamma)\cap(\displaystyle\bigcup_{k\geq 0}(\Omega\setminus U_k)\times\Omega)$ has the desired properties. □

The above Lemma yields us to introduce the following notion:

Definition 5.3. The couple $(x,y)\in\Omega\times\Omega$ is a γ-*regular point* if $x\in\gamma^{-1}(B(y,r))$ is a Lebesgue point of this set for any positive r.

Notice that any element of the set $R(\gamma)$ of Lemma 5.2 is a γ-regular point.

For future use, we introduce a suitable notation to indicate a cone: let $x_0,\xi\in\mathbb{R}^d$ with $|\xi|=1$ and let $\delta\in[0,2]$. Then we define

$$C(x_0,\xi,\delta):=\{x\in\mathbb{R}^d\setminus\{x_0\}\;:\;\frac{x-x_0}{|x-x_0|}\cdot\xi\geq 1-\delta\}.$$

Notice that if $\delta = 0$, $C(x_0, \xi, 0)$ degenerates to a half-line, while in the case $\delta = 2$, $C(x_0, \xi, 2)$ is $\mathbb{R}^d \setminus \{x_0\}$.

We now remark the following property for the regular points of a transport plan:

Proposition 5.4. *Let (x_0, y_0) be a γ-regular point, $r > 0$, $\alpha \in (0,1)$ and $\delta > 0$. Then for $\varepsilon > 0$ sufficiently small the set of points $x \in \gamma^{-1}(B(y_0, r))$ such that $x \in C(x_0, \xi, \delta) \cap (B(x_0, \varepsilon) \setminus B(x_0, \alpha\varepsilon))$ has positive $\mathcal{L}^d$ measure.*

Proof. It is enough to remark that $x_0 \in \mathrm{Leb}(\gamma^{-1}(B(y_0, r)))$ and then

$$\lim_{\varepsilon \to 0} \frac{\mathcal{L}^d((B(x_0,\varepsilon) \setminus B(x_0, \alpha\varepsilon)) \cap C(x_0, \xi, \delta) \cap \gamma^{-1}(B(y_0, r)))}{\mathcal{L}^d(B(x_0, \varepsilon))} = c(\alpha, \delta),$$

where $c(\alpha, \delta) := \dfrac{\mathcal{L}^d(((B(x_0, 1) \setminus B(x_0, \alpha)) \cap C(x_0, \xi, \delta))}{\mathcal{L}^d(B(x_0, 1))} > 0.$ □

5.2. **Existence of an optimal transport map.** Our main result in this subsection is the following theorem, which states that under the hypothesis that μ is absolutely continuous with respect to the Lebesgue measure $\mathcal{L}^d$, any optimal infinitely cyclically monotone transport plan for (P_∞) is induced by a transport map. This generalizes the corresponding result for the problem (P_p) when $p \in]1, \infty[$: if one assumes that

$$\mu(B) = 0 \quad \text{whenever} \quad \mathcal{H}^{d-1}(B) < \infty \tag{5.1}$$

then any p-cyclically monotone transport plan is induced by a transport map (see Remark 4.7 in [20]).

It is in doubt whether (5.1) is sufficient to ensure that the conclusion of Theorem 5.5 below holds. In the case $p = 1$, the hypothesis (5.1) is not sufficient to ensure that any 1-cyclically monotone transport plan is induced by a transport map. Even worse, for $\mu = \mathcal{L}^1 \lfloor_{[0,1]}$ and $\nu = \frac{1}{2}\mathcal{L}^1 \lfloor_{[0,2]}$ there exists an optimal (and hence 1-cyclically monotone) transport plan which is not induced by a transport map, see [2, p. 125].

We now state the main result of this section, and refer to §5.4 for further comments.

Theorem 5.5. *Assume that $\mu << \mathcal{L}^d$ and let $\gamma \in \Pi(\mu, \nu)$ be an infinitely cyclically monotone transport plan. Then there exists a Borel transport map $T : \Omega \to \Omega$ such that $\gamma = (id \times T)_\# \mu$.*

Proof. By Proposition 2.1 in [1], it is sufficient to prove that γ is concentrated on a γ-measurable graph. In view of Lemma 5.2, it is then sufficient to prove that $R(\gamma)$ is included in a graph, or more generally that if (x_0, y_0) and (x_0, y_0') are both γ-regular points then $y_0 = y_0'$.

We divide the proof into two parts, and first show that $|x_0 - y_0| = |x_0 - y_0'|$. Arguing by contradiction, we assume that $|x_0 - y_0| < |x_0 - y_0'|$ and suppose for the time being that $x_0 \neq y_0$. Let $\xi' = \frac{y_0' - x_0}{|y_0' - x_0|}$, $0 < \varepsilon < |x_0 - y_0|$ and $0 < r < |y_0' - x_0| - |y_0 - x_0|$. We claim that for $\delta := 1 - \frac{|x_0 - y_0|}{|x_0 - y_0'|}$ one has

$$\max\{|x - y_0'|, |x_0 - y|\} < \max\{|x - y|, |x_0 - y_0'|\} \tag{5.2}$$

for any (x,y) such that $x \in C(x_0,\xi',\delta) \cap B(x_0,\varepsilon) \setminus B(x_0,\frac{1}{2}\varepsilon)$ and $y \in B(y_0,r)$. Indeed, take (x,y) as above, it then follows from the choice of r that $|x_0-y| < |x_0-y_0'|$, while on the other hand

$$\begin{aligned}|x-y_0'|^2 &= |x-x_0|\left(|x-x_0| - 2\frac{x-x_0}{|x-x_0|}\cdot(y_0'-x_0)\right) + |x_0-y_0'|^2 \\ &\leq |x-x_0|\left(|x_0-y_0| - 2(1-\delta)|x_0-y_0'|\right) + |x_0-y_0'|^2 \ < \ |x_0-y_0'|^2.\end{aligned}$$

This proves the claim. We now infer from Proposition 5.4 that the set of points $x \in \gamma^{-1}(B(y_0,r))$ such that $x \in C(x_0,\xi',\delta) \cap B(x_0,\varepsilon) \setminus B(x_0,\frac{1}{2}\varepsilon)$ has positive measure when ε is small enough. In particular, this set is non-empty for small ε, and (5.2) then clearly contradicts the infinite cyclical monotonicity of γ. As a consequence, $|x_0-y_0| = |x_0-y_0'|$ in the case $x_0 \neq y_0$.

If $x_0 = y_0$, we repeat the argument above with the choices $0 < \varepsilon < \frac{1}{4}|x_0-y_0'|$, $0 < r < \frac{1}{4}|y_0'-x_0|$ and $\delta = \frac{1}{2}$. Then for any (x,y) such that $x \in C(x_0,\xi',\delta) \cap B(x_0,\varepsilon) \setminus B(x_0,\frac{1}{2}\varepsilon)$ and $y \in B(y_0,r)$, we clearly have $|x_0-y| < |x_0-y_0'|$ and the other inequality $|x-y_0'|^2 < |x_0-y_0'|^2$ follows as above.

We now prove by contradiction that $y_0 = y_0'$. Note that since we already know that $|x_0-y_0| = |x_0-y_0'|$, we may assume that $|x_0-y_0| = |x_0-y_0'| > 0$. If $y_0 \neq y_0'$, we can find $\xi \in \mathbb{R}^d$ such that

$$\xi \cdot \frac{x_0-y_0}{|x_0-y_0|} < 0 \quad \text{and} \quad \xi \cdot \frac{x_0-y_0'}{|x_0-y_0'|} > 0.$$

Next we choose $r > 0$ such that

$$\sup\left\{\xi \cdot \frac{x_0-y}{|x_0-y|} \ : \ y \in B(y_0,r)\right\} \ < \ 0$$

and $\delta > 0$ such that

$$\alpha \ := \ \inf\left\{\frac{x_0-x}{|x_0-x|}\cdot\frac{x_0-y_0'}{|x_0-y_0'|} \ : \ x \in C(x_0,\xi,\delta)\right\} \ > \ 0 \tag{5.3}$$

as well as

$$\sup\left\{\frac{x_0-x}{|x_0-x|}\cdot\frac{x_0-y}{|x_0-y|} \ : \ x \in C(x_0,\xi,\delta),\ y \in B(y_0,r)\right\} \ < \ 0. \tag{5.4}$$

We now claim that for $\varepsilon > 0$ small enough, (5.2) holds for any (x,y) such that $x \in C(x_0,\xi,\delta) \cap B(x_0,\varepsilon) \setminus B(x_0,\frac{1}{2}\varepsilon)$ and $y \in B(y_0,r)$. Notice that this claim concludes the proof of $y_0 = y_0'$ modulo applying Proposition 5.4 as before. To verify that (5.2) holds, we first notice that (5.4) implies

$$\text{for all } x \in C(x_0,\xi,\delta),\ y \in B(y_0,r), \qquad |x_0-y| \ < \ |x-y| \tag{5.5}$$

since $|x-y|^2 = |x-x_0|^2 - 2(x_0-x)\cdot(x_0-y) + |x_0-y|^2$. We can also infer from (5.3) that

$$|x-y_0'|^2 \ \leq \ |x-x_0|\left(|x_0-x| - 2\alpha|x_0-y_0'|\right) + |x_0-y_0'|^2$$

for any $x \in C(x_0,\xi,\delta)$. It follows that

$$\text{for all } x \in C(x_0,\xi,\delta) \cap B(x_0,\varepsilon) \qquad |x-y_0'| \ < \ |x_0-y_0'| \tag{5.6}$$

whenever $0 < \varepsilon < 2\alpha|x_0 - y'_0|$. We then get (5.2) from (5.5) and (5.6), which concludes the proof. □

5.3. **Uniqueness of the infinitely cyclically monotone transport map.** We now consider the question of the uniqueness of the infinitely cyclically monotone transport map obtained in the preceding section. We recall that when (5.1) holds and $p \in]1, \infty[$, problem (P_p) admits a unique (up to a μ negligible set) p-cyclically monotone transport map (see for example [20] and §5.4). Notice that in contrast this result does not hold for $p = 1$, not even under the stronger hypothesis that $\mu << \mathcal{L}^d$, as shown by Example 1.3 in [1]: when $\mu = \mathcal{L}^1 \lfloor_{[0,1]}$ and $\nu = \mathcal{L}^1 \lfloor_{[1,2]}$, both transport maps $t \mapsto t + 1$ and $t \mapsto 2 - t$ are optimal.

In the case of problem (P_∞), the question of uniqueness is largely open. At the moment, we only have the following partial result stating the uniqueness of infinitely cyclically monotone transport map under the hypothesis that ν is purely atomic with finite support.

Theorem 5.6. *Suppose that $\mu << \mathcal{L}^d$ and $\nu = \sum_{i=0}^{k} a_i\, \delta_{y_i}$ for some $(y_i)_{0 \le i \le k} \subset \Omega$ and positive numbers $a_0, \dots, a_k$. Then there exists a unique (up to a μ-negligible set) infinitely cyclically monotone Borel transport map T from μ to ν.*

Proof. Suppose that there are two distinct infinitely cyclically monotone Borel transport maps T and $\tilde{T}$, and let us introduce the sets

$$U^i_j = T^{-1}(y_j) \cap \tilde{T}^{-1}(y_i).$$

We first claim that it is possible to define a sequence of integers $(i(p))_{p \ge 0}$ such that

$$\text{for all} \quad p \ge 0, \qquad i(p) \ne i(p+1) \quad \text{and} \quad \mu\left(U^{i(p+1)}_{i(p)}\right) > 0.$$

Indeed, the fact that the two transport maps are distinct means that it is possible to choose two indices $i(0) \ne i(1)$ such that $\mu\left(U^{i(1)}_{i(0)}\right) > 0$. Next we notice that since $\tilde{T}$ maps μ to ν,

$$\nu(\{y_{i(1)}\}) \;=\; \sum_{p=0}^{k} \mu\left(U^{i(1)}_p\right) \;\ge\; \mu\left(U^{i(1)}_{i(0)}\right) + \mu\left(U^{i(1)}_{i(1)}\right) \;>\; \mu\left(U^{i(1)}_{i(1)}\right).$$

Since T also maps μ to ν, we infer from the above inequality that there exists $p \ne i(1)$ such that $\mu\left(U^p_{i(1)}\right) > 0$: we then set $i(2) = p$ and start again from $U^{i(2)}_{i(1)}$. By repeating the above argument we can build recursively the sequence $(i(p))_{p \ge 0}$ with the desired properties.

Since the sequence $(i(p))_{p \ge 0}$ takes its values in the finite set $\{0, \dots, k\}$, we may assume that there exists some $m \ge 2$ such that $i(m) = i(0)$. For any $p \in \{0, \dots, m-1\}$, the set $U^{i(p+1)}_{i(p)}$ has non-zero Lebesgue measure so we may choose a Lebesgue point x_p of $U^{i(p+1)}_{i(p)}$ for which $|y_{i(p+1)} - x_p| \ne |y_{i(p)} - x_p|$, and then set $x_m = x_0$. By definition,

$$T(x_p) = y_{i(p)} \quad \text{and} \quad \tilde{T}(x_p) = y_{i(p+1)} \qquad \text{for all } p \in \{0, \dots, m-1\}.$$

Since T and $\tilde{T}$ are infinitely cyclically monotone, we have

$$\max_{0\le p\le m-1}|y_{i(p)}-x_p| \;\le\; \max_{0\le p\le m-1}|y_{i(p+1)}-x_p| \;\le\; \max_{0\le p\le m-1}|y_{i(p)}-x_p|,$$

so that

$$\max_{0\le p\le m-1}|y_{i(p)}-x_p| \;=\; \max_{0\le p\le m-1}|y_{i(p+1)}-x_p|. \tag{5.7}$$

Let then $I := \{q : |y_{i(q)}-x_q| = \max_{0\le p\le m-1}|y_{i(p)}-x_p|\}$. We infer from (5.7) and the choice of the points x_p that for any $q\in I$ one has $|y_{i(q)}-x_q| > |y_{i(q+1)}-x_q|$. Since x_q is a Lebesgue point of $U_{i(q)}^{i(q+1)}$, there exists $\tilde{x}\in U_{i(q)}^{i(q+1)}$ arbitrarily close to x_q for which

$$|y_{i(q)}-\tilde{x}| \;>\; |y_{i(q)}-x_q|,$$

and thus we can choose $\tilde{x}\in U_{i(q)}^{i(q+1)}$ such that

$$|y_{i(q)}-\tilde{x}| \;>\; \max\{|y_{i(q+1)}-\tilde{x}|, |y_{i(q)}-x_q|\}. \tag{5.8}$$

We now set $\tilde{x}_q := \tilde{x}$ as well as $\tilde{x}_p := x_p$ for $p\neq q$, and notice that (5.7) should in fact also hold for this new choice of elements $\tilde{x}_p$ in $U_{i(p)}^{i(p+1)}$. This leads to a contradiction since we can infer from the definition of I, the choice of q and (5.8) that

$$\begin{aligned}\max_{0\le p\le m-1}|y_{i(p)}-\tilde{x}_p| \;&=\; |y_{i(q)}-\tilde{x}| \\ &>\; \max\{|y_{i(q+1)}-\tilde{x}|, |y_{i(q)}-x_q|\} \;\ge\; \max_{0\le p\le m-1}|y_{i(p+1)}-\tilde{x}_p|.\end{aligned}$$

This is contradiction with (5.7), and concludes the proof of the theorem. □

Remark 5.7. The construction of the sequence $(x_p)_{0\le p\le m}$ is close to that proposed in the proof of Theorem 3.4, but it is easier since we don't need that it loops at x_0. Indeed, in the above proof we don't really need that $x_m = x_0$, and we only assume this for convenience of notations, while in the course of the proof of Theorem 3.4 we intended to use *Claim 1* and then had to start from the special x_0 found there.

At the moment we are not able to generalize this result to the case where ν is any probability measure on Ω. On the other hand, it is clear that the above uniqueness theorem requires that μ does not concentrate, as the following example shows.

Example 5.8. Assume that $\mu := \mathcal{H}^1\lfloor_{[0,1]\times\{0\}}$ while $\nu := \frac{1}{2}\left(\delta_{(0,-1)}+\delta_{(0,1)}\right)$. Then any transport map T (i.e. any μ-measurable function for which $T_\#\mu=\nu$) is an infinitely cyclically monotone optimal transport map from μ to ν.

5.4. **Comments around duality.** For $1\le p<\infty$, the mass transport problem (P_p) may be rewritten as

$$(P_p) \qquad W_p^p(\mu,\nu) \;=\; \inf\left\{\mathcal{C}_p^p(\lambda) \;=\; \int_{\Omega^2}|y-x|^p d\lambda(x,y) \;:\; \lambda\in\Pi(\mu,\nu)\right\}.$$

In this form, the objective functional $\lambda\mapsto\mathcal{C}_p^p(\lambda)$ is linear over the compact convex set $\Pi(\mu,\nu)$, and it is then quite natural to associate to (P_p) its dual problem

$$(D_p) \qquad \sup\left\{\int_\Omega\phi(x)d\mu(x)+\int_\Omega\psi(y)d\nu(y) \;:\; \phi(x)+\psi(y)\le|y-x|^p\right\},$$

where $\phi \in L^1(d\mu)$, $\psi \in L^1(d\nu)$ and the constraint is required to hold for μ a.e. x and ν a.e. y. Due to the regularity of the integrand $c_p(x,y) := |y-x|^p$, the supremum of (D_p) is achieved for a couple (φ, φ^{c_p}) where the Kantorovich potential φ is continuous and c_p-concave. We refer for example to §3 of [2], Part I of [20], §3.3 of [26] or §2.4 of [32] for more on the related concepts and results.

The Kantorovich dual problem (D_p) appears to be a fundamental tool in understanding and solving the problems of the characterization, existence and uniqueness for an optimal transport map for (P_p). For example, the notion of p-cyclical monotonicity (3.1) naturally appears via the equality $\sup(D_p) = \inf(P_p)$, see e.g. the proof of Theorem 3.2 in [2] (an alternative and direct proof is for example that of Theorem 2.3 in [20]). Moreover, a key point in the construction of an optimal transport map $T_p : \mathrm{supp}(\mu) \to \mathrm{supp}(\nu)$ for (P_p) is to identify the directions of transportation (known as *transport rays* for $p = 1$), that is to associate to μ-almost every $x \in \mathrm{supp}(\mu)$ the direction $\frac{T_p(x)-x}{|T_p(x)-x|}$ to which the mass present at x is transferred. It is now well understood that this direction may be obtained as the adequate c_p-gradient of an optimal Kantorovich potential φ (see e.g. [20], [19], [27] or §2.4 of [32]). In fact, the definition as well as the regularity properties of the transport rays are deeply linked with the fact that the support of an optimal transport plan γ_p for (P_p) is p-cyclically monotone, and thus inherits good properties from being included in the subdifferential of a c_p-concave function (which in turn turns out to be a Kantorovich potential, see e.g. §2.4 of [32]).

In light of the preceding discussion, it is natural to try to develop a duality theory for the problem (P_∞) as well. We hereafter informally discuss this issue.

First, in view of the Example 3.3, it does not seem realistic that one could obtain useful information from an intrinsic approach. Indeed, the optimal transport plan λ proposed in Example 3.3 is not induced by any transport map, so we cannot expect that a dual problem directly associated to (P_∞) via some general construction gives information on the geometry of the optimal transport plans. On the other hand, notice that Theorems 5.5 and 5.6 above do indicate that there exists a unique infinitely cyclically monotone optimal transport map for the particular problem of Example 3.3.

In view of the results of the two preceding sections, and since the notion of infinitely cyclically monotone plan was at first obtained via a limiting argument, one is led to study the asymptotic behavior of the family of dual problems (D_p) as $p \to \infty$. But as mentioned above, (D_p) is not directly related to (P_p) but to a reformulation of (P_p) which requires taking the p-power of the objective functional $\mathcal{C}_p$. As a consequence, one should in fact take the $\frac{1}{p}$-power of the objective functional of (D_p) and then study the limiting problem as $p \to \infty$; unfortunately, our research in this direction has been unfruitful up to now. Finally, since the functional $\mathcal{C}_p$ is not convex in λ, the convex duality theory does not apply directly to (P_p). But one may wonder whether it is possible to overcome this difficulty and associate to (P_p) a dual problem with a structure similar to that of (D_p): this is a quite involved question known as Dudley's problem (see e.g. (1.1.10) and Remark 2.6.2 in [26]), and it is out of the scope of the present study.

Despite the above difficulties, we believe that developing a duality theory for the problem (P_∞) is an important issue since it would yield a deeper understanding of the problem of the existence and uniqueness for a particular optimal transport map.

Acknowledgments

The authors would like to thank Guy Bouchitté for proposing the problem and for several fruitful discussions. Part of this research was conducted while the authors were visiting each other at their respective institutions. In particular Thierry Champion's visit to Pisa in November 2006 was supported by the I.N.D.A.M project "Traffic flows and optimization on complex networks" and by the G.N.A.M.P.A project "Fenomeni di evoluzione non lineari suggeriti dalla Fisica e dalla Biologia", and his visit to Jyväskylä in June 2007 was supported by the ESF program "Global and geometric properties of solutions of nonlinear partial differential equations". The second author is partially supported by the italian M.I.U.R. project "Metodi variazionali nella teoria del trasporto ottimo di massa e nella teoria geometrica della misura" and by the "Fondo di ateneo per la ricerca" of the University of Pisa. The third author is supported by the Academy of Finland project 108374. The authors acknowledge these supports.

References

[1] Ambrosio, L., Lecture Notes on Optimal Transportation Problems, Mathematical aspects of evolving interfaces (Funchal, 2000), Lecture Notes in Math., **1812**, Springer, Berlin, 2003, 1–52.

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

[3] Aronsson, G., *Minimization problems for the functional* $\sup_x F(x, f(x), f'(x))$, Ark. Mat. 6 (1965) 33–53.

[4] Aronsson, G., *Extension of functions satisfying Lipschitz conditions*, Ark. Mat. 6 (1967), 551–561.

[5] Aronsson, G., Crandall, M. G., and Juutinen, P., *A tour of the theory of absolutely minimizing functions*, Bull. Amer. Math. Soc. (N.S.) 41 (2004), no. 4, 439–505

[6] E. N. Barron, R. R. Jensen, and C. Y. Wang, *The Euler equation and absolute minimizers of L^∞ functionals*, Arch. Ration. Mech. Anal. 157 (2001), no. 4, 255–283.

[7] Bernot, M., Caselles, V., Morel, J. M., *Are there infinite irrigation trees?*, J. Math. Fluid Mech. 8 (2006), no. 3, 311–332.

[8] Bouchitté, G., Jimenez, C., Rajesh, M., *An L^∞ estimate in Mass Transportation*, Proc. Amer. Math. Soc., posted on July 3, 2007, PII S 0002-9939(07)08877-6 (to appear in print).

[9] Brancolini, A., Buttazzo, G., Santambrogio, F., *Path functionals over Wasserstein spaces*, J. Eur. Math. Soc. (JEMS) 8 (2006), no. 3, 415–434.

[10] Benamou, J. D., Brenier, Y., *A computational fluid mechanics solution to the Monge–Kantorovich mass transfer problem*, Numer. Math., **84** (2000), 375–393.

[11] Buttazzo, G., Pratelli, A., and Stepanov, E., *Optimal pricing policies for public transportation networks*, SIAM J. Optim. 16 (2006), no. 3, 826–853.

[12] Carrillo, J. A., Di Francesco, M., Lattanzio, C., *Contractivity and asymptotics in Wasserstein metrics for viscous nonlinear scalar conservation laws*, Boll. Unione Mat. Ital. Sez. B Artic. Ric. Mat. (8) 10 (2007), no. 2, 277–292.

[13] Carrillo, J.A., Gualdani, M. P., Toscani, G., *Finite speed of propagation in porous media by mass transportation methods*, C. R. Math. Acad. Sci. Paris 338 (2004), no. 10, 815–818.

[14] Champion, T., and De Pascale L., *A principle of comparison with distance functions for absolute minimizers*, J. Convex Anal., 14 (2007), No. 3., 515–541.

[15] Champion, T., and De Pascale L., in preparation.

[16] De Pascale, L., Pratelli, A., *Regularity properties for Monge transport density and for solutions of some shape optimization problem*, Calc. Var. Partial Differ. Equ., **14** (2002), no. 3, 249–274.

[17] Dines, L. L., *On positive solutions of a system of linear equations*, Ann. of Math. (2) 28 (1926/27), no. 1-4, 386–392.

[18] Evans, L. C., *Partial Differential Equations and Monge–Kantorovich Mass Transfer*, Current developments in mathematics, 1997 (Cambridge, MA), Int. Press, Boston, MA, (1999), 65–126.

[19] Evans, L. C., Gangbo, W., *Differential Equations Methods for the Monge–Kantorovich Mass Transfer Problem*, Mem. Amer. Math. Soc., Vol. **137** (1999).

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

[21] Hauray, M., Jabin, P.-E., *N-particles approximation of the Vlasov equations with singular potential*, Arch. Ration. Mech. Anal. 183 (2007), no. 3, 489–524.

[22] Hauray, M., *Personal communication.*

[23] Maddalena, F., Morel, J. M., Solimini, S., *A variational model of irrigation patterns*, Interfaces Free Bound. 5 (2003), 391–415.

[24] McCann, R. J., *Stable rotating binary stars and fluid in a tube*, Houston J. Math. 32 (2006), no. 2, 603–631.

[25] Pratelli, A., *On the sufficiency of c-cyclical monotonicity for optimality of transport plans*, Math. Z., to appear; DOI 10.1007/s00209-007-0191-7.

[26] Rachev, S., Rüschendorf, L., Mass transportation problems. Vol. I. Theory. Probability and its Applications, Springer-Verlag, New York (1998).

[27] Rüschendorf, L., *Optimal solutions of multivariate coupling problems*, Appl. Math. (Warsaw) 23 (1995), no. 3, 325–338.

[28] Rüschendorf, L., *On $c-$optimal random variables*, Statistic & Probability letters, **27** (1996), 267–270.

[29] Santambrogio, F., *Variational problems in transport theory with mass concentration*, PhD Thesis, Scuola Normale Superiore, Pisa (2007), (available at http://cvgmt.sns.it).

[30] Savin, O., *C^1 regularity for infinity harmonic functions in two dimensions*, Arch. Ration. Mech. Anal. 176 (2005), no. 3, 351–361.

[31] Schachermayer, W., and Teichmann, J., *Characterization of optimal Transport Plans for the Monge-Kantorovich-Problem*, Proc. Amer. Math. Soc., to appear.

[32] Villani, C., Topics in optimal transportation. Graduate Studies in Mathematics, 58, American Mathematical Society (2003)

[33] Villani, C., Optimal Transport, Old and New. Available at http://www.umpa.ens-lyon.fr/~cvillani/surveys.html#oldnew

[34] Xia, Q., *Optimal paths related to transport problems*, Commun. Contemp. Math. 5, 251–279 (2003).

T.C. Laboratoire d'Analyse Non Linéaire Appliquée, U.F.R. des Sciences et Techniques, Université du Sud Toulon-Var, Avenue de l'Université, BP 20132, 83957 La Garde cedex, FRANCE

L.D.P. Dipartimento di Matematica Applicata, Universitá di Pisa, Via Buonarroti 1/c, 56127 Pisa, ITALY

P.J. Department of Mathematics and Statistics, P.O.Box 35 (MaD), FI-40014 University of Jyväskylä, Finland