HOW TO SHOW THAT SOME RAYS ARE MAXIMAL TRANSPORT RAYS IN MONGE PROBLEM

ALDO PRATELLI

ABSTRACT. In this paper we prove a result, extending Lemma 8.1 in [2], which allows to proof that a set of segments is the set of the maximal transport rays for a transport problem. This is particularly useful to build non-trivial examples of transport maps, then in particular to provide specific examples (or counterexamples) in mass transportation. We also give some of these examples.

1. INTRODUCTION

1.1. **The mass transportation problem.** We are interested in the mass transportation problem, first proposed by Monge [10] in 1781; in today's language, we can state it as follows. The ambient space is the closure Ω of an open, bounded and convex subset $\Omega^\circ \subseteq \mathbb{R}^2$, and we are given two probability measures f^+ and f^- on Ω. A *transport map* from f^+ to f^- is a Borel map $t : \Omega \to \Omega$ such that $t_\# f^+ = f^-$, where the push-forward $t_\# : \mathcal{M}^+(\Omega) \to \mathcal{M}^+(\Omega)$ is defined as $t_\# f^+(B) := f^+\big(t^{-1}(B)\big)$ for any Borel set $B \subseteq \Omega$. If we think that f^+ and f^- measure respectively the distribution of some mass and the depth of a hole, then any transport map is a "strategy" to move all the mass inside the hole, covering the latter. To any transport map we associate a cost given by

$$\mathcal{C}(t) := \int_\Omega |t(x) - x| \, df^+(x), \tag{1.1}$$

and the mass transportation problem consists in finding an *optimal transport map*, that is, a transport map minimizing the cost $\mathcal{C}$. In general, there may be no optimal transport maps, or even no transport maps at all; however, one can consider a sort of relaxation of this problem, as first proposed by Kantorovich ([8, 9]) in the '40s: we call *transport plan* between f^+ and f^- any probability measure $\gamma \in \mathcal{P}(\Omega \times \Omega)$ such that the two projections $\pi_1\gamma$ and $\pi_2\gamma$ of γ on Ω coincide with f^+ and f^- respectively. Moreover, to any transport plan γ one associates the cost

$$\mathcal{C}(\gamma) := \iint_{\Omega\times\Omega} |y - x| \, d\gamma(x, y). \tag{1.2}$$

Any transport map t is "naturally" associated to the transport plan $\gamma_t = (1, t)_\# f^+$ (in fact, it is trivially checked that γ_t is a transport plan whenever t is a transport map, as well as $\mathcal{C}(t) = \mathcal{C}(\gamma_t)$). However, there are much more transport plans than transport maps, and it can be shown that, unlike the transport maps, there always exist transport plans (for instance, $f^+ \otimes f^-$ is such a plan), and there always exist also *optimal transport plans*, that is, transport plans minimizing $\mathcal{C}$ (since $\mathcal{C}$ is easily a weakly* l.s.c. functional and the set of the transport plans is weakly* compact in $\mathcal{M}^+(\Omega \times \Omega)$). The mass transportation problem is widely studied, and in its more general statements the ambient space Ω is replaced by a subset of $\mathbb{R}^N$, or a manifold, or even a general metric space, and the Euclidean distance $|\cdot|$ appearing in (1.1) and (1.2) is replaced by a more general norm, or by a distance, or even by a generic l.s.c. functions: there is a huge literature about this field, for a general reference we only quote [12, 1, 11].

We consider now the maximization problem, usually referred to as *shape optimization problem*, which consists in maximizing the functional

$$I(u) := \int_\Omega u(x)\, d(f^+ - f^-)(x) \tag{1.3}$$

among the $1-$Lipschitz functions $u : \Omega \to \mathbb{R}$; any maximizer is called *optimal shape*. It is well known (see for instance [4, 5, 1]) the following result, which states a deep (and at first glance unexpected) connection between the two problems.

Theorem 1.1. $\inf(1.2) = \sup(1.3)$ *and both the extremals are reached. Moreover, given any optimal shape u and any optimal plan γ, one has $u(x) - u(y) = |y - x|$ for $\gamma-$a.e. $(x, y) \in \Omega \times \Omega$.*

As the Theorem above enlightens, the segments on which the mass is moved, *i.e.* the segments xy with $(x, y) \in \mathrm{spt}\gamma$, have particular properties. To understand them, we need some standard notation (see for instance [6, 2, 11]): given an optimal transport plan γ, we say that

- an open oriented segment xy is a *transport ray* if $(x, y) \in \mathrm{spt}\gamma$;
- an open oriented segment xy is a *maximal transport ray* if any $z \in xy$ belongs to the closure of some transport ray contained into xy, and xy is not strictly contained into any segment with the same property;
- a point z is a *doubling point* if it belongs to the closure of at least two different maximal transport rays;
- $T := \big\{z \in \mathbb{R}^2 :\ z \text{ is contained into some maximal transport ray}\big\}$ is the *transport set*.

It is to be remarked that a maximal transport ray need not to be a transport ray; we also notice the following facts.

Remark 1.2. *Thanks to Theorem 1.1, for any optimal shape u and any optimal transport plan γ, u is decreasing at maximal slope on any maximal transport ray; this implies, in particular, that different maximal transport rays can not cross (recall that the maximal transport rays are open segments).*

By the linearity of the functional (1.2) as well as of the constraints $\pi_1\gamma = f^+$, $\pi_2\gamma = f^-$, it is clear that any convex combination of optimal transport plans is still an optimal transport plan. Therefore, the above Remark says, more generally, that two different maximal transport rays can not cross, even if they are obtained from two different optimal transport plans. In fact, usually the maximal transport rays do not depend on the choice of an optimal transport plan, see Remark 2.12.

From now on, we will consider the maximal transport rays as oriented following the slope of u (by Theorem 1.1, this orientation does not depend on the choice of the optimal shape u): therefore, whenever $(x, y) \in \mathrm{spt}\gamma$ we have that xy belongs to some maximal transport ray and $x > y$. Finally, from the definition it follows also that any doubling point does not belong to T, since it is an endpoint of all the maximal transport rays to the closure of which it belongs; more precisely, it is either the upper endpoint of all the maximal transport rays to the closure of which it belongs, or the lower endpoint of all them.

We conclude this section with an easy but important consequence of Theorem 1.1; first we give a straightforward

Corollary 1.3. *Given a transport plan γ and a $1-$Lipschitz function u, one has that both γ and u are optimal (for (1.2) and (1.3) respectively) if and only if $\mathcal{C}(\gamma) = I(u)$.*

Then we can sharpen the above claim as follows:

Lemma 1.4. *Given a transport plan γ and a $1-$Lipschitz function u, one has that both γ and u are optimal (for (1.2) and (1.3) respectively) if and only if γ is concentrated on the set $\big\{(x,y)\in\Omega\times\Omega : u(x)-u(y)=|y-x|\big\}$.*

Proof. An implication is already given by Theorem 1.1; concerning the other one, taken u and γ as in the statement one can evaluate

$$\begin{aligned}\mathcal{C}(\gamma) &= \iint_{\Omega\times\Omega} |y-x|\, d\gamma(x,y) = \iint_{\Omega\times\Omega} u(x)-u(y)\, d\gamma(x,y)\\ &= \iint_{\Omega\times\Omega} u(x)\, d\gamma(x,y) - \iint_{\Omega\times\Omega} u(y)\, d\gamma(x,y) = \int_\Omega u(x)\, df^+(x) - \int_\Omega u(y)\, df^-(y)\\ &= \int_\Omega u(x)\, d(f^+-f^-)(x) = I(u)\end{aligned}$$

recalling that the projections of γ are $f^\pm$. Therefore, the thesis follows by Corollary 1.3 □

1.2. **Discerning the maximal transport rays.** As usual, also in the mass transportation is it of great interest to exhibit examples (and even more counterexamples!); however, except for situations with a broad symmetry, it is not easy to find explicit examples: to be more precise, it is often easy to build examples where one guesses the right optimal transport plan or optimal transport map, but it may be not so easy actually to show their optimality. Therefore, it can be useful a tool allowing to claim in many situations that a particular plan or map is optimal.

In view of Lemma 1.4, we know that a transport plan is optimal if and only if it has "the right maximal transport rays", that is, if all its maximal transport rays are segments of maximal slope for *some* optimal shape (equivalently, for *all* the optimal shapes): we will for simplicity refer to these segments as *admissible maximal transport rays.* Hence, an useful tool could be a result asserting which are all the admissible maximal transport rays: having this at hand, to show that a plan is optimal one could simply check that all its maximal transport rays are admissible. Moreover, as we will discuss in Remark 2.12, the maximal tranpsort rays usually do not depend on the choice of the optimal transport plan.

A tool as the one we are looking for is already present in the literature, namely Lemma 8.1 in [2]:

Lemma 1.5 (Horizontal transport rays)**.** *Denote $K=[0,1]$, and let $f^\pm$ be concentrated respectively in $K\times K$ and $[5,6]\times K$. We assume that*

$$f^+([0,1]\times[0,t]) = f^-([5,6]\times[0,t]) \qquad \forall t\in K. \tag{1.4}$$

Then the admissible maximal transport rays are the horizontal segments.

In words, this result says that a transport plan is optimal if and only if all its maximal transport rays are horizontal segments: this Lemma can be applied quite often, basically whenever one wants to build examples where an optimal transport plan moves everything horizontally. For instance, in [2] it was used to build a counterexample giving a negative answer to the following question, that was open: is it true that $f^+\ll\mathcal{H}^s$ with $1<s<2$ implies the existence of an optimal transport map? The question was interesting, since it was already known that the answer was negative for $s\le 1$ and affirmative for $s=2$, and that the answer was affirmative also for $1<s<2$ replacing the Euclidean distance $|\cdot|$ in (1.1) by its square $|\cdot|^2$.

Our main result is a generalization of Lemma 1.5, which allows to consider much more situations; in fact, as we will discuss before Example 3.5, our results works in a broad range of cases. Before to state it, we need some notation; first of all, we give the following

Definition 1.6. *A* ray configuration *is a family of pairwise disjoint open segments covering Ω°, each of them with extremes in $\partial\Omega$. Given a ray configuration $\mathscr{R} = \{R_i,\, i \in I\}$, we define its set of* doubling points *as the set $\mathscr{D}$ of all the points belonging to the closure of more than one R_i, $i \in I$ (of course, $\mathscr{D} \subseteq \partial\Omega$). Finally, a* curve transversal to $\mathscr{R}$ *(or simply a* transversal curve*) is a C^1 injective curve on Ω intersecting each segment R_i exactly once and whose gradient at any point is not parallel to the direction of the unique ray passing through that point.*

It is an easy geometric property that, given any ray configuration $\mathscr{R} = \{R_i,\, i \in I\}$, it is possible to find a transversal curve H, and it is an open curve with both the endpoints in $\partial\Omega$, so that it divides Ω in two parts, that we will refer to arbitrarily as the "left" and the "right" part, each of them intersecting all the R_i's; we equip then each segment R_i with the orientation going from the left part to the right part of the segment. We give now the following last definitions:

Definition 1.7. *Given any Borel set $S \subseteq \Omega$, we define*

$$F(S) := \bigcup \big\{\overline{R}_i :\ i \in I,\ \overline{R}_i \cap S \neq \emptyset\big\};$$

then, we say that the configuration $\mathscr{R}$ is balanced *if for any $S \subseteq \Omega$ the following holds:*

$$f^+(S) \leq f^-\big(F(S)\big), \qquad\qquad f^-(S) \leq f^+\big(F(S)\big). \tag{1.5}$$

We will deduce many fruitful informations about the behaviour of $f^\pm$ with respect to the ray configuration $\mathscr{R}$ in the balanced case in Lemma 2.5. We also point out that, in fact, the two possible orientations "left-right" on the rays R_i are determined uniquely by $\mathscr{R}$, and not depend on the particular choice of a transversal curve H. Finally, we state the

Theorem A. *Let $\mathscr{R} = \{R_i,\, i \in I\}$ be a balanced ray configuration such that there exists two disjoint Borel sets $\Gamma^\pm$ in Ω on which $f^\pm$ are concentrated with the property that, for all $i \in I$, $\sup R_i \cap \Gamma^+ \leq \inf R_i \cap \Gamma^-$. Then, the admissible maximal transport rays are all the segments R_i, $i \in I$ and all their open subsegments.*

We point out that Lemma 1.5 is the particular case when $\Omega = [0,6] \times [0,1]$ and $\mathscr{R} = \{[0,6] \times \{i\}, i \in [0,1]\}$, since of course in the hypotheses of Lemma 1.5 the configuration $\mathscr{R}$ is balanced.

In Section 2 we will prove Theorem A, while in Section 3 we will provide some examples using it. In particular, Example 3.5 shows that there are non-trivial situations with presence of doubling point in the interior of $\mathrm{spt} f^+$, that may be not obvious at first glance.

2. Proof of Theorem A

This Section is entirely devoted to show Theorem A: to do that, we will exhibit a particular transport plan γ and a particular shape, *i.e.* 1$-$Lipschitz function, u; then, making use of Lemma 1.4, we will check that they are both optimal and we will deduce the position of all the admissible maximal transport rays. In Subsection 2.1 we will define u, in Subsection 2.2 we will define γ and finally in Subsection 2.3 we will give the proof of the Theorem.

First of all, we fix a transversal curve H and we fix arbitrarily on it one of the two possible orientations. Notice that, by the obvious bijection between H and I, this allows to equip I with an order, and that $I \approx \mathbb{R}$ with this order; therefore, in a purely formal way that will be quite useful in the sequel, we will consider also the indeces $\pm\infty$ on I corresponing to the empty rays $R_{\pm\infty}$, but we will also intend

$\overline{R}_{+\infty}$ and $\overline{R}_{-\infty}$ to be the two closed sets consisting of the single points $\sup H$ and $\inf H$ in $\partial\Omega$, and we will consider also those two points as elements of H.

Whenever C and D belong to H, we denote by $\widetilde{CD}$ the part of the curve H between C and D, that is $\widetilde{CD} = \big\{P \in H: \ C \leq P \leq D\big\}$ if $C \leq D$ and $\widetilde{CD} = \big\{P \in H: \ C \geq P \geq D\big\}$ if $C \geq D$. Therefore, we will write for brevity

$$\overrightarrow{CD} := \begin{cases} \mathcal{H}^1\big(\widetilde{CD}\big) & \text{if } C > D; \\ 0 & \text{if } C = D; \\ -\mathcal{H}^1\big(\widetilde{CD}\big) & \text{if } C < D. \end{cases}$$

Analogously, whenever A and B belong to a same $R_i \in \mathscr{R}$ (recall that all the R_i are oriented segments), we set

$$\overrightarrow{AB} := \begin{cases} \overline{AB} & \text{if } A > B; \\ 0 & \text{if } A = B; \\ -\overline{AB} & \text{if } A < B. \end{cases}$$

In addition, we will write $\int_C^D g(s)\,ds$ (resp. $\int_A^B g(s)\,ds$) to denote the integral of any function g on the curve $\widetilde{CD}$ (resp. on the segment AB) with respect to $\mathcal{H}^1$, multiplied by 1 or -1 if $C < D$ of $C > D$ (resp. if $A < B$ or $A > B$): formally,

$$\int_C^D g(s)\,ds := \overrightarrow{CD} \fint_{\widetilde{CD}} g(s)\,d\mathcal{H}^1 \llcorner H(s), \qquad \int_A^B g(s)\,ds := \overrightarrow{AB} \fint_{AB} g(s)\,d\mathcal{H}^1(s). \tag{2.1}$$

2.1. **Definition of u.** Here we define the function u, and we check that it is $1-$Lipschitz; we give first a

Definition 2.1. *Let $i : \Omega \setminus \mathscr{D} \to I$ the function which associates to any $x \in \Omega \setminus \mathscr{D}$ the unique $i = i(x) \in I$ such that $\overline{R}(i(x)) \ni x$. Moreover, we set $\varphi : \Omega \setminus \mathscr{D} \to H$ as*

$$\varphi(x) := R_{i(x)} \cap H,$$

so that $\varphi(x)$ is the intersection of the ray passing through x with H. Finally, we let $\theta : \Omega \setminus \mathscr{D} \to (0, \pi)$ be the function so that $\theta(x)$ is the angle between the (oriented) segment $R_{i(x)}$ and the (oriented) tangent to H at $R_{i(x)} \cap H$.

We can finally define the function u as we claimed: we fix a point $O \in H$, deciding $u(O) := 0$; then, we set –keep in mind (2.1)–

$$\begin{cases} u(x) := -\displaystyle\int_0^x \cos\big(\theta(s)\big)\,ds & \forall x \in H, \\ u(x) := u\big(\varphi(x)\big) - \overrightarrow{\varphi(x)x}, & \forall x \in \Omega^\circ \setminus H. \end{cases} \tag{2.2}$$

Remark 2.2. *By immediate geometrical arguments, the functions φ and θ are continuous on $\Omega \setminus \mathscr{D}$; notice that $\theta(x)$ belongs to the open interval $(0, \pi)$ by definition of transversal curve, and that it is constant on each ray R_i, $i \in I$. By (2.2), then, we infer that u is continuous on Ω°. Notice that,* a priori, *u could happen not to be continuously extendable on $\partial\Omega$, in particular on the set $\mathscr{D}$. However, we will show that u is $1-$Lipschitz, so that it can also be extended to the whole Ω, still remaining $1-$Lipschitz.*

The continuity of u at Ω° stated in the above remark is of course not satisfactory: since we mean to show that u is an optimal shape, we have first to check that u is $1-$Lipschitz.

Lemma 2.3. *The function u defined in (2.2) is $1-$Lipschitz and it decreases with slope 1 in all the rays $R_i \in \mathscr{R}$.*

Proof. The second part of the claim is immediate by the definition (2.2), the difficult part is the first one. Assume that this is not true: then, there must be two points A and B in Ω° such that

$$u(A) - u(B) > \lambda |A - B|$$

with some $\lambda > 1$. Consider now the segment connecting A and B: of course, there is some point P in this segment with the property that

$$\limsup_{Q \to P,\, Q \in AB} \frac{|u(Q) - u(P)|}{|Q - P|} > \lambda; \tag{2.3}$$

we claim that $P \notin H$. Indeed, by the continuity of θ noticed in Remark 2.2 and by the definition (2.2) of u, it is clear that $\nabla u(S)$ exists at each $S \in H$, and moreover $|\nabla u(S)| = 1$ and $\nabla u(S)$ is parallel to $R_{i(S)}$, so that (2.3) ensures P not to be inside H. Without loss of generality, we assume P to be in the right part of Ω; take now a point Q close to P with the property that

$$|u(Q) - u(P)| > \lambda \overline{PQ} : \tag{2.4}$$

we will show that this leads to a contradiction provided Q is sufficiently close to P. To do that we give now some definitions, as in Figure 1 where of course we assume that our conventions for left and right, and for up and down are the obvious ones: we set $\widetilde{P} := \varphi(P)$, $\widetilde{Q} := \varphi(Q)$, $\varepsilon := \overrightarrow{\widetilde{P}\widetilde{Q}}$ and $\widehat{Q}$ the point of

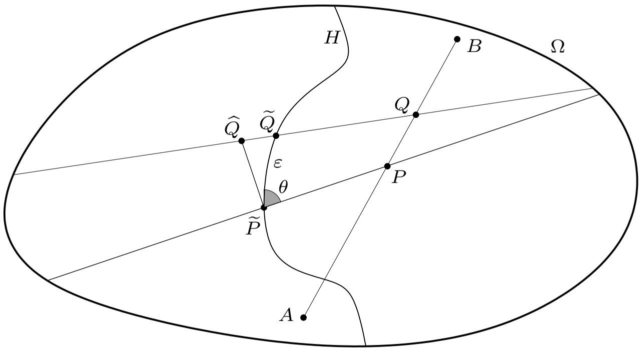

Figure 1. Position of points in Lemma 2.3

$R_{i(Q)}$ whose orthogonal projection on $R_{i(P)}$ is $\widetilde{P}$. Notice that $i(P) \neq i(Q)$ by (2.4), since u has exactly slope 1 along each ray R_i, so that $\varepsilon \neq 0$. Moreover there is a suitable constant $\rho > 0$ such that

$$\overline{PQ} \geq \rho |\varepsilon| \tag{2.5}$$

provided $\overline{PQ}$ is sufficiently small: recalling that P belongs to Ω° by convexity, this follows immediately via a similitude argument since $R_{i(P)}$ and $R_{i(Q)}$ do not intersect in Ω°, and of course the constant ρ depends on the angle between the segment AB and the gradient of H at $\widetilde{P}$, as well as on the distance between P and $\partial\Omega$ (ρ becomes smaller when P gets closer to $\partial\Omega$). As a consequence, $\varepsilon \to 0$ when $Q \to P$. By (2.2), we evaluate $u(P) = u(\widetilde{P}) - \overline{P\widetilde{P}}$ and $u(Q) = u(\widetilde{Q}) - \overline{Q\widetilde{Q}}$ recalling that P and Q are on the right of H (concerning Q, this is of course true provided $|Q - P| \ll 1$). Therefore, by (2.4) one infers

$$\left| u(\widetilde{Q}) - u(\widetilde{P}) + \overline{P\widetilde{P}} - \overline{Q\widetilde{Q}} \right| > \lambda \overline{PQ}. \tag{2.6}$$

Again by (2.2), we know that

$$u(\widetilde{Q}) = u(\widetilde{P}) - \int_{\widetilde{P}}^{\widetilde{Q}} \cos\bigl(\theta(s)\bigr)\, ds :$$

since, as noticed in Remark 2.2, the function θ is continuous, by the definition of ε we deduce

$$u(\widetilde{Q}) = u(\widetilde{P}) - \varepsilon\cos(\theta) + o(\varepsilon),$$

where we write θ in place of $\theta(\widetilde{P})$ for shortness. Substituting this estimate in (2.6), one finds

$$\left|\overline{P\widetilde{P}} - \overline{Q\widetilde{Q}} - \varepsilon\cos(\theta) + o(\varepsilon)\right| > \lambda\overline{PQ}. \tag{2.7}$$

Again by the continuity of θ at $\widetilde{P}$, we can also clearly estimate

$$\overline{Q\widetilde{Q}} = \overline{Q\widehat{Q}} + \overrightarrow{\widetilde{Q}\widehat{Q}}, \qquad \overrightarrow{\widetilde{Q}\widehat{Q}} = -\varepsilon\cos(\theta) + o(\varepsilon), \qquad \overline{\widetilde{P}\widehat{Q}} = \varepsilon\sin(\theta) + o(\varepsilon); \tag{2.8}$$

putting then together (2.7) and (2.8), we have

$$\left|\overline{P\widetilde{P}} - \overline{Q\widehat{Q}} + o(\varepsilon)\right| > \lambda\overline{PQ}. \tag{2.9}$$

To recover a contradiction with $|\varepsilon|$ sufficiently small, we recall that by the triangle inequality

$$\overline{Q\widehat{Q}} \leq \overline{PQ} + \overline{P\widehat{Q}}; \tag{2.10}$$

moreover by Pitagora's Theorem we can evaluate, using also (2.8) and recalling that $P \neq \widetilde{P}$ because $P \notin H$,

$$\overline{P\widehat{Q}} = \sqrt{\overline{P\widetilde{P}}^2 + \overline{\widetilde{P}\widehat{Q}}^2} = \sqrt{\overline{P\widetilde{P}}^2 + \big(\varepsilon\sin(\theta) + o(\varepsilon)\big)^2} = \overline{P\widetilde{P}} + o(\varepsilon)$$

that together with (2.10) gives

$$\overline{Q\widehat{Q}} \leq \overline{PQ} + \overline{P\widetilde{P}} + o(\varepsilon). \tag{2.11}$$

In the very same way, the triangle inequality ensures also

$$\overline{P\widetilde{P}} \leq \overline{PQ} + \overline{Q\widetilde{P}}; \tag{2.12}$$

and again Pithagora's Theorem, recalling the continuity of θ, gives $\overline{Q\widetilde{P}} = \overline{Q\widehat{Q}} + o(\varepsilon)$, so that from (2.12) one has

$$\overline{P\widetilde{P}} \leq \overline{PQ} + \overline{Q\widehat{Q}} + o(\varepsilon). \tag{2.13}$$

Putting together (2.11) and (2.13), and substituting into (2.9) gives

$$\overline{PQ} > \lambda\overline{PQ} + o(\varepsilon),$$

that gives the desired contradiction when $|\varepsilon|$ is sufficiently small keeping in mind (2.5). □

Remark 2.4. *The function u has been defined only on Ω°; however, by the previous Lemma it can be uniquely extended, remaining $1-$Lipschitz, to the whole Ω.*

2.2. **Definition of γ.** In this subsection we will define the measure γ with them we will prove the Theorem. First, let us briefly recall some well-known facts about the disintegration of the measures (for a formal treatment of this subject and for the proof of the claims below one can refer to [3]): given the spaces X and Y, and measures $\mu \in \mathcal{M}^+(X)$, $\mu_y \in \mathcal{P}(X)$ for $y \in Y$ and $\eta \in \mathcal{M}(Y)$, we write $\mu = \mu_y \otimes \eta$ if for any Borel set $\Delta \subseteq X$ the function $y \mapsto \mu_y(\Delta)$ is $\eta-$measurable and one has

$$\mu(\Delta) = \int_{y\in Y} \mu_y(\Delta)\, d\eta(y).$$

A particularly interesting situation when such a decomposition of μ is possible, is when one is given a function $\alpha : X \to Y$: in this case, the Disintegration Theorem ensures that there exists probability measures $\mu_y \in \mathcal{P}(X)$, each of them concentrated in the corresponding $\alpha^{-1}(\{y\})$, such that $\mu = \mu_y \otimes \alpha_\# \mu$; moreover, these measures μ_y are uniquely determined for $\alpha_\#\mu-$a.e. $y \in Y$. Finally, if there is a space Z

and two functions $\beta : X \to Z$ and $\psi : Z \to Y$ such that $\alpha = \psi \circ \beta$, then the decomposition of $\beta_\# \mu$ with respect to γ is given by

$$\beta_\# \mu = \beta_\# \mu_y \otimes \psi_\# \big(\beta_\# \mu\big). \tag{2.14}$$

Now, our aim is to define a transport plan γ with the property that each of its maximal transport rays is contained in some of the $R_i \in \mathscr{R}$. We point out that, without any assumption on the ray configuration $\mathscr{R}$, if such γ exists then $\mathscr{R}$ is balanced, so that the balance property is also necessary for the maximal transport rays. To notice this fact, just fix $S \subseteq \Omega$ and recall that

$$f^+(S) = \gamma\big(\big\{(x,y) :\ x \in S\big\}\big) \leq \gamma\big(\big\{(x,y) :\ y \in F(S)\big\}\big) = f^-\big(F(S)\big);$$

indeed, by definition of maximal transport rays, whenever $(x, y) \in \mathrm{spt}\gamma$ and $x \in S$ then xy is contained in some ray $R_i \in \mathscr{R}$ intersecting S (for instance, at x!), so $y \in F(S)$. Therefore, the first inequality in (1.5) is verified, and the second can be shown in the very same way.

For any $i \in I$, now, we consider the "upper set"

$$\Lambda_i := \bigcup_{j\geq i} \overline{R}_j,$$

and we prove some useful consequences of the balance property:

Lemma 2.5. *Under the assumptions of Theorem A, take $i \in I$ and write $R_i =: AB \in \mathscr{R}$; if both A and B do not belong to $\mathscr{D}$, then $f^+(\Lambda_i) = f^-(\Lambda_i)$. More precisely, if there is no $j < i$ such that A is the left endpoint of R_j (resp. B is the right endpoint of R_j) then $f^+(\Lambda_i) \leq f^-(\Lambda_i)$ (resp. $f^-(\Lambda_i) \leq f^+(\Lambda_i)$).*

Proof. Defining $S = \Lambda_i \setminus \{B\}$, if A is not the left endpoint of some R_j with $j < i$ then by construction $F(S) = \Lambda_i$; therefore, since by the hypotheses of the Theorem clearly $B \notin \Gamma^+$, (1.5) gives

$$f^+(\Lambda_i) = f^+(S) \leq f^-\big(F(S)\big) = f^-(\Lambda_i).$$

The case when B is not the right endpoint of R_j for any $j < i$ is of course completely similar; and the first claim, in which A and B are not doubling points of $\mathscr{R}$, is now a consequence. $\square$

Remark 2.6. *For each $i \in I$, at least one of the claims in the above Lemma applies. Indeed, since the rays in $\mathscr{R}$ cover Ω° and are disjoint, it is immediately checked that, in case that both A and B belong to $\mathscr{D}$, then either A is the left endpoint only of rays R_j with $j \geq i$ and B is the right endpoint only of rays R_j with $j \leq i$, or A is the left endpoint only of rays R_j with $j \leq i$ and B is the right endpoint only of rays R_j with $j \geq i$.*

Take now a point $P \in \Gamma^+ \cap \mathscr{D}$: by geometrical arguments, there are $i_{\min} < i_{\max} \in I$, depending on P, such that P is the left endpoint of all and only the rays R_i with $i_{\min} \leq i \leq i_{\max}$. We claim that the function $\delta : (i_{\min}, i_{\max}] \to \mathbb{R}$ given by

$$\delta(j) := f^-(\Lambda_j) - f^+\big(\Lambda_j \setminus \{P\}\big)$$

is decreasing. Indeed, for $i_{\min} < j_1 < j_2 \leq i_{\max}$ we set $S = \cup\{R_j : j_1 \leq j < j_2\}$, for which $F(S) = \Lambda_{j_1} \setminus \Lambda_{j_2} \cup \{P\}$; by the hypotheses, $P \notin \Gamma^-$, hencete also that again by (1.5) one immediately obtains that $0 \leq \delta \leq f^+(\{P\})$ in $(i_{\min}, i_{\max}]$ –we also underline that the fact that the interval is open in the left is fundamental to derive this estimate. Therefore, we extend δ to the whole $\mathbb{R}$ by setting $\delta(j) \equiv f^+(\{P\})$ if $j \leq i_{\min}$ and $\delta(j) \equiv 0$ if $j > i_{\max}$, and we define the measure $\mu_P \in \mathcal{M}^+(I)$ given by $\mu_P := -D\delta$: since δ is a bounded decreasing function, μ_P is a positive measure of mass $f^+(\{P\})$ concentrated in $[i_{\min}, i_{\max}]$.

For any $i \in I$, we also define $P(n,i)$ as the point in $R_i = AB$ such that $\overline{AP(n,i)} = \overline{AB}/n$. Finally, we introduce the measures

$$f_n^+ := f^+ \llcorner \left(\Gamma^+ \setminus \mathscr{D}\right) + \sum_{P \in \mathscr{D} \cap \Gamma^+} \int_{i=i_{\min}(P)}^{i_{\max}(P)} \delta_{P(n,i)} d\mu_P(i). \tag{2.15}$$

The meaning of μ_P and f_n^+ is worth to be understood: μ_P says how the part of f^+ contained in the point P should be distributed among the different rays of $\mathscr{R}$ in order to have a precise mass balance. The measures f_n^+ simply perform this distribution for all the doubling points, shifting a corresponding part of mass in the interior of the rays, but remaining arbitrarily close to the doubling points in the limit $n \to +\infty$. Notice that the definition (2.15) makes sense since the function $i \mapsto P(n,i)$ is continuous and the set $\mathscr{D} \cap \Gamma^+$ is at most countable. Our last ingredient is the following measure.

Definition 2.7. *We define the measure $\nu \in \mathcal{P}(I)$ setting $\nu\big([i,+\infty)\big)$ for any $i \in \mathbb{R}$ as follows:*

$$\nu\big([i,+\infty)\big) := \begin{cases} f^+(\Delta_i) & \text{if the left endpoint of } R_i \text{ is not left endpoint of any } R_j \text{ with } j < i; \\ f^-(\Delta_i) & \text{if the right endpoint of } R_i \text{ is not right endpoint of any } R_j \text{ with } j < i. \end{cases}$$

Notice that the two definitions agree where they apply simultaneously thanks to Lemma 2.5 (recall also Remark 2.6).

Remark 2.8. *We point out that, if $\mathscr{D}$ is f^+-negligible, then $\varphi : \Omega \setminus \mathscr{D} \to H$ is in fact defined f^+-a.e., so that it is possible to define the measure $\varphi_\# f^+$. Moreover, by Definition 2.7 and by construction, it is clearly $\nu = \varphi_\# f^+$; in the very same way, $\nu = \varphi_\# f^-$ if $f^-(\mathscr{D}) = 0$.*

Finally, we are ready to prove the fundamental

Lemma 2.9. *The measures f_n^+ are probability measures weakly* converging to f^+; moreover, there are probability measures $f_{n,i}^+$ concentrated on R_i for $n \in \mathbb{N}$, $i \in I$ such that one has the disintegration $f_n^+ = f_{n,i}^+ \otimes \nu$. Finally, for any $i \in I$ one has $f_{n,i}^+ \xrightharpoonup{*} f_i^+$, and $f^+ = f_i^+ \otimes \nu$ with f_i^+ concentrated in $\overline{R}_i$. The same statement allows to write $f^- = f_i^- \otimes \nu$ with f_i^- concentrated in $\overline{R}_i$.*

Proof. Since f^+ is a probability measure and f_n^+ coincides with f^+ up to splitting the mass of any $P \in \Gamma^+ \cap \mathscr{D}$ in the corresponding rays (recall $\|\mu_P\| = f^+(\{P\})$), then also f_n^+ is a probability measure. Moreover, since the points $P(n,i)$ uniformly converge to the left endpoints of the rays R_i, then the convergence $f_n^+ \xrightharpoonup{*} f^+$ follows. Notice now that by construction f_n^+ is concentrated in $\Omega \setminus \mathscr{D}$, and as noticed in Remark 2.8 it is clear that $\varphi_\# f_n^+ = \nu$. Therefore, we can disintegrate obtaining $f_n^+ = f_{n,i}^+ \otimes \nu$ with $f_{n,i}^+$ concentrated in R_i.

By construction, for any $i \in I$ the sequence of measures $n \mapsto f_{n,i}^+$ is the sum of a constant measure and a Dirac mass in a point moving to the left extreme of R_i, so that $f_{n,i}^+ \xrightharpoonup{*} f_i^+$ with f_i^+ concentrated in $\overline{R}_i$; since the $f_{n,i}$'s are probability measures, the Dominated Convergence Theorem ensures that then $f_{n,i}^+ \otimes \nu \xrightharpoonup{*} f_i^+ \otimes \nu$, so that also the second part of our claim follows.

The result for f^- can be derived in the same way, replacing in the previous argument and in the definition of δ the left with the right and f^+ with f^-. $\square$

Remark 2.10. *We like to mention that the previous claim,* i.e. *that one can write $f^\pm = f_i^\pm \otimes \nu$ with probability measures f_i^+ on $\overline{R}_i$, is all we need to perform our construction of γ; therefore, the whole discussions we made in this subsection until now would be completely unnecessary under the assumption that $\mathscr{D}$ is negligible with respect to both f^+ and f^-, since in that case one could directly decompose $f^\pm$ with respect to φ. Notice also that by construction $\mathscr{D}$ contains countable many points, so that it would* *have been sufficient to add to Theorem A the hypothesis of $f^\pm$ being non-atomic. However, since our result is true without this assumption, we decided to present this more involved argument.*

Finally, we can give our construction of the transport plan γ.

Lemma 2.11. *There exists a transport plan γ between f^+ and f^- such that any maximal transport ray of γ is contained in some R_i, $i \in I$.*

Proof. We will define
$$\gamma := \gamma_i \otimes \nu \tag{2.16}$$
being γ_i a transport plan between f_i^+ and f_i^- (we write $f^\pm = f_i^\pm \otimes \nu$ in view of Lemma 2.9); any choice of γ_i works, provided of course that $i \mapsto \gamma_i$ is a $\nu-$measurable measure valued map, in the sense of [1]; in other words, since defining $\gamma := \gamma_i \otimes \nu$ means that for any $v \in C_b(\Omega \times \Omega)$ one sets
$$\langle \gamma, v\rangle = \int_I \langle \gamma_i, v\rangle \, d\nu(i),$$
we need to check that the preceding integral is well defined, that is that $i \mapsto \langle \gamma_i, v\rangle$ is a $\nu-$measurable real map for any continuous and bounded function v on $\Omega \times \Omega$. To solve easily this problem, it suffices to define
$$\gamma_i := f_i^+ \otimes f_i^- : \tag{2.17}$$
γ_i is of course a transport plan between f_i^+ and f_i^-, and $i \mapsto \gamma_i$ is easily seen to be a $\nu-$measurable measure valued map because so are $i \mapsto f_i^+$ and $i \mapsto f_i^-$ by the properties of disintegration; indeed, $i \mapsto \langle \gamma_i, v\rangle$ is clearly $\nu-$measurable if v is the characteristic function of a set $\Omega_1 \times \Omega_2$ where Ω_1 and Ω_2 are Borel subsets of Ω, and by a standard density argument one concludes for a general v.

Having defined γ via (2.16) and (2.17), we need to check that it satisfies our claim; in fact, by (2.14) one has $\pi_1\gamma = \pi_1\gamma_i \otimes \nu = f_i^+ \otimes \nu = f^+$ and analogously $\pi_2\gamma = f^-$, so that γ is a transport plan between f^+ and f^-. Moreover, $\gamma-$a.e. pair (x, y) in $\Omega \times \Omega$ is contained in $\mathrm{spt}\gamma_i$ for some $i \in I$, then both x and y are in $\overline{R}_i$; this immediately implies that all the transport rays are contained in some R_i, and then also the maximal transport rays do the same (recall that transport rays and maximal transport rays are open segments). □

2.3. **Proof of the Theorem.** Here we can show Theorem A.

Proof. of Theorem A: We denote by u the function defined in (2.2) and by γ the transport plan built in Lemma 2.11: by construction, all the maximal transport rays of γ are contained in some of the rays R_i, and on the other hand we know by Lemma 2.3 that u decreases with slope 1 in all these rays (recall also Remark 2.4). Hence, we can apply Lemma 1.4 to derive that u is an optimal shape and γ is an optimal transport plan. Recalling the discussions at the beginning of Section 1.2, we know that then the admissible maximal transport rays are exactly those segments on which u decreases at slope 1. Since the rays $R_i \in \mathscr{R}$ cover the whole Ω°, we derive that these segments are exactly all the subsegments of some $R_i \in \mathscr{R}$, and the proof is completed. □

Remark 2.12. *Theorem A says which are the admissible maximal transport rays, and thanks to the discussions in Section 1.2 this tells us which of the transport plans are optimal and which are not. However, usually something more is true, that is, that the maximal transport rays are the same for all the different optimal transport plans: indeed, if $\mathscr{D}$ is $f^\pm-$negligible (in fact, it suffices that $\mathscr{D}$ is negligible with respect to just one between f^+ and f^-), then the maximal transport rays are exactly all the maximal* *subsegments of the R_i's intersecting the supports of f^+ and f^-. If $\mathscr{D}$ is not negligible, the unicity of the maximal transport rays is still true under the assumptions of our Theorem, since as we proved in Section 2.2 it can be precisely determined how much of each doubling point must be moved along each ray. But in the completely general case, this unicity may not be true, since the different optimal transport plans could split differently the doubling points in the different admitted directions. For instance, in Example 3.2 the maximal transport rays associated to γ_0 are the vertical segments AC and DB, the maximal transport rays associated to γ_1 are the horizontal segments AB and DC, and the maximal transport rays associated to γ_s with $0 < s < 1$ are all the four segments.*

3. Examples

In this section, we give several examples to comment our Theorem and to show some of its possible applications. First of all, we give an easy example to enlighten the necessity of introducing the sets $\Gamma^\pm$ in the claim of the Theorem.

Example 3.1. *Let f^+ and f^- be concentrated in two polygons as in Figure 2.a; more precisely, let $f^\pm$ be absolutely continuous with respect to the Lebesgue measure $\mathcal{L}$, with densities given by*

$$f^+_{\mathcal{L}}(x,y) := \begin{cases} 1 & \text{in } [0,1]\times[0,1/2]; \\ 1/2 & \text{in } [0,2]\times[1/2,1]; \end{cases} \qquad f^-_{\mathcal{L}}(x,y) := \begin{cases} 1/2 & \text{in } [1,3]\times[0,1/2]; \\ 1 & \text{in } [2,3]\times[1/2,1]. \end{cases}$$

Then one can apply Theorem A with the horizontal rays and with $\Gamma^\pm = \mathrm{spt} f^\pm \setminus R_0$, being $R_0 := [0,3]\times\{1/2\}$, and it follows that the admissible maximal transport rays are the horizontal segments (thus this could have been recovered also through a stronger version of Lemma 1.5). Notice that the condition $\sup R_i \cap \Gamma^+ \le \inf R_i \cap \Gamma^-$ would have not been satisfied replacing $\Gamma^\pm$ with $\mathrm{spt} f^\pm$: this shows the importance of introducing the sets $\Gamma^\pm$ in our statement also for very simple situations.

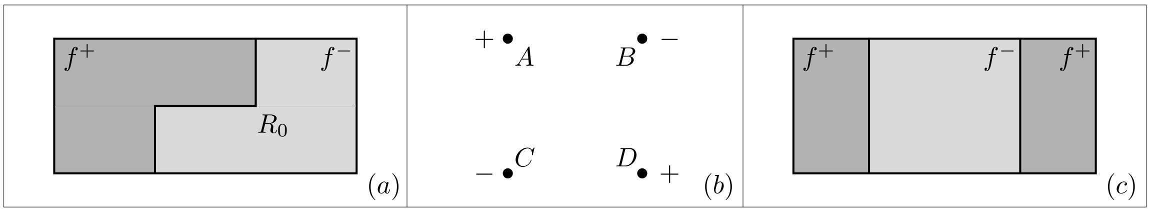

Figure 2. Examples 3.1, 3.2 and 3.3

Now, consider once more the meaning of the Theorem: roughly speaking, if one has the insight of which are the maximal transport rays, it allows to show it formally, and this seems to be the easiest way to give explicit examples of optimal transport plans (or maps), except than in trivial cases. In other words, there is a certain class of pairs $f^\pm$ (and of corresponding sets Ω) for which the Theorem can be applied to reveal the maximal transport rays –of course, after that one has guessed them. This class was very little after Lemma 1.5 in [2], since it covered only some situations when all the transport rays are horizontal; with our result at hand, this class has become much larger, as we will try to show mainly through Examples 3.4 and 3.5. But of course this class can not cover *all* the possible configurations $f^\pm$, as we can understand with the following two examples.

Example 3.2. *Consider the situation when $f^+ = 1/2(\delta_A + \delta_D)$ and $f^- = 1/2(\delta_B + \delta_C)$ and the four point A, B, C and D are the vertexes of a square as in Figure 2.b. Then it is clear that the transport plan* *are exactly the measures* $\gamma_s = s\big(\delta_A\otimes\delta_B+\delta_D\otimes\delta_C\big)+(1-s)\big(\delta_A\otimes\delta_C+\delta_D\otimes\delta_B\big)$ *with* $0\le s\le 1$. *Observe that, for any convex set* Ω *containing the square, it is not possible to find a ray configuration satisfing the conditions of the Theorem: in particular, there are balanced configurations (for instance, the one made by all horizontal lines), but the condition* $\sup R_i\cap\Gamma^+\le\inf R_i\cap\Gamma^-$ *can not be satisfied because whenever* A *is on the left of its ray then* D *is on the right of its one. Thus, the Theorem can not be applied.*

A less critical situation in which the Theorem still can not be applied is the following one:

Example 3.3. *Consider, as in Figure 2.c, the situation when* f^+ *and* f^- *are (up to the constant* $1/2$ *to ensure* $\|f^\pm\|=1$*) the Lebesgue measure restricted on the sets* $\big([0,1]\cup[3,4]\big)\times[0,1]$ *and* $[1,3]\times[0,1]$ *respectively: then, it is easy to understand that the maximal transport rays are all the open segments* $(0,2)\times\{s\}$ *and* $(2,4)\times\{s\}$ *for* $0\le s\le 1$*, so one can not find any ray configuration with segments having the endpoints in* $\partial\Omega$.

Now we give a simple example, to show a situation that can be handled with Theorem A and could have not been solved with Lemma 1.5: (this example has been already presented in [7] because it admits an optimal transport map which is not continuous).

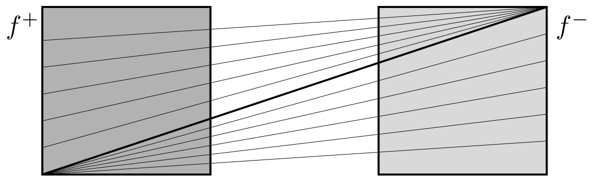

FIGURE 3. Example 3.4

Example 3.4. *Let* $\Omega:=[0,3]\times[0,1]$*, and let* $\mathscr{R}:=\{R_i,\ i\in[0,1]\}\cup\{\widetilde{R}_i,\ i\in[0,1]\}$*, being* R_i *(resp.* $\widetilde{R}_i$*) the open segment connecting* $(0,0)$ *and* $(3,i)$ *(resp.* $(0,i)$ *and* $(3,1)$*); Figure 3 shows some of these rays. We let* f^+ *and* f^- *be concentrated on* $[0,1]\times[0,1]$ *and* $[2,3]\times[0,1]$ *and absolutely continuous with respect to the Lebesgue measure, with densities given by*

$$f^+_{\mathcal{L}}(x,y):=\left\{\begin{array}{ll} C^+(i) & \text{in } [0,1]\times[0,1]\cap R_i;\\ \widetilde{C}^+(i) & \text{in } [0,1]\times[0,1]\cap\widetilde{R}_i; \end{array}\right. \qquad f^-_{\mathcal{L}}(x,y):=\left\{\begin{array}{ll} C^-(i) & \text{in } [2,3]\times[0,1]\cap R_i;\\ \widetilde{C}^-(i) & \text{in } [2,3]\times[0,1]\cap\widetilde{R}_i \end{array}\right.$$

and where the Borel functions $C^\pm,\ \widetilde{C}^\pm:[0,1]\to\mathbb{R}^+$ *can be arbitrarily chosen subject to the constraint*

$$C^+\equiv 5C^-,\qquad\qquad \widetilde{C}^-\equiv 5\widetilde{C}^+. \tag{3.1}$$

We claim that Theorem A can be applied to this situation, and hence that the maximal transport rays are exactly the R_i *and the* $\widetilde{R}_i$*. To show this assertion, one has only to check that the mass balance condition holds, then by symmetry we limit ourselves to show that* $f^+(T)=f^-(T)$ *for the triangle* T *with vertexes* $(0,0)$*,* $(3,0)$ *and* $(3,i)$ *for a generic* $i\in[0,1]$*. A trivial calculation ensures that* $f^+(T)=1/6\int_0^i C^+(t)\,dt$ *and* $f^-(T)=5/6\int_0^i C^-(t)\,dt$*, so that thanks to (3.1), our claim follows.*

We give now our last example, which will be a more involved application of Theorem A and it will also give an interesting counterexample. To show its meaning, we first underline that our Theorem can not clearly be applied whenever there is some doubling point in the interior of the support of f^+ or of f^-, as it happens in Example 3.3. From this example and other simple situations, it could seem reasonable that whenever $f^\pm$ are sufficiently regular and with supports convex and disjoint, there may not be doubling points in the interior of the supports: notice that this would imply that the Theorem can be applied in all these situations. On the countrary, in the example below we present a situation where the supports of $f^\pm$ are convex and disjoint and $f^\pm$ are regular, but there is a doubling point in the interior of $\mathrm{spt} f^+$. We remark that as a consequence the Theorem can not be applied in this situation: nevertheless, we will prove our claim with a careful application of our result in some subset of the supports.

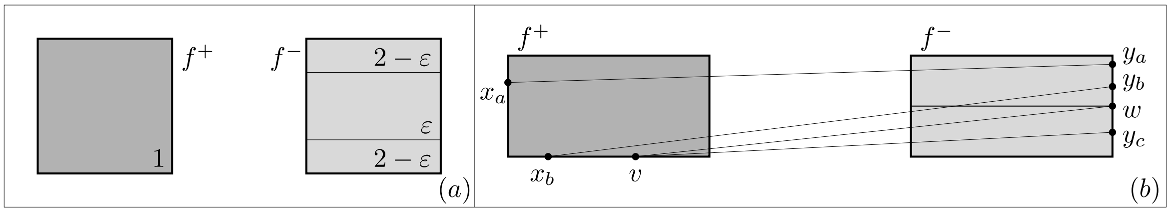

FIGURE 4. Example 3.5

Example 3.5. *Let f^+ and f^-, as in Figure 4.a, be two absolutely continuous measures concentrated on the squares $[0,2]\times[-1,1]$ and $[4,6]\times[-1,1]$, whose densities (up to the rescaling factor $4=\mathcal{L}(\mathrm{spt} f^+)=\mathcal{L}(\mathrm{spt}(f^-))$) are given by*

$$f^+_{\mathcal{L}}(x,y):=1, \qquad\qquad f^-_{\mathcal{L}}(x,y):=\begin{cases} 2-\varepsilon & \text{in } [4,6]\times\big([-1,-1/2]\cup[1/2,1]\big),\\ \varepsilon & \text{in } [4,6]\times[-1/2,1/2],\end{cases}$$

where ε is a small positive number. An immediate symmetry argument shows that whenever γ is an optimal transport plan and $(x,y)\in\mathrm{spt}\gamma$, if $x_2>0$ then $y_2>0$ and if $x_2<0$ then $y_2<0$. Therefore, we consider the transport problem associated to $\tilde f^+ := f^+ \llcorner \{x_2\geq 0\}$ and $\tilde f^- := f^-\llcorner\{x_2\geq 0\}$, and the admissible maximal transport rays for the original problem will be all the admissible maximal transport rays for the restricted problem and all their symmetric images with respect to the axis $\{x_2=0\}$. We will apply Theorem A to the restricted configuration. To this aim, for any $0\leq s\leq 3$ and any $0\leq t\leq 1$ we define the points $x_s\in\partial\mathrm{spt}\tilde f^+$ and $y_t\in\partial\mathrm{spt}\tilde f^-$ as

$$x_s:=\begin{cases}\big(0,1-s\big) & \text{if } 0\leq s\leq 1,\\ \big(s-1,0\big) & \text{if } 1\leq s\leq 3;\end{cases} \qquad\qquad y_t:=\big(6,1-t\big);$$

notice that the set of the points $\{x_s,\, 0\leq s\leq 3\}$ is the left and the bottom part of the boundary of $\mathrm{spt}\tilde f^+$, while the set of the points $\{y_t,\, 0\leq t\leq 1/2\}$ is the right part of the boundary of $\mathrm{spt}\tilde f^-$. We also define the function $\varphi:[0,3]\times[0,1]\to\mathbb{R}^2$ as follows: the first (resp. the second) component of $\varphi(s,t)$ is the measure w.r.t. $\tilde f^+$ (resp. to $\tilde f^-$) of the region above the segment connecting x_s and y_t. We state now the
Claim. *There is a strictly increasing continuous functions $\tau:[0,1/2]\to[0,3]$ with $\tau(0)=0$ with the property that $\varphi_1(\tau(t),t)=\varphi_2(\tau(t),t)$ for any $0\leq t\leq 1/2$.*
Before to show the claim, notice that an immediate consequence is that $\tau(1/2)<3$ since $\varphi_1(3,1/2)=2<\varphi_2(3,1/2)$: consistently, in Figure 4.b we drawn $x_a=\tau(y_a)$, $x_b=\tau(y_b)$ and $v=\tau(w)$ with $w\equiv(6,1/2)$. Extend now the function τ to $(1/2,1]$ setting $\tau(t):=\tau(1/2)$ for $1/2<t\leq 1$, and consider the ray configuration $\mathscr{R}$ made by all the segments $x_{\tau(t)}y_t$ for $0\leq t\leq 1$ (as the segment vy_c in the figure): since, for any $t>1/2$, in the triangle with vertexes v, w and $(6,0)$ the densities of $\tilde f^+$ and $\tilde f^-$ (where positive) are costantly 1 and ε, by the Claim it follows that $\mathscr{R}$ is balanced. Hence, we can clearly apply the Theorem, finding that the elements of $\mathscr{R}$ are exactly the maximal transport rays for the problem with data $\tilde f^+$ and $\tilde f^-$. As noticed before, a symmetry tells us also which are the maximal transport rays for

the original problem: therefore, we found that the point v is a doubling point inside the interior of $\mathrm{spt} f^+$*! We remark also that v clearly converges to $(2,0)$ when $\varepsilon \to 0$; moreover, the discontinuity of (the density of) f^- on the lines $\{x_2 = \pm 1/2\}$ is clearly not important in this example, and one could easily replace $f^\pm$ with probability measures having smooth densities. We conclude now this example showing the claim.*

Proof of the Claim: *One could easily obtain the proof through trivial but very boring calculations; nevertheless, we prefer to give now a simple abstract proof. First of all, notice that for any $0 < t \le 1/2$ one has clearly $\varphi_1(0,t) < \varphi_2(0,t)$ but $\varphi_1(3,t) = 2 > \varphi_2(3,t)$: therefore, by continuity one has that for any t there exists some s with $\varphi_1(s,t) = \varphi_2(s,t)$ (also for $t = 0$ since of course $\varphi(0,0) = (0,0)$). We want to show that for any t there is a unique such $s := \tau(t)$, and for any s there is at most a unique $t \le 1/2$ such that $\varphi_1(s,t) = \varphi_2(s,t)$: as a consequence, the claim will immediately follow.*

Concerning the uniqueness of s for a given t, assume by contradiction the existence of $0 \le t \le 1/2$ and $0 \le s_1 < s_2 \le 3$ with $\varphi_1(s_i,t) = \varphi_2(s_i,t)$ for both $i = 1, 2$. Assume first that $s_2 \le 1$, and denote by T the triangle with vertexes x_{s_1}, x_{s_2} and y_t: it should be $\tilde f^+(T) = \tilde f^-(T)$, but by similitude one immediately has $\mathcal{L}\big(T \cap \mathrm{spt}\tilde f^+\big) > 2\mathcal{L}\big(T \cap \mathrm{spt}\tilde f^-\big)$, and since the density of $\tilde f^+$ on $T \cap \mathrm{spt}\tilde f^+$ is costantly 1 while the density of $\tilde f^+$ on $T \cap \mathrm{spt}\tilde f^+$ is everywhere less than 2, one finds the contradiction, then it must be $s_2 > 1$.

We can then assume the existence of $1 \le \bar s < s_2 \le 3$ with

$$\varphi_1(\bar s,t) \ge \varphi_2(\bar s,t), \qquad\qquad \varphi_1(s_2,t) = \varphi_2(s_2,t): \tag{3.2}$$

indeed, if $s_1 \ge 1$ the choice $\bar s := s_1$ clearly works and the "$\ge$" above is in fact "$=$", while if $s_1 < 1$ then the same similitude argument as before ensures that the choice $\bar s := 1$ works and the "$\ge$" is a "$>$". We will obtain that (3.2) is absurd by showing that $\varphi_1(\cdot,t)$ and $\varphi_2(\cdot,t)$ are respectively concave and convex in the interval $[\bar s, 3]$ and recalling that $\varphi_1(3,t) = 2 > \varphi_2(3,t)$. To obtain this, just call T_s the triangle with vertexes y_t, $x_{\bar s}$ and x_s for a generic $\bar s \le s \le 3$: then the concavity of φ_1 follows by the obvious geometric fact that $s \mapsto \mathcal{L}\big(T_s \cap \mathrm{spt}\tilde f^+\big) = \tilde f^+\big(T_s\big)$ is (strictly) concave on $[\bar s, 3]$. On the other hand, the convexity of φ_2 is true since in $[\bar s, 3]$ the map

$$s \mapsto \mathcal{L}\big(T_s \cap \{p \in \mathbb{R}^2 : \text{ the density of } \tilde f^- \text{ at } p \text{ is } 2-\varepsilon\}\big)$$

is linear, while the map

$$s \mapsto \mathcal{L}\big(T_s \cap \{p \in \mathbb{R}^2 : \text{ the density of } \tilde f^- \text{ at } p \text{ is } \varepsilon\}\big)$$

is (strictly) convex.

Finally, concerning the uniqueness of t for a given s, assume that there exists $s \in [0,3]$ and $0 \le t_1 < t_2 \le 1/2$ with $\varphi_1(s,t_i) = \varphi_2(s,t_i)$ for both $i = 1, 2$. The conclusion follows in a very similar way as before: we have that $\varphi_1(s,t_1) = \varphi_2(s,t_1)$, and clearly $\varphi_1(s,1) \le \varphi_2(s,1)$ (the inequality is strict if and only if $s < 1$). The absurd (which concludes the proof of the claim and the example) is given by the fact that on $[t_1, 1]$ the function $\varphi_1(s,\cdot)$ is linear, while $\varphi_2(s,\cdot)$ is strictly concave on $[t_1, 1/2]$ and linear on $[1/2, 1]$.

References

[1] L. Ambrosio, Lecture Notes on Optimal Transport Problems, in "Mathematical Aspects of Evolving Interfaces", Lecture Notes in Mathematics, LNM **1812**, Springer (2003), 1–52.

[2] L. Ambrosio & A. Pratelli, Existence and stability results in the L^1 theory of optimal transportation, in "Optimal Transportation and Applications", Lecture Notes in Mathematics, LNM **1813**, Springer (2003), 123–160.

[3] L. Ambrosio, N. Fusco & D. Pallara, Functions of Bounded Variation and Free Discontinuity Problems, Oxford University Press (2000).

[4] G. Bouchitté & G. Buttazzo, Characterization of optimal shapes and masses through Monge–Kantorovich Equation, J. Eur. Math. Soc., **3** (2001), 139–168.

[5] G. Bouchitté, G. Buttazzo & P. Seppecher, Shape optimization solutions via Monge-Kantorovich equation, C.R. Acad. Sci. Paris, **324-I** (1997), 1185–1191.

[6] L.C. Evans & W. Gangbo, Differential equations methods for the Monge–Kantorovich mass transfer problem, Memoirs of the A.M.S., Vol. **137**, Number 653, (1999).

[7] I. Fragalà, M.S. Gelli & A. Pratelli, Continuity of an optimal transport in Monge problem, preprint, Scuola Normale Superiore, Pisa, Italy (2004).

[8] L.V. Kantorovich, On the transfer of masses, Dokl. Akad. Nauk. SSSR, **37** (1942), 227–229.

[9] L.V. Kantorovich, On a problem of Monge, Uspekhi Mat. Nauk., **3** (1948), 225–226.

[10] G. Monge, Memoire sur la Theorie des Déblais et des Remblais, Hist. de l'Acad. des Sciences de Paris (1781).

[11] A. Pratelli, Existence of optimal transport maps and regularity of the transport density in mass transportation problems, Ph.D. Thesis, Scuola Normale Superiore, Pisa, Italy (2003). Avalaible on http://cvgmt.sns.it/ .

[12] S.T. Rachev & L. Rüschendorf, Mass Transportation Problems, Springer–Verlag (1998).

Scuola Normale Superiore, Piazza dei Cavalieri 7, 56126 Pisa, Italy
E-mail address: a.pratelli@sns.it