

**L^p BOUNDS FOR BOUNDARY-TO-BOUNDARY TRANSPORT DENSITIES,
AND $W^{1,p}$ BOUNDS FOR THE BV LEAST GRADIENT PROBLEM IN 2D**

SAMER DWEIK AND FILIPPO SANTAMBROGIO

ABSTRACT. The least gradient problem (minimizing the total variation with given boundary data) is equivalent, in the plane, to the Beckmann minimal-flow problem with source and target measures located on the boundary of the domain, which is in turn related to an optimal transport problem. Motivated by this fact, we prove L^p summability results for the solution of the Beckmann problem in this setting, which improve upon previous results where the measures were themselves supposed to be L^p . In the plane, we carry out all the analysis for general strictly convex norms, which requires to first introduce the corresponding optimal transport tools. We then obtain results about the $W^{1,p}$ regularity of the solution of the anisotropic least gradient problem in uniformly convex domains.

1. INTRODUCTION

A classical problem in calculus of variations, which is of interest in image processing and which is also related to minimal surfaces, is the so-called least gradient problem, considered for instance in [4, 24, 25, 28, 30, 39]. This is the problem of minimizing the total variation of the vector measure ∇u among all BV functions u defined on an open domain Ω with given boundary datum. To be more general, we will directly consider the anisotropic case ([27, 29]), using an arbitrary strictly convex norm φ in \mathbb{R}^d . Consider

$$(1.1) \quad \inf \left\{ \int_{\Omega} \varphi(\nabla u) : u \in BV(\Omega), u|_{\partial\Omega} = g \right\},$$

where $u|_{\partial\Omega}$ denotes the trace of u in the sense of BV functions and $\int_{\Omega} \varphi(\nabla u)$ is the φ -total variation measure of ∇u (i.e., $\int_{\Omega} \varphi(\nabla u) := \int_{\Omega} \varphi(\frac{d\nabla u}{|d\nabla u|}(x)) d|\nabla u|$), this problem relaxes into (see, for instance, [29, 30])

$$(1.2) \quad \min \left\{ \int_{\Omega} \varphi(\nabla u) + \int_{\partial\Omega} |u|_{\partial\Omega} - g| \varphi(\mathbf{n}) d\mathcal{H}^{d-1} : u \in BV(\Omega) \right\},$$

where \mathbf{n} is the outward normal to $\partial\Omega$. This can also be expressed in the following way: extend g into a BV function \tilde{g} defined on a larger domain Ω' , and then consider

$$(1.3) \quad \min \left\{ \int_{\Omega} \varphi(\nabla u) : u \in BV(\Omega'), u = \tilde{g} \text{ on } \Omega' \setminus \Omega \right\}.$$

The boundary datum g should be taken as a possible trace of BV functions, i.e. in $L^1(\partial\Omega)$. Yet, the fact that the (a) solution u to (1.2) and (1.3) satisfies or not $u|_{\partial\Omega} = g$ could depend on g (and on the domain). In case we have $u|_{\partial\Omega} = g$, then u is also a solution of (1.1). In the Euclidean case (i.e., when $\varphi = |\cdot|$), the author of [24] proves existence of solutions to (1.1) for boundary data in $BV(\partial\Omega)$, while, in [37], the authors give an example of a function g such that

(1.1) has no solution (g was chosen to be the characteristic function of a certain fat Cantor set, which does not lie in $BV(\partial\Omega)$).

In this paper the anisotropic least gradient problem will only be considered in the planar case $\Omega \subset \mathbb{R}^2$, and the boundary datum g will be at least in $BV(\partial\Omega)$ (this makes perfectly sense, since $\partial\Omega$ is a closed curve, and we only need to consider BV functions in 1D). Following [25], we can see that there is a one-to-one correspondence between vector measures ∇u in (1.3) (considered as measures on $\bar{\Omega}$, so that we also include the part of the derivative of u which is on the boundary, i.e. the possible jump from $u|_{\partial\Omega}$ to g) and vector measures w satisfying, in $\bar{\Omega}$, $\nabla \cdot w = f$ where $f := f^+ - f^-$ is the measure obtained as the tangential derivative of $g \in BV(\partial\Omega)$: we set $f = \partial g / \partial \mathbf{t}$, where $\mathbf{t} := R_{-\frac{\pi}{2}} \mathbf{n}$ (R_θ denotes a rotation with angle θ around the origin) is the tangent vector to $\partial\Omega$; f^+ and f^- are then defined as the positive and negative parts of f , respectively, and we have $\int_{\bar{\Omega}} f^+ = \int_{\bar{\Omega}} f^-$. Moreover, the mass of ∇u and of w are the same. Indeed, one just needs to take $w = R_{\frac{\pi}{2}} \nabla u$, and w solves the following problem introduced by Beckmann in [3]

$$(1.4) \quad \inf \left\{ \|w\|(\bar{\Omega}) : w \in \mathcal{M}^2(\bar{\Omega}), \nabla \cdot w = f \right\},$$

where $\mathcal{M}^d(\bar{\Omega})$ is the space of finite vector measures on $\bar{\Omega}$ valued in \mathbb{R}^d (here, $d = 2$), $\|\cdot\|$ is the *rotated norm* of φ (i.e. $\|v\| := \varphi(R_{-\frac{\pi}{2}} v)$ for every $v \in \mathbb{R}^2$) and, $\|w\|$ denotes the variation measure associated with the vector measure w , i.e., $\|w\|(E) := \sup \{ \sum_i \|w(A_i)\| \}$, for every measurable set $E \subset \Omega$, where the supremum is taken over all partitions $E = \bigcup_i A_i$ into a countable number of disjoint measurable subsets. If we identify $f = \partial g / \partial \mathbf{t}$ with its restriction to the boundary, we can also write the condition $\nabla \cdot w = f$ as $\nabla \cdot w = 0$ in Ω , $w \cdot \mathbf{n} = f$ on $\partial\Omega$ (in general, when we write $\nabla \cdot w = f$ we mean $\int \nabla \phi \cdot dw = - \int \phi df$ for every smooth test function ϕ , without imposing ϕ to have compact support, i.e. we also include boundary conditions).

The study of the anisotropic least gradient problem can consequently be done by studying (1.4). The question whether (1.1) has a solution becomes whether the solution to (1.4) gives mass to the boundary or not.

The important point is that the Beckmann problem is strongly related to optimal transport theory, and is in some sense equivalent to the Monge problem

$$(1.5) \quad \min \left\{ \int_{\bar{\Omega} \times \bar{\Omega}} \|x - y\| d\gamma : \gamma \in \mathcal{M}^+(\bar{\Omega} \times \bar{\Omega}), (\Pi_x)_\# \gamma = f^+ \text{ and } (\Pi_y)_\# \gamma = f^- \right\},$$

where f^\pm represent the positive and negative parts of f , i.e. two positive measures with the same mass (see for instance [35, Chapter 4] for details about this equivalence in the Euclidean case, i.e. when $\|\cdot\| = |\cdot|$, or Section 2 for the case of a general strictly convex norm $\|\cdot\|$). The scalar measure $\sigma = \|w\|$ obtained from an optimal w is called *transport density* (see Section 2).

Because of its many connections to shape optimization problems ([5, 6]), traffic congestion ([8]), image processing ([33]), ..., many results are available in the literature about the transport density, and in particular its summability (even if most of these results concern only the Euclidean case). Of course, L^p summability of σ is equivalent to $W^{1,p}$ regularity of the optimal u .

In the Euclidean case, L^p summability results on the transport density have been analyzed in [16, 17, 18, 34]: in dimension d , for $p < d/(d-1)$, we have $w \in L^p$ as soon as at least one between f^+ or f^- is in L^p , while for general p (including $p = \infty$), this is proven when both f^+ and f^- belong to $L^p(\Omega)$. However, in the case of interest for applications to the least gradient problem, the measures f^\pm are singular (they are concentrated on the negligible set $\partial\Omega$). The only result obtained so far with measures concentrated on the boundary is the one that we presented in [19], where we considered the case where f^- is the projection of f^+ on $\partial\Omega$. On the other hand, this is far from the setting that we want to study now, since, first, in [19], only one of the two measures is on $\partial\Omega$, and second, it is not an arbitrary measure but it is chosen to be the projection of the other.

We can say that, so far, the L^p summability of w in the case where both the measures f^+ and f^- are concentrated on the boundary is unknown (a counterexample is presented in a particular case in [19]). In particular, we do not know whether the optimal flow w belongs or not to $L^p(\Omega)$ provided $f^\pm \in L^p(\partial\Omega)$. The goal of the present paper is exactly to investigate this kind of L^p summability results under suitable assumptions on the domain Ω , and then applying them to the $W^{1,p}$ regularity of the solution of (1.1).

The paper is organized as follows. In Section 2, we adapt some well-known facts concerning the Monge-Kantorovich problem to the precise setting of transport from the boundary to the boundary. In Section 3, we show positive results on the L^p summability of σ in the transport between measures on the boundary of a uniformly convex domain, for $d = 2$ in the case of a general strictly convex norm $\|\cdot\|$, or in arbitrary dimension $d \geq 2$ with $\|\cdot\| = |\cdot|$. In particular, we see that $f^\pm \in L^p(\partial\Omega) \Rightarrow \sigma \in L^p(\Omega)$ holds for $p \leq 2$ (to go beyond L^2 summability one needs extra regularity of the data) in the case $d = 2$ with a general norm $\|\cdot\|$, or for every $d \geq 2$ with $\|\cdot\| = |\cdot|$. Section 4 gives indeed a counter-example where $f \in L^\infty$ but $\sigma \notin L^p$ for any $p > 2$. Finally, Section 5 summarizes the applications, in the case $d = 2$, of these results to the anisotropic least gradient problem.

Many of the results that we recover were already known, at least for the case of the Euclidean norm, thanks to different methods, but we believe that the connection with optimal transport and the technique we develop are interesting in themselves. Moreover, the generalizations to the anisotropic case are non-trivial via standard methods, while they are essentially straightforward via the present approach. As a last interesting point, to the best of our knowledge, the following statement is novel even in the Euclidean case: if Ω is a uniformly convex domain in dimension 2, $p \leq 2$ and $g \in W^{1,p}(\partial\Omega)$, then the solution to (1.1) exists, is unique, and belongs to $W^{1,p}(\Omega)$ (this is our Theorem 5.3).

2. MONGE-KANTOROVICH AND BECKMANN PROBLEMS

Let $\|\cdot\|$ be an arbitrary strictly convex norm in \mathbb{R}^d . Given two finite positive Borel measures f^+ and f^- on a compact convex domain $\Omega \subset \mathbb{R}^d$ (the closure of a non-empty convex open bounded set; note that from now on, to enlighten the notation, we will call Ω the closed domain, and not the open one), satisfying the mass balance condition $f^+(\Omega) = f^-(\Omega)$, we consider the following minimization problem

$$(2.1) \quad \min \left\{ \int_{\Omega \times \Omega} \|x - y\| d\gamma : \gamma \in \Pi(f^+, f^-) \right\},$$

where

$$\Pi(f^+, f^-) := \left\{ \gamma \in \mathcal{M}^+(\Omega \times \Omega) : (\Pi_x)_\# \gamma = f^+, (\Pi_y)_\# \gamma = f^- \right\}.$$

This is a relaxation of the classical Monge optimal transportation problem [32], which is the

following

$$\inf \left\{ \int_{\Omega} \|x - T(x)\| \, d f^+ : T_{\#} f^+ = f^- \right\}.$$

Actually, these two problems are equivalent as soon as one can prove that there exists an optimal γ in (2.1) which is concentrated on the graph of a measurable map T , i.e. $\gamma = (id, T)_{\#} f^+$. This is the case whenever $f^+ \ll \mathcal{L}^d$: the existence of an optimal map T in this problem (or the fact that an optimal γ is of the form $(id, T)_{\#} f^+$) has been a matter of active study between the end of the '90s and the beginning of this century, and we cite in particular [1, 2, 20, 36, 10] for the case of the Euclidean norm. For different norms, see [11, 14] when the norm is strictly convex, and [15] for the general case. However, since we are interested in transport problems where f^+ is concentrated on the negligible set $\partial\Omega$, we will discuss later a specific technique in this particular case.

We underline that, by a suitable inf-sup exchange procedure, it is possible to see that the following maximization problem

$$(2.2) \quad \max \left\{ \int_{\Omega} \phi \, d(f^+ - f^-) : \|\nabla\phi\|_{\star, \infty} \leq 1 \right\},$$

is the dual of (2.1) (its value equals $\min(2.1)$), where

$$\|\nabla\phi\|_{\star, \infty} = \sup_{x \in \Omega} \|\nabla\phi(x)\|_{\star}$$

and

$$\|v\|_{\star} := \sup \left\{ v \cdot \xi : \|\xi\| \leq 1 \right\}, \text{ for every } v \in \mathbb{R}^d.$$

Note that, as soon as the domain Ω is convex, the condition $\|\nabla\phi\|_{\star, \infty} \leq 1$ is nothing but the fact that ϕ is 1-Lip with respect to $\|\cdot\|$. Moreover, the equality of the two optimal values implies that optimal γ and ϕ satisfy $\phi(x) - \phi(y) = \|x - y\|$ on the support of γ but also that, whenever we find some admissible γ and ϕ satisfying $\int_{\Omega \times \Omega} \|x - y\| \, d\gamma = \int_{\Omega} \phi \, d(f^+ - f^-)$, they are both optimal. The maximizers in (2.2) are called *Kantorovich potentials*. Given a maximizer ϕ , we call *transport ray* any maximal segment $[x, y]$ such that $\phi(x) - \phi(y) = \|x - y\|$, and the important fact is that whenever a point (x_0, y_0) belongs to the support of an optimal γ , then x_0 and y_0 must belong to a common transport ray. In addition, two different transport rays cannot intersect at an interior point of either of them (this strongly uses the strict convexity of the norm, see, for instance, [35, 10, 15]).

In optimal transport theory it is classical to associate with any optimal transport plan γ a positive measure σ_{γ} on Ω , called *transport density*, which represents the amount of transport taking place in each region of Ω . This measure σ_{γ} is defined via (this is in fact an adaptation of the definition of the transport density, see for instance [35], which is usually given in the Euclidean case)

$$(2.3) \quad \langle \sigma_{\gamma}, \phi \rangle = \int_{\Omega \times \Omega} d\gamma(x, y) \int_0^1 \phi(\omega_{x,y}(t)) \|\omega'_{x,y}(t)\| \, dt \quad \text{for all } \phi \in C(\Omega)$$

where $\omega_{x,y}$ is a parameterization of the geodesic connecting x to y (which is, thanks to the strict convexity of the norm $\|\cdot\|$, the straight line segment between x and y , and we take $\omega_{x,y}(t) = (1-t)x + ty$). Notice in particular that one can write, when the norm $\|\cdot\|$ is the

Euclidean norm $|\cdot|$,

$$(2.4) \quad \sigma_\gamma(A) = \int_{\Omega \times \Omega} \mathcal{H}^1([x, y] \cap A) d\gamma(x, y) \quad \text{for every Borel set } A$$

and similar formulas exist for other norms if one computes the \mathcal{H}^1 measure with respect to the induced distance. This means that, for a subregion A , $\sigma_\gamma(A)$ stands for “how much” the transport takes place in A , if particles move from their origin x to their destination y on straight lines. We can also define a vector version of σ_γ , as a vector measure w_γ on Ω defined by

$$(2.5) \quad \langle w_\gamma, \xi \rangle = \int_{\Omega \times \Omega} d\gamma(x, y) \int_0^1 \xi(\omega_{x,y}(t)) \cdot \omega'_{x,y}(t) dt \quad \text{for all } \xi \in C(\Omega, \mathbb{R}^d).$$

One can show that this vector measure w_γ solves the following problem (the so-called *continuous transportation model* proposed by Beckmann in [3]):

$$(2.6) \quad \min \left\{ \|w\|(\bar{\Omega}) : w \in \mathcal{M}^d(\bar{\Omega}), \nabla \cdot w = f \right\}.$$

Indeed, we can see easily that w_γ is admissible in (2.6) by using a gradient test function ξ . Moreover, we have $\|w_\gamma\| \leq \sigma_\gamma$, which implies $\|w_\gamma\|(\Omega) \leq \sigma_\gamma(\Omega) = \min(2.1) = \sup(2.2) \leq \min(2.6)$, where the last inequality follows from the fact that, if $\nabla \cdot w = f$ and $\|\nabla \phi\|_{*,\infty} \leq 1$, then we have $\int_\Omega \phi df = -\int_\Omega \nabla \phi \cdot dw \leq \|w\|(\Omega)$. This yields that w_γ is an optimal flow for (2.6).

As we said, most of the analysis of the Beckmann problem has been performed so far in the Euclidean case, and we will summarize here the main achievements. In this case, it is well-known that every solution of the problem (2.6) is of the form $w = w_\gamma$ for an optimal transport plan γ for (2.1) (see, for instance, [35, Chapter 4]). As in general Problem (2.1) can admit several different solutions, also (2.6) can have non-unique solutions. Yet, it is possible to prove that when either f^+ or f^- are absolutely continuous measures, then all different optimal transport plans γ induce the same vector measure w_γ . The following result also includes summability estimates:

Proposition 2.1. *Suppose that $\|\cdot\|$ is the Euclidean norm and $f^+ \ll \mathcal{L}^d$. Then, the transport density σ_γ is unique (i.e., it does not depend on the choice of the optimal transport plan γ) and $\sigma_\gamma \ll \mathcal{L}^d$. If $f^+ \in L^p(\Omega)$ with $p < d/(d-1)$, then $\sigma_\gamma \in L^p(\Omega)$. Moreover, for arbitrary $p \in [1, \infty]$, if both $f^+, f^- \in L^p(\Omega)$, then σ_γ also belongs to $L^p(\Omega)$.*

These properties are well-known in the literature, and we refer to [19], [21], [34], [18] and [34], as well as to [35, Chapter 4].

We want now to prove that, even in the case of a general strictly convex norm $\|\cdot\|$, any optimal flow for (2.6) comes from an optimal transport plan for (2.1). This requires to adapt the proofs which were done in the Euclidean case, and, following [35, Chapter 4], we first need to introduce some objects that generalize both σ_γ and w_γ . Let \mathcal{C} be the set of absolutely continuous curves $\omega : [0, 1] \mapsto \Omega$. We call *traffic plan* any positive measure Q on \mathcal{C} with total mass equal to $f^+(\Omega) = f^-(\Omega)$, and

$$\int_{\mathcal{C}} L(\omega) dQ(\omega) < +\infty,$$

where $L(\omega)$ is the length of the curve ω , i.e. $L(\omega) = \int_0^1 \|\omega'(t)\| dt$ (note that the length is

measured according to the norm $\|\cdot\|$. We define the *traffic intensity* $i_Q \in \mathcal{M}^+(\Omega)$ as follows

$$\int_{\Omega} \phi \, di_Q = \int_{\mathcal{C}} \left(\int_0^1 \phi(\omega(t)) \|\omega'(t)\| \, dt \right) dQ(\omega) \quad \text{for all } \phi \in C(\Omega).$$

This definition (which is taken from [12] and adapted to the case of a general norm) is a generalization of the notion of transport density σ_γ (2.3). The interpretation is - again - the following: for a subregion A , $i_Q(A)$ represents the total cumulated traffic in A induced by Q , i.e., for every path we compute “how long” it stays in A , and then we average on paths. We also associate a vector measure w_Q (called *traffic flow*) with any traffic plan Q via

$$\int_{\Omega} \xi \cdot dw_Q = \int_{\mathcal{C}} \left(\int_0^1 \xi(\omega(t)) \cdot \omega'(t) \, dt \right) dQ(\omega) \quad \text{for all } \xi \in C(\Omega, \mathbb{R}^d).$$

Taking a gradient field $\xi = \nabla\phi$ in the previous definition yields

$$\int_{\Omega} \nabla\phi \cdot dw_Q = \int_{\mathcal{C}} (\phi(\omega(1)) - \phi(\omega(0))) \, dQ(\omega) = \int_{\Omega} \phi \, d((e_1)_{\#}Q - (e_0)_{\#}Q),$$

where e_t is the evaluation map at time t , i.e. $e_t(\omega) := \omega(t)$, for all $\omega \in \mathcal{C}$, $t \in [0, 1]$. From now on, we will restrict our attention to admissible traffic plans Q , i.e. traffic plans such that $(e_0)_{\#}Q = f^+$ and $(e_1)_{\#}Q = f^-$, since, in this case, w_Q will be an admissible flow in (2.6), i.e. one has

$$\nabla \cdot w_Q = f^+ - f^-.$$

Lemma 2.2. *Let w be a flow such that $\nabla \cdot w = f^+ - f^-$. Then, there is an admissible traffic plan Q such that $\|w - w_Q\|(\Omega) + i_Q(\Omega) = \|w\|(\Omega)$.*

Proof. The result is just a variant of what is presented in [35, Section 4.2.3], and the proof will also follow the same lines. Following [35, Section 4.2.3], we first assume that f^+ , f^- and w are smooth with $f^+, f^- > 0$, and obtain existence of an admissible traffic plan Q with $w_Q = w$ and $i_Q = \|w\|$ via a Dacorogna-Moser construction.

For the general case: following again [35, Theorem 4.10], convolve w (resp. f^+ and f^-) with a Gaussian kernel η_ε and take care of the boundary behavior (for more details, see [35, Lemma 4.8]); we obtain smooth vector fields w_ε and strictly positive smooth densities f_ε^\pm with $\nabla \cdot w_\varepsilon = f_\varepsilon^+ - f_\varepsilon^-$ such that $w_\varepsilon \rightharpoonup w$ (but also $\|w_\varepsilon\| \rightharpoonup \|w\|$ because of standard properties of convolutions) and $f_\varepsilon^\pm \rightharpoonup f^\pm$. Let $(Q_\varepsilon)_\varepsilon$ be the sequence of traffic plans such that, for every $\varepsilon > 0$, $w_{Q_\varepsilon} = w_\varepsilon$ and $i_{Q_\varepsilon} = \|w_\varepsilon\|$. The measures Q_ε were constructed so that $(e_0)_{\#}Q_\varepsilon = f_\varepsilon^+$ and $(e_1)_{\#}Q_\varepsilon = f_\varepsilon^-$, which implies, at the limit, that $Q_\varepsilon \rightharpoonup Q$ (since $\int_{\mathcal{C}} L(\omega) \, dQ_\varepsilon(\omega) = \|w_\varepsilon\|(\Omega) \leq C$ and so Q_ε is tight) with $(e_0)_{\#}Q = f^+$ and $(e_1)_{\#}Q = f^-$. Moreover, it is not difficult to check that Proposition 4.7 in [35] is still true if we replace the Euclidean norm $|\cdot|$ by an arbitrary norm $\|\cdot\|$. In particular, we have

$$\begin{aligned} \int_{\Omega} \xi \cdot dw &= \lim_{\varepsilon} \int_{\Omega} \xi \cdot dw_\varepsilon = \lim_{\varepsilon} \int_{\mathcal{C}} \left(\int_0^1 \xi(\omega(t)) \cdot \omega'(t) \, dt \right) dQ_\varepsilon(\omega) \\ &\geq \int_{\Omega} \xi \cdot dw_Q + \|\xi\|_{\star, \infty} (i_Q(\Omega) - \|w\|(\Omega)), \end{aligned}$$

for all $\xi \in C(\Omega, \mathbb{R}^d)$. Hence, $\|w - w_Q\|(\Omega) + i_Q(\Omega) \leq \|w\|(\Omega)$. Yet, the other inequality is always true since $\|w_Q\| \leq i_Q$. Then, we get that $\|w - w_Q\|(\Omega) + i_Q(\Omega) = \|w\|(\Omega)$. \square

Proposition 2.3. *Let w be an optimal flow for (2.6), then there is an optimal transport plan γ for (2.1) such that $w = w_\gamma$.*

Proof. From Lemma 2.2, there is an admissible traffic plan Q such that $\|w - w_Q\|(\Omega) + i_Q(\Omega) = \|w\|(\Omega)$. The optimality of the flow w and the fact that $\|w_Q\| \leq i_Q$ imply that $w_Q = w$ and $i_Q = \|w\|$. Hence,

$$\|w\|(\Omega) = i_Q(\Omega) = \int_{\mathcal{C}} L(\omega) dQ(\omega) \geq \int_{\mathcal{C}} \|\omega(0) - \omega(1)\| dQ(\omega) = \int_{\Omega \times \Omega} \|x - y\| d(e_0, e_1)_{\#} Q \geq \min(2.1).$$

Yet, the equalities $\|w\|(\Omega) = \min(2.6) = \min(2.1)$ imply that the above inequalities are actually equalities. This means that Q must be concentrated on segments (thanks to the strict convexity of the norm $\|\cdot\|$). Also, the measure $\gamma = (e_0, e_1)_{\#} Q$, which belongs to $\Pi(f^+, f^-)$, must be optimal in (2.1) and, we have $w = w_Q = w_{\gamma}$. \square

We also introduce an easy stability result that we will need later on.

Proposition 2.4. *Suppose $f^+ \in \mathcal{M}^+(\Omega)$ is fixed and $f_n^- \rightharpoonup f^-$. Let γ_n be an optimal transport plan between f^+ and f_n^- . Then, up to a subsequence, $\gamma_n \rightharpoonup \gamma$, where γ is an optimal transport plan between f^+ and f^- . Moreover, if all the plans γ_n are induced by transport maps T_n and γ is induced by a map T , then we have $T_n \rightarrow T$ in $L^2(f^+)$.*

Proof. Firstly, we see easily that, up to a subsequence, γ_n admits a weak limit γ in the sense of measures. The condition $\gamma_n \in \Pi(f^+, f_n^-)$ passes to the limit, thus giving $\gamma \in \Pi(f^+, f^-)$. Moreover, for each n , there is a corresponding Kantorovich potential ϕ_n , which is 1-Lip according to $\|\cdot\|$ with $\min \phi_n = 0$ (this may be done if we note that adding a constant to ϕ_n is always possible since $\int_{\Omega} f^+ = \int_{\Omega} f_n^-$), such that

$$\int_{\Omega} \phi_n d(f^+ - f_n^-) = \int_{\Omega \times \Omega} \|x - y\| d\gamma_n.$$

Up to a subsequence, we can suppose $\phi_n \rightarrow \phi$ uniformly in Ω , where ϕ is also a 1-Lip function with respect to $\|\cdot\|$. Then, passing to the limit in the above equality, we get

$$\int_{\Omega} \phi d(f^+ - f^-) = \int_{\Omega \times \Omega} \|x - y\| d\gamma,$$

which is sufficient to infer that γ is actually an optimal transport plan between f^+ and f^- , and ϕ is the corresponding Kantorovich potential.

The last part of the statement, when plans are induced by maps, can be deduced by the weak convergence of the plans. Using $\gamma_n = (id, T_n)_{\#} f^+$ and $\gamma_n \rightharpoonup \gamma := (id, T)_{\#} f^+$ and testing the weak convergence against the test function $\phi(x, y) = \xi(x) \cdot y$ we obtain

$$\int \xi(x) \cdot T_n(x) df^+(x) \rightarrow \int \xi(x) \cdot T(x) df^+(x),$$

which means that we have the weak convergence $T_n \rightharpoonup T$ in $L^2(f^+)$. We can now test against $\phi(x, y) = |y|^2$ and obtain

$$\int |T_n(x)|^2 df^+(x) \rightarrow \int |T(x)|^2 df^+(x),$$

which proves the convergence of the L^2 norm. This gives strong convergence in $L^2(f^+)$. \square

We have now to consider the case where the measures f^+ and f^- are concentrated on the boundary. As we said, the theory of existence of optimal maps, even for general norms, is now well-developed, but requires at least f^+ to be absolutely continuous. Hence, it cannot be applied here. Moreover, uniqueness of the optimal map is in general not guaranteed.

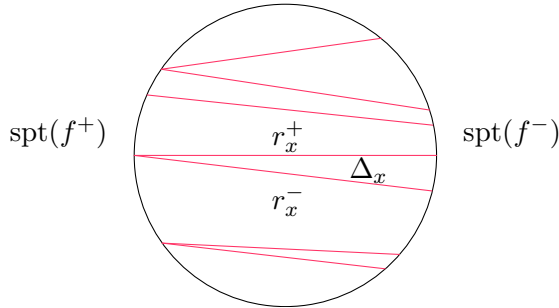


FIGURE 1

Surprisingly, the case of measures concentrated on the boundary of a strictly convex domain turns out to be easier. We can indeed prove in some cases that any optimal γ between measures on the boundary is induced by a transport map, which also implies uniqueness of γ and of σ_γ . This deserves to be compared to the results in [22], where the authors prove existence of a unique optimal transport plan γ (which is also induced by a map T) for (1.5) between two measures f^+ and f^- supported on the boundaries of two domains $\Omega, \Omega' \subset \mathbb{R}^d$ with f^+ absolutely continuous with respect to the surface measure $\mathcal{H}^{d-1} \llcorner \partial\Omega$ if Ω is strictly convex; yet, their result only concerns the quadratic case (i.e., when the transport cost is $c(x, y) = |x - y|^2$). Moreover, the authors of [31] even studied the Hölder continuity of this unique optimal map T between arbitrary measures supported on Euclidean spheres.

In our case we need to study the cost induced by a norm (and not the square of a norm), and we first start from the case $d = 2$ (with a general strictly convex norm $\|\cdot\|$), which is easier to deal with.

From now on we will suppose the condition that f^+ and f^- have no common mass, which means that there exist two disjoint sets A^+ and A^- contained in $\partial\Omega$ with f^\pm concentrated on A^\pm (beware that these sets are not necessarily the two supports of f^+ and f^-).

Proposition 2.5. *Suppose that Ω is strictly convex, and $d = 2$. Then, if f^+ is atomless (i.e., $f^+(\{x\}) = 0$ for every $x \in \partial\Omega$) and f^+ and f^- have no common mass, there is a unique optimal transport plan γ for (2.1), between f^+ and f^- , and it is induced by a map T .*

Proof. Let γ be an optimal transport plan between f^+ and f^- . Let \mathcal{D} be the set of double points, that is those points whose belong to several transport rays. Take $x \in \mathcal{D}$ and let r_x^\pm be two different transport rays starting from x . Let $\Delta_x \subset \Omega$ be the region delimited by r_x^+, r_x^- and $\partial\Omega$ (see Figure 1). As Ω is strictly convex, then we see easily that $|\Delta_x| > 0$ and the interior parts of all these sets $\Delta_x, x \in \mathcal{D}$, are disjoint (thanks also to the strict convexity of the norm $\|\cdot\|$). This implies that the set \mathcal{D} is at most countable and so $f^+(\mathcal{D}) = 0$ as f^+ is atomless. On the other hand, for every $x \in A^+ \setminus \mathcal{D}$ there is a unique transport ray r_x starting from x , and this ray r_x intersects A^- in - at most - one point, which will be denoted by $T(x)$. We note that this map T is also measurable since its graph is measurable (see, for instance, [9]). Moreover, we have $\gamma = (id, T)_\# f^+$, which is equivalent to saying that γ is, in fact, induced by a map T . The uniqueness follows in the usual way: if two plans γ and γ' optimize (2.1), the same should be true for $(\gamma + \gamma')/2$. Yet, for this measure to be induced by a map, it is necessary to have $\gamma = \gamma'$. \square

The higher-dimensional counterpart of the above result should replace the assumption that f^+ is atomless with the assumption that f^+ gives no mass to $(d - 2)$ -dimensional sets (i.e. sets

of codimension 1 within the boundary). Yet, this seems more complicated to prove, and we will just stick to an easier result, in the case where f^+ is absolutely continuous w.r.t. to the \mathcal{H}^{d-1} measure on $\partial\Omega$ (that we simply write $f^+ \in L^1(\partial\Omega)$). Unfortunately, the easy proof that we provide here below only works when the norm $\|\cdot\|$ is the Euclidean norm $|\cdot|$.

Proposition 2.6. *Suppose that Ω is strictly convex and $d \geq 2$. Then, if $f^+ \in L^1(\partial\Omega)$ and f^- and f^+ have no common mass, there is a unique optimal transport plan γ for (2.1) with the Euclidean cost $|x - y|$, and it is induced by a map T .*

Proof. Let γ be an optimal transport plan between f^+ and f^- . According to the strategy above, it is enough to prove that for f^+ -a.e. $x \in A^+$ there is at most a unique point $y \in A^-$ such that $(x, y) \in \text{spt } \gamma$. We will parametrize A^\pm via variables $s^\pm \in \mathbb{R}^{d-1}$. To do this, we first remove from A^+ and A^- a common point x_0 , which is always possible since f^+ and f^- have no atoms, hence removing a point does not change the measure. Then, $\partial\Omega \setminus \{x_0\}$, as a proper subset of the boundary, is homeomorphic to a subset of \mathbb{R}^{d-1} , via an homeomorphism Φ which can also be chosen to be locally bi-Lipschitz. Up to removing a negligible set, we can also assume that Φ is differentiable everywhere, hence we will call this map diffeomorphism by abuse of terminology. Under this parameterization, we face a new transport problem in \mathbb{R}^{d-1} , with a new cost function $c(s^+, s^-) := |\Phi(s^+) - \Phi(s^-)|$.

Using standard arguments from optimal transport theory (see [35, Chapter 1]) one can see that the Kantorovich potentials in this new transport problem are locally Lipschitz continuous, and hence differentiable a.e. This can be seen either by composing the Kantorovich potential for the Euclidean cost $|x - y|$ with the parameterization Φ , or by using a c -transform formula

$$\phi^c(s^+) = \inf_{s^-} c(s^+, s^-) - \phi(s^-) = \inf_y |\Phi(s^+) - y| - \tilde{\phi}(y),$$

which shows that the regularity of any c -transform in the s^+ variable is given by the composition of a 1-Lipschitz function with the parameterization Φ , which provides local Lipschitz continuity.

As a consequence of this, it is enough to check that c satisfies the twist condition to prove that γ is necessarily induced by a map T , and that it is unique. Computing the gradient of c w.r.t. the variable s^+ one gets

$$\nabla_{s^+} c(s^+, s^-) = \frac{\Phi(s^+) - \Phi(s^-)}{|\Phi(s^+) - \Phi(s^-)|} D\Phi(s^+),$$

where $D\Phi(s^+)$ is the Jacobian matrix of the diffeomorphism Φ . We need to prove that this expression is injective in s^- . Having two different values of s^- (say, s_0^- and s_1^-) where these expressions coincide means, using that $s^+ \mapsto \Phi(s^+)$ is a diffeomorphism, that the two unit vectors $\Phi(s^+) - \Phi(s_i^-) / |\Phi(s^+) - \Phi(s_i^-)|$ have the same projection onto the tangent space to $\partial\Omega$ at $\Phi(s^+)$ (note that, from $A^+ \cap A^- = \emptyset$, we can assume $\Phi(s^+) \neq \Phi(s_i^-)$). Since they are unit vectors, and they both point to the interior of Ω , which is convex, then they should fully coincide. But this means that the direction connecting $\Phi(s^+)$ to the points $\Phi(s_i^-)$ is the same, and since all these points lie on the boundary of a strictly convex domain, we have $\Phi(s_0^-) = \Phi(s_1^-)$. \square

For the sake of the next section, we also want stability results on the transport density. Suppose that f^+ and f^- are fixed, and that a unique optimal transport plan γ exists in the transportation from f^+ to f^- . In this case we will directly write σ instead of σ_γ , if no ambiguity arises. Given the optimal transport plan γ , let us define the measure f_t via

$$(2.7) \quad f_t = (\Pi_t)_\#(|x - y| \cdot \gamma)$$

where $\Pi_t(x, y) := (1 - t)x + ty$. From (2.3), the transport density σ may be easily written as

$$\sigma = \int_0^1 f_t dt.$$

We also define a sort of partial transport density that will be useful in the sequel: given $\tau \leq 1$, set

$$(2.8) \quad \sigma^{(\tau)} = \int_0^\tau f_t dt.$$

Note that $\sigma^{(\tau)}$ really depends on γ , i.e., differently from σ , it is not in general true that different optimal plans γ induce the same $\sigma^{(\tau)}$. On the other hand, we will only use this partial transport density in cases where the optimal γ is unique. In this case we can also obtain:

Proposition 2.7. *Suppose $f^+ \in \mathcal{M}^+(\Omega)$ is fixed and $f_n^- \rightharpoonup f^-$. Let γ_n be an optimal transport plan between f^+ and f_n^- and suppose that there is a unique optimal transport plan between f^+ and f^- . Fix $\tau \leq 1$ and define $\sigma_n^{(\tau)}$ according to (2.7) and (2.8) using γ_n , and $\sigma^{(\tau)}$ using γ . Then, we have $\sigma_n^{(\tau)} \rightharpoonup \sigma^{(\tau)}$.*

Proof. This is a simple consequence of Proposition 2.4, of the continuity of the function $(x, y) \mapsto \|x - y\|$, and of the uniqueness of the optimal γ . \square

3. L^p SUMMABILITY OF BOUNDARY-TO-BOUNDARY TRANSPORT DENSITIES

In all that follows, Ω is a compact and uniformly convex domain in \mathbb{R}^d , f^+ and f^- are two positive Borel measures concentrated on the boundary, and at least one of them will belong to $L^1(\partial\Omega)$. Since we are only interested in the transport density between these two measures, we can always assume that they have no common mass, as the transport density only depends on the difference $f^+ - f^-$ and common mass can be subtracted to both of them.

Uniform convexity of Ω is crucial to obtain our result, so we need to be precise about its definition, in particular since we do not restrict our analysis to smooth domains Ω . We will consider the following assumption on Ω : there exists $R < \infty$ such that, for every $x \in \partial\Omega$ and every unit vector $-\mathbf{n}$ in the exterior normal cone to Ω at x we have $\Omega \subset B(z, R)$ with $z = x + R\mathbf{n}$. For smooth domains, this condition is equivalent to the fact that all principal curvatures of $\partial\Omega$ are larger than $\kappa := R^{-1}$. Moreover, we note that this condition implies, for every $x, y \in \partial\Omega$, the inequality $(y - x) \cdot \mathbf{n} \geq \frac{1}{2R}|y - x|^2$, which we will use. Indeed, we can write $|y - (x + R\mathbf{n})|^2 \leq R^2$ which, expanded, gives $|y - x|^2 - 2R(y - x) \cdot \mathbf{n} \leq 0$.

In the case $d = 2$, by Proposition 2.5, there will exist one unique optimal transport plan between these two measures, while Proposition 2.6 gives the existence of a unique optimal transport plan between them in arbitrary dimension $d \geq 2$, but under the assumption that the norm $\|\cdot\|$ is the Euclidean one $|\cdot|$. For simplicity of exposition, we say that the assumption **(UA)** (standing for ‘‘Uniqueness Assumption’’) holds if

$$(UA) \quad \text{either } d = 2 \text{ or } d \geq 2 \text{ and } \|\cdot\| \text{ is the Euclidean norm.}$$

We will make use of the transport density σ and of $\sigma^{(\tau)}$, defined in (2.8) and provide estimate on them. The main point is the following estimate.

Proposition 3.1. *Suppose that the domain $\Omega \subset \mathbb{R}^d$ is uniformly convex, take $p > 1$ and $f^+ \in L^p(\partial\Omega)$. If $p > 2$ also suppose $\int f^+(x)^p d(x, \text{spt}(f^-))^{2-p} d\mathcal{H}^{d-1}(x) < +\infty$. If **(UA)** holds, then there exists a constant $C = C(\kappa, \text{diam}(\Omega))$ such that we have*

$$\int_{\Omega} |\sigma^{(\tau)}|^p dx \leq C \left(\int_0^\tau \frac{1}{(1-t)^{(d-1)(p-1)}} dt \right) \int_{\partial\Omega} f^+(x)^p D(x)^{2-p} d\mathcal{H}^{d-1}(x),$$

where $D(x) := |x - T(x)|$ is the distance between each point $x \in \partial\Omega$ and its image $T(x)$ in the optimal transport map which induces the optimal plan γ .

Proof. Following the same strategy as in [34], we first assume that the target measure f^- is finitely atomic (the points $(x_j)_{j=1,\dots,m}$ being its atoms). Let T be the optimal transport map from f^+ onto f^- . For all $j \in \{1, \dots, m\}$, consider $T^{-1}(\{x_j\}) \subset \partial\Omega$; we can partition it in finitely many smaller parts, so that each of them can be represented in a single smooth (Lipschitz) chart parameterizing a part of $\partial\Omega$. We will call $(\chi_i)_{i=1,\dots,n}$ these parts. Let us call Ω_i the union of all transport rays starting from points in χ_i , all these rays pointing to a common point $x_{j(i)}$ (but we will write x_i for simplicity). Call $\Omega_i^{(\tau)}$ the set of points of the form $(1-t)x + tx_i$, with $x \in \chi_i$ and $t \leq \tau$. The sets Ω_i (and hence also $\Omega_i^{(\tau)}$) are essentially disjoint (the mutual intersections between them are Lebesgue-negligible).

Set $\sigma_i^{(\tau)} := \sigma^{(\tau)} \llcorner \Omega_i$, for every $i \in \{1, \dots, n\}$. Of course, $\sigma_i^{(\tau)}$ is concentrated on $\Omega_i^{(\tau)}$. In order to get L^p estimates on $\sigma^{(\tau)}$, we want to give an explicit formula for each $\sigma_i^{(\tau)}$. Fix $i \in \{1, \dots, n\}$ and let α_i be a regular function such that, up to choosing a suitable system of coordinates, χ_i is contained in the graph of $s \mapsto \alpha_i(s)$, with $s \in \tilde{\chi}_i \subset \mathbb{R}^{d-1}$ (hence, the sets $\tilde{\chi}_i$ are the $(d-1)$ -dimensional domains where the charts are defined). For every $y \in \Omega_i^{(\tau)}$, there exist a unique point $x = (s, \alpha_i(s)) \in \chi_i$ and $t \in [0, \tau]$ such that

$$y := (y', y_d) = (1-t)x + tx_i = ((1-t)s + tx'_i, (1-t)\alpha_i(s) + tx_{i,d}),$$

where we write $x_i := (x'_i, x_{i,d})$ by separating the last (vertical) coordinate from the others. For all $\varphi \in C(\Omega_i)$, we get

$$\begin{aligned} \int_{\Omega_i} \varphi(y) d\sigma_i^{(\tau)}(y) &= \int_{\chi_i} \int_0^\tau \varphi((1-t)x + tx_i) \|x - x_i\| dt df^+(x) \\ &= \int_{\Omega_i^{(\tau)}} \varphi(y) \frac{\|(s, \alpha_i(s)) - x_i\| f^+(s, \alpha_i(s)) \sqrt{1 + |\nabla \alpha_i(s)|^2}}{J_i(t, s)} dy, \end{aligned}$$

where $J_i(t, s) := |\det(D_{(s,t)}(y', y_d))|$. Hence, we get

$$(3.1) \quad \sigma^{(\tau)}(y) = \frac{\|(s, \alpha_i(s)) - x_i\| f^+(s, \alpha_i(s)) \sqrt{1 + |\nabla \alpha_i(s)|^2}}{J_i(t, s)}, \quad \text{for a.e. } y \in \Omega_i^{(\tau)}.$$

We then have

$$\begin{aligned} \|\sigma^{(\tau)}\|_{L^p(\Omega)}^p &= \sum_{i=1}^n \int_{\tilde{\chi}_i} \int_0^\tau \sigma^{(\tau)}((1-t)x + tx_i)^p J_i(t, s) dt ds \\ &= \sum_{i=1}^n \int_{\chi_i} \int_0^\tau \frac{\|x - x_i\|^p f^+(x)^p (1 + |\nabla \alpha_i(s)|^2)^{\frac{p-1}{2}}}{J_i(t, s)^{p-1}} dt d\mathcal{H}^{d-1}(x). \end{aligned}$$

Compute

$$D_{(s,t)}(y', y_d) = \begin{pmatrix} (1-t)\mathbf{I} & x'_i - s \\ (1-t)\nabla \alpha_i(s) & x_{i,d} - \alpha_i(s) \end{pmatrix},$$

where \mathbf{I} is the $(d-1) \times (d-1)$ identity matrix. Fix $x \in \chi_i$ a point where the normal vector

to $\partial\Omega$ is well-defined (the boundary is not supposed to be smooth, but it is locally a Lipschitz graph, hence \mathcal{H}^{d-1} -a.e. point is a differentiability point). It is easy to see that the quantity

$$\frac{J_i(t, s)}{\sqrt{1 + |\nabla\alpha_i(s)|^2}}$$

should be an intrinsic geometric quantity, independent of the parameterization. Hence, choosing a coordinate system in such a way that the vertical axis is parallel to the normal vector at a point $x = (s, \alpha_i(s))$, we can compute the above ratio as if we had $\nabla\alpha_i(s) = 0$. This gives as a result $J_i(t, s) = (1-t)^{d-1}(x_{i,d} - \alpha_i(s))$ and $(1 + |\nabla\alpha_i(s)|^2)^{1/2} = 1$. Anyway, it is also possible to explicitly compute the Jacobian, thus getting $J_i(t, s) = (1-t)^{d-1}(x_i - x) \cdot \mathbf{n}(x) \sqrt{1 + |\nabla\alpha_i(s)|^2}$, where $x = (s, \alpha_i(s))$ and $\mathbf{n}(x)$ is the unit inward normal vector to the graph of α_i at x (i.e. we have $(-\nabla\alpha_i(s), 1) = \mathbf{n}(x) \sqrt{1 + |\nabla\alpha_i(s)|^2}$). This allows to write the change-of-variable coefficients in an intrinsic way, and thus obtain

$$\|\sigma^{(\tau)}\|_{L^p(\Omega)}^p = \sum_{i=1}^n \int_{\chi_i} \int_0^\tau \frac{\|x - x_i\|^p f^+(x)^p}{(1-t)^{(d-1)(p-1)} ((x_i - x) \cdot \mathbf{n}(x))^{p-1}} dt d\mathcal{H}^{d-1}(x),$$

where $\mathbf{n}(x)$ is the inward normal vector to $\partial\Omega$ at x . Uniform convexity of Ω yields

$$(x_i - x) \cdot \mathbf{n}(x) \geq c|x - x_i|^2.$$

Using then $x_i = T(x)$ for $x \in \chi_i$ and the equivalence between the two norms $\|\cdot\|$ and $|\cdot|$, this provides the desired formula

$$\|\sigma^{(\tau)}\|_{L^p(\Omega)}^p \leq C \int_{\partial\Omega} \int_0^\tau \frac{|x - T(x)|^{2-p} f^+(x)^p}{(1-t)^{(d-1)(p-1)}} dt d\mathcal{H}^{d-1}(x).$$

This proves the claim when f^- is atomic. If not, take a sequence $(f_n^-)_n$ of atomic measures converging to f^- and concentrated on $\text{spt}(f^-)$. Call T_n the optimal maps from f^+ to f_n^- and $D_n(x) = |x - T_n(x)|$. By Proposition 2.7, the partial transport densities $\sigma_n^{(\tau)}$ converge to the corresponding partial transport density $\sigma^{(\tau)}$ and by Proposition 2.4 the optimal transport maps T_n also converge a.e. to the optimal transport map T inducing γ (up to extracting a subsequence, since $L^2(f^+)$ convergence implies a.e. convergence up to a subsequence). Moreover, we have $\int_{\partial\Omega} D_n(x)^{2-p} f^+(x)^p d\mathcal{H}^{d-1}(x) \rightarrow \int_{\partial\Omega} D(x)^{2-p} f^+(x)^p d\mathcal{H}^{d-1}(x)$ by dominated convergence, using either $p \leq 2$ and $f^+ \in L^p(\partial\Omega)$ or $\int_{\partial\Omega} d(x, \text{spt}(f^-))^{2-p} f^+(x)^p d\mathcal{H}^{d-1}(x) < +\infty$, according to our assumptions. Using semicontinuity on the left hand side, we get

$$\|\sigma^{(\tau)}\|_{L^p(\Omega)}^p \leq \liminf_n \|\sigma_n^{(\tau)}\|_{L^p(\Omega)}^p \leq C \left(\int_0^\tau \frac{1}{(1-t)^{(d-1)(p-1)}} dt \right) \int_{\partial\Omega} D(x)^{2-p} f^+(x)^p d\mathcal{H}^{d-1}(x)$$

and the result is proven in general. \square

From the estimate in Proposition 3.1 we can deduce many integrability results.

Proposition 3.2. *Let Ω be a uniformly convex domain in \mathbb{R}^d and suppose $f^+ \in L^p(\partial\Omega)$ with $p < d/(d-1)$. If (UA) holds, then the transport density σ between f^+ and any $f^- \in \mathcal{M}^+(\partial\Omega)$ is in $L^p(\Omega)$.*

Proof. Note that our assumption on p implies $p \leq 2$. To prove this result it is enough to use Proposition 3.1 with $\tau = 1$, since in this case the integral $\int_0^1 \frac{1}{(1-t)^{(d-1)(p-1)}} dt$ converges, and the term $D(x)^{2-p}$ is bounded since $p \leq 2$. \square

Proposition 3.3. *Let Ω be a uniformly convex domain in \mathbb{R}^d and suppose that $f^+, f^- \in L^p(\partial\Omega)$ with $p \leq 2$. If **(UA)** holds, then the transport density σ between these two measures is in $L^p(\Omega)$.*

Proof. In this case the integral in the estimate of $\sigma = \sigma^{(\tau)}$ with $\tau = 1$ can diverge, so we need to adapt our strategy. Following again [34], we write

$$\sigma = \sigma^+ + \sigma^-,$$

where $\sigma^+ = \sigma^{(1/2)}$ and $\sigma^- = \sigma - \sigma^{(1/2)}$. In this case the L^p summability of f^+ guarantees that of σ^+ since $p \leq 2$ implies that $D(x)^{2-p}$ is bounded. Symmetrically, the L^p summability of f^- guarantees that of σ^- .

Note that, thanks to Propositions 2.5 and 2.6, we do not face the same difficulties as in [34], where it was not obvious to glue together estimates on σ^+ obtained by approximating f^- and estimates on σ^- coming from the approximation of f^+ . \square

We will see in Section 4 that the same result is false for $p > 2$, and that in order to obtain higher integrability we need to assume much more on f^+ and f^- .

Remark 3.4. We do not discuss it here in details, but the summability result also works for Orlicz spaces with growth less than quadratic, i.e. we have, for every convex and superlinear function $\Psi = \mathbb{R}_+ \rightarrow \mathbb{R}_+$ with $\Psi(s) \leq C(s^2 + 1)$,

$$\int_{\Omega} \Psi(\sigma(x)) dx \leq C \int_{\partial\Omega} \Psi(|f(x)|) d\mathcal{H}^{d-1}(x) + C.$$

This can be proven in similar ways with suitable manipulations on the function Ψ . In particular, this implies that $f \in L^1(\partial\Omega) \Rightarrow \sigma \in L^1(\Omega)$.

Proposition 3.5. *Suppose that $f^+, f^- \in C^{0,\alpha}(\partial\Omega)$ for $0 < \alpha \leq 1$. If **(UA)** holds, then the transport density σ between these two measures is in $L^p(\Omega)$ for $p = 2/(1 - \alpha)$ (with $p = \infty$ for $\alpha = 1$).*

Proof. First, we check that we can apply Proposition 3.1, since in this case we need to use $p = 2/(1 - \alpha) > 2$. Consider a point x with $f^+(x) > 0$, and take a point $y \in \text{spt}(f^-)$ with $|x - y| = d(x, \text{spt}(f^-))$. Then we have $f^+(y) = 0$ (since f^+ and f^- have no mass in common) and $f^+(x) = |f^+(x) - f^+(y)| \leq C|x - y|^\alpha$. This provides $d(x, \text{spt}(f^-))^{2-p} f^+(x)^p \leq C d(x, \text{spt}(f^-))^{2-p+p\alpha}$. With our choice of p , this quantity is bounded since the exponent is non-negative (for $\alpha < 1$ the choice $p = 2/(1 - \alpha)$ provides a zero exponent; for $\alpha = 1$ this exponent is equal to 2 for any p). This in particular guarantees $\int_{\partial\Omega} d(x, \text{spt}(f^-))^{2-p} f^+(x)^p d\mathcal{H}^{d-1}(x) < +\infty$. Of course, the same can be performed on f^- . Then, the same strategy as in Proposition 3.3 shows

$$\|\sigma\|_{L^p(\Omega)}^p \leq C \left(\int_{\partial\Omega} D(x)^{2-p} f^+(x)^p d\mathcal{H}^{d-1}(x) + \int_{\partial\Omega} D^-(x)^{2-p} f^-(x)^p d\mathcal{H}^{d-1}(x) \right),$$

where $D^-(x) := |x - T^{-1}(x)|$ is defined as $D(x)$, but relatively to f^- . Using $D(x) = |x - T(x)| \geq d(x, \text{spt}(f^-))$ and the fact that the exponent $2 - p$ is negative, the quantity $|D(x)|^{2-p} f^+(x)^p \leq C D(x)^{2-p+p\alpha}$ is bounded. Since a similar argument can be performed on f^- , we obtain finiteness of the norm $\|\sigma\|_{L^p(\Omega)}$ (and for $\alpha = 1$ we obtain $\sigma \in L^\infty$ by passing to the limit $p \rightarrow \infty$). \square

4. COUNTER-EXAMPLE TO THE $L^{2+\varepsilon}$ SUMMABILITY

In this section, we show that the L^p estimates for the transport density, in the case where $p > 2$, fail even if we assume $f^\pm \in L^\infty(\partial\Omega)$. More precisely, we will construct an example of f^\pm , where $f^\pm \in L^\infty(\partial\Omega)$, but the transport density σ between them does not belong to $L^{2+\varepsilon}(\Omega)$ for any $\varepsilon > 0$. For simplicity, this will be done in dimension $d = 2$ and the norm $\|\cdot\|$ will be the Euclidean one.

Let Ω be a disk and let $(\chi_n^\pm)_n$ be a sequence of arcs in $\partial\Omega$ such that $\mathcal{H}^1(\chi_n^\pm) = \varepsilon_n$, for some sequence ε_n to be chosen later. We will put these arcs one after the other, so that they only have endpoints in common, and we assume that they are ordered in the following way: χ_{n-1}^+ , χ_{n-1}^- , χ_n^- , χ_n^+ , χ_{n+1}^+ , χ_{n+1}^- , for all n (see Figure 2).

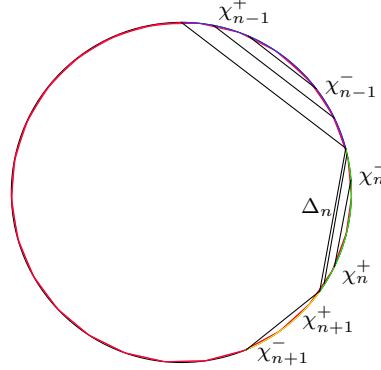


FIGURE 2

Set $f^\pm = \mathbb{1}_{\chi^\pm}$, where $\chi^\pm = \cup_n \chi_n^\pm$, and let T be the optimal transport map from f^+ to f^- . Correspondingly, let σ be the transport density. We see easily that the restriction of T to χ_n^+ is the optimal transport map T_n between f_n^+ and f_n^- , where f_n^\pm is the restriction of f^\pm to χ_n^\pm . Moreover, if we denote by Δ_n the union of all transport rays from f_n^+ onto f_n^- , then the restriction of the transport density σ to Δ_n is the transport density σ_n between f_n^+ and f_n^- . We want to compute this density σ_n . Let $s \mapsto \alpha_n(s)$ be a parameterization of $\chi_n^+ \cup \chi_n^-$ where $s = 0$ corresponds to the boundary point between the two arcs. It is clear that $T_n(s, \alpha_n(s)) = (-s, \alpha_n(s))$, for all $s \in [0, \varepsilon_n]$. Then, for every $y \in \Delta_n$, there is a unique $(t, s) \in [0, 1] \times [0, \varepsilon_n]$ such that

$$y := (y_1, y_2) = (1-t)(s, \alpha_n(s)) + t(-s, \alpha_n(s)) = ((1-2t)s, \alpha_n(s)).$$

Hence, for every $\varphi \in C(\Delta_n)$, we have

$$\begin{aligned} \int_{\Delta_n} \varphi(y) \sigma_n(y) dy &= \int_{\chi_n^+} \int_0^1 \varphi((1-t)x + tT_n(x)) |x - T_n(x)| f_n^+(x) dt dx \\ &= \int_0^{\varepsilon_n} \int_0^1 \varphi((1-2t)s, \alpha_n(s)) 2s \sqrt{1 + \alpha_n'(s)^2} dt ds \\ &= \int_{\Delta_n} \varphi(y) \frac{2s \sqrt{1 + \alpha_n'(s)^2}}{J_n(t, s)} dy, \end{aligned}$$

where $J_n(t, s) := |\det(D_{(t,s)}(y_1, y_2))|$ on Δ_n . This provides

$$\sigma_n(y) = \frac{2s\sqrt{1 + \alpha'_n(s)^2}}{J_n(t, s)}, \quad \text{for a.e. } y \in \Delta_n.$$

Consequently, we obtain

$$\begin{aligned} \|\sigma\|_{L^p(\Omega)}^p &= \sum_{n=1}^{\infty} \int_0^{\varepsilon_n} \int_0^1 \sigma_n((1-2t)s, \alpha_n(s))^p J_n(t, s) dt ds \\ &\approx \sum_{n=1}^{\infty} \int_0^{\varepsilon_n} \int_0^1 \frac{s^p}{J_n(t, s)^{p-1}} dt ds. \end{aligned}$$

Computing

$$D_{(t,s)}(y_1, y_2) = \begin{pmatrix} -2s & 1-2t \\ 0 & \alpha'_n(s) \end{pmatrix},$$

we get

$$J_n(t, s) = 2s \alpha'_n(s) \approx s^2.$$

Finally, we have

$$\|\sigma\|_{L^p(\Omega)}^p \approx \sum_{n=1}^{\infty} \varepsilon_n^{3-p}.$$

This immediately shows that with this construction we cannot have $\sigma \in L^3$. Moreover, it is enough to choose a sequence ε_n satisfying

$$\sum_{n=1}^{\infty} \varepsilon_n < +\infty, \quad \sum_{n=1}^{\infty} \varepsilon_n^\beta = +\infty$$

for all $\beta < 1$, to prove $\sigma \notin L^p(\Omega)$ for all $p > 2$. Take for instance $\varepsilon_n = \frac{1}{n(\log(1+n))^2}$.

5. APPLICATIONS TO THE BV LEAST GRADIENT PROBLEM

We collect in this section some corollaries of the results of the previous sections, which give interesting proofs for some properties of the BV least gradient problem in dimension $d = 2$. We need to restrict to $d = 2$ because only in this framework we can use rotated gradients, and they have prescribed divergence. As we saw, in dimension $d = 2$ all our results are valid for arbitrary strictly convex norms, and hence apply to the anisotropic least gradient problem.

In all the cases, we will suppose $g \in BV(\partial\Omega)$. Note that this assumption is required to apply the classical theory of optimal transport to $f = \partial_t g$; this requires to transport a measure onto another one with the same total mass, since f is zero-mass. If g was only in $L^1(\partial\Omega)$, then f would be the (one-dimensional) derivative of an L^1 function, i.e. an element of the dual of Lipschitz functions (since $W^{-1,1} = (W^{1,\infty})'$). It is not surprising that a Monge-Kantorovich theory is also possible in this case (because formula (2.2) characterizes the transport cost as the dual norm to the Lipschitz norm), see [7], but no estimates are possible.

Proposition 5.1. *If $\Omega \subset \mathbb{R}^2$ is strictly convex and $g \in BV(\partial\Omega)$, then Problem (1.1) has a solution (i.e. Problems (1.2) and (1.3) have a solution whose trace is g).*

Proof. We have already discussed the fact that we just need to exclude that the solution of (1.2) or (1.3) has a part of its distributional derivative on the boundary. After the rotation, this means that its trace agrees with g if and only if $\sigma(\partial\Omega) = 0$. In strictly convex domains, the transport density gives no mass to the boundary, because of the representation formula (2.3). \square

Proposition 5.2. *If $\Omega \subset \mathbb{R}^2$ is strictly convex, and $g \in (BV \cap C^0)(\partial\Omega)$, then Problem (1.1) has a unique solution.*

Proof. Using again the rotation trick, we just need to prove uniqueness of the solution of the Beckmann Problem. The condition $g \in C^0$ implies that its tangential derivative has no atoms, and we can apply Proposition 2.5. \square

The following result is probably the main contribution of this paper to the understanding of the anisotropic least gradient problem, as we are not aware of similar results already existing in the literature.

Theorem 5.3. *If $\Omega \subset \mathbb{R}^2$ is uniformly convex, and $g \in W^{1,p}(\partial\Omega)$ with $p \leq 2$, then the unique solution of Problem (1.1) belongs to $W^{1,p}(\Omega)$.*

Proof. Setting $f = \partial_{\mathbf{t}}g$ and using f^+ and f^- as its positive and negative parts, the condition $g \in W^{1,p}(\partial\Omega)$ implies $f^\pm \in L^p(\partial\Omega)$. Hence, Proposition 3.3 implies $\sigma \in L^p(\Omega)$, and then $\nabla u \in L^p(\Omega, \mathbb{R}^2)$. \square

Proposition 5.4. *Even if $\Omega \subset \mathbb{R}^2$ is a disk, for every $p > 2$ there exists $g \in \text{Lip}(\partial\Omega)$ such that the unique u solution of Problem (1.1) is not in $W^{1,p}(\Omega)$.*

Proof. It is enough to take g as the antiderivative of the function $f = f^+ - f^-$ of the counterexample of Section 4. \square

Proposition 5.5. *If $\Omega \subset \mathbb{R}^2$ is uniformly convex, and $g \in C^{1,\alpha}(\partial\Omega)$ with $\alpha < 1$, then the unique solution of Problem (1.1) belongs to $W^{1,p}(\Omega)$ for $p = 2/(1 - \alpha)$.*

Proof. This is a consequence of Proposition 3.5. \square

Remark 5.6. Note that the above $W^{1,p}$ regularity also implies Hölder bounds, since in dimension $d = 2$ we have $W^{1,p} \subset C^{0,1-2/p}$. In particular, using $p = 2/(1 - \alpha)$, we get $g \in C^{1,\alpha}(\partial\Omega) \Rightarrow u \in C^{0,\alpha}(\Omega)$. Yet, this bound is not optimal, as it is known (see, for instance, [39]) that we have $g \in C^{1,\alpha}(\partial\Omega) \Rightarrow u \in C^{0,(\alpha+1)/2}(\Omega)$. It is interesting to note that one would obtain exactly the desired $C^{0,(\alpha+1)/2}$ behavior if it was possible to use the Sobolev injection of $W^{1,p}$ corresponding to dimension 1 instead of dimension 2. This seems reasonable, using the fact that level lines of u are transport rays in the transport problem from f^+ to f^- , hence are line segments, but it is not easy to justify and goes beyond the scopes of this paper.

Proposition 5.7. *If $\Omega \subset \mathbb{R}^2$ is uniformly convex, and $g \in C^{1,1}(\partial\Omega)$, then the unique solution of Problem (1.1) is Lipschitz continuous.*

Proof. This is also a consequence of Proposition 3.5. \square

Remark 5.8. Note that the above Lipschitz result is optimal, and perfectly coherent with the theory involving the bounded slope condition (see, for instance [38, 13]), since $C^{1,1}$ functions on the boundary of uniformly convex domains satisfy the bounded slope condition (see [26]).

We finish this section with two last remarks.

Remark 5.9. The strict convexity assumption on φ is crucial in this framework in order to obtain the uniqueness of the minimizers (Proposition 5.2) and also for the approximation procedures performed in Section 3 which allow to translate the estimates in the atomic case into estimates which are valid in the general case. Yet, the L^p bounds obtained in Section 3 do not depend on how much the norm $\|\cdot\|$ is strictly convex, and, whenever $c_0\|\cdot\| \leq \|\cdot\| \leq c_1\|\cdot\|$, the constants in the estimates only depend on c_0 and c_1 . Hence, given an arbitrary norm φ (which is of course equivalent to the Euclidean one), it is possible, just by approximating it with $\varphi_\varepsilon(z) = \varphi(z) + \varepsilon|z|$, to obtain the results of Theorem 5.3, Proposition 5.5 and Proposition 5.7, but in this case the estimates will be true for at least one minimizer (the one selected by this approximation), as there is no more any guarantee of uniqueness.

Remark 5.10. We observe that we have not used Proposition 3.2 in this Section. Indeed, in the framework of the least gradient problem assuming assumptions on f^+ (i.e. on the positive part of the tangential derivative of the boundary datum) are not natural at all. Proposition 3.2 has been inserted in Section 3 just because it was an easy consequence of Proposition 3.1. Also consider that a simple result which could have been proven in Section 3 was the implication $f^\pm \in L^p(\partial\Omega) \Rightarrow \sigma \in L^p(\Omega)$ for arbitrary p (including $p > 2$) under the assumption $\text{spt}(f^+) \cap \text{spt}(f^-) = \emptyset$, but we did not consider it because this assumption, in terms of g , is very unnatural: it would mean that g has some flat regions separating those with positive and negative derivatives.

REFERENCES

- [1] L. AMBROSIO, Lecture Notes on Optimal Transport Problems, in *Mathematical Aspects of Evolving Interfaces*, Lecture Notes in Mathematics (1812) (Springer, New York, 2003), pp. 1–52.
- [2] L. AMBROSIO AND A. PRATELLI, Existence and stability results in the L^1 theory of optimal transportation, in *Optimal transportation and applications*, Lecture Notes in Mathematics (CIME Series, Martina Franca, 2001) 1813, L.A. Caffarelli and S. Salsa Eds., 123–160, 2003.
- [3] M. BECKMANN, A continuous model of transportation, *Econometrica*, 20, 643–660, 1952.
- [4] E. BOMBIERI, E. DE GIORGI AND E. GIUSTI, Minimal cones and the Bernstein problem, *Invent. Math.*, 7 (3) (1969) 243–268.
- [5] G. BOUCHITTÉ AND G. BUTTAZZO, Characterization of optimal shapes and masses through Monge-Kantorovich equation *J. Eur. Math. Soc.*, 3 (2), 139–168, 2001.
- [6] G. BOUCHITTÉ, G. BUTTAZZO AND P. SEPPECHER, Shape optimization solutions via Monge-Kantorovich equation, *C. R. Acad. Sci. Paris Sér. I Math.*, 324 (10), 1185–1191, 1997.
- [7] G. BOUCHITTÉ, T. CHAMPION AND C. JIMENEZ, Completion of the space of measures in the Kantorovich norm, proc. of “Trends in the Calculus of Variations”, Parma, 2004, E.D. Acerbi and G.R. Mingione Editors, *Rivista di Matematica della Università di Parma*, serie 7 (4*), pp. 127–139.
- [8] L. BRASCO, G. CARLIER AND F. SANTAMBROGIO, Congested traffic dynamics, weak flows and very degenerate elliptic equations, *J. Math. Pures et Appl.*, 93 (6), 652–671, 2010.
- [9] J. J. BUCKLEY, Graphs of measurable functions, *Proc. Amer. math. Soc.*, 44, 1974, 78–80.
- [10] L. CAFFARELLI, M. FELDMAN, R. MCCANN, Constructing optimal maps for Monge’s transport problem as a limit of strictly convex costs, *Journal of the American Mathematical Society*, 15 (1), 1–26, 2002.
- [11] L. CARAVENNA, A proof of Sudakov theorem with strictly convex norms, *Math. Z.*, 268, 371–407, 2011.
- [12] G. CARLIER, C. JIMENEZ AND F. SANTAMBROGIO, Optimal transportation with traffic congestion and Wardrop equilibria, *SIAM J. Control Optim.*, 47, 1330–1350 (2008).
- [13] A. CELLINA, On the bounded slope condition and the validity of the Euler Lagrange equation, *SIAM J. Contr. Opt.*, Vol 40 (4), 1270–1279, 2001.
- [14] T. CHAMPION AND L. DE PASCALE, The Monge problem for strictly convex norms in \mathbb{R}^d , *Journal of the European Mathematical Society* Vol 12 (6), 1355–1369, 2010.
- [15] T. CHAMPION AND L. DE PASCALE, The Monge problem in \mathbb{R}^d , *Duke Math. J.*, 157, 3 (2011), 551–572.
- [16] L. DE PASCALE, L. C. EVANS AND A. PRATELLI, Integral estimates for transport densities, *Bull. of the London Math. Soc.*, 36, n. 3, pp. 383–395, 2004.
- [17] L. DE PASCALE AND A. PRATELLI, Regularity properties for Monge Transport Density and for Solutions of some Shape Optimization Problem, *Calc. Var. Par. Diff. Eq*, 14, n. 3, pp. 249–274, 2002.

- [18] L. DE PASCALE AND A. PRATELLI, Sharp summability for Monge Transport density via Interpolation, *ESAIM Control Optim. Calc. Var.*, 10, n. 4, pp. 549–552, 2004.
- [19] S. DWEIK AND F. SANTAMBROGIO, Summability estimates on transport densities with Dirichlet regions on the boundary via symmetrization techniques, *ESAIM Control Optim. Calc. Var.*, 24(3), 1167–1180, 2018.
- [20] L. C. EVANS AND W. GANGBO, Differential equations methods for the Monge-Kantorovich mass transfer problem, *Mem. Amer. Math. Soc.*, **137**, No. 653 (1999).
- [21] M. FELDMAN AND R. MCCANN, Uniqueness and transport density in Monge’s mass transportation problem, *Calc. Var. Par. Diff. Eq.*, 15, n. 1, pp. 81–113, 2002.
- [22] W. GANGBO AND R. J. MCCANN, Shape recognition via Wasserstein distance, *Quart. Appl. Math.* 58, 705–737 (2000).
- [23] D. GILBARG, N. S. TRUDINGER, Elliptic Partial Differential Equations of Second Order, Grundlehren Math. Wiss. 224, Springer, Berlin (1983).
- [24] W. GÓRNY, Planar least gradient problem: existence, regularity and anisotropic case, *arXiv:1608.02617*.
- [25] W. GÓRNY, P. RYBKA AND A. SABRA, Special cases of the planar least gradient problem, *Nonlinear Analysis*, 151, pp. 66 – 95, 2017.
- [26] P. HARTMAN, On the bounded slope condition, *Pacific J. Math.*, Vol. 18 (3), 495–511, 1966.
- [27] G. MERCIER, Continuity results for TV-minimizers, *Indiana University Math. J.*, 67, 1499–1545, 2018.
- [28] A. MORADIFAM, A. NACHMAN AND A. TAMASAN, Uniqueness of minimizers of weighted least gradient problems arising in conductivity imaging, to appear in *Calc. Var. Par. Diff. Eq.*
- [29] J. M. MAZON, The Euler-Lagrange equation for the anisotropic least gradient problem, *Nonlinear Analysis: Real World Applications*, 31 (2016), pp. 452–472.
- [30] J. M. MAZON, J. D. ROSSI AND S. S. DE LEON, Functions of least gradient and 1-harmonic functions, *Indiana Univ. J. Math.*, 63, pp. 1067–1084.
- [31] R. MCCANN AND M. SOSIO, Hölder continuity for optimal multivalued mappings, *SIAM J. Math. Anal.*, 43, p. 1855–1871 (2011).
- [32] G. MONGE, Mémoire sur la théorie des déblais et des remblais, *Histoire de l’Académie Royale des Sciences de Paris* (1781), 666–704.
- [33] J. LELLMANN, D.A. LORENZ, C. SCHOENLIEB AND T. VALKONEN, Imaging with Kantorovich- Rubinstein discrepancy, *SIAM J. Imaging Sciences*, 7 (4), 2833–2859, 2014.
- [34] F. SANTAMBROGIO, Absolute continuity and summability of transport densities: simpler proofs and new estimates, *Calc. Var. Par. Diff. Eq.*, (2009) 36: 343–354.
- [35] F. SANTAMBROGIO, *Optimal Transport for Applied Mathematicians*, in *Progress in Nonlinear Differential Equations and Their Applications* 87, Birkhäuser Basel (2015).
- [36] N. TRUDINGER AND X.-J. WANG, On the Monge mass transfer problem, *Calculus of Variations and Partial Differential Equations*, 13, 19–31, 2001.
- [37] G. SPRADLIN AND A. TAMASAN, Not all traces on the circle come from functions of least gradient in the disk, *Indiana University Mathematics Journal*, 63 (2014), 1819–1837.
- [38] G. STAMPACCHIA, On some regular multiple integral problems in the calculus of variations, *Comm. Pure Appl. Math.*, 16, 383–421, 1963.
- [39] P. STERNBERG, G. WILLIAMS, W. P. ZIEMER, Existence, uniqueness, and regularity for functions of least gradient, *J. Reine Angew. Math.*, 430 (1992), pp. 35–60.

S.D.: LABORATOIRE DE MATHÉMATIQUES D’ORSAY, UNIV. PARIS-SUD, CNRS, UNIVERSITÉ PARIS-SACLAY, 91405 ORSAY CEDEX, FRANCE; **F.S.:** INSTITUT CAMILLE JORDAN, UNIVERSITÉ CLAUDE BERNARD LYON 1, 43 BOULEVARD DU 11 NOVEMBRE 1918, 69622 VILLEURBANNE CEDEX, FRANCE

E-mail address: `samer.dweik@math.u-psud.fr`, `santambrogio@math.univ-lyon1.fr`