Rend. Istit. Mat. Univ. Trieste
Volume 43 (2011), 31–51

A Proof of Monge Problem in $\mathbb{R}^n$ by Stability

Laura Caravenna

Abstract. *The Monge problem in $\mathbb{R}^n$, with a possibly asymmetric norm cost function and absolutely continuous first marginal, is generally underdetermined. An optimal transport plan is selected by a secondary variational problem, from a work on crystalline norms. In this way the mass still moves along lines. The paper provides a quantitative absolute continuity push forward estimate for the translation along these lines: the consequent area formula, for the disintegration of the Lebesgue measure w.r.t. the partition into these $1D$-rays, shows that the conditional measures are absolutely continuous, and yields uniqueness of the optimal secondary transport plan non-decreasing along rays, recovering that it is induced by a map.*

Keywords: Monge Problem, Area Estimates, Disintegration of Measures.
MS Classification 2010: 49J45, 49K30, 28A50, 49Q20

1. Introduction

Topic of this note is a sharp area push forward estimate relative to a solution to the Kantorovich problem in $\mathbb{R}^n$, when the cost function is given by a possibly asymmetric norm $\|\cdot\|$ and only the first marginal is assumed to be absolutely continuous, without assuming strict convexity of the norm. In particular, this provides a proof of existence of solutions to the Monge problem which is based on a 1-dimensional disintegration technique relying on the stability of a particular solution of the problem. Given two Borel probability measures $\mu, \nu \in \mathscr{P}(\mathbb{R}^n)$, we study the minimization of the functional

$$\mathcal{I}_M(t) = \int_{\mathbb{R}^n} \|t(x) - x\| \, d\mu(x) \tag{MP}$$

among the Borel maps $t : \mathbb{R}^n \to \mathbb{R}^n$ whose image measure of μ is ν. We prove that the additional optimality conditions choosen in [2] determine a unique optimal transport map, selected also in [20], under the natural assumption that μ is absolutely continuous w.r.t. the Lebesgue measure $\mathcal{L}^n$. This assumption is necessary, as shown in Section 8 of [3].

The strategy is by reduction to 1-dimensional transport problems. It consists mainly of an area formula for the Lebesgue measure which allows the reduction: we prove a regularity of the disintegration along rays of the limit plan by the one of the approximations — once convergence is established. The limiting procedure is not based on Hopf-Lax formula of potential functions but on a uniqueness criterion. It is a particular case of a more general result in a forthcoming work by Bianchini and Daneri. The technique has been used in [9] in 2007, and then [17], improved simplifying the basic estimate in [7, 18].

Before introducing this work, we present a brief review of the main literature.

1.1. An Account on the Literature

The original Monge problem arose in 1781 for continuous masses μ, ν supported on compact, disjoint sets in dimension 2, 3 and with the cost defined by the Euclidean norm ([31]). Monge himself conjectured important features of the transport, such as, with the Euclidean norm, the facts that two transport rays may intersect only at endpoints and that the directions of the transported particles form a family of normals to some family of surfaces.

Investigated first in [4, 21], the problem was left apart for a long period.

A fundamental improvement in the understanding came with the relaxation of the problem in the space of probability measures ([26, 27]), consisting in the Kantorovich formulation. Instead of looking at maps in $\mathbb{R}^n$, ones considers the following minimization problem in the space $\Pi(\mu, \nu)$ of couplings between μ and ν: minimizing the linear functional

$$\mathcal{I}_K(\pi) = \int_{\mathbb{R}^n \times \mathbb{R}^n} \|y - x\| \, d\pi(x, y) \tag{KP}$$

among the *transport plans* π, defined as members of the set

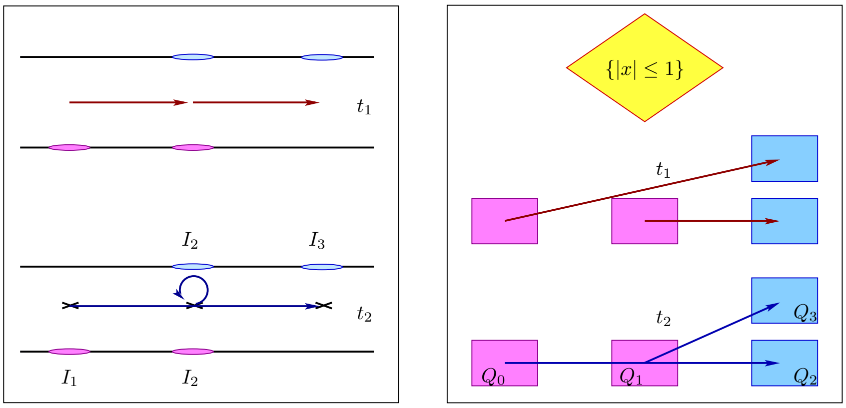
$$\Pi(\mu, \nu) = \{\pi \in \mathcal{M}^+ : \; p^x_\sharp \pi = \mu, \; p^y_\sharp \pi = \nu\},$$

where p^x, p^y are respectively the projections on the first and on the second factor space of $\mathbb{R}^n \times \mathbb{R}^n$. Notice that $\Pi(\mu, \nu)$ is convex, w^*-compact.

In particular, minimizers to (KP) always exist by the direct method of Calculus of Variations. The formulation (KP) is indeed a generalization of the model, allowing that mass at some point can be split to more destinations. Therefore, a priori the minimum value in (MP) is higher than the one in (KP), and the minimizers of the latter are not suitable for the former.

A standard approach to (MP) consists in showing that at least one of the optimizers to (KP) is concentrated on the graph of a function.

This is plainly effective when the cost is given by the squared Euclidean distance instead of $\|y - x\|$: by the uniform convexity there exists a unique



(a) One dimensional example.

Let μ be the Lebesgue measure on $I_1 \cup I_2 \subset \mathbb{R}$ and ν the Lebesgue measure on $I_2 \cup I_3$. Both the maps t_1 translating I_1 to I_2, I_2 to I_3 and the map t_2 translating I_1 to I_3 and leaving I_2 fixed are optimal. Moreover, any convex combination of the two transport plans induced by t_1, t_2 is again a minimizer for (KP), but clearly it is not induced by a map.

(b) Two dimensional example.

The unit ball of $\|\cdot\|$ is given by the rhombus. Let μ be the Lebesgue measure $Q_0 \cup Q_1 \subset \mathbb{R}^2$ and ν the Lebesgue measure on $Q_2 \cup Q_3$. Both the maps t_1, t_2 translating one of the first two squares to one of the second to squares are optimal, and they transport mass in different directions.

Figure 1: The optimal transport map with a generic norm is not unique.

optimizer π to (KP) of the form $\pi = (\mathrm{Id}, \mathrm{Id} - \nabla\phi)_\sharp\mu$ for a semiconcave function ϕ, the Kantorovich potential. Therefore, when $\mu \ll \mathcal{L}^n$ and $\nu \ll \mathcal{L}^n$, the optimal map is μ-a.e. defined by $x \mapsto x - \nabla\phi(x)$ and it is one-to-one ([12, 13, 28] are the first results, extended to *uniformly convex* functions of the distance e.g. in [32, 30, 25, 15]).

However, even in the case of the Euclidean norm, it is well known that this approach presents difficulties: at $\mathcal{L}^n$-a.e. point the Kantorovich potential fixes the direction of the transport, but not the precise point where the mass goes to. This is a feature of the problem, also in dimension one (see the example in Figure 1a).

The data are not sufficient to determine a single transport map, since there is no uniqueness. Uniqueness can be recovered with the further requirement of monotonicity along transport rays ([24]).

The situation becomes even more complicated with a generic norm cost function, instead of the Euclidean one. The symmetry of the norm plays no role, but the loss in strict convexity of the unit ball is relevant, since the transport may not occur along lines and the direction of the transport can vary (see the example in Figure 1b).

The Euclidean case, and thus the one proposed by Monge, has been rigorously solved only around 2000 in [22, 35, 3, 14].

Roughly, the approaches in the last three papers is at least partially based on a decomposition of the domain into 1-dimensional invariant regions for the transport, called transport rays. Due to the strict convexity of the unit ball, these regions are 1-dimensional convex sets. Due to regularity assumptions on the unit ball and a clever countable partition of the ambient space, it is moreover possible to reduce to the case where the directions of these segments is Lipschitz continuous. This, by Area or Coarea formula, allows to disintegrate the Lebesgue measure w.r.t. the partition in transport rays, obtaining absolutely continuous conditional probabilities on the 1-dimensional rays. In turn, this suffices to perform a reduction argument, that we also use in the present paper, which yields the thesis: indeed, one can fix within each ray an optimal transport map, uniquely defined imposing monotonicity within each ray. However, as in [9, 17, 16], we do not rely on any Lipschitz regularity of the vector field of directions for deriving an Area formula.

This kind of approach was introduced already in 1976 by Sudakov ([34]), in the more generality of a possibly asymmetric norm — which actually is the case we are considering. However, its argument remains incomplete: a regularity property of the disintegration of the Lebesgue measure w.r.t. decompositions of the space into affine regions was not proved correctly, and, actually, stated in a form which does not hold ([1]). Indeed, there exists a compact subset of the unit square having measure 1 and made of disjoint segments, with Borel direction, such that the disintegration of the Lebesgue measure w.r.t. the partition in segments has atomic conditional measures ([29], in [2] improved by Alberti et al.). The reduction argument described above requires instead absolutely continuous conditional measures, in order to solve the 1-dimensional transport problems, and therefore a regularity of the partition in transport rays must be proved. In the case of a strictly convex norm the affine regions reduce to lines and Sudakov argument was completed in [17]. In this paper we follow the alternative 1-dimensional decomposition selected by the additional variational principles, instead of the affine one considered by Sudakov. We choose the selection of [2], chosen also in [20].

The method in [22] is based on PDEs and they introduce the concept of transport density, widely studied since there — the very first works are [23, 1, 11, 24]. In [33] one finds more references as well as summability estimates obtained by interpolation and a limiting procedure of the kind also of this note; these are proved for the Euclidean distance, but they should work as well in this setting. Given a Kantorovich potential u for the transport problem between two absolutely continuous measures with compactly supported and smooth densities f^+, f^-, they define as transport density a nonnegative function a supported on the family of transport rays and satisfying

$$-\operatorname{div}(a\nabla u) = f^+ - f^-$$

in distributional sense. The above equation was present already in [5] with different motivation. It allows a generalization to measures, and an alternative definition introduced first in [10] for $\rho := a\mathcal{L}^n$ is given by the Radon measure defined on $A \in \mathcal{B}(\mathbb{R}^n)$ as

$$\rho(A) := \int_{\mathbb{R}^n\times\mathbb{R}^n} \mathcal{H}^1 \llcorner (A \cap [\![x, y]\!])\, d\pi(x, y), \tag{1}$$

where π is an optimal transport plan.

When the unit ball in not *strictly* convex, the first results available were given in [2] for the 2-dimensional case, completely solved, and for crystalline norms. Their strategy is to fix both the direction of the transport and the transport map by imposing additional optimality conditions, and then to carry out a Sudakov-type argument on the selected transports.

We follow the same strategy, and the disintegration technique from [6, 9].

A different proof of existence for general norms, with a selection based on the same optimality conditions, has been presented in [20], improving their argument for strictly convex norms in [19]. It does not arrive to disintegration of measures, it is more concerned with the regularity of the transport density. Also their argument is based on the geometric constraint that c_{s}-monotonicity impose on c_{s}-optimal transference plans, and an intermediate step is to prove that the set of initial points of secondary rays of a limit plan π, of the same maps we consider, is Lebesgue negligible. This important observation was also used for the solution in the special 2-dimensional case in [2], and generalized in more dimensions in [6, 9].

1.2. Topic of this Paper

By a *possibly asymmetric norm* $\|\cdot\|$ we mean a continuous function $\mathbb{R}^n \to [0, +\infty)$ having convex sublevel sets, containing the origin in the interior, and which is positively homogeneous ($\lambda\|x\| = \|\lambda x\|$ for $\lambda \geq 0$ and $x \in \mathbb{R}^n$). The study of this paper lies in the context of the following general problem, difficult due to the degeneracy and non-smoothness of the norm.

Primary Transport Problem. *Consider the Monge-Kantorovich optimal transport problem*

$$\min_{\pi\in\Pi(\mu,\nu)} \int \|y - x\|\, \mathrm{d}\pi(x, y) \tag{2}$$

between two positive Radon measures μ, ν with the same total variation, assuming that $\mu \ll \mathcal{L}^n$.

In order to avoid triviality we suppose that there exists a transport plan with finite cost. Since there is no uniqueness by the lack of *strict* convexity, one considers the family $\mathcal{O}_{\mathrm{p}} \subset \Pi(\mu, \nu)$ of minimizers to the primary problem. We call the members of $\mathcal{O}_{\mathrm{p}}$ the *optimal primary transport plans.* Let ϕ be a Kantorovich potential for this primary problem, by which we mean a function $\phi : \mathbb{R}^n \to \mathbb{R}$ such that

$$\phi(x) - \phi(y) \leq \|y - x\| \qquad \forall (x, y) \in \mathbb{R}^n \times \mathbb{R}^n \tag{3a}$$

$$\phi(x) - \phi(y) = \|y - x\| \qquad \text{for } \pi\text{-a.e. } (x, y), \quad \forall \pi \in \mathcal{O}_{\mathrm{p}}. \tag{3b}$$

We select then particular minimizers by the following secondary problem.

Secondary Transport Problem. *Consider a strictly convex norm* $|\cdot|$*. Study*

$$\min_{\pi \in \mathcal{O}_{\mathrm{p}}} \int |y - x| \, \mathrm{d}\pi(x, y) = \min_{\pi \in \Pi(\mu, \nu)} \int c_{\mathrm{s}}(x, y) \, d\pi(x, y) \tag{4}$$

where the secondary cost function c_{s} *is defined by*

$$c_{\mathrm{s}}(x, y) := \begin{cases} |y - x| & \text{if } \ \phi(x) - \phi(y) \leq \|y - x\|, \\ +\infty & \text{otherwise.} \end{cases}$$

This selection criterion has been applied first to the case of crystalline norms in [2], where in Section 4 one can also find the equivalence of the two minimizations in (4), by a general a variational argument based on Γ-convergence (applied also in the proof of Proposition 7.1 of [3]). The point of this paper is to show how to adapt the disintegration technique from [6, 9] in order to provide an area formula for the disintegration w.r.t. the rays of a plan which is optimal also for the secondary transport problem. Then one can apply the Sudakov-type argument to deduce existence and uniqueness of the optimal transport plan π monotone along rays which solves the secondary problem (4). In particular, this provides a different and simple proof of the existence result in [20]. We try to sketch it after some statements, the proofs are in Section 2.

We obtain more precisely the following. Let $\varepsilon \to 0^+$ and let t_ε be the optimal transport map, non decreasing along rays, between μ and ν for the strictly convex norm

$$c_\varepsilon(x, y) := \|y - x\| + \varepsilon |y - x|.$$

This map satisfies an absolutely continuity push forward estimate (below) that we want to prove in the limit. Restrict the attention for example to any part S of the domain $\{x \cdot \mathrm{e} \leq h^-\}$ where the map t_ε is valued in $\{x \cdot \mathrm{e} \geq h^+\}$. Lemma 2.17 of [17] proves that the maps satisfy the following area estimate:

$$\left(\frac{h^+ - t}{h^+ - s} \right)^{n-1} \mathcal{H}^{n-1}(\sigma_\varepsilon^s S) \leq \mathcal{H}^{n-1}(\sigma_\varepsilon^t S) \leq \left(\frac{t - h^-}{s - h^-} \right)^{n-1} \mathcal{H}^{n-1}(\sigma_\varepsilon^s S), \tag{5}$$

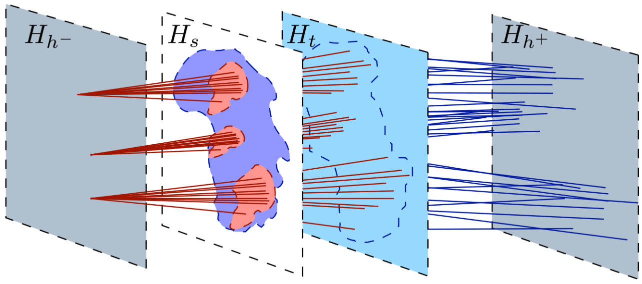


Figure 2: Area estimates of sections. The worst and better cases are obtained when the transport rays (right strip) are rays of cones (left and intermediate strips). Indeed, the proof can be made by cone approximations and a limiting procedure.

where $\sigma_\varepsilon^s S$, $\sigma_\varepsilon^t S$ are the intersections of the segments $[\![x, t_\varepsilon(x)]\!]$, for $x \in S$, with the hyperplanes $H_s = \{x \cdot \mathrm{e} = s\}$, $H_t = \{x \cdot \mathrm{e} = t\}$ and $h^- < s \leq t < h^+$. This estimate means that, moving a transversal section along rays, the area can either increase or decrease at most as if we were moving between H_{h^-} and H_{h^+} along cones with vertices respectively on H_{h^-} or H_{h^+}. See Figure 2.

Theorem 1.1. *The maps t_ε converge in measure to* the c_{s}*-optimal transport map monotone along rays.*

This implies pointwise convergence up to subsequence, and then by the minimality condition one has immediately that $t_\varepsilon - t$ converges to zero in $L^1(\mu)$. By the Γ-convergence argument quoted above ([2]), t and t_ε should moreover satisfy the asymptotic expansion

$$\int \|t(x) - x\|_\varepsilon d\mu = \int \|t_\varepsilon(x) - x\| d\mu + \varepsilon \int |t_\varepsilon(x) - x| d\mu + o(\varepsilon).$$

We sketch now the proof. If μ_ε, ν_ε are finite Radon measures w^*-converging to μ, ν, by the theory of Γ-convergence any w^*-limit π of c_ε-optimal transport plans π_ε in $\Pi(\mu_\varepsilon, \nu_\varepsilon)$ is a c_{s}-optimal transport plan in $\Pi(\mu, \nu)$ (e.g. Th. 4.1 in [3]). The convergence in μ-measure of a sequence of maps t_ε to a map t is equivalent to the w^*-convergence of the plans $(\mathrm{Id}, t_\varepsilon)_\sharp \mu$ to $(\mathrm{Id}, t)_\sharp \mu$: then the theorem follows

- providing the stated regularity of one limit plan π of π_ε (Proposition 1.4);
- observing that it is the unique c_{s}-optimal transport plan monotone along rays (Lemma 1.6).

Remark 1.2. *By uniqueness, the convergence holds also for different approximations, e.g. approximating contemporary* ν *by finitely many masses* ν_ε w^**-converging to* ν*.*

Before stating these auxiliary results, we remind some standard notations.

Recall 1.3. *By* transport set *associated to a set* $\Gamma \subset \mathbb{R}^n \times \mathbb{R}^n$ *we refer to the set*

$$\mathcal{T} = \big\{z : \quad z \in [\![x,y]\!],\ (x,y) \in \Gamma\big\},$$

where $[\![x,y]\!]$ *denotes the segment from* x *to* y *with endpoints,* $(\!|x,y|\!)$ *without. The transport set associated to a transport plan* π *is then a transport set associated to some* Γ *such that* $\pi(\Gamma) = 1$*. This definition is motivated by the fact that the optimal transport w.r.t. a strictly convex norm cost moves the mass along straight lines, as a consequence of the fact that the triangular inequality is strict when points are not aligned.*

A set $\Gamma \subset \mathbb{R}^n \times \mathbb{R}^n$ *is* c_{s}-monotone *if for all finite number of points* $\{(x_i, y_i)\}_{i=1,\dots,M}$ *belonging to* Γ *one has the following inequality:*

$$c_{\mathrm{s}}(x_0,y_0) + \cdots + c_{\mathrm{s}}(x_M,y_M) \leq c_{\mathrm{s}}(x_0,y_1) + \cdots + c_{\mathrm{s}}(x_{M-1},y_M) + c_{\mathrm{s}}(x_M,y_0).$$

In turn a transport plan π *is* c_{s}*-monotone if there exists a* c_{s}*-monotone set* Γ *such that* $\pi(\Gamma) = 1$*. In this case, secondary rays of* π *are those (nontrivial) segments* $[\![x,y]\!]$ *such that* $(x,y) \in \Gamma$*, where one can assume that*

$$(x,y) \in \Gamma,\ [\![z,w]\!] \subset [\![x,y]\!] \quad \Longrightarrow (z,w) \in \Gamma, \tag{6a}$$

$$(x,y),(z,w) \in \Gamma,\ [\![z,w]\!] \cap [\![x,y]\!] = [\![z,y]\!] \quad \Longrightarrow (x,w) \in \Gamma. \tag{6b}$$

Initial/terminal points of secondary rays are usually intended to be the initial/terminal points of secondary rays maximal w.r.t. inclusion. A plan π *is monotone (non decreasing) along rays if there exists* $\overline{\Gamma}$*,* $\pi(\overline{\Gamma}) = 1$*, such that if* $(x,y),(z,w) \in \overline{\Gamma}$ *then* $[\![z,w]\!] \not\subset (\!|x,y|\!)$*, or equivalently* $(y-w)\cdot(x-z) \geq 0$ *when aligned.*

Proposition 1.4. *There exists a* c_{s}*-optimal transport* $\pi \in \Pi(\mu,\nu)$ *such that the disintegration of the Lebesgue measure w.r.t. the rays of* π *has conditional probabilities equivalent to the Hausdorff* 1*-dimensional measure on the rays. The area estimates* (5) *hold.*

Corollary 1.5. *A* c_{s}*-optimal transport* $\pi \in \Pi(\mu,\nu)$ *is induced by a Borel map.*

The corollary is based on the 1-dimensional result: by the disintegration the transport in $\mathbb{R}^n$ reduces to transports on the 1-dimensional rays. The new measures to be transported are the conditional probabilities of μ and ν: then by Proposition 1.4 each conditional probability providing the first marginal on the relative ray is absolutely continuous and therefore the optimal transport problem can be solved by a map (a full proof is e.g. in [17], Th. 3.2). In particular the absolutely continuous disintegration implies that the (measurable) set of initial points is Lebesgue negligible, because its Hausdorff 1-dimensional measure on each ray is 0.

Having Proposition 1.4, at μ-a.e. point there is a unique outgoing secondary ray of π, because by c_s-monotonicity there can be more outgoing rays only at initial points and the set of initial points is negligible. Then one can see by considering convex combinations that any other c_s-optimal transport plan π' must have that same vector field of secondary rays direction, so that Lemma 1.6 below applies yielding the uniqueness stated in Theorem 1.1.

Lemma 1.6. *If there exists a Borel vector field fixing the direction of the secondary transport ray of any c_s-optimal transport plan $\pi' \in \Pi(\eta, \xi)$ at η-a.e. point, there exists at most one c_s-optimal plan in $\Pi(\eta, \xi)$ monotone along rays.*

Proofs are provided in Section 2, here we just sketch some ideas. We remark that by [7] the absolute continuity of Proposition 1.4 follows by the simplified area push forward estimate

$$\mu(A) > 0 \qquad \Rightarrow \qquad \mu(A_t) > 0 \text{ for a } \mathcal{L}^1\text{-positive set of times } t,$$

where A_t is the set of points in A translated of length t along any secondary ray they belong to. In this classical setting one can also prove the quantitative full estimate (5), which is clearly stronger.

This push forward area estimate for the secondary rays of the selected transport plan π is the main issue of this paper. It is derived by a compactness argument, and estimates (5) on approximating maps. In particular, it can be obtained as in the literature if one knows that there is just one c_s-optimal transference plan monotone along rays (Section 2.2). It is not difficult however to establish that the optimal transport monotone along rays is unique if the direction e.g. of terminal points is fixed for almost every point w.r.t. the target measure (Lemma 2.1, Section 2.1).

We then split our transport plan into partial transports, restrictions of π on suitable regions of $\mathbb{R}^n \times \mathbb{R}^n$. Moreover, we split them into fictional intermediate ones which are easily seen to be the unique c_s-optimal transport plans, monotone along rays, among their marginals. More precisely, for a model transport from $\mathbf{B}_{1/2}(0)$ to $\{x \cdot \mathrm{e} \geq 1\}$, the first fictional transport goes from $\mathbf{B}_{1/2}(0)$ to an intermediate section of the transport set, transversal to the rays, with an hyperplane H_λ, while the second from H_λ to $\{x \cdot \mathrm{e} \geq 1\}$. Let η be the target measure of the first one, and the source of the second one (see Figure 3). By standard geometric considerations implied by c_s-optimality and then c_s-monotonicity, one can deduce that the direction of the first transport is fixed at η-a.e. point. This allows a reduction to the previous cases, yielding the area push forward estimate for these partial transports, restrictions of π.

Having the area push forward estimate everything is done. We recover that the set of initial points is Lebesgue negligible (e.g. Lemma 2.20 in [17], coming from [9]). This implies again the uniqueness of the optimal transport plan monotone along rays (Lemma 1.6 below), and thus full estimates.

In Section 3 we stress some standard consequences of the disintegration result, and of the quantitative estimates. Namely, they provide some regularity of the divergence of rays directions vector field — a kind of Green-Gauss formula holds on special sets — and it allows moreover an explicit expression for the transport density. We give finally an example of the fact that the global optimal Kantorovich potential for the secondary problem with the cost c_{s} does not exist in general, but only on countably many sets which partition μ-all of $\mathbb{R}^n$.

2. Proof

We show the convergence of the optimal maps t_ε, non decreasing along rays, for the transport problem

$$c_\varepsilon(x,y) := \|y-x\| + \varepsilon|y-x|$$

to the optimal map non decreasing along rays for the transport problem

$$\min_{\pi\in\Pi(\mu,\nu)} \int c_{\mathrm{s}}(x,y)\, d\pi(x,y), \quad c_{\mathrm{s}}(x,y) := \begin{cases} |y-x| & \text{if } \phi(x)-\phi(y) \leq \|y-x\|, \\ +\infty & \text{otherwise.} \end{cases}$$

Section 2.1 remarks a uniqueness criterion relying on the fact that the direction of the transport is fixed at almost every point w.r.t. the source or target measure, that we state for the case we are considering.

We prove then by stability absolutely continuity area estimates for a limit plan. The estimates are quantitative as in [9], we indeed follow the same basic argument. In Section 2.2 we first prove this estimate in a simpler case, assuming that the support of the w^*-limit of $(\mathrm{Id}, t_\varepsilon)_\sharp\mu$ is c_{s}-cyclically monotone and that μ, ν are concentrated on disjoint balls. Then it is generalized in Section 2.3 by a countable partition satisfying some uniform estimates, and by the uniqueness of these c_{s}-optimal partial transports among their marginals.

2.1. A Uniqueness Remark

Consider two probability measures $\eta, \xi \in \mathcal{P}(\mathbb{R}^n)$. We stress uniqueness of the c_{s}-optimal transport plan (monotone along rays) when the direction of the transport is fixed e.g. at terminal points, by disintegration along rays and uniqueness of the optimal transport map in dimension one — known fact that does not require absolutely continuity assumptions.

LEMMA 2.1. *If there exists a Borel vector field fixing the direction of the secondary transport ray of any c_{s}-optimal transport plan $\pi' \in \Pi(\eta,\xi)$ at η-a.e. point, then there is a unique c_{s}-optimal plan in $\Pi(\eta,\xi)$ monotone along rays.*

Proof. Given any c_{s}-optimal transport plans $\bar{\pi}, \pi' \in \Pi(\eta,\xi)$, consider a transport set $\mathcal{T}$ of $(\pi'+\bar{\pi})/2$.

By c_{s}-monotonicity the secondary rays composing $\mathcal{T}$ can bifurcate only at endpoints. Moreover, by definition of terminal points both π, π' leave them fixed, and thus they coincide there; let us directly assume that the set of terminal points is η-negligible. Since by assumption at η-a.e. initial point of $\mathcal{T}$ there is precisely one secondary transport ray, then the vector field of secondary rays r'_{q} is single valued η-a.e.: secondary rays $\{\mathsf{r}_{\mathsf{q}}\}_{\mathsf{q}}$ partition η-almost all $\mathcal{T}$.

We recall that rays can be parametrized w.l.o.g. by countably many compact subsets of hyperplanes. Let h denote the quotient projection, and $\theta = h_\sharp\eta$ the quotient measure. Consider the disintegration of η, $\bar{\pi}$, π'

$$\eta = \int \eta_{\mathsf{q}}\theta(\mathsf{q}), \quad \bar{\pi} = \int \bar{\pi}_{\mathsf{q}}\theta(\mathsf{q}), \quad \pi' = \int \pi'_{\mathsf{q}}\theta(\mathsf{q})$$

respectively w.r.t. secondary transport rays $\{\mathsf{r}_{\mathsf{q}}\}_{\mathsf{q}}$ and w.r.t. the partition $\{\mathsf{r}_{\mathsf{q}} \times \mathbb{R}^n\}_{\mathsf{q}}$. In particular, we show that $\bar{\pi}_{\mathsf{q}}$ and π'_{q} have the same second marginal ξ_{q} for θ-a.e. q: indeed for any θ-measurable A and Borel S

$$\begin{aligned}
\xi(h^{-1}(A)\cap S) &\overset{\pi'\in\Pi(\eta,\xi)}{=} \pi'([0,1]\times(h^{-1}(A)\cap S)) \\
&= \int_A \pi'_{\mathsf{q}}([0,1]\times S)\theta(d\mathsf{q}) \\
&= \int_A \xi'_{\mathsf{q}}(S)\theta(d\mathsf{q}),
\end{aligned}$$

and the same holds for $\bar{\pi}$ with the second marginals $\bar{\xi}_{\mathsf{q}}$. Therefore, it must be $\bar{\xi}_{\mathsf{q}} = \xi'_{\mathsf{q}}$ and $\bar{\pi}_{\mathsf{q}}$, π'_{q} are monotone, 1-dimensional, $|y-x|$-optimal transports in $\Pi(\mu_{\mathsf{q}},\xi_{\mathsf{q}})$, even if we do not know whether $\xi_{\mathsf{q}} \ll \mathcal{H}^1 \llcorner_{\mathsf{r}_{\mathsf{q}}}$.

By the uniqueness of transport plans in dimension one (see e.g. Prop. 4.5 in [2]), then π and $\bar{\pi}$ must coincide. □

2.2. Example

Before treating the general case, we consider an example where we assume that the support of the limit plan is c_{s}-monotone. Even if by the theory of Γ-convergence the limit plan π is c_{s}-optimal and then c_{s}-monotone, since the cost c_{s} is just l.s.c. this hypothesis is indeed a restriction: we have $\pi(\Gamma) = 1$ for some c_{s}-monotone set Γ, which can be taken σ-compact but in general not closed. Moreover, if we restrict π to a compact set, then we loose in general the information that it is obtained by a limit.

If we knew the uniqueness of the c_{s}-optimal transport plan π, monotone along rays, then we could instead restrict π to a compact subset of Γ and we could apply to this restriction π' e.g. the statement below. Indeed, this restriction is still the unique c_{s}-optimal transport between its marginals: the transport plans induced by the c_ε-optimal transport maps between the marginals of π' would necessarily w^*-converge to π'.

Focus on an elementary domain. Let $\mu = f\mathcal{L}^n \llcorner_C$ for a compact $C \subset \mathbf{B}_{1/2}(0)$ and $f > 0$ on C, let e be a unit vector. Consider a sequence of continuous transport maps $t_{\varepsilon_j} : C \to \{x \cdot \mathrm{e} \geq 1\}$ which are c_{ε_j}-optimal and such that $(\mathrm{Id}, t_{\varepsilon_j})_\sharp \mu$ is weakly* convergent to a plan π.

We denote by Γ the support of π, extended by (6), and by $\mathcal{T}$ the relative transport set. The flux on secondary rays of π is the multivalued map $x \mapsto \sigma^t(x)$ which moves points along rays defined by

$$\sigma^t(x) = \big\{z : \ (x,z) \in \Gamma, \quad z \cdot \mathrm{e} = x \cdot \mathrm{e} + t\big\}.$$

The domain $\mathrm{Dom}(\sigma^h)$ of σ^h is the set of x such that $\sigma^h(x)$ is nonempty.

Lemma 2.2. *If the support of π is c_{s}-monotone, the transport set $\mathcal{T}$ satisfies the estimate*

$$\left(\frac{1-t}{1-s}\right)^{n-1} \mathcal{H}^{n-1}\big(S\big) \leq \mathcal{H}^{n-1}\big(\sigma^{t-s} S\big) \tag{7}$$

for all compact $S \subset C \cap \{x \cdot \mathrm{e} = s\} \cap \mathrm{Dom}(\sigma^{t-s})$ and $s < t \leq 1$. The symmetric estimate holds similarly.

Corollary 2.3. *If there exists a unique c_{s}-optimal transport plan π, it is induced by a transport map satisfying the absolutely continuous push forward estimates of the kind* (5)*.*

The corollary follows by the elementary restriction argument above.

Proof of Lemma 2.2. Since π is concentrated on the compact set Γ, by the weak convergence for every $\delta > 0$

$$\begin{aligned} 0 &= \pi\big(\big\{(x,y) : \ \mathrm{dist}((x,y),\Gamma) \geq \delta\big\}\big) \\ &\geq \limsup_{\varepsilon_j \to 0} \Big\{ \big((\mathrm{Id}, t_{\varepsilon_j})_\sharp \mu\big)\big(\big\{(x,y) : \ \mathrm{dist}((x,y),\Gamma) \geq \delta\big\}\big)\Big\} \geq 0. \end{aligned}$$

Thus $\mu(\{x : \ \mathrm{dist}((x, t_{\varepsilon_j}(x)),\Gamma) \geq \delta\})$ tends to 0 as $\varepsilon_j \to 0$. Up to a subsequence, one can require then

$$\mu(J_j) < 2^{-j} \qquad \text{with } J_j := \big\{x : \ \mathrm{dist}((x, t_{\varepsilon_j}(x)),\Gamma) > 2^{-j}\big\}.$$

Define the intermediate hyperplanes

$$H_\lambda = \{x : \ x \cdot \mathrm{e} = \lambda\}.$$

Notice that $\mathcal{H}^{n-1}(J_j \cap H_\lambda)$ converges to 0 for $\mathcal{L}^1$-a.e. λ, being $1\!\!1_{J_i}$ converging to 0 $\mathcal{L}^n$-a.e.

Up to a translation, we set $s = 0$. Let S be any compact subset of $C \cap \{x \cdot \mathrm{e} = 0\} \cap \mathrm{Dom}(\sigma^t)$. Since secondary transport rays are identified by the compact, c_{s}-monotone set Γ, which is the support of π suitably extended by (6), and since S is also compact, notice then that $\sigma^t(S)$ is compact, too.

Let $h \to \sigma_j^h$ be the analogous flux along secondary rays of t_{ε_j} and set

$$K^j := \sigma_j^t(S \setminus J_j) \quad \subset \quad \{x \cdot \mathrm{e} = t\}.$$

Being compact, K^j converges in the Hausdorff distance, up to subsequence, to a compact set K. Moreover, since by construction $d((x, t_{\varepsilon_j}(x)), \Gamma) \leq 2^{-j}$ out of J_j, we have that $K \subset \overline{\sigma^t(S)} = \sigma^t(S)$. By the u.s.c. of the Hausdorff measure and the regularity of the approximating vector field we conclude

$$\begin{aligned}\mathcal{H}^{n-1}(\sigma^t(S)) \geq \mathcal{H}^{n-1}(K) \geq \limsup_j \mathcal{H}^{n-1}(K^j) \\ \geq \lim_j \big\{(1-t)^{n-1}\mathcal{H}^{n-1}(S \setminus J_j)\big\} = (1-t)^{n-1}\mathcal{H}^{n-1}\big(S\big),\end{aligned}$$

where the last equality holds if $\mathcal{H}^{n-1}(J_j \cap H_\lambda)$ goes to 0. The thesis holds as well also for the remaining ($\mathcal{L}^1$-negligible) values of s by the lower semicontinuity of $\mathcal{H}^{n-1}$, being for λ decreasing to s

$$\mathcal{H}^{n-1}(\mathbf{B}_\delta(S) \cap H_s) \leq \liminf_{\lambda \downarrow s} \mathcal{H}^{n-1}(\mathbf{B}_\delta(\sigma^{\lambda - s} S) \cap H_\lambda). \qquad \square$$

2.3. Proof of Proposition 1.4

We disintegrate here the Lebesgue measure on the transport set $\mathcal{T}$, associated to any w^*-limit π of $(\mathrm{Id}, t_{\varepsilon_j})_\sharp \mu$, w.r.t. the partition into secondary rays: we show by a quantitative area push forward estimate that the conditional probabilities are absolutely continuous. We basically reduce to the case of the example in Lemma 2.2.

The idea is the following. In the model case of a transport $\mathbf{B}_{1/2}(0) \to \{x \cdot \mathrm{e} \geq 1\}$, we see the optimal transport plan we selected as a composition of two other optimal transport plans: $\mathbf{B}_{1/2}(0) \to H_\lambda$ and $H_\lambda \to \{x \cdot \mathrm{e} \geq 1\}$, where H_λ is an intermediate section transversal to the rays. The two intermediate transports should still be c_{s}-optimal, and moreover they basically share the same secondary rays as their composite plan. The terminal points of secondary rays of the first, coinciding with the initial ones of the second, should then be

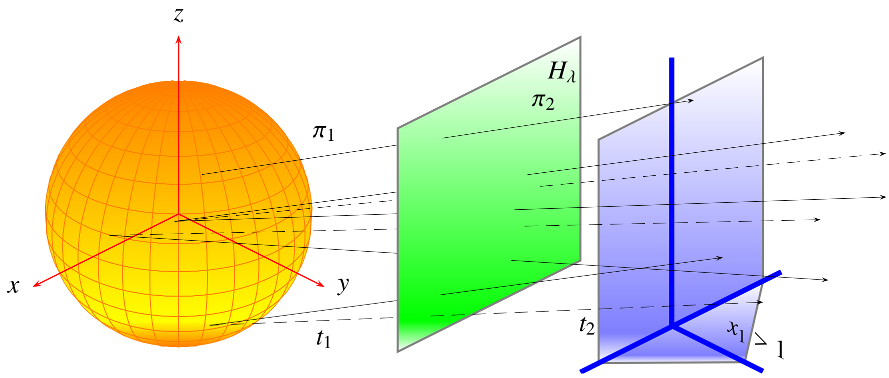


Figure 3: Decomposition of the transport.

fixed — because of c_{s}-monotonicity applied to the composite plans. Then the intermediate transports we selected are the unique optimal transports between their marginals by Lemma 1.6. In particular, Lemma 2.2 applies for these transports yielding the area estimate, which holds both for the intermediate and composite transports.

Since the property of the proposition is local, up to a countable partition and similarity transformations we are allowed to decompose π into countably many restrictions of it, that we place on a model set. Notice moreover that, up to the present purpose, one does not need to consider the restriction of π to the diagonal, i.e. the fixed points, because also removing them the secondary rays of the transport set remain the same. They matter only in order to determine later the transport map solving the Monge problem with the given marginal.

Renewing the notations for the marginals of these partial plans, we assume

$$\mu \ll \mathcal{L}^n \llcorner_{\mathbf{B}_{1/2}(0)} \qquad \text{and} \qquad \nu(\{x \cdot \mathrm{e} \geq 1\}) = 1.$$

We now denote by t_ε the c_ε-optimal maps monotone along ray from μ to ν. If, by Lusin theorem, we also assume that the maps t_ε are continuous by restricting them to suitable compact sets, then we are in a setting where the second marginal is in general different from ν, w^*-converging to it:

$$t_\varepsilon : \mathbf{B}_{1/2}(0) \to \{x \cdot \mathrm{e} \geq 1\} \qquad \text{and} \qquad \nu_\varepsilon := (t_\varepsilon)_\sharp \mu.$$

The w^*-limit π of a sequence $(\mathrm{Id}, t_\varepsilon)_\sharp \mu$, after showing the uniqueness, will turn out to be precisely one of the restrictions originally considered.

View each map t_ε as the composition of two c_ε-optimal transports (Figure 3): for $\lambda \in (1/2, 1)$

$$t^1_\varepsilon(x) := x + \lambda[(\mathrm{e} - x) \cdot d_\varepsilon(x)]d_\varepsilon(x), \qquad t^2_\varepsilon(x) := t_\varepsilon((t^1_\varepsilon)^{-1}(x)),$$

$$d_\varepsilon(x) := \frac{t_\varepsilon(x) - x}{|t_\varepsilon(x) - x|}.$$

Let $\xi_\varepsilon := (t^1_\varepsilon)_\sharp\mu$ be the intermediate measure on $H_\lambda := \{x : \; x \cdot \mathrm{e} = \lambda\}$, which is the source of t^2_ε and the target of t^1_ε. Notice that t^1_ε is injective, by the c_ε-monotonicity, so that t^2_ε is well defined ξ_ε-a.e.

By compactness, the transport plans π^2_ε, π^1_ε associated to t^1_ε, t^2_ε w^*-converge, up to a subsequence, to plans $\pi^1 \in \Pi(\mu, \xi)$, $\pi^2 \in \Pi(\xi, \nu)$ — where ξ, ν are the w^*-limit of ξ_ε, ν_ε. By the theory of Γ-convergence (see e.g. Th. 4.1 in [3]) π^1, π^2 are c_s-optimal transport plans. Moreover, since for all $\hat\pi \in \Pi(\mu + \xi_\varepsilon, \nu_\varepsilon + \xi_\varepsilon)$

$$\int c_\varepsilon(x, t^1_\varepsilon(x))\mu(dx) + \int c_\varepsilon(z, t^2_\varepsilon(z))\xi(dz) = \int c_\varepsilon(x, t_\varepsilon(x))\mu(dx) \leq \int c_\varepsilon \hat\pi,$$

by the c_ε-optimality of $(\mathrm{Id}, t^1_\varepsilon)_\sharp\mu + (\mathrm{Id}, t^2_\varepsilon)_\sharp\xi_\varepsilon$ also $\pi^1 + \pi^2$ *is* c_s*-optimal*.

Since (maximal) secondary rays of $\pi^1 + \pi^2$ go from $\mathbf{B}_{1/2}(0)$ to $\{x \cdot \mathrm{e} \geq 1\}$, *the direction of the transport is unique at* ξ*-a.e. point.* The measurability of this vector field follows from the fact that there is a representative with a σ-compact graph, because $\pi^1 + \pi^2$ is concentrated on a σ-compact set. Observing that $\pi^1 \ll \pi^1 + \pi^2$ and that π^1 is monotone along rays, then Lemma 1.6 states that π^1 is the only c_s-optimal transport plan from μ to ξ monotone along rays.

The uniqueness of π^1 yields the basic push forward estimate (7) for π^1 by Corollary 2.3. However, this estimate coincides with the basic push forward estimate also for the c_s-optimal transports $\pi^1 + \pi^2 \in \Pi(\mu + \xi, \nu + \xi)$ and for

$$\Pi(\mu, \nu) \ni \pi(dx, dy) := \int \pi^2_z(dy)\pi^1(dx, dz), \text{ where } \pi^2(dz, dy) = \int \pi^2_z(dy)\xi(dz),$$

which share the same transport set $\mathcal{T}$.

This yields a one-sided estimate, but it is enough in order to deduce by a density argument that the set of initial points of π is negligible (precisely Lemma 2.20 in [17], proved before in [9]). Indeed, if we could take a point $\bar x$ of density one for the set of initial points $\mathcal{E}$, then we would reach an absurd: on one hand moving them along the ray directions, close to e, they are no more initial points and therefore they do not belong to the set $\mathcal{E}$; on the other hand the one-sided estimate implies that $\bar x$ is a point of positive density for these translated points.

As remarked in Section 2.1, then uniqueness hold and one finds by the limiting procedure the full estimates giving the statements of the proposition.

3. Some Remarks

The first corollary of the previous computations is the following disintegration.

THEOREM 3.1. *The family of secondary transport rays* $\{\mathsf{r}_\mathsf{q}\}_{\mathsf{q}\in\mathsf{Q}}$ *can be parameterized by a Borel subset* Q *of countably many hyperplanes. The transport set* $\mathcal{T} = \cup_{\mathsf{q}\in\mathsf{Q}}\mathsf{r}_\mathsf{q}$ *is Borel and there exists a Borel function* γ *such that the following disintegration of* $\mathcal{L}^n \llcorner_\mathcal{T}$ *holds:* $\forall\varphi$ *either integrable or positive*

$$\int_\mathcal{T} \varphi(x) d\mathcal{L}^n(x) = \int_\mathsf{Q} \left\{ \int_{\mathsf{r}_\mathsf{q}} \varphi(s)\gamma(s) d\mathcal{H}^1(s) \right\} d\mathcal{H}^{n-1}(\mathsf{q})$$

The set of endpoints of rays is Lebesgue negligible.

Denote by d the unit vector field of secondary rays directions, defined $\mathcal{L}^n$-a.e. on $\mathcal{T}$. The quantitative estimates, as in the previous works in $\mathbb{R}^n$, imply a further regularity of the density γ. We omit the proof and precise formulas, analogous to the one of Proposition 4.17 in [16].

LEMMA 3.2. *For* $\mathcal{L}^n$*-a.e.* x *the real function*

$$\lambda \mapsto \gamma(x + \lambda\mathsf{d}(x))$$

is locally Lipschitz, with locally finite total variation, for x, $x + \lambda\mathsf{d}(x)$ *belonging to a same ray.*

Moreover, as e.g. in [16] the function

$$\frac{\partial_\lambda \gamma(x + \lambda\mathsf{d}(x))}{\gamma(x + \lambda\mathsf{d}(x))} =: \big(\operatorname{div}\mathsf{d}\ \big)_{\text{a.c.}}(x)$$

is of particular interest, as we explain below motivating the abuse of notation.

3.1. Divergence of the Rays Directions Vector Field

Consider a compact subset $\mathcal{Z}$ of the transport set $\mathcal{T}$ made of secondary transport rays which intersect an hyperplane H at points which are in the relative interior of the rays. Then the divergence of the vector field $\mathsf{d}\mathbb{1}_\mathcal{Z}$ is a Radon measure.

LEMMA 3.3. *There exist nonnegative measures* η^+, η^-*, concentrated respectively on initial and terminal points of secondary rays of* $\mathcal{Z}$*, such that*

$$\operatorname{div}(\mathsf{d}\mathbb{1}_\mathcal{Z}) = \big(\operatorname{div}\mathsf{d}\ \big)_{\text{a.c.}}(x)\mathcal{L}^n \llcorner_\mathcal{Z} - \eta^+ + \eta^-.$$

If the initial points are on a same hyperplane H^- *and the terminal points on a same hyperplane* H^+ *orthogonal to a unit direction* e*, then the measures* $\mu^\pm$ *are just* $|\mathsf{d}\cdot\mathrm{e}|\mathcal{H}^{n-1}$ *on the relative hyperplane.*

The proof relies on the disintegration theorem for reducing the integrals $\int_{\mathcal{Z}} \nabla\varphi \cdot \mathrm{d}\mathcal{L}^n$ on the rays, where φ is a test function. The factor γ appears in the area formula, and *on each ray* by the estimates providing BV regularity one can integrate by parts (see the proof of Lemma 2.30 in [17]).

It follows then that the distributional divergence of d is a series of measures. We remark however that in general the divergence of d is just a distribution, and it may fail to be a measure (see e.g. Examples 4.2, 4.3 in [17]). As well, the function $\big(\operatorname{div}\mathsf{d}\big)_{\text{a.c.}}$ could fail to be locally integrable.

3.2. Transport Density

We now stress another known consequence of the disintegration theorem: one can write the expression of the transport density, vanishing approaching initial points along secondary transport rays, relative to optimal secondary transport plans in terms of the conditional measures μ_{q}, ν_{q}, $\mathsf{q} \in \mathsf{Q}$ of μ, ν for the ray equivalence relation. In particular, one can see its absolute continuity. It does not vanish approaching terminal points — see Example 3.5 below taken from [24]. We omit the verification, since it is quite standard (see e.g. Section 8 in [9]).

Let f, the Radon-Nycodim derivative of μ w.r.t. $\mathcal{L}^n$, and γ, introduced in the disintegration, be Borel functions from $\mathbb{R}^n$ to $\mathbb{R}$ such that

$$\mu \llcorner \mathcal{T} = \int_{\mathsf{Q}} \mu_{\mathsf{q}}\, d\mathcal{H}^{n-1}(\mathsf{q}) = \int_{\mathsf{Q}} (f\gamma\mathcal{H}^1 \llcorner_{r_{\mathsf{q}}})\, d\mathcal{H}^{n-1}(\mathsf{q}) \qquad \nu \llcorner \mathcal{T} = \int_{\mathsf{Q}} \nu_{\mathsf{q}}\, d\mathcal{H}^{n-1}(\mathsf{q}).$$

Let $\mathsf{q} : \mathcal{T} \to \mathsf{Q}$ be the Borel multivalued quotient projection. Set $d = 0$ where $\mathcal{D}$ is multivalued.

Lemma 3.4. *A solution $\rho \in \mathcal{M}^+_{\mathrm{loc}}(\mathbb{R}^n)$ to the transport equation*

$$\operatorname{div}(\rho\mathsf{d}) = \mu - \nu$$

is given by

$$\rho(x) = \frac{(\mu_{\mathsf{q}(x)} - \nu_{\mathsf{q}(x)})((\!| a(x), x |\!))}{\gamma(x)} \mathcal{L}^n(x) \llcorner \mathcal{T} = \left(\frac{1\!\!\!\!\not{}\,\mathcal{T}(x)}{\gamma(x)} \int_{(\!|\tilde{t}^{-1}(x), x|\!)} f\gamma\, d\mathcal{H}^1 \right) \mathcal{L}^n(x). \tag{8}$$

Example 3.5 (Taken from [24]). *Consider in $\mathbb{R}^2$ the measures $\mu = 2\mathcal{L}^2 \llcorner_{\mathbf{B}_1}$ and $\nu = \frac{1}{2|x|^{3/2}}\mathcal{L}^2\mathbf{B}_1$, where $|\cdot|$ here denotes the Euclidean norm. A Kantorovich potential is provided by $|x|$. The transport density is $\rho = (|x|^{-\frac{1}{2}} - |x|)\mathcal{L}^2 \llcorner_{\mathbf{B}_1}$. While vanishing towards $\partial\mathbf{B}_1$, the density of ρ blows up towards the origin. Concentrating ν at the origin, the density would be instead $\rho = -|x|^2 \llcorner_{\mathbf{B}_1}$.*

3.3. Example of no Global Secondary Potential

We show here that in general there exists no function ϕ_{s} which satisfies on the whole transport set relative to two measures μ, ν

$$\phi_{\mathrm{s}}(x) - \phi_{\mathrm{s}}(y) = c_{\mathrm{s}}(x, y) \qquad \text{for } x, y \text{ on a same secondary ray} \tag{9a}$$
$$\phi_{\mathrm{s}}(x) - \phi_{\mathrm{s}}(y) \leq c_{\mathrm{s}}(x, y) \qquad \forall (x, y) \in \mathcal{T} \times \mathcal{T}. \tag{9b}$$

It would otherwise provide a global Kantorovich potential for the secondary transport problem, which exists only up to a countable partition of the domain. The secondary cost function was defined by

$$c_{\mathrm{s}}(x, y) := \begin{cases} |y - x| & \text{if } \phi(x) - \phi(y) = \|y - x\|, \\ +\infty & \text{otherwise.} \end{cases}$$

Consider in $\mathbb{R}^2$ the norm $\|x\| = |P_1 x| + |P_2 x|$, where P_1, P_2 are the projections on the first and second component, and let $|\cdot|$ denote the Euclidean norm. We show for simplicity of notations a transport problem with atomic marginals, the example can then be adapted spreading the mass as in the pictures of Figure 4. Consider the transport among the measures

$$\mu = \sum_{i=1}^{\infty} \sum_{j=-1}^{4/h_i} (h_i/12)^2 \delta_{w_{ij}} + (1/24)^2 \delta_{w_\infty} \quad \nu = \sum_{i=1}^{\infty} \sum_{j=-1}^{4/h_i} (h_i/12)^2 \delta_{z_{ij}} + (1/24)^2 \delta_{z_\infty}$$

where $h_i = 2^{-i-1}$, $w_\infty = (-1.5, 0)$, $z_\infty = (-1.5, 1)$ and

$$\begin{cases} w_{1,-1} = (0,\, 0) \\ w_{i,-1} = \big(-\sum_{k=1}^{i,-1} 2h_k,\, 0 \big) \\ w_{ij} = w_{i,-1} + (-h_i,\, jh_i/4) \end{cases} \qquad \begin{cases} z_{i,-1} = w_{i,-1} + (0, 1) & i \in \mathbb{N} \\ z_{ij} = w_{ij} + (h_i/2,\, 0) & j = 0, \dots, 4/h_i. \end{cases}$$

Let π be the transport plan induced by the map t which translates each w_{ik} to z_{ik}. It is an optimal one for the primary problem: one can see it for example by duality, noticing that the function

$$\phi(x) = \|x - z_{1,-1}\|$$

is a Kantorovich potential. Moreover, one can immediately verify that it is also c_{s}-optimal. In this case, one can take

$$c_{\mathrm{s}}(x, y) := \begin{cases} |y - x| & \text{if } \{P_1 y \geq P_1 x,\ P_2 y \geq P_2 x\}, \\ +\infty & \text{otherwise.} \end{cases}$$

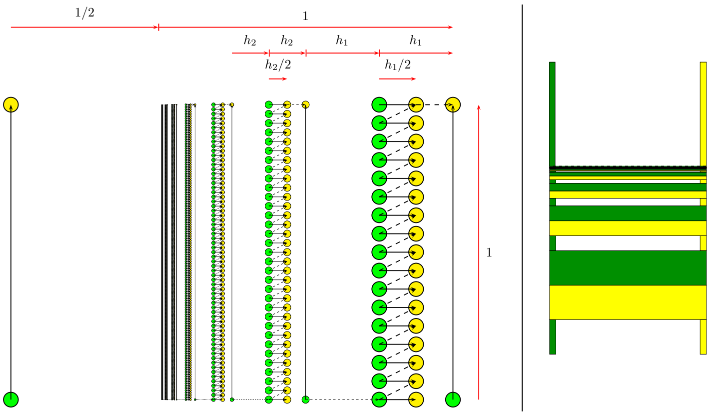


Figure 4: Non existence of a global secondary potential. LHS: back on the path along the arrows the secondary potential must go to $-\infty$. RHS (rotated): the mass is spread so that the primary potential is unique, up to constants.

However, no c_{s}-monotone carriage Γ of π is contained in the c_{s}-subdifferential of a c_{s}-monotone function ϕ_{s}, which by definition would satisfy (9). Indeed, suppose the contrary. Then, considering the path in the figure and applying repeatedly the maximal growth equality (9a) (full line) and the Lipschitz inequality (9b) (dashed line), one finds

$$\begin{aligned}\phi_{\mathrm{s}}(w_{(i+1),-1}) &\leq \phi_{\mathrm{s}}(w_{i,-1}), -1 + \frac{h_i}{2} + \frac{h_i}{2}\left(\frac{\sqrt{5}}{2} - 1\right) \cdot \frac{4}{h_i} + \frac{3h_i}{2} \\ &= \phi_{\mathrm{s}}(w_{i,-1}) + \sqrt{5} - 3 + 2h_i.\end{aligned}$$

For every potential ϕ_{s} finite on $w_{1,-1}$, we find therefore that $\phi_{\mathrm{s}}(w_{i,-1}) \to -\infty$ for $i \to \infty$, as well as every other $\phi_{\mathrm{s}}(w_{ij})$. This implies that ϕ_{s} must be $-\infty$ on w_∞: for all i, j

$$\phi_{\mathrm{s}}(w_\infty) \leq \phi_{\mathrm{s}}(w_{ij}) + \|w_{ij} - w_\infty\|,$$

which implies $\phi_{\mathrm{s}}(w_\infty) = -\infty$.

REMARK 3.6. *One could think that the problem is that the primary potential we have chosen is not the right one. However, this is not the case. A completely similar behavior happens spreading the mass as in the second picture of Figure 4 (rotated of $-\pi/2$), but there the primary potential is unique, up to constants.*

Acknowledgements. The author wishes to thank in particular professor Stefano Bianchini for several valuable discussions and the referee for his constructive remarks.

References

[1] L. Ambrosio, *Lecture notes on optimal transport problems*, Lecture Notes in Mathematics, volume 1812, Springer, Berlin (2003), pp. 1–52.

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

[3] L. Ambrosio and A. Pratelli, *Existence and stability results in the* L^1 *theory of optimal transportation*, Lecture Notes in Mathematics, volume 1813, Springer, Berlin (2003), pp. 123–160.

[4] P. Appell, *Mémoire sur les déblais et les remblais des systèmes continus ou discontinus*, Mémoires preésentes par divers Savants à l'Académie des Sciences de l'Institut de France **29** (1887), 1–208.

[5] M. Beckmann, *A continuous model of transportation*, Econometrica **20** (1952), 643–660.

[6] S. Bianchini, *On the Euler-Lagrange equation for a variational problem*, Discrete Contin. Dyn. Syst. **17** (2007), 449–480.

[7] S. Bianchini and F. Cavalletti, *The Monge problem for distance cost in geodesic spaces*, arXiv:1103.2796.

[8] S. Bianchini and S. Daneri, to appear.

[9] S. Bianchini and M. Gloyer, *On the Euler Lagrange equation for a variational problem: the general case II*, Math. Z. **265** (2010), 889–923.

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

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

[12] Y. Brenier, *Décomposition polaire et réarrangement monotone des champs de vecteurs*, C. R. Acad. Sci. Paris Sér. I Math. **305** (1987), 805–808.

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

[14] L. Caffarelli, M. Feldman and R. McCann, *Constructing optimal mass for Monge's transport problem as a limit of strictly convex costs*, J. Amer. Math. Soc. (2002), 1–26.

[15] L.A. Caffarelli, *Boundary regularity of maps with convex potentials. II*, Ann. Math. **144** (1996), 453–496.

[16] L. Caravenna and S. Daneri, *The disintegration of the Lebesgue measure on the faces of a convex function.*, J. Funct. Anal. **258** (2010), 3604–3661.

[17] L. Caravenna, *A proof of Sudakov theorem with strictly convex norms*, Math. Z. **268** (2011), 371–407.

[18] F. Cavalletti, *A strategy for non-strictly convex distance transport cost and the obstacle problem*, arXiv:1103.2797.

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

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

[21] C. Dupin, *Applications de la géométrie et de la mécanique*, Bachelier, Paris (1822).

[22] L.C. Evans and W. Gangbo, *Differential equations methods for the Monge-Kantorovich mass transfer problem*, Current Developments in Mathematics (1997), 65–126.

[23] M. Feldman and R. McCann, *Monge's transport problem on a Riemannian manifold*, Trans. Amer. Math. Soc. (2002), 1667–1697.

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

[25] W. Gangbo, *An elementary proof of the polar factorization of vector-valued functions*, Arch. Rational Mech. Anal. **128** (1994), 381–399.

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

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

[28] M. Knott and C. Smith, *On the optimal mapping of distributions*, J. Optim. Theory Appl. **43** (1984), , 39–49.

[29] D.G. Larman, *A compact set of disjoint line segments in $\mathbb{R}^3$ whose end set has positive measure.*, Mathematika **18** (1971), 112–125.

[30] R.J. McCann, *Existence and uniqueness of monotone measure-preserving maps*, Duke Math. J. **80** (1995), 309–323.

[31] G. Monge, *Mémoire sur la theorie des déblais et des remblais*, Histoire de l'Acad. de Sciences de Paris (1781), 666–704.

[32] L. Rüschendorf and S.T. Rachev, *A characterization of random variables with minimum L^2-distance*, J. Multivariate Anal. **32** (1990), 48–54.

[33] F. Santambrogio, *Absolute continuity and summability of transport densities: simpler proofs and new estimates*, Calc. Var. Partial Differential Equations **36** (2008), 343–354.

[34] V.N. Sudakov, *Geometric problems in the theory of infinite-dimensional probability distributions*, Proc. Steklov Inst. Math. **2** (1979), 1–178 (in English); Trudy Mat. Inst. Steklov **141** (1976), 3–191 (in Russian).

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

Author's address:

Laura Caravenna
CRM De Giorgi, Collegio Puteano
Scuola Normale Superiore
Piazza dei Cavalieri 3, 56100 Pisa, Italy
E-mail: `laura.caravenna@sns.it`

Received December 14, 2010
Revised March 21, 2011