Lectures on quantitative stability of optimal transport

Cyril Letrouit

May 28, 2025

Abstract

Optimal transport plays a prominent role across numerous disciplines, including analysis, probability, statistics, geometry, and machine learning. For many reasons, not only existence and uniqueness of optimal transport maps, but also their stability with respect to variations of the marginal distributions is of fundamental importance. Qualitative stability results have long been established, but quantitative estimates are often needed both for numerical and theoretical purposes. We review recent theoretical advances in this emerging and flourishing field. We also discuss a range of applications, including embedding of subsets of the Wasserstein space into Hilbert spaces, linearized optimal transport, statistical optimal transport and the random matching problem. These notes are based on my Cours Peccot delivered at the Collège de France in May-June 2025.

Preliminary version, comments welcome. Email: prenom.nom@universite-paris-saclay.fr

Contents

Program of the lectures:

- May 14th: General introduction (Section 1), statement of the main results (Section 2).
- May 21st: The Kantorovich functional (Section 3). This lecture is mostly based on a work by Alex Delalande and Quentin Mérigot [31], revisited in [54].
- May 28th: Gluing techniques (Section 4), examples and counterexamples (Section 5), stability of maps (Section 6). This lecture is mostly based on my joint work with Quentin Mérigot [54].
- June 4th: Generalizations (Section 7), applications (Section 8) and non-optimal transport of measures (Transformers, diffusion models). This lecture is based on various works by many authors.

It is my great pleasure to thank Quentin Mérigot, who introduced me to this topic. I am also thankful to Guillaume Carlier, Vincent Divol, William Ford, Michael Goldman, Antoine Julia, Jun Kitagawa, Tudor Manole, Jonathan Niles-Weed, Pierre Pansu, Aram Pooladian and Philippe Rigollet for various discussions.

To lighten the presentation, most references are gathered at the end of every chapter in a bibliographical paragraph.

1 General introduction

1.1 The optimal transport problem

The nearly 250 years old Monge transportation problem consists in finding the optimal way to transport mass from a given source to a given target probability measure, while minimizing an integrated cost.

Let ρ be a probability measure on a Polish (i.e., complete, separable metric) space $\mathcal{X}$ and μ be a probability measure on a Polish space $\mathcal{Y}$. For simplicity, one may assume $\mathcal{X},\mathcal{Y}\subset\mathbb{R}^d$. Let c be a non-negative measurable function on $\mathcal{X}\times\mathcal{Y}$. An *admissible mass transport plan* is an element γ of the space $\mathcal{P}(\mathcal{X}\times\mathcal{Y})$ of probability measures over $\mathcal{X}\times\mathcal{Y}$ whose marginals coincide with ρ and μ, i.e., for all measurable sets $A\subset\mathcal{X}$, $B\subset\mathcal{Y}$,

$$\gamma(A\times\mathcal{Y})=\rho(A)\qquad\text{and}\qquad\gamma(\mathcal{X}\times B)=\mu(B).\tag{1.1}$$

These conditions mean that for any $x\in\mathcal{X}$, $y\in\mathcal{Y}$, the amount of mass taken from x coincides with $\mathrm{d}\rho(x)$, and the amount of mass arriving at y coincides with $\mathrm{d}\mu(y)$. The set of all admissible transport plans is

$$\Pi(\rho,\mu)=\{\gamma\in\mathcal{P}(\mathcal{X}\times\mathcal{Y})\mid(1.1)\text{ holds}\}.$$

It is non-empty and convex. The optimal transport problem with cost c is the minimization problem

$$\inf_{\gamma\in\Pi(\rho,\mu)}\int_{\mathcal{X}\times\mathcal{Y}}c(x,y)\mathrm{d}\gamma(x,y).\tag{1.2}$$

A solution to (1.2) is called an optimal transport plan. In the particular case where $\mathcal{X},\mathcal{Y}\subset\mathbb{R}^d$ and $c(x,y)=|x-y|^2$, one finds the *quadratic optimal transport problem*

$$\inf_{\gamma\in\Pi(\rho,\mu)}\int_{\mathbb{R}^d\times\mathbb{R}^d}|x-y|^2\mathrm{d}\gamma(x,y),\tag{1.3}$$

which will be our main focus during the largest part of these lectures. The case of p-costs $c(x,y)=|x-y|^p$ with $p\geq 1$ is also of interest, and gives rise to the p-Wasserstein distance defined as

$$W_p(\rho,\mu)=\left(\inf_{\gamma\in\Pi(\rho,\mu)}\int_{\mathbb{R}^d\times\mathbb{R}^d}|x-y|^p\mathrm{d}\gamma(x,y)\right)^{1/p}.$$

Optimal transport and Wasserstein distances are used in an incredible number of fields. Here is a very incomplete list of fields, with one or two applications and/or names for each:

- Engineering: move mass from one place to another while minimizing a total transportation cost (Monge 1781).
- Economics: optimal allocation of resources between m production stations and n consumption stations (Kantorovich 1942).
- Mathematical physics and modelling: interpretation of the Euler equation of fluid mechanics via a least action principle in the space of diffeomorphisms (Arnold 1966, Brenier 1989); interpretation of the heat equation as a gradient descent of entropy in the geometry of mass transport (Otto 1998); construction of the semigeostrophic model in atmospheric sciences (Cullen and Purser 1980's); kinetic theory (Tanaka 1970's).
- Mathematics: analysis of the Monge-Ampère partial differential equation $\det(D^2f)=g$ (Caffarelli 1990's); convex geometry; functional inequalities; definition of geometric and topological properties in spaces without smooth structures, e.g., synthetic theory of Ricci curvature (Lott-Sturm-Villani 2006-2009).
- Image processing: measure distance between images (image retrieval and comparison); color transfer; image interpolation; super-resolution and denoising.
- Statistics: rate of convergence of empirical probability measures μ_n to their limit μ (Dudley 1969); estimation of coupling between data.
- Machine learning: generative modeling; interpolation of multiple data distributions (e.g. samples, images, domains, etc) using Wasserstein barycenters; analysis of the training dynamics of neural networks; analysis of sampling algorithms such as the Langevin Monte Carlo algorithm.

Although the distance cost $c(x,y)=|x-y|$ might seem more physical at first sight than the quadratic cost (it is the natural cost in the Monge problem for instance), the quadratic cost $c(x,y)=|x-y|^2$ is actually the one which is most useful in the above examples due to Brenier's theorem recalled below, the link with the W_2 distance which gives a Riemannian structure to the space of probability measures, its smoothness which makes it suitable for optimization, its computational advantages in relation to Sinkhorn's algorithm, etc. The distance cost does not come with as nice properties as the quadratic cost.

A solution to (1.3) (or (1.2)) exists under mild assumptions: for instance that $\mathcal{X},\mathcal{Y}$ are Polish spaces (i.e., complete and separable metric spaces) and that c is lower semi-continuous. However, the solution to (1.3) (or (1.2)) is not unique in general. For instance, if $A=(1,0)$, $B=(-1,0)$, $C=(0,1)$ and $D=(0,-1)$ are the vertices of a square in $\mathbb{R}^2$, there is an infinite number of solutions to (1.3) when $\rho=\frac{1}{2}(\delta_A+\delta_B)$ and $\mu=\frac{1}{2}(\delta_C+\delta_D)$: for any $a\in[0,1]$,

$$\gamma=\frac{1}{2}\left(a\delta_{(A,C)}+(1-a)\delta_{(A,D)}+(1-a)\delta_{(B,C)}+a\delta_{(B,D)}\right)$$

is an admissible transport plan which is a solution of (1.3). Notice that in this example, the mass leaving A is split into one part going to C and one part going to D.

Let us pause for a moment and ask what would happen if we would not allow mass-splitting, i.e., if if we replace the infimum in (1.3) by a minimization over the admissible transport plans $\gamma \in \Pi(\rho, \mu)$ which are supported on the graph of a univalued map $T : \mathcal{X} \to \mathcal{Y}$: in other words, all the mass at $x \in \mathcal{X}$ is sent into $T(x) \in \mathcal{Y}$. The condition $\gamma \in \Pi(\rho, \mu)$ then turns into the condition that for any measurable $U \subset \mathcal{Y}$, $\rho(T^{-1}(U)) = \mu(U)$, i.e., $T_\#\rho = \mu$ where $\#$ denotes the pushforward operation. The associated admissible transport plan is $\gamma = (\mathrm{Id}, T)_\#\rho$. We obtain the so-called Monge problem:

$$\inf_{\substack{S:\mathcal{X}\to\mathcal{Y}\\ S_\#\rho=\mu}} \int_{\mathbb{R}^d} |x - S(x)|^2 \mathrm{d}\rho(x). \tag{1.4}$$

A solution to (1.4) is called an optimal transport map. Notice that without the absolute continuity assumption on ρ, the Monge problem does not necessarily have a solution. If ρ is a sum of Dirac masses but μ is not, then there does not exist any $S : \mathcal{X} \to \mathcal{Y}$ such that $S_\#\rho = \mu$.

There exists a simple assumption which guarantees that the solution to (1.3) is unique: Brenier showed that the absolute continuity of the source measure ρ is a sufficient condition for a unique solution to (1.3) to exist. And even more: he shows that in this case, the Monge problem (1.4) has a unique solution T, and that these solutions to the two problems are related by $\gamma = (\mathrm{Id}, T)_\#\rho$.

In the sequel, $\mathcal{P}(\mathcal{X})$ denotes the set of probability measures on $\mathcal{X} \subset \mathbb{R}^d$, and $\mathcal{P}_p(\mathcal{X})$ is the set of probability measures on $\mathcal{X}$ with finite p-th moment:

$$\mathcal{P}_p(\mathcal{X}) = \left\{ \rho \in \mathcal{P}(\mathcal{X}) \mid \int_{\mathcal{X}} |x|^p \mathrm{d}\rho(x) < +\infty \right\}.$$

The weak topology on $\mathcal{P}(\mathcal{X})$ (or topology of weak convergence, or narrow topology) is induced by convergence against $C_b(\mathcal{X})$, i.e., bounded continuous functions.

Theorem 1.1 (Brenier). *Let $\rho, \mu \in \mathcal{P}_2(\mathbb{R}^d)$ and $c(x, y) = |x - y|^2$ be the quadratic cost on $\mathbb{R}^d$. Assume that ρ is absolutely continuous with respect to the Lebesgue measure. Then there exists between ρ and μ a ρ-a.e. unique optimal transport map T and a unique optimal transport plan γ, and these solutions are related by $\gamma = (\mathrm{Id}, T)_\#\rho$. Furthermore, the map T is the gradient of a convex function $\phi : \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$, and if $(\nabla f)_\#\rho = \mu$ for some other convex function f, then $\nabla f = \nabla\phi$ ρ-a.e.*

If the support $\mathcal{X}$ of ρ is the closure of a bounded connected open set, ϕ is uniquely determined on $\mathcal{X}$ up to additive constants. As a consequence of Brenier's theorem, for any convex function ϕ and any absolutely continuous $\rho \in \mathcal{P}_2(\mathbb{R}^d)$, the map $\nabla\phi$ is the optimal transport map from ρ to $(\nabla\phi)_\#\rho$.

To turn (1.3) (or (1.4)) into a well-posed problem in the sense of Hadamard, there only remains to show *stability of the solution* T with respect to perturbations of ρ and μ. The question of stability is fundamental both from the theoretical and the numerical point of view. Soft (compactness) arguments provide without any difficulty a qualitative stability result presented in Section 1.2. However, *quantitative results* are needed in most applications, and for this more difficult problem, tools have started to emerge only in the last five years. The purpose of these notes is to review the recent theoretical advances in this now fastly developing field, and to discuss applications to various problems.

1.2 Stability of optimal transport

Recall that weak convergence of measures is understood against continuous bounded test functions. The following general qualitative stability result is true.

Proposition 1.2. *Let $(\rho_k)_{k\in\mathbb{N}}$ converge weakly to ρ and $(\mu_k)_{k\in\mathbb{N}}$ converge weakly to μ. For each $k \in \mathbb{N}$, let γ_k be an optimal transport plan between ρ_k and μ_k, and assume that*

$$\liminf_{k\in\mathbb{N}} \int_{\mathcal{X}\times\mathcal{Y}} |x-y|^2 \mathrm{d}\gamma_k(x,y) < +\infty.$$

Then the optimal transport cost between ρ and μ is finite and, up to extraction of a subsequence, γ_k converges weakly to some optimal transport plan γ between ρ and μ.

The proof relies on the Prokhorov theorem (to extract a converging subsequence) and on a characterization of optimal transport plans as cyclically monotone sets. Proposition 1.2 actually holds in general Polish spaces $\mathcal{X}$ and $\mathcal{Y}$, with a continuous cost function $c : \mathcal{X} \times \mathcal{Y} \to \mathbb{R}$ such that $\inf c > -\infty$.

In these lectures, we will fix the source measure ρ and consider stability with respect to the target measure only. The problem we are interested in reads

If μ and ν are quantitatively close, prove that $T_{\rho\to\mu}$ and $T_{\rho\to\nu}$ are quantitatively close

where $T_{\rho\to\mu}$ (resp. $T_{\rho\to\nu}$) is the optimal transport map from ρ to μ (resp. ρ to ν) given by Brenier's theorem. There are several reasons for this choice of fixing the source measure:

- first, because the mapping $\mu \mapsto T_{\rho\to\mu}$ may be used to embed the Wasserstein space (or part of it) into the Hilbert space $L^2(\rho)$ with a controlled distortion, as explained in Section 1.4. This is important in its own.

- Second, because $T_{\rho\to\mu}$ and $T_{\rho\to\nu}$ are in $L^2(\rho)$ according to Brenier's Theorem 1.1, and thus we may measure their distance simply in $L^2(\rho)$, whereas if we had ρ and ρ' as source measures, measuring distances between the maps would be less easy (instead, one would probably measure the Wasserstein distance between optimal transport plans).

- Third, because in some applications, ρ is a perfectly known probability density, e.g. a standard Gaussian.

- Finally, it is sometimes possible to deduce stability with respect to both marginals from the proof techniques.

To summarize, in these lecture notes, some $\rho \in \mathcal{P}_2(\mathbb{R}^d)$, assumed to be absolutely continuous with respect to the Lebesgue measure, is fixed. Therefore, we may drop in the notation the reference to this source measure, and given $\mu \in \mathcal{P}_2(\mathbb{R}^d)$ we call

- the *optimal transport map* and denote by $T_\mu \in L^2(\rho)$ the unique solution to (1.4).

- the *Kantorovich potential* the unique convex function $\phi_\mu \in L^2(\rho)$ such that $T_\mu = \nabla\phi_\mu$ and $\int_{\mathcal{X}} \phi_\mu \mathrm{d}\rho = 0$.

In the context of these lectures, the Kantorovich potential is always uniquely defined. This uniqueness may fail, however, if the support of ρ consists of multiple connected components.

The source measure ρ being now fixed, we formulate the qualitative stability of optimal transport maps as follows:

Proposition 1.3. *The map $\mu \mapsto T_\mu$ from $(\mathcal{P}_2(\mathbb{R}^d), W_2)$ to $L^2(\rho)$ is continuous.*

Proof. Let $(\mu_n)_{n\in\mathbb{N}}$ and μ be in $\mathcal{P}_2(\mathbb{R}^d)$ such that $W_2(\mu_n, \mu) \to 0$. Then $W_2(\rho, \mu_n) \to W_2(\rho, \mu)$ by the triangle inequality, hence

$$\int_{\mathbb{R}^d} |x - T_{\mu_n}(x)|^2 \mathrm{d}\rho(x) \to \int_{\mathbb{R}^d} |x - T_\mu(x)|^2 \mathrm{d}\rho(x). \tag{1.5}$$

Therefore, (T_{μ_n}) is bounded in $L^2(\rho)$.

Let

$$K_\varepsilon = \{x \in \mathbb{R}^d \mid |x| \leq \varepsilon^{-1},\ \rho(x) \geq \varepsilon,\ \mathrm{dist}(x, \partial\Omega) \geq \varepsilon\}$$

where $\Omega = \mathrm{supp}(\rho)$. Let us prove that for any $\varepsilon > 0$,

$$\sup_n \|T_{\mu_n}\|_{L^\infty(K_\varepsilon)} < +\infty. \tag{1.6}$$

For this we rely on the fact that for any convex function f over $\mathbb{R}^d$, any $x \in \mathbb{R}^d$ and $\eta > 0$,

$$\|\partial f\|_{L^\infty} \leq \frac{6}{\beta_d \eta^d} \int_{B(x,4\eta)} |\nabla f| \mathrm{d}\lambda \tag{1.7}$$

where λ denotes the Lebesgue measure on $\mathbb{R}^d$. The proof of (1.7) is provided in Appendix A. Let us deduce (1.6) from (1.7). For any x such that $B(x, 4\varepsilon) \subset K_\varepsilon$,

$$\frac{\beta_d \varepsilon^d}{6} \|T_{\mu_n}\|_{L^\infty(B(x,\varepsilon))} \leq \int_{B(x,4\varepsilon)} |T_{\mu_n}| \mathrm{d}\lambda \leq \varepsilon^{-1} \Big(\int_{B(x,4\varepsilon)} |T_{\mu_n}|^2 \mathrm{d}\rho \Big)^{1/2} \tag{1.8}$$

by applying (1.7) to ϕ_{μ_n}, using that $\rho(x) \geq \varepsilon$ on K_ε, and finally applying the Cauchy-Schwarz inequality. Since (T_{μ_n}) is bounded in $L^2(\rho)$, the right-hand side in (1.8) for fixed $\varepsilon > 0$ is uniformly bounded in n. Therefore $\|T_{\mu_n}\|_{L^\infty(K'_\varepsilon)}$ is uniformly bounded in n for $K'_\varepsilon = \{x \in K_\varepsilon \mid B(x, 4\varepsilon) \subset K_\varepsilon\}$. Sending ε to 0, this implies that $\sup_n \|T_{\mu_n}\|_{L^\infty(K)} < +\infty$ for any compact set K included in the interior of the support of ρ. In particular, this implies (1.6).

From now on, we normalize ϕ_{μ_n} in a way that $\phi_{\mu_n}(0) = 0$. By Arzelà-Ascoli, up to extraction of a subsequence omitted in the notation, (ϕ_{μ_n}) converges toward some ϕ uniformly over any K_ε. Of course, ϕ is convex. Passing to the limit $n \to +\infty$ in the inequality $\phi_{\mu_n}(y) \geq \phi_{\mu_n}(x) + \langle y - x, \nabla\phi_{\mu_n}(x)\rangle$ yields that any limit point of $(\nabla\phi_{\mu_n}(x))$ is in $\partial\phi(x)$. This proves that at any point x of differentiability of ϕ, $(\nabla\phi_{\mu_n})$ converges to $\nabla\phi$. Since ϕ is convex, it is differentiable almost everywhere, thus $T_{\mu_n}(x) \to T(x)$ for ρ-almost every x, where $T = \nabla\phi$. We deduce using (1.6) and Lebesgue's dominated convergence theorem that

$$(T_{\mu_n}) \text{ converges (strongly) to } T \text{ in } L^2(\rho, K_\varepsilon) \text{ for any } \varepsilon > 0. \tag{1.9}$$

Also, since (T_{μ_n}) is bounded in $L^2(\rho)$, it converges weakly to some $T' \in L^2(\rho)$, and we deduce from (1.9) that $T' = T$. Therefore $\langle \mathrm{Id}, T_{\mu_n}\rangle_{L^2(\rho)} \to \langle \mathrm{Id}, T\rangle_{L^2(\rho)}$, and plugging into (1.5) we obtain that $\|T_{\mu_n}\|_{L^2(\rho)} \to \|T\|_{L^2(\rho)}$. This proves that (T_{μ_n}) in fact converges *strongly* to T in $L^2(\rho)$.

Finally, let us observe that $T_\#\rho = \mu$ since (T_{μ_n}) converges a.e. to T and (T_{μ_n}) is locally uniformly bounded according to (1.6). Since T is the gradient of a convex function, Brenier's theorem implies that $T = T_\mu$ is the optimal transport map from ρ to μ. We conclude that the full sequence (T_{μ_n}) converges strongly to T in $L^2(\rho)$. □

The main problem under consideration in these notes will (almost) be the following one: for a given absolutely continuous $\rho \in \mathcal{P}_2(\mathbb{R}^d)$, do there exist constant $C, \alpha > 0$ such that for all μ, ν with finite second moment,

$$\|T_\mu - T_\nu\|_{L^2(\rho)} \leq C W_2(\mu, \nu)^\alpha \tag{1.10}$$

holds? More generally, replacing W_2 by W_p for some $p \geq 1$, we will consider inequalities of the type

$$\|T_\mu - T_\nu\|_{L^2(\rho)} \leq C W_p(\mu, \nu)^\alpha, \tag{1.11}$$

the strongest one being for $p = 1$ (since $W_p \leq W_q$ for $p \leq q$) and α as large as possible. The largest possible α is sometimes called the stability exponent (associated to ρ) in the sequel.

An important observation is that the reverse inequality

$$\|T_\mu - T_\nu\|_{L^2(\rho)} \geq W_2(\mu, \nu) \tag{1.12}$$

always holds: indeed, $\gamma = (T_\mu, T_\nu)_\# \rho$ is an admissible transport plan between μ and ν, and its cost

$$\left(\int_{\mathbb{R}^d \times \mathbb{R}^d} |x - y|^2 \mathrm{d}\gamma(x, y) \right)^{1/2} = \left(\int_{\mathbb{R}^d} |T_\mu(x) - T_\nu(x)|^2 \mathrm{d}\rho(x) \right)^{1/2} = \|T_\mu - T_\nu\|_{L^2(\rho)}$$

is by definition not lower than the cost $W_2(\mu, \nu)$ of an optimal transport plan between μ and ν.

Put together, the inequalities (1.10) and (1.12) imply that the mapping $\mu \mapsto T_\mu$ is a bi-Hölder embedding of the Wasserstein space into $L^2(\rho)$. However, as we shall discuss in more details in Section 1.4, it is known that if $d \geq 3$, then (1.10) cannot hold uniformly over all probability measures μ and ν on $\mathbb{R}^d$ with finite second moment (the case $d = 2$ seems open). In fact, it is not possible to embed the Wasserstein space into any L^p space, even in a very weak sense. Nevertheless, what we will discuss in depth in these lectures is that a stability bound such as (1.10) can hold if one restricts to slightly smaller families of measures μ, ν. For instance, we will show that under some assumptions on ρ, for any compact set $\mathcal{Y} \subset \mathbb{R}^d$, there exist $C, \alpha > 0$ such that (1.10) holds for any μ, ν supported in $\mathcal{Y}$.

In these notes, we will also be interested in quantitative stability estimates for Kantorovich potentials, which take the form

$$\|\phi_\mu - \phi_\nu\|_{L^2(\rho)} \leq C' W_{p'}(\mu, \nu)^{\alpha'}. \tag{1.13}$$

Actually the stability of optimal transport maps (1.11) will be deduced from the stability of Kantorovich potentials (1.13), as explained in detail in Section 6. Kantorovich potentials are interesting objects on their own, for many reasons. First, many numerical methods used to solve optimal transport problem, for instance semi-discrete optimal transport and dual gradient methods, rely on solving first the dual formulation of the problem, discussed in Section 3.1. In these methods, one computes the Kantorovich potentials first, before taking the gradient to obtain the optimal transport map. Also, the Sinkhorn algorithm, which is one of the best ways to compute solutions of (regularized) optimal transport problems, computes the entropic version of the Kantorovich potentials (discussed in Section 7.1). Finally, Kantorovich potentials have an economic interpretation which may help understand their meaning (see [72, Chapter 5]).

Let us already mention that although these lecture notes are focused on the quadratic cost in $\mathbb{R}^d$ given by $c(x, y) = |x - y|^2$, most results remain valid for more general costs, for instance p-costs $c(x, y) = |x - y|^p$, $p > 1$ (see Section 7.2) and the quadratic cost $c(x, y) = \frac{1}{2}\mathrm{dist}(x, y)^2$ on Riemannian manifolds (see Section 7.3).

It is clear that inequalities like (1.11) and (1.13) are useful to justify the theoretical consistance of "plugin methods" to compute optimal transport: if we want to compute the optimal transport map or the Kantorovich potential from ρ to μ but do not know exactly μ (due to some noise for instance) and have only access to some approximation ν of μ, these inequalities tell us how close we may expect T_ν to be from T_μ (and ϕ_ν from ϕ_μ), depending on some Wasserstein distance between μ and ν.

We shall not discuss numerical methods and algorithms used to compute optimal transport in practice. The computational errors that they induce are another interesting subject for mathematical analysis, not covered in these lecture notes.

1.3 The one-dimensional case

The case where $d = 1$, i.e., ρ, μ, ν are probability measures on $\mathbb{R}$, is particularly simple. Indeed, as soon as ρ is absolutely continuous on $\mathbb{R}$, the mapping $\mu \mapsto T_\mu$ is an isometric embedding:

$$\|T_\mu - T_\nu\|_{L^p(\rho)} = W_p(\mu, \nu) \tag{1.14}$$

for any $p \geq 1$. The stability problem is thus completely solvable in this case: the bound (1.10) holds with $C = \alpha = 1$. To prove (1.14) it is sufficient to observe that

$$\gamma_{\text{opt}} = (T_\mu, T_\nu)_\# \rho \tag{1.15}$$

is an optimal transport plan between μ and ν. Indeed, (1.14) then follows immediately:

$$W_p^p(\mu, \nu) = \int_{\mathbb{R}^2} |x - y|^p \mathrm{d}\gamma_{\text{opt}}(x, y) = \int_{\mathbb{R}} |T_\mu(x) - T_\nu(x)|^p \mathrm{d}\rho(x) = \|T_\mu - T_\nu\|^p_{L^p(\rho)}.$$

It is clear that γ_{opt} is an admissible transport plan between μ and ν since $(T_\mu)_\#\rho = \mu$ and $(T_\nu)_\#\rho = \nu$. The difficulty is to show that it is optimal.

Optimal transport plans in 1d are always monotone. This means that if γ is an optimal transport plan between two 1d probability measures, and (x, y) and (x', y') are in the support of γ and $x < x'$, then necessarily $y \leq y'$. This due to the convexity of the quadratic cost. Indeed, for any $x < x'$ and $y \leq y'$, the inequality

$$|x - y|^2 + |x' - y'|^2 \leq |x - y'|^2 + |x' - y|^2$$

holds, which means that it is always less costly to transport mass from x to y and from x' to y' than to "cross trajectories" and transport mass from x to y' and from x' to y.

Applying this to the transport from ρ to μ, it is possible to give a completely explicit expression for T_μ. Let us verify that

$$T_\mu(m) = \inf\{x \in \mathbb{R} \mid F_\mu(x) \geq F_\rho(m)\} \tag{1.16}$$

where

$$F_\mu : x \mapsto \mu((-\infty, x])$$

denotes the cumulative distribution function. We first check that T_μ pushes forward ρ to μ. To prove this, we observe that $T_\mu(m) \leq x$ if and only if $F_\mu(x) \geq F_\rho(m)$. Setting $\widehat{\mu} = T_{\mu\#}\rho$, we thus have

$$\begin{aligned}\widehat{\mu}((-\infty, x]) &= \rho(T_\mu^{-1}((-\infty, x])) = \rho(\{m \mid T_\mu(m) \leq x\}) \\ &= \rho(\{m \mid F_\rho(m) \leq F_\mu(x)\}) = F_\mu(x) = \mu((-\infty, x])\end{aligned}$$

hence $\widehat{\mu} = \mu$. Moreover, T_μ is optimal since it is the only transport map from ρ to μ which is monotone. The optimal transport map T_ν from ρ to ν is of course given by an analogous expression to (1.16).

Now, since T_μ and T_ν are monotone, it is immediate to check that γ_{opt} is also monotone. But there is only one monotone admissible transport plan between μ and ν, and thus γ_{opt} is optimal.

Finally, what can be said about stability of Kantorovich potentials in 1d? If ρ satisfies the Poincaré inequality, i.e., if there exists $C > 0$ such that

$$\int_{\mathcal{X}} f \, \mathrm{d}\rho = 0 \Rightarrow \int_{\mathcal{X}} f^2 \, \mathrm{d}\rho \leq C \int_{\mathcal{X}} |\nabla f|^2 \, \mathrm{d}\rho,$$

then it follows from (1.14) (with $p = 2$) that $\|\phi_\mu - \phi_\nu\|_{L^2(\rho)} \leq C W_2(\mu, \nu)$. As we shall see in Section 5.3 (see notably Remark 5.3), this stability inequality for Kantorovich potentials is no longer guaranteed if ρ does not satisfy the Poincaré inequality, even if the support of ρ is an interval (in which case Kantorovich potentials are unique).

1.4 Applications: embedding of the Wasserstein space and linearized optimal transport

In this section we describe one important application of quantitative stability estimates. Further applications are discussed in Section 8.

Let $\rho \in \mathcal{P}_2(\mathbb{R}^d)$ be absolutely continuous. When $T = \nabla\phi$ is the gradient of a convex function, the curve

$$[0,1] \ni t \mapsto ((1-t)\mathrm{Id} + tT)_{\#}\, \rho \in \mathcal{P}_2(\mathbb{R}^d) \tag{1.17}$$

is a Wasserstein geodesic from ρ to $T_\# \rho$, meaning that it is a curve which minimizes the W_2-distance between any two of its points. The Wasserstein space $(\mathcal{P}_2(\mathbb{R}^d), W_2)$ may then be viewed formally as an "infinite-dimensional Riemannian manifold". Its tangent space at ρ is naturally defined as the set

$$\mathrm{Tan}_\rho \mathcal{P}_2(\mathbb{R}^d) = \overline{\{\nabla\phi - \mathrm{Id} \mid \phi \text{ convex}, \ \phi \in C_c^\infty(\mathbb{R}^d)\}}^{L^2(\rho)} \tag{1.18}$$

where $\nabla\phi - \mathrm{Id}$ is the (initial) tangent vector to the Wasserstein geodesic given by (1.17). In (1.18), it is natural to take the closure: in analogy, the (solid) tangent cone at a boundary point x of a closed convex set $C \subset \mathbb{R}^d$ is defined as the *closure* of the cone formed by all half-lines emanating from x and intersecting C in at least one point distinct from x. On this Riemannian manifold, the exponential map with base-point ρ is nothing else than

$$\mathrm{Tan}_\rho \mathcal{P}_2(\mathbb{R}^d) \ni T \mapsto T_\# \rho \in \mathcal{P}_2(\mathbb{R}^d).$$

And the map $\mu \mapsto T_\mu - \mathrm{Id}$ from $\mathcal{P}_2(\mathbb{R}^d)$ to $L^2(\rho)$ is the analog of the Riemannian logarithm. It is an injective map, with image the tangent space (1.18).

If the stability inequality (1.10) holds for any $\mu, \nu \in \mathcal{P}_2(\mathbb{R}^d)$, this means that $\mu \mapsto T_\mu$ is a bi-Hölder embedding from $(\mathcal{P}_2(\mathbb{R}^d), W_2)$ to the Hilbert space $L^2(\rho)$ (using the reverse inequality (1.12)). For instance, the previous section showed that in 1d, $\mu \mapsto T_\mu$ is an *isometric* embedding. However, in dimension $d > 1$, it is known that Wasserstein spaces do not embed into any Banach space, even for much coarser notions of embedding. Therefore, we will aim at establishing (1.10) for strict subsets of $\mathcal{P}_2(\mathbb{R}^d)$, for instance for target probability measures μ, ν supported in a fixed compact set, or with bounds on some moments. Working with this embedding is equivalent to endow $\mathcal{P}_2(\mathbb{R}^d)$ with the "ρ-based" distance

$$W_{2,\rho}(\mu, \nu) = \|T_\mu - T_\nu\|_{L^2(\rho)}. \tag{1.19}$$

Due to the linear structure of the Hilbert space $L^2(\rho)$, the logarithm map $\mu \mapsto T_\mu$ is also used as a way to "linearize" optimal transport. For instance, to compute an "average" between two measures μ and ν in the Wasserstein space, one usually resorts to the notion of Wasserstein barycenter (or McCann interpolation), defined as a minimizer of

$$\inf_{\chi\in\mathcal{P}_2(\mathbb{R}^d)} \frac{1}{2}\left(W_2(\mu,\chi)^2 + W_2(\nu,\chi)^2\right)$$

Solving this optimization problem is often complicated, but one may get an approximate solution $\widehat{\chi}$ by first fixing an absolutely continuous $\rho \in \mathcal{P}_2(\mathbb{R}^d)$, then computing T_μ, T_ν and their average $\frac{1}{2}(T_\mu + T_\nu)$, and finally considering

$$\widehat{\chi} = \left(\frac{1}{2}(T_\mu + T_\nu)\right)_{\#} \rho.$$

Notice that $\frac{1}{2}(T_\mu + T_\nu)$ is simply the average of the initial tangent vectors giving rise to the geodesics from ρ to μ and ρ to ν. Then $\widehat{\chi}$, which is the endpoint of the geodesic with this tangent vector, is an approximation of the midpoint between μ and ν. It is also the midpoint of the so-called generalized Wasserstein geodesic (in the terminology of Ambrosio-Gigli-Savaré) between μ and ν defined as the curve

$$[0,1] \ni t \mapsto ((1-t)T_\mu + tT_\nu)_{\#}\, \rho \in \mathcal{P}_2(\mathbb{R}^d).$$

In case $\mu = \rho$, the generalized geodesic between μ and ν coincides with the Wasserstein geodesic between μ and ν.

More generally, since it is often difficult to perform computations in Wasserstein spaces, which are curved (and infinite dimensional), it is a current practice in applications to first make computations in the Hilbert space $L^2(\rho)$, i.e., on the side of T_μ, before pushing forward ρ by the result of the computations in $L^2(\rho)$.

1.5 Bibliographical notes

§1.1: There are many great books about optimal transport, in particular: the two books by Villani [71] and [72], the one by Santambrogio for "applied mathematicians" [66], the book by Peyré-Cuturi about computational aspects of optimal transport [64], and the very recent book by Chewi-Niles Weed-Rigollet about statistical optimal transport [23]. To write the present text, I also took inspiration from lecture notes by Quentin Mérigot at IHP, available on his webpage, and from the PhD thesis of Delalande [29]. Brenier presented his theorem in a short note [14] and gave details in an extended paper [15].

§1.2: The proof of Proposition 1.2 can be found in [72, Theorem 5.20]. Proposition 1.3 is a consequence of [15, Theorem 1.3] together with [72, Theorem 6.9], at least when $\mathcal{X}$ is smooth, bounded, and ρ is bounded above and below on $\mathcal{X}$ by positive constants. The idea of the proof we provide was kindly communicated to us by Guillaume Carlier. The impossibility of embedding the Wasserstein space in Hilbert and Banach spaces is studied for instance in [4]. The precise statement is the following: if $p > 1$ and $d \geq 3$, then $\mathcal{P}_p(\mathbb{R}^d)$ does not admit a coarse embedding into any Banach space of nontrivial type, and in particular does not admit a coarse embedding into Hilbert space.

§1.3: For a more complete treatment of the 1d case, see Chapter 2 in Santambrogio's book [66].

§1.4: Wasserstein geodesics are the main subject of the book by Ambrosio-Gigli-Savaré [3]. For a quick view on the subject, see [66, Chapter 5.4]. The interpretation of W_2 as a (pseudo) Riemannian manifold is due to Otto [63], who used it to study the long-time behavior of the porous medium equation. McCann introduced the concept of displacement interpolation in [57].

The paper [73] introduced the linearized optimal transport distance $W_{2,\rho}$ defined in (1.19) and used it for pattern recognition in images. Since then, this distance has been used for instance to perform super-resolution of highly corrupted images [51] and to detect and visualize phenotypic differences between classes of cells [7].

Wasserstein barycenters have been introduced in [1], generalizing the concept of displacement interpolation of McCann. This notion of barycenter has found many successful applications, for instance in image processing [65], geometry processing [69], statistics [68] or machine learning [28]. The book chapters [64, Chapter 9.2], [23, Chapter 8] survey the topic.

2 Main results

In this chapter, we state the main results which will be covered in these lecture notes.

2.1 Warm-up: stability around regular optimal transport maps

The earliest quantitative stability result for optimal transport maps, due to Gigli, addressed stability in the vicinity of a sufficiently regular map.

Theorem 2.1 (Gigli, Stability near regular OT maps)**.** *Let ρ be a probability measure on $\mathbb{R}^d$, absolutely continuous with respect to the Lebesgue measure, and with compact support. Let $\mathcal{Y} \subset \mathbb{R}^d$ be compact, and $K > 0$. Let $\mu, \nu \in \mathcal{P}(\mathcal{Y})$. If the optimal transport map T_μ from ρ to μ is K-Lipschitz, then*

$$\|T_\mu - T_\nu\|_{L^2(\rho)} \leq C W_1(\mu, \nu)^{1/2}$$

where $C = (2K \operatorname{diam}(\operatorname{supp}(\rho)))^{1/2}$.

We provide a complete proof of Theorem 2.1 in Section 3.2, close in spirit to the other proofs presented in these notes. This is not the original proof of Gigli.

The important weakness of Theorem 2.1 is that the assumption that T_μ is K-Lipschitz is very strong. First, it implies that the support of μ is connected. Second, to prove that T_μ is Lipschitz one has to invoke the regularity theory for optimal transport maps, which requires very strong assumptions on μ. The Lipschitz regularity of the optimal transport map, studied by many authors starting with Caffarelli, is indeed only known under restrictive assumptions: Caffarelli proved this property under the assumption that the source and target measure have bounded support, are bounded above and below by positive constants on their support, and that the support of the target is convex; since this seminal result, some improvements and extensions have been obtained, but the spirit remains the same. And it is also known that continuity of the optimal transport map fails in some cases, even when the target has connected support: Caffarelli gave the example of a source measure ρ supported on a 2d domain $\mathcal{X}$ obtained by connecting two half disks by a thin corridor.

There is a whole line of research, notably in the statistical optimal transport community (see Section 8.1), working under this kind of regularity assumptions on T_μ. They have established stronger stability results (in terms of exponents) than what we present in these notes. For instance, under the assumption that T_μ is bi-Lipschitz, it is known that $\|T_\mu - T_\nu\|_{L^2(\rho)} \lesssim W_2(\mu, \nu)$, where the hidden constant depends on the Lipschitz constants of T_μ and T_μ^{-1}. We shall explain a bit the proof techniques in Section 8.1.

2.2 Main results

The discussion of the previous paragraph motivates us to look for results in which much weaker assumptions are made on the measures, than those ensuring regularity of the optimal transport map. Our main results state various assumptions on ρ under which we are able to prove quantitative stability inequalities of the form (1.11)-(1.13), with nearly no assumption on the target measures μ and ν. The discussion about the sharpness of these assumptions and the resulting stability inequalities is pretty long, and therefore we decided to devote Section 5 to this subject (see also a preliminary example in Section 2.4). In a nutshell, let us already mention that

the results presented in these notes are pretty sharp for Kantorovich potentials, but we still have less understanding of the stability of optimal transport maps.

The field is progressing fast. Our understanding so far is that stability of Kantorovich potentials is related to some Poincaré inequality on ρ, while stability of optimal transport maps should hold under weaker (but still mysterious) assumptions. Notice that if the Poincaré inequality holds for ρ, then

$$\|\phi_\mu - \phi_\nu\|_{L^2(\rho)} \leq C\|T_\mu - T_\nu\|_{L^2(\rho)}. \tag{2.1}$$

Hence any stability inequality on optimal transport maps immediately implies a stability inequality on Kantorovich potentials! However, with our present knowledge, we are not able to prove stability inequalities on optimal transport maps directly, except under regularity assumptions as in Theorem 2.1. Therefore, we will have to proceed differently.

The first main result we discuss in these notes is the following:

Theorem 2.2 (Log-concave case)**.** *Let $\rho = e^{-U-F}$ be a probability density on $\mathbb{R}^d$, with $D^2U \geq \kappa\mathrm{Id}$, $\kappa > 0$, and $F \in L^\infty(\mathbb{R}^d)$. Then for any compact set $\mathcal{Y}$, there exists $C > 0$ such that for any μ, ν supported in $\mathcal{Y}$,*

$$\|\phi_\mu - \phi_\nu\|_{L^2(\rho)} \leq CW_1(\mu,\nu)^{1/2}(1 + |\log W_1(m,\nu)|^{1/2}). \tag{2.2}$$

If moreover $D^2U \leq \kappa'\mathrm{Id}$, then there exists $C > 0$ such that for any μ, ν supported in $\mathcal{Y}$,

$$\|T_\mu - T_\nu\|_{L^2(\rho)} \leq CW_1(\mu,\nu)^{1/9}. \tag{2.3}$$

Up to the logarithmic loss in (2.2), the inequality (2.2) is sharp, as discussed in Section 5.2. The additional assumption $D^2U \leq \kappa'\mathrm{Id}$ made to prove (2.3) is probably only technical, but we have not been able to avoid it.

Let us turn to the second main result of these notes, which handles the case of source measures ρ with bounded support. Recall that a domain is a non-empty, bounded and connected open set.

Theorem 2.3 (Non-degenerate densities on bounded domains)**.** *Let ρ be a probability density on a John domain $\mathcal{X} \subset \mathbb{R}^d$, and assume that ρ is bounded above and below on $\mathcal{X}$ by positive constants. Then for any compact set $\mathcal{Y} \subset \mathbb{R}^d$, there exists $C > 0$ such that for any $\mu, \nu \in \mathcal{P}(\mathcal{Y})$,*

$$\|\phi_\mu - \phi_\nu\|_{L^2(\rho)} \leq CW_1(\mu,\nu)^{1/2}. \tag{2.4}$$

If moreover $\partial\mathcal{X}$ has a finite $(d-1)$-dimensional Hausdorff measure, then there exists $C > 0$ such that for any $\mu, \nu \in \mathcal{P}(\mathcal{Y})$,

$$\|T_\mu - T_\nu\|_{L^2(\rho)} \leq CW_1(\mu,\nu)^{1/6}. \tag{2.5}$$

We do not know whether the assumption that $\partial\mathcal{X}$ has finite $(d-1)$-dimensional Hausdorff measure is technical or not. John domains are a vast family of domains which contains in particular all bounded connected Lipschitz domains, but also some fractal domains like the Koch snowflake.

Definition 2.4. *A bounded open subset $\mathcal{X}$ of a metric space is called a John domain if there exist $x_0 \in \mathcal{X}$ and a constant $\eta > 0$ such that, for every $x \in \mathcal{X}$, there is $T > 0$ and a rectifiable curve $\gamma : [0,T] \to \mathcal{X}$ parametrized by the arclength (and whose length T depends on x) such that $\gamma(0) = x$, $\gamma(T) = x_0$, and for any $t \in [0,T]$,*

$$\mathrm{dist}(\gamma(t), \mathcal{X}^c) \geq \eta t \tag{2.6}$$

where $\mathcal{X}^c$ denotes the complement of $\mathcal{X}$.

In Theorem 2.3, the target measures are assumed to be supported in a large compact set $\mathcal{Y}$; it is possible to relax this assumption, and work only under moment constraints, as done in [31].

Theorem 2.3 also holds when $\mathbb{R}^d$ is replaced by an arbitrary smooth connected Riemannian manifold M, and optimal transport is considered with respect to the quadratic cost $c(x, y) = \frac{1}{2}\mathrm{dist}(x, y)^2$ where dist denotes the Riemannian distance on M. In case M is compact without boundary (e.g., the sphere), then we may choose $\mathcal{X} = \mathcal{Y} = M$. We shall detail a bit more this generalization to Riemannian manifolds in Section 7.3.

2.3 Comments

There are two important directions to improve and generalize the results presented above:

- proving/disproving stability inequalities for a wider range of probability densities ρ
- improving the stability exponents (1/9 in (2.3), 1/6 in (2.5)) for the source measures ρ considered in our main results.

To make progress on the second direction, which is blocked at the time of writing, new ideas are needed. Therefore, we comment only on the first direction. Indeed, our proof strategy is robust enough to handle other cases of interest. In all the following cases we are able to prove stability inequalities for Kantorovich potentials and optimal transport maps (we do not discuss stability exponents here, they are all dimension-free except for (2.7)):

(i) Degenerate densities ρ in bounded domains. The assumption in Theorem 2.3 that ρ is bounded above and below on $\mathcal{X}$ is not always necessary. We illustrate this on two examples which we find particularly relevant in applications. The first example is given by source probability densities satisfying

$$c_1 \mathrm{dist}(x, \partial\mathcal{X})^\delta \leq \rho(x) \leq c_2 \mathrm{dist}(x, \partial\mathcal{X})^\delta$$

for some $\delta > -1$ and $c_1, c_2 > 0$, when $\mathcal{X}$ is a bounded Lipschitz domain. These densities blow-up or decay near $\partial\mathcal{X}$. The second example is the source probability density

$$\rho(x) = \frac{c_d}{|x|^{d-1}} \mathbf{1}_{B(0,1)} \tag{2.7}$$

on $\mathbb{R}^d$, with c_d is a normalising constant. This probability density is sometimes called the spherical uniform distribution, and has been used in the literature to define multivariate quantiles. The stability inequality is relevant in this application, see Section 8.3.

(ii) Source measures ρ on $\mathbb{R}^d$ which decay polynomially at infinity:

$$\rho(x) = f(x)(1 + |x|)^{-\beta} \tag{2.8}$$

with $0 < m \leq f(x) \leq M < +\infty$ uniformly over $x \in \mathbb{R}^d$, and $\beta > d + 2$ so that ρ has finite second moment. The reason why we find this family of source probability measures interesting is that it is not possible to use the same proof strategy as for the families of probability measures covered by Theorems 2.2 and 2.3, see Section 4.2.

(iii) Source measures with disconnected support. If we replace the beginning of the statement of Theorem 2.3 by "Let ρ be a probability density on a finite union of John domains", then (2.5) still holds. Some modified version of (2.4) also holds, but one needs to be careful since Kantorovich potentials are not unique when the support of ρ is not connected.

Regarding the fact that the targets are assumed to be compactly supported in Theorems 2.2 and 2.3, we do not believe that this is a fundamental assumption. In [31], the assumption that was used is that they have p-th moment for some $p > d$ (for $p < d$, there exist unbounded Brenier potentials). We guess that our proof techniques may also cover this case, but shall not pursue this here.

As we explain in Section 4.4, the strategy we use to prove Theorem 2.2, Theorem 2.3 and point (i) above allows us to recover the known fact that for any ρ satisfying the assumptions of one of these results, the Poincaré inequality holds:

$$\int_{\mathcal{X}} \left(f - \int_{\mathcal{X}} f \mathrm{d}\rho\right)^2 \mathrm{d}\rho \leq C \int_{\mathcal{X}} |\nabla f|^2 \mathrm{d}\rho \tag{2.9}$$

(for ρ as in Theorem 2.2, $\mathcal{X}$ has to be replaced by $\mathbb{R}^d$). The examples and counterexamples of Section 5 show an analogy, but not an equivalence, between the fact that the Poincaré inequality holds for ρ and the fact that a stability inequality for Kantorovich potentials holds.

2.4 An elementary example

In this paragraph, we show on a simple example that one cannot hope in general to have a better exponent than 1/2 in (2.5).

Let $\rho = \rho(x)dx = \frac{1}{\pi}\mathbf{1}_{\mathbb{D}}(x)dx$ is the uniform probability on the unit disk $\mathbb{D} \subset \mathbb{R}^2$. For $\theta \in \mathbb{R}/2\pi\mathbb{Z}$, we set $x_\theta = (\cos(\theta), \sin(\theta)) \in \mathbb{R}^2$ and define the probability measure

$$\mu_\theta = \frac{1}{2}(\delta_{x_\theta} + \delta_{x_{\theta+\pi}}).$$

The ρ-a.e. unique optimal transport map T_{μ_θ} from ρ to μ_θ for the quadratic cost is explicit:

$$T_{\mu_\theta}(x) = \begin{cases} x_\theta & \text{if } \langle x, x_\theta \rangle \geq 0 \\ x_{\theta+\pi} & \text{if } \langle x, x_\theta \rangle < 0 \end{cases}$$

for $x \in \mathbb{D}$. In other words, each point $x \in \mathbb{D}$ is sent to the closest point among x_θ and $x_{\theta+\pi}$. This cuts the disk into two (equal) halves, see Figure 1.

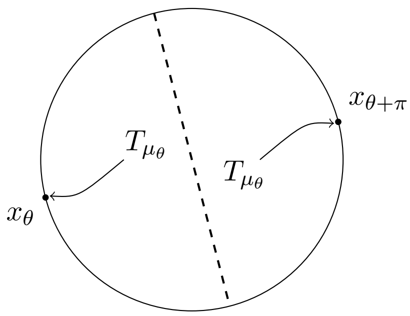

Figure 1: The optimal transport T_{μ_θ} from ρ to μ_θ.

Fix $\theta \in \mathbb{R}/2\pi\mathbb{Z}$, close to 0. Then, $\mathbb{D}$ may be written as $\mathbb{D} = A \sqcup B$ where A is the set of points whose images under T_{μ_0} and T_{μ_θ} are at angular distance θ, and B is the set of points whose images under T_{μ_0} and T_{μ_θ} are at angular distance $\pi - \theta$. We find $\rho(A) = 1 - \frac{\theta}{\pi}$ and $\rho(B) = \frac{\theta}{\pi}$, hence as $\theta \to 0$,

$$\|T_{\mu_\theta} - T_{\mu_0}\|^2_{L^2(\rho)} = |2\sin(\theta/2)|^2\rho(A) + |2\sin((\pi-\theta)/2)|^2\rho(B) \underset{\theta \to 0}{\sim} \frac{4|\theta|}{\pi}. \tag{2.10}$$

On the other hand, for θ close enough to 0 and $p \geq 1$ arbitrary, the W_p distance between μ_0 and μ_θ is obviously achieved by the map which sends x_0 to x_θ and x_π to $x_{\theta+\pi}$. Its p-cost is

$$W_p(\mu_0, \mu_\theta) = |2\sin(\theta/2)| \underset{\theta \to 0}{\sim} |\theta|. \tag{2.11}$$

Putting together (2.10) and (2.11) for $p = 2$, we see that we cannot hope in this case to have a better exponent than 1/2 in (2.5).

In this example, it is not difficult either to compute the difference in L^2-norm between Kantorovich potentials. For this, we denote by $D_\theta \subset \mathbb{R}^2$ the line through the origin which is perpendicular to the segment $[x_\theta, x_{\theta+\pi}]$ (the dashed line on Figure 1) and observe that

$$\phi_{\mu_\theta}(x) = \mathrm{dist}(x, D_\theta) - C$$

for some constant C independent of θ (simply equal to the integral of $\mathbb{D} \ni (x, y) \mapsto |x|/\pi$). It is not difficult to see that

$$\|\phi_{\mu_\theta} - \phi_{\mu_0}\|^2_{L^2(\rho)} = \int_{\mathbb{D}} (|x_1 \cos(\theta) + x_2 \sin(\theta)| - |x_1|)^2 \mathrm{d}x = \theta^2 \int_{\mathbb{D}} x_2^2 \mathrm{d}x + O(\theta^3)$$

where $x = (x_1, x_2)$. Therefore, one cannot hope in this case to have a better exponent than 1 in (2.4).

The computations presented above can easily be generalized to any dimension and more general sources than the uniform probability on the disk. Further examples where explicit computations can be made will be discussed in Section 5.

2.5 Bibliographical notes

§2.1: Theorem 2.1 is due to [36] and another proof has been given in [59]. The regularity theory of the Monge-Ampère equation and its link to regularity of optimal transport maps is explained in the survey [32]. The counterexample to the continuity of the optimal transport map is due to Caffarelli, see [18].

§2.2: Berman [8] was the first to obtain quantitative stability estimates without assuming regularity of the OT map. He derived dimension-dependent stability exponents for ρ bounded above and below on a compact, convex domain, using complex geometry. Then, Delalande and Mérigot [31] improved his stability exponent, making it dimension-free, under the same assumptions on ρ. But more importantly, they introduced a robust proof technique based on the study of the Kantorovich functional, see Chapter 3.

John domains were named in honor of F. John who introduced them in his work on elasticity [45]; Martio and Sarvas [56] introduced this terminology. They appear also in the theory of quasi-conformal mappings and in geometric measure theory.

§2.4: The example in this section is due to [59].

3 The Kantorovich functional

In this chapter, we introduce properly the Kantorovich relaxation of the Monge problem and its dual formulation in terms of Kantorovich potentials. Then, we study the so-called Kantorovich functional, whose strong convexity implies the stability of Kantorovich potentials.

3.1 The dual formulation of optimal transport

Monge formulated in 1781 the optimal transport problem as

$$\inf_{\substack{S:\mathcal{X}\to\mathcal{Y}\\ S_{\#}\rho=\mu}} \int_{\mathcal{X}} |x - S(x)| \mathrm{d}\rho(x) \tag{3.1}$$

where $\mathcal{X}, \mathcal{Y} \subset \mathbb{R}^d$, ρ is a probability measure on $\mathcal{X}$, and μ a probability measure on $\mathcal{Y}$. For many reasons already explained, it is natural to put a square on the $|x - S(x)|$ term, thus yielding (1.4). It is only in 1942 that Kantorovich introduced what is now known as the Kantorovich relaxation already mentioned in (1.3) and which we recall:

$$\inf_{\gamma\in\Pi(\rho,\mu)} \int_{\mathcal{X}\times\mathcal{Y}} |x-y|^2 \mathrm{d}\gamma(x,y). \tag{3.2}$$

We show in this paragraph how to solve (3.2), and introduce along the way the primal and dual Kantorovich potentials, which play a prominent role in these notes. At some places, we remain at an informal level; complete references are given at the end of the section.

First we notice that for $\gamma \in \Pi(\rho,\mu)$,

$$\int_{\mathcal{X}\times\mathcal{Y}} |x-y|^2 \mathrm{d}\gamma(x,y) = \int_{\mathcal{X}} |x|^2 \mathrm{d}\rho(x) + \int_{\mathcal{Y}} |y|^2 \mathrm{d}\mu(y) - 2\int_{\mathcal{X}\times\mathcal{Y}} \langle x,y\rangle \ \mathrm{d}\gamma(x,y).$$

Since the first two terms in the right-hand side do not depend on γ, the quadratic optimal transport problem (1.3) is equivalent to

$$\sup_{\gamma\in\Pi(\rho,\mu)} \int_{\mathcal{X}\times\mathcal{Y}} \langle x,y\rangle \ \mathrm{d}\gamma(x,y). \tag{3.3}$$

We denote this supremum by $\mathcal{I}(\rho,\mu)$.

3.1.1 The dual problem

The constraint $\gamma \in \Pi(\rho,\mu)$ may be written under the form of Lagrange multipliers. Let $\mathcal{M}_+(\mathcal{X}\times\mathcal{Y})$ denote the set of positive Radon bounded measures supported over $\mathcal{X}\times\mathcal{Y}$, and let $C_b(E)$ denote the set of bounded continuous functions over a set E. For $\gamma \in \mathcal{M}_+(\mathcal{X}\times\mathcal{Y})$.

$$\inf_{\substack{\phi\in C_b(\mathcal{X})\\ \psi\in C_b(\mathcal{Y})}} \int_{\mathcal{X}} \phi(x) \ \mathrm{d}\rho(x) + \int_{\mathcal{Y}} \psi(y) \ \mathrm{d}\mu(y) - \int_{\mathcal{X}\times\mathcal{Y}} \phi\oplus\psi \ \mathrm{d}\gamma(x,y) = \begin{cases} 0 & \text{if } \gamma\in\Pi(\rho,\mu) \\ -\infty & \text{otherwise} \end{cases}$$

where $\phi\oplus\psi : (x,y) \mapsto \phi(x) + \psi(y)$. Therefore (3.3) is equivalent to

$$\sup_{\gamma\in\mathcal{M}_+(\mathcal{X}\times\mathcal{Y})} \inf_{\substack{\phi\in C_b(\mathcal{X})\\ \psi\in C_b(\mathcal{Y})}} \int_{\mathcal{X}\times\mathcal{Y}} (\langle x,y\rangle - \phi\oplus\psi) \ \mathrm{d}\gamma(x,y) + \int_{\mathcal{X}} \phi(x) \ \mathrm{d}\rho(x) + \int_{\mathcal{Y}} \psi(y) \ \mathrm{d}\mu(y).$$

The duality principle consists in exchanging the sup and the inf; we get

$$\inf_{\substack{\phi\in C_b(\mathcal{X})\\ \psi\in C_b(\mathcal{Y})}} \left[\int_{\mathcal{X}} \phi(x)\ \mathrm{d}\rho(x) + \int_{\mathcal{Y}} \psi(y)\ \mathrm{d}\mu(y) + \sup_{\gamma\in\mathcal{M}_+(\mathcal{X}\times\mathcal{Y})} \int_{\mathcal{X}\times\mathcal{Y}} (\langle x,y\rangle - \phi\oplus\psi)\ \mathrm{d}\gamma(x,y)\right] \tag{3.4}$$

(after rearranging the terms). As we will see, this new optimization problem is equivalent to (3.3), which may seem surprising at first sight. But let us first simplify (3.4) a bit. The supremum inside the brackets can itself be seen as a constraint: it is equal to 0 if $\phi\oplus\psi \geq \langle x,y\rangle$ for any $x\in\mathcal{X}, y\in\mathcal{Y}$, and equal to $+\infty$ otherwise. Therefore we end-up with the dual problem

$$\inf\left\{\int_{\mathcal{X}} \phi(x)\ \mathrm{d}\rho(x) + \int_{\mathcal{Y}} \psi(y)\ \mathrm{d}\mu(y) \mid \phi\in C_b(\mathcal{X}), \psi\in C_b(\mathcal{Y}), \phi\oplus\psi\geq\langle x,y\rangle\right\}. \tag{3.5}$$

We denote this infimum by $\mathcal{J}(\rho,\mu)$.

It is immediate to check that $\mathcal{J}(\rho,\mu)\geq\mathcal{I}(\rho,\mu)$. Indeed, for any ϕ,ψ such that $\phi\oplus\psi\geq\langle x,y\rangle$, and any $\gamma\in\Pi(\rho,\mu)$ we have

$$\int_{\mathcal{X}\times\mathcal{Y}} \langle x,y\rangle\ \mathrm{d}\gamma(x,y) \leq \int_{\mathcal{X}\times\mathcal{Y}} \phi\oplus\psi\ \mathrm{d}\gamma(x,y) = \int_{\mathcal{X}} \phi(x)\ \mathrm{d}\rho(x) + \int_{\mathcal{Y}} \psi(y)\ \mathrm{d}\mu(y).$$

What is more surprising is that the reverse inequality $\mathcal{J}(\rho,\mu)\leq\mathcal{I}(\rho,\mu)$ is also true, but we shall not show it here. The resulting equality

$$\mathcal{I}(\rho,\mu) = \mathcal{J}(\rho,\mu) \tag{3.6}$$

is often referred to as the "Kantorovich duality", or "strong duality". It holds in great generality, for general costs for instance.

The problem (3.3) has a maximizer under mild conditions (see [72, Theorem 5.10]), and the standard proof of this fact goes as follows: consider a maximizing sequence (γ_n), use compactness in the set of probability measures (Prokhorov theorem) to extract a converging subsequence, and conclude that the limit is a maximizer using lower semi-continuity of the cost. Similarly, the problem (3.5) admits a minimizer under mild assumptions, and a pair (ϕ,ψ) which minimizes (3.5) is called a pair of Kantorovich potentials. But for both (3.3) and (3.5), the optimizers are not necessarily unique.

3.1.2 Support of optimizers

Let us prove that if γ is a maximizer in (3.3) and (ϕ,ψ) is a minimizer in (3.5), then

$$\mathrm{support}(\gamma) \subset \{(x,y)\in\mathcal{X}\times\mathcal{Y} \mid \phi(x)+\psi(y) = \langle x,y\rangle\}. \tag{3.7}$$

For this we use (3.6) which yields

$$\int_{\mathcal{X}\times\mathcal{Y}} ((\phi\oplus\psi)(x,y) - \langle x,y\rangle)\ \mathrm{d}\gamma(x,y) = 0.$$

Since $\phi\oplus\psi\geq\langle x,y\rangle$, this gives (3.7). The converse is true: if (3.7) holds for some $\gamma\in\Pi(\rho,\mu)$ and (ϕ,ψ) such that $\phi\oplus\psi\geq\langle x,y\rangle$, then γ and (ϕ,ψ) are solutions of their respective optimization problems.

3.1.3 Semi-dual formulation

It is possible to give an equivalent unconstrained formulation of (3.5). Recall the definition of the Legendre transform:

$$\phi^*(y) = \sup_{x\in\mathbb{R}^d} \langle y, x\rangle - \phi(x).$$

If a function ϕ is defined only over a subset of $\mathbb{R}^d$, we first extend ϕ by $+\infty$ outside this subset to define its Legendre transform. As a consequence, $\phi^*(y) = \sup_{x\in\mathcal{X}} \langle y, x\rangle - \phi(x)$ if $\phi \in C_b(\mathcal{X})$.

For a given ϕ, the smallest possible ψ that one may choose to satisfy the constraints in (3.5) is $\psi = \phi^*$. Similarly, for a given ψ, the smallest possible ϕ that one may choose is $\phi = \psi^*$. Therefore, one has

$$\mathcal{J}(\rho,\mu) = \inf\left\{\int_{\mathcal{X}} \phi \,\mathrm{d}\rho + \int_{\mathcal{Y}} \phi^* \,\mathrm{d}\mu \mid \phi \in C_b(\mathcal{X})\right\} = \inf\left\{\int_{\mathcal{X}} \psi^* \,\mathrm{d}\rho + \int_{\mathcal{Y}} \psi \,\mathrm{d}\mu \mid \psi \in C_b(\mathcal{Y})\right\}.$$

This leads us to the so-called semi-dual formulation of (3.5):

$$\inf_{\psi\in C^0(\mathcal{Y})} \int_{\mathcal{X}} \psi^* \,\mathrm{d}\rho + \int_{\mathcal{Y}} \psi \,\mathrm{d}\mu \tag{3.8}$$

(recall that $\mathcal{Y}$ is compact, so $C^0(\mathcal{Y}) = C_b(\mathcal{Y})$).

If we want to solve (3.8), it seems natural to write the first and second-order optimality conditions with respect to ψ. This will be done in Section 3.3: the second integral is linear in ψ, hence easy to differentiate, but the first part is non-linear in ψ.

3.1.4 Convex functions and proof of Brenier's theorem

For any ϕ, ψ such that $\phi \oplus \psi \geq \langle x, y\rangle$, we have $\psi \geq \phi^*$ and $\phi \geq \phi^{**}$, hence

$$\int_{\mathcal{X}} \phi \,\mathrm{d}\rho + \int_{\mathcal{Y}} \psi \,\mathrm{d}\mu \geq \int_{\mathcal{X}} \phi^{**} \,\mathrm{d}\rho + \int_{\mathcal{Y}} \phi^* \,\mathrm{d}\mu.$$

Recall that a convex function is called proper if it has a non-empty domain, it never takes on the value $-\infty$ and also it is not identically equal to $+\infty$. By the Fenchel-Moreau theorem, as soon as ϕ is a proper lower semi-continuous convex function, $\phi^{**} = \phi$. Therefore, in (3.5) we may restrict the infimum to the set of pairs (ϕ, ϕ^*) of proper lower semi-continuous conjugate functions on $\mathbb{R}^d$. We do not discuss here the fact that in this case ϕ and ϕ^* cannot be bounded (unless constant) since they are convex, hence strictly speaking one would need to justify that the result is the same as in (3.5) where the infimum is restricted to bounded continuous functions.

We recall that if $\phi : \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$ is a convex function, then its subdifferential at $x \in \mathbb{R}^d$ is defined as

$$\partial\phi(x) = \{v \in \mathbb{R}^d \mid \forall z \in \mathbb{R}^d,\ \phi(z) \geq \phi(x) + \langle z - x, v\rangle\}.$$

The graph of the subdifferential is

$$\partial\phi = \bigcup_{x\in\mathbb{R}^d} \{x\} \times \partial\phi(x).$$

If $\phi : \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$ is a proper lower semi-continuous convex function, then for all $x, y \in \mathbb{R}^d$,

$$\phi(x) + \phi^*(y) = \langle x, y\rangle \Leftrightarrow y \in \partial\phi(x) \Leftrightarrow x \in \partial\phi^*(y). \tag{3.9}$$

Indeed,

$$\begin{aligned}\phi(x) + \phi^*(y) = \langle x, y\rangle &\Leftrightarrow \langle x, y\rangle \geq \phi(x) + \phi^*(y)\\ &\Leftrightarrow \forall z \in \mathbb{R}^d, \quad \langle x, y\rangle \geq \phi(x) + \langle y, z\rangle - \phi(z)\\ &\Leftrightarrow \forall z \in \mathbb{R}^d, \quad \phi(z) \geq \phi(x) + \langle y, z - x\rangle\\ &\Leftrightarrow y \in \partial\phi(x).\end{aligned}$$

By symmetry, the other equivalence follows, since $\phi^{**} = \phi$ by the Fenchel-Moreau theorem (which uses the assumptions on ϕ).

Combining (3.9) and (3.7), we have obtained:

Proposition 3.1. *Any optimal transport plan (i.e., any solution of* (3.3)*) has its support contained in the graph of the subdifferential of a convex function.*

The proof of Brenier's theorem 1.1 is now straightforward.

Proof of Brenier's theorem. Let γ be an optimal transport plan, and take ψ a proper lower semi-continuous function on $\mathbb{R}^d$ that solves (3.8). For any (x_0, y_0) in the support of γ such that ψ^* is differentiable at x_0, one has $y_0 = \nabla\psi^*(x_0)$ according to Proposition 3.1. Since ψ^* is a proper lower semi-continuous convex function on $\mathbb{R}^d$, it is differentiable ρ-almost everywhere. This implies that $\gamma = (\mathrm{Id}, \nabla\psi^*)_{\#}\rho$. But ψ and γ have been chosen independently, so for any other optimal transport plan $\widetilde{\gamma}$ there also holds $\widetilde{\gamma} = (\mathrm{Id}, \nabla\psi^*)_{\#}\rho$, in other words there is a unique optimal transport plan γ. If f is another proper lower semi-continuous convex function such that $\mu = (\nabla f)_{\#}\rho$, then $\widetilde{\gamma} = (\mathrm{Id}, \nabla f)$ satisfies

$$\begin{aligned}\int_{\mathbb{R}^d\times\mathbb{R}^d} \langle x, y\rangle \,\mathrm{d}\gamma(x, y) \leq \int_{\mathbb{R}^d} f(x)\,\mathrm{d}\rho(x) + \int_{\mathbb{R}^d} f^*(y)\,\mathrm{d}\mu(y) &= \int_{\mathbb{R}^d} (f(x) + f^*(\nabla f(x)))\,\mathrm{d}\rho(x)\\ &= \int_{\mathbb{R}^d\times\mathbb{R}^d} \langle x, y\rangle \,\mathrm{d}\tilde{\gamma}(x, y)\end{aligned}$$

where the last equality comes from the Fenchel-Young equality case (3.9). Since γ is the unique maximizer of (3.3), we get $\widetilde{\gamma} = \gamma$, and thus $\nabla f = \nabla\psi^*$ ρ-a.e. □

3.1.5 Kantorovich-Rubinstein formula

We conclude this section with the following important formula.

Theorem 3.2 (Kantorovich-Rubinstein duality formula)**.** *For any μ, ν probability measures on $\mathcal{Y}$ with finite first moment,*

$$W_1(\mu, \nu) = \sup\left\{\int_{\mathcal{Y}} f\mathrm{d}\mu - \int_{\mathcal{Y}} f\mathrm{d}\nu \mid \mathrm{Lip}(f) \leq 1\right\}. \tag{3.10}$$

In these notes, we will only need the inequality $\geq$, which is easy to prove. For any $\gamma \in \Pi(\mu, \nu)$ and any 1-Lipschitz function f,

$$\int_{\mathcal{Y}} f\,\mathrm{d}\mu - \int_{\mathcal{Y}} f\,\mathrm{d}\nu = \int_{\mathcal{X}\times\mathcal{Y}} (f(x) - f(y))\,\mathrm{d}\gamma(x, y) \leq \int_{\mathcal{X}\times\mathcal{Y}} |x - y|\,\mathrm{d}\gamma(x, y).$$

Taking the infimum over $\gamma \in \Pi(\mu, \nu)$ and the supremum over 1-Lipschitz functions f, we get that in (3.10) the left-hand side is larger than the right-hand side. For the converse inequality, which is another instance of Kantorovich duality, we provide references at the end of this section.

3.2 Proof of Theorem 2.1

With the Kantorovich potentials at hand, we will be able to give in this section a short proof of Theorem 2.1. We start with a classical result which turns the Lipschitzness assumption on T_μ into strong convexity of ψ_μ. For this we need the following definitions.

Definition 3.3. *For $K > 0$, a function $\phi : \mathbb{R}^d \to \mathbb{R}$ is K-smooth if $\frac{K}{2}|x|^2 - \phi(x)$ is convex, i.e.,*

$$\phi((1-t)x_0 + tx_1) + \frac{Kt(1-t)}{2}|x_0 - x_1|^2 \geq (1-t)\phi(x_0) + t\phi(x_1),$$

for any $x_0, x_1 \in \mathbb{R}^d$ and $t \in [0,1]$. For $\lambda > 0$, a function $\psi : \mathbb{R}^d \to \mathbb{R}$ is λ-strongly convex if $\psi(y) - \frac{\lambda}{2}|y|^2$ is convex, i.e., for any $y_0, y_1 \in \mathbb{R}^d$ and $t \in [0,1]$,

$$\psi((1-t)y_0 + ty_1) + \frac{\lambda t(1-t)}{2}|y_0 - y_1|^2 \leq (1-t)\psi(y_0) + t\psi(y_1). \tag{3.11}$$

When $\phi, \psi \in C^2$, the first condition is simply $D^2\phi \leq K$, and the second condition is $D^2\psi \geq \lambda$. Now, if ϕ and ϕ^* are C^2 convex functions, differentiating the identity $\nabla\phi(\nabla\phi^*) = \mathrm{Id}$ (which comes from (3.9)) we get $(\nabla^2\phi^*)^{-1} = \nabla^2\phi(\nabla\phi^*)$, and thus $\|\nabla^2\phi\|_{L^\infty} \leq K$ if and only if $\nabla^2\phi^* \geq \frac{1}{K}$, i.e., ϕ is K-smooth if and only if ϕ^* is K^{-1}-strongly convex. This result is actually true without assuming that ϕ, ϕ^* are C^2 (see [55, Lemma 2.2] for a proof):

Lemma 3.4. *Let $\phi : \mathbb{R}^d \to \mathbb{R}$ be a convex function. Then ϕ is K-smooth if and only if ϕ^* is λ-strongly convex for $\lambda = K^{-1}$.*

In the sequel, if $f \in C^0(\mathcal{Y})$ and μ is a Radon measure on $\mathcal{Y}$, then we set $\langle f \mid \mu\rangle = \int_{\mathcal{Y}} f \mathrm{d}\mu$. The (distinct) notation $\langle\cdot,\cdot\rangle$ denotes the Euclidean scalar product in $\mathbb{R}^d$.

The proof of Theorem 2.1 mainly relies on the following inequality:

Lemma 3.5. *Under the assumptions of Theorem 2.1, there holds*

$$\langle \psi_\mu - \psi_\nu \mid \nu - \mu\rangle \geq \frac{1}{2K}\|T_\nu - T_\mu\|^2_{L^2(\rho)} \tag{3.12}$$

where (ϕ_μ, ψ_μ) (resp. (ϕ_ν, ψ_ν)) is a pair of Kantorovich potentials associated to the transport from ρ to μ (resp. ρ to ν).

Proof of Lemma 3.5. We take $\mathcal{X} = \mathbb{R}^d$. Using that $\mu = (T_\mu)_{\#}\rho$ and $\nu = (T_\nu)_{\#}\rho$ we have

$$\langle \psi_\mu \mid \nu - \mu\rangle = \int_{\mathbb{R}^d} (\psi_\mu(T_\nu(x)) - \psi_\mu(T_\mu(x)))\mathrm{d}\rho(x).$$

Since T_μ is K-Lipschitz, ϕ_μ is K-smooth and therefore $\psi_\mu = \phi_\mu^*$ is K^{-1}-strongly convex by Lemma 3.4. Hence, letting $t \to 0$ in (3.11), we get that $\psi_\mu(y) - \psi_\mu(z) \geq \langle y - z, v\rangle + \frac{1}{2K}|y - z|^2$ for any $v \in \partial\psi_\mu(z)$. Now we fix x in the support of ρ and choose $y = T_\nu(x)$ and $z = T_\mu(x)$. Therefore $z \in \partial\phi_\mu(x)$ (see Proposition 3.1) and (3.9) yields $x \in \partial\psi_\mu(z)$. We deduce

$$\langle \psi_\mu \mid \nu - \mu\rangle \geq \int_{\mathbb{R}^d} \Big(\langle T_\nu(x) - T_\mu(x), x\rangle + \frac{1}{2K}|T_\nu(x) - T_\mu(x)|^2\Big)\mathrm{d}\rho(x).$$

Since ψ_ν is also convex (but not necessarily strongly convex), choosing $y = T_\mu(x)$ and $z = T_\nu(x)$ we obtain similarly

$$\langle \psi_\nu \mid \mu - \nu\rangle \geq \int_{\mathbb{R}^d} \langle T_\mu(x) - T_\nu(x), x\rangle \, \mathrm{d}\rho(x).$$

Adding the two previous inequalities we get (3.12). □

End of the proof of Theorem 2.1. Using the Kantorovich-Rubinstein duality formula (3.10), we get that the left-hand side of (3.12) is bounded above by $\mathrm{Lip}(\psi_\mu - \psi_\nu)W_1(\mu,\nu)$. Finally, to conclude the proof, it remains to observe that $\psi_\mu - \psi_\nu$ is $\mathrm{diam}(\mathrm{supp}(\rho))$-Lipschitz. Essentially this is due to the fact that $\partial\psi_\mu$ and $\partial\psi_\nu$ are subsets of the support of ρ, and here is a formal proof: if $y, y' \in \mathcal{Y}$, let x, x' in the support of ρ such that $\psi_\mu(y) = \langle y, x\rangle - \phi_\mu(x)$ and $\psi_\nu(y') = \langle y', x'\rangle - \phi_\nu(x')$. These points exist due to (3.7) and (3.9). Then

$$\begin{aligned}(\psi_\mu - \psi_\nu)(y) - (\psi_\mu - \psi_\nu)(y') &= \psi_\mu(y) - \psi_\mu(y') + \psi_\nu(y') - \psi_\nu(y) \\ &\leq \langle y - y', x\rangle + \langle y' - y, x'\rangle \leq |y - y'|\mathrm{diam}(\mathrm{supp}(\rho)).\end{aligned}$$

Exchanging the roles of y and y' we get the Lipschitz bound and Theorem 2.1 follows. □

3.3 The Kantorovich functional: definition and derivatives

The Kantorovich functional is defined as

$$\mathcal{K}_\rho : \psi \mapsto \int_{\mathcal{X}} \psi^* \mathrm{d}\rho \tag{3.13}$$

for $\psi \in C^0(\mathcal{Y})$. We prove here (some kind of) strong convexity of this functional under some assumptions on ρ, and explain how it implies stability properties for Kantorovich potentials. The Kantorovich functional $\mathcal{K}_\rho$ is one part of the quantity appearing in the semidual formulation (3.8). The other part is linear in ψ and thus does not affect the convexity properties of $\mathcal{K}_\rho$. Also, it is immediate to see that $\mathcal{K}_\rho$ is convex, since it is a convex combination of the convex functions $\psi \mapsto \psi^*(x)$.

Let us explain on a basic example how one can deduce stability from strong convexity. Let $f : \mathbb{R}^d \to \mathbb{R}$ with $\mathrm{Hess}(f) \geq c\mathrm{Id}$ for some $c > 0$. Then for any $x_1, x_2 \in \mathbb{R}^d$,

$$c\|x_1 - x_2\|^2 \leq \langle \nabla f(x_1) - \nabla f(x_2), x_1 - x_2\rangle \leq |x_1 - x_2||\nabla f(x_1) - \nabla f(x_2)| \tag{3.14}$$

hence $|x_1 - x_2| \leq c^{-1}|\nabla f(x_1) - \nabla f(x_2)|$. In particular, if $\nabla f(x_1)$ and $\nabla f(x_2)$ are close to each other, then x_1 and x_2 are close to each other. In other words, strong convexity of f implies that the map $\nabla f(x) \mapsto x$ is well-defined and stable.

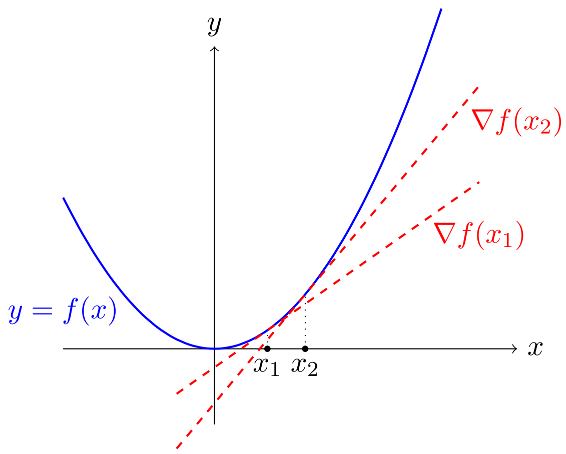

Figure 2: For a strongly convex function f, if $\nabla f(x_1)$ and $\nabla f(x_2)$ are close to each other, then x_1 and x_2 are close to each other.

To prove our main results, we shall develop an analogous computation for $f = \mathcal{K}_\rho$ the Kantorovich functional, defined on $C^0(\mathcal{Y})$ instead of $\mathbb{R}^d$. We will see that if $\mathcal{K}_\rho$ is taken as (3.13) below, then for $\psi \in C^0(\mathcal{Y})$ the gradient $\nabla \mathcal{K}_\rho(\psi)$ is a measure, and

$$\nabla \mathcal{K}_\rho(\psi) = -(\nabla \psi^*)_{\#}\rho. \tag{3.15}$$

In particular, $\nabla \mathcal{K}_\rho(\psi_\mu) = -\mu$ for ψ_μ the dual Kantorovich potential from ρ to μ. The above computations show (heuristically) that if one is able to prove that $\mathcal{K}_\rho$ is strongly convex in some sense, then one gets a stability inequality of the form

$$\mu, \nu \text{ close to each other} \Rightarrow \psi_\mu, \psi_\nu \text{ close to each other.}$$

If we consider again Lemma 3.5 under this new light, we see that it plays the role of the left-hand side inequality in (3.14). And the Kantorovich-Rubinstein duality formula replaces the Cauchy-Schwarz inequality used in (3.14), and yields an upper bound on

$$\langle \psi_\mu - \psi_\nu \mid \nu - \mu \rangle$$

since ψ_μ and ψ_ν are Lipschitz when $\mathcal{X}$ is bounded. However, this application of the Kantorovich-Rubinstein duality formula does not always yield sharp results, we shall comment on this again in Remark 3.10.

To study the convexity of $\mathcal{K}_\rho$, we compute its first two derivatives. The equality (3.16) below is a formal writing of (3.15).

Lemma 3.6. *Let $\phi_0, \phi_1 \in C^2(\mathbb{R}^d)$ be strongly convex functions. Define $\psi_0 = \phi_0^*$, $\psi_1 = \phi_1^*$, and $v = \psi_1 - \psi_0$. For $t \in [0,1]$, define $\psi_t = \psi_0 + tv$, and finally $\phi_t = \psi_t^*$. Then, ϕ_t is a strongly convex function, belongs to $C^2(\mathbb{R}^d)$, and*

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{K}_\rho(\psi_t) = -\int_{\mathcal{X}} v(\nabla\phi_t(x))\mathrm{d}\rho(x) \tag{3.16}$$

$$\frac{\mathrm{d}^2}{\mathrm{d}t^2}\mathcal{K}_\rho(\psi_t) = \int_{\mathcal{X}} \langle \nabla v(\nabla\phi_t(x)), D^2\phi_t(x)\cdot\nabla v(\nabla\phi_t(x))\rangle \mathrm{d}\rho(x). \tag{3.17}$$

Proof. The maximum in

$$\max_{y\in\mathcal{Y}} \langle x, y\rangle - \psi_t(y)$$

is attained at $y_x \in \mathcal{Y}$ for which $x = \nabla\psi_t(y_x)$, which is equivalent to $y_x = \nabla\psi_t^*(x)$ according to (3.9). Therefore, by the envelope theorem,

$$\psi_{t+\varepsilon}^*(x) = \max_{y\in\mathcal{Y}} \langle x, y\rangle - \psi_t(y) - \varepsilon v(y) = \psi^*(x) - \varepsilon v(\nabla\psi_t^*(x)) + o(\varepsilon) \tag{3.18}$$

as $\varepsilon \to 0$. In other words $\frac{\mathrm{d}}{\mathrm{d}t}\psi_t^*(x) = -v(\nabla\psi_t^*(x))$, and integrating against ρ we get (3.16).

For (3.17), applying (3.16) to ψ_t we see that we need to evaluate

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_{\mathcal{X}} v(\nabla\phi_t(x))\mathrm{d}\rho(x).$$

Using the chain rule and the fact that $\frac{\mathrm{d}}{\mathrm{d}t}\nabla\phi_t = \nabla\frac{\mathrm{d}}{\mathrm{d}t}\phi_t = -D^2\phi_t(x)\cdot\nabla v(\nabla\phi_t(x))$ due to (3.18), we get (3.17). □

3.4 Variance inequality in compact convex sets

In this section, we prove a "variance inequality", i.e., an upper bound on the variance of the difference of two Kantorovich potentials corresponding to two different target measures. It reflects a form of strong convexity of $\mathcal{K}_\rho$. The estimate (3.19) below is of fundamental importance in the rest of these notes.

Recall that a probability density σ on a convex set $Q \subset \mathbb{R}^d$ is called logarithmically concave, or log-concave, if there exists a convex function $V : Q \to \mathbb{R}$ such that $\sigma = e^{-V}$. The Hessian of V is denoted by D^2V. Also, recall that the variance of a function f with respect to a probability measure ρ on a set $\mathcal{X}$ is defined as

$$\mathrm{Var}_\rho(f) = \int_{\mathcal{X}} \left(f - \int_{\mathcal{X}} f \mathrm{d}\rho \right)^2 \mathrm{d}\rho.$$

Theorem 3.7. *Let $Q \subset \mathbb{R}^d$ be a compact convex set with non-empty interior, let σ be a log-concave probability density over Q and let ρ be another probability density over Q satisfying $m_\rho \sigma \leq \rho \leq M_\rho \sigma$ for some constants $M_\rho \geq m_\rho > 0$. Let $\mathcal{Y} \subset \mathbb{R}^d$ be a compact set, and let $R_{\mathcal{Y}} = \max_{y \in \mathcal{Y}} \|y\|$. Then, for all $\psi_0, \psi_1 \in \mathcal{C}^0(\mathcal{Y})$,*

$$e^{-1} \frac{m_\rho}{M_\rho} \frac{1}{R_{\mathcal{Y}} \mathrm{diam}(Q)} \mathrm{Var}_\rho(\psi_1^* - \psi_0^*) \leq \langle \psi_1 - \psi_0 \mid \nabla \psi_{0\#}^* \rho - \nabla \psi_{1\#}^* \rho \rangle. \tag{3.19}$$

An example to keep in mind is when σ is the characteristic function of Q, normalized to be a probability density. Another important example is when ρ itself is log-concave, in which case we may take $\sigma = \rho$ and $m_\rho / M_\rho = 1$.

Notice that the above estimate, in particular the constant in the left-hand side of (3.19), is dimension-free. Also, the inequality (3.19) is not exactly a strong convexity estimate on $\mathcal{K}_\rho$ since primal (and not dual) Kantorovich potentials appear in the LHS. In the original proof of [31], a true strong convexity estimate with dual potentials has been obtained. However it is not strong enough to imply Theorem 2.2, contrarily to (3.19).

The fundamental tool on which our proof of Theorem 3.7 relies is the Brascamp-Lieb inequality, which is a kind of Poincaré inequality with respect to log-concave densities.

Theorem 3.8 (Brascamp-Lieb inequality). *Let $\mathcal{X} \subset \mathbb{R}^d$ be a compact, convex set and let $\rho_0 = e^{-V} \mathrm{d}x$ be a probability measure on $\mathcal{X}$, where $V \in C^2(\mathcal{X})$ is assumed to be strictly convex. Then every smooth function f on $\mathcal{X}$ verifies*

$$\mathrm{Var}_{\rho_0}(f) \leq \int_{\mathcal{X}} \langle \nabla f, (D^2 V)^{-1} \nabla f \rangle \mathrm{d}\rho_0$$

We provide a proof of the Brascamp-Lieb inequality in Appendix B. The strength of this inequality is that the Poincaré constant is 1, for an arbitrary strictly convex $V \in C^2(\mathcal{X})$.

Sketch of proof of Theorem 3.7. Fix $\psi_0, \psi_1 \in \mathcal{C}^0(\mathcal{Y})$. Let $v = \psi_1 - \psi_0$ and $\psi_t = \psi_0 + tv = (1-t)\psi_0 + t\psi_1$ for $t \in [0,1]$. Set also $\phi_t = \psi_t^*$. In this sketch of proof, we assume that ϕ_t has all the nice properties which make the involved objects well-defined. The approximation arguments which allow to assume this are skipped. Then

$$\begin{aligned}
\langle \psi_1 - \psi_0 \mid \nabla \psi_{0\#}^* \rho - \nabla \psi_{1\#}^* \rho \rangle &= \frac{\mathrm{d}}{\mathrm{d}t} \mathcal{K}_\rho(\psi_t) \Big|_{t=1} - \frac{\mathrm{d}}{\mathrm{d}t} \mathcal{K}_\rho(\psi_t) \Big|_{t=0} \\
&= \int_0^1 \frac{\mathrm{d}^2}{\mathrm{d}t^2} \mathcal{K}_\rho(\psi_t) \mathrm{d}t \\
&= \int_0^1 \int_Q \langle \nabla v(\nabla \phi_t), D^2 \phi_t \cdot \nabla v(\nabla \phi_t) \rangle \mathrm{d}\rho \mathrm{d}t
\end{aligned}$$

according to Lemma 3.6. We introduce $w_t(x) = v(\nabla\phi_t(x))$, then $\nabla w_t = D^2\phi_t \cdot \nabla v(\nabla\phi_t)$, and we get

$$\int_Q \langle \nabla v(\nabla\phi_t), D^2\phi_t \cdot \nabla v(\nabla\phi_t)\rangle \mathrm{d}\rho = \int_Q \langle \nabla w_t, (D^2\phi_t)^{-1} \cdot \nabla w_t\rangle \mathrm{d}\rho \tag{3.20}$$

($D^2\phi_t$ is invertible because thanks to some regularization argument we may assume that ψ_0, ψ_1 are C^2).

There is a first approach to lower bound the last expression by directly applying Brascamp-Lieb (with $\rho_0 = e^{-\phi_t}$, properly normalized). This approach has the drawback that it yields a constant in (3.19) much worse than the one stated, in particular not good enough to prove Theorem 2.2.

To circumvent this, we write $\sigma = e^{-V}$ and set $\rho_t = Z_t^{-1} e^{-V-\phi_t}$ where Z_t is a normalizing constant, so that ρ_t is a probability measure. The idea is to apply Brascamp-Lieb with ρ_t. For this, we need to replace ρ by ρ_t in the RHS of (3.20). We denote by m_t and M_t the minimum and the maximum of $\phi_t(x)$ over $x \in Q$, and let $r = \sup_{t\in[0,1]} M_t - m_t$. We get

$$\begin{aligned}
\rho_t(x) &\geq Z_t^{-1} e^{-M_t}\sigma \geq Z_t^{-1} e^{-M_t} M_\rho^{-1}\rho, \\
\rho_t(x) &\leq Z_t^{-1} e^{-m_t}\sigma \leq Z_t^{-1} e^{-m_t} m_\rho^{-1}\rho.
\end{aligned}$$

In particular,

$$\mathrm{Var}_{\rho_t}(f) = \int_{\mathcal{X}} |f - \overline{f}|^2 \mathrm{d}\rho_t \geq \alpha \int_{\mathcal{X}} |f - \overline{f}|^2 \mathrm{d}\rho \geq \alpha \mathrm{Var}_\rho(f)$$

where $\alpha = Z_t^{-1} e^{-M_t} M_\rho^{-1}$ and $\overline{f} = \int_{\mathcal{X}} f \mathrm{d}\rho_t$. Then

$$\begin{aligned}
\int_Q \langle \nabla w_t, (D^2\phi_t)^{-1} \cdot \nabla w_t\rangle \mathrm{d}\rho &\geq \int_Q \langle \nabla w_t, (D^2\phi_t + D^2 V)^{-1} \cdot \nabla w_t\rangle \mathrm{d}\rho \\
&\geq Z_t e^{m_t} m_\rho \int_Q \langle \nabla w_t, (D^2\phi_t + D^2 V)^{-1} \cdot \nabla w_t\rangle \mathrm{d}\rho_t \\
&\geq Z_t e^{m_t} m_\rho \mathrm{Var}_{\rho_t}(w_t) \\
&\geq e^{m_t - M_t} \frac{m_\rho}{M_\rho} \mathrm{Var}_\rho(w_t) \\
&\geq e^{-r} \frac{m_\rho}{M_\rho} \mathrm{Var}_\rho(w_t).
\end{aligned}$$

Integrating this inequality over $t \in [0,1]$, there remains to lower bound $\int_0^1 \mathrm{Var}_\rho(w_t)\mathrm{d}t$. We notice that $\frac{\mathrm{d}}{\mathrm{d}t}\phi_t(x) = -v(\nabla\phi_t(x)) = -w_t(x)$ due to the same computation as in (3.18). Therefore we deduce from Minkowski's inequality

$$\int_0^1 \mathrm{Var}_\rho(w_t)\mathrm{d}t \geq \mathrm{Var}_\rho\left(\int_0^1 w_t \ \mathrm{d}t\right) = \mathrm{Var}_\rho\left(\int_0^1 \frac{\mathrm{d}\phi_t}{\mathrm{d}t}\mathrm{d}t\right) = \mathrm{Var}_\rho(\phi_1 - \phi_0).$$

All in all,

$$\langle \psi_1 - \psi_0 \mid \nabla\psi^*_{0\#}\rho - \nabla\psi^*_{1\#}\rho\rangle \geq e^{-r} \frac{m_\rho}{M_\rho} \mathrm{Var}_\rho(\phi_1 - \phi_0).$$

A short scaling argument that we do not detail allows to replace e^{-r} by $1/er$. Finally, there remains to control r. Let $x, x' \in Q$, and $y \in \mathcal{Y}$ such that $\phi_t(x) = \langle x, y\rangle - \phi_t^*(y)$. Then

$$\phi_t(x') \geq \langle x', y\rangle - \phi_t^*(y) = \langle x' - x, y\rangle + \phi_t(x) \geq -\mathrm{diam}(Q) R_{\mathcal{Y}} + \phi_t(x). \tag{3.21}$$

Therefore $M_t - m_t \leq \mathrm{diam}(Q) R_{\mathcal{Y}}$, and taking the supremum over $t \in [0,1]$ we get that r has the same upper bound, which concludes. □

Remark 3.9. *At this point, it is possible to conclude the proof of the stability of Kantorovich potentials when ρ is supported on a compact,* convex *set, and bounded above and below on its support (i.e., (2.4) when $\mathcal{X}$ is assumed to be convex). This recovers, with an improved constant, a result due to Delalande and Mérigot [31] (after anterior work by Berman, see the bibliographical notes in Section 3.6). For this, one just needs to take $\psi_0 = \psi_\mu$, $\psi_1 = \psi_\nu$, and upper bound the right-hand side in (3.19) thanks to the Kantorovich-Rubinstein duality formula. This is very similar to the argument at the end of the proof of Theorem 2.1 in Section 3.2.*

3.5 Stability for log-concave sources: proof of the first part of Theorem 2.2

To prove (2.2), we truncate the primal Kantorovich potentials in large balls, apply Theorem 3.7 and show that we do not lose too much by this truncation argument.

Let ϕ_μ, ϕ_ν be the Kantorovich potentials from ρ to μ and ρ to ν respectively. For $r > 0$ we set $\mathcal{B}_r = B(0, r)$ and we denote by $\phi_{\mu,r}, \phi_{\nu,r}$ the restriction of ϕ_μ and ϕ_ν to $\mathcal{B}_r$, extended by $+\infty$ outside $\mathcal{B}_r$. Then we set

$$\rho_r = \frac{\rho_{|\mathcal{B}_r}}{\rho(\mathcal{B}_r)}, \qquad \mu_r = (\nabla\phi_{\mu,r})_\#\rho_r, \qquad \nu_r = (\nabla\phi_{\nu,r})_\#\rho_r.$$

We also consider the dual Kantorovich potentials $\psi_{\mu,r} = \phi^*_{\mu,r}$, $\psi_{\nu,r} = \phi^*_{\nu,r}$ and recall that $\psi^*_{\mu,r} = \phi_{\mu,r}$ and $\psi^*_{\nu,r} = \phi_{\nu,r}$.

We apply Theorem 3.7 to ρ_r, taking for σ the unique probability density over $\mathcal{B}_r$ whose density is proportional to e^{-U} (recall that $\rho = e^{-U-F}$). This gives

$$\mathrm{Var}_{\rho_r}(\phi_{\mu,r} - \phi_{\nu,r}) \le C_{\rho,\mathcal{Y}} r\langle\psi_{\nu,r} - \psi_{\mu,r} \mid \mu_r - \nu_r\rangle.$$

Since $\psi_{\mu,r} - \psi_{\nu,r}$ is r-Lipschitz (by a similar computation to (3.21)), we obtain by the Kantorovich-Rubinstein duality formula (3.10) that

$$\mathrm{Var}_{\rho_r}(\phi_{\mu,r} - \phi_{\nu,r}) \le C_{\rho,\mathcal{Y}} r^2 W_1(\mu_r, \nu_r).$$

Using various truncation estimates which we leave to the reader, we get

$$\mathrm{Var}_{\rho}(\phi_{\mu} - \phi_{\nu}) \le C_{\rho,\mathcal{Y}} \left(r^2 W_1(\mu, \nu) + r^2 m_0(r) + m_1(r)^2 + m_2(r)\right)$$

where

$$m_\ell(r) = \int_{\mathbb{R}^d\setminus\mathcal{B}_r} |x|^\ell d\rho(x) \le C_{\rho,\ell} r^{d+\ell-2} e^{-\frac{1}{2}\kappa r^2}$$

(the last estimate uses $D^2U \ge \kappa\mathrm{Id}$, $\kappa > 0$). We may assume $W_1(\mu, \nu) < 1$, and optimize over r, by taking

$$r = (4\kappa^{-1}|\log W_1(\mu, \nu)|)^{1/2}.$$

This yields (2.2).

Remark 3.10. *Going back to the intuitions given at the beginning of Section 3.3, we see that compared to the Cauchy-Schwarz inequality in (3.14), we sort of lose one factor $|x - y|$ on the right when using Kantorovich-Rubinstein. However, we prove in Section 5.2 that the $1/2$ exponent in (2.2) is sharp (but probably not the log loss).*

3.6 Bibliographical notes

§3.1: Monge's original paper is [60]. The dual formulation of the Monge problem was introduced by Kantorovich, the founding father of linear programming, in [47]. His goal was to solve concrete problems for the Russian industry. The foreword to the English translation of his paper [47], written by an American scientist in the journal Management Science in 1958, is an historical gem: "[...] It is to be noted, however, that the problem of determining an effective method of actually acquiring the solution of a specific problem is not solved in this paper. In the category of development of such methods we seem to be currently, ahead of the Russians."

For a smooth introduction to duality in optimal transport, we refer to [71, Chapter 1]. Brenier's theorem was proved in [15]. The strong duality theorem (3.6) can be formulated for general costs, see [72, Theorem 5.10]. The Kantorovich-Rubinstein formula is a particular case of this strong duality, when the cost is the distance function: $c(x, y) = |x - y|$.

§3.2: The first result similar to Theorem 2.1 is due to Gigli [36], and in the form stated above it is due to Mérigot-Delalande-Chazal [59].

§3.3: For a rigorous proof of Lemma 3.6, see [31, Proposition 2.2].

§3.4: Berman [8] was the first to prove an inequality of the form (3.19), using complex geometry. In his result corresponding to the variance inequality (3.19), the right-hand side is raised to the power $1/2^{d-1}$, which makes it non-optimal. Inspired by his paper, Mérigot-Delalande-Chazal proved in [59] the inequality $\|T_\mu - T_\nu\|_{L^2(\rho)} \leq CW_p(\mu, \nu)^{2/15}$, using a very instructive proof technique. Their arguments are a sort of "discrete version" of the arguments later developed in [31]. They first rely on an approximation argument allowing them to assume that μ, ν are discrete, and then they leverage specific features of semi-discrete optimal transport, notably the geometry of Laguerre cells. In their paper, the Brunn-Minkowski inequality plays the same role as the Brascamp-Lieb in [31] and in the proof of Theorem 3.7 presented above. In the paper [31], Delalande and Mérigot prove the same inequality as in Theorem 3.7, except with a worse constant than ours (notably because they do not compare ρ to a log-concave measure σ). Inspired by the paper [61] by Mischler and Trevisan, which proves a variance inequality with a good constant for log-concave source measures, Mérigot and myself found in [54] the simple proof of Theorem 3.7 presented in Section 3.4, which shortcuts several arguments of [31]. The approximation arguments which are not presented in the proof of Theorem 3.7 above are written in detail in [54].

Another approach to variance inequalities is possible, using entropic optimal transport and the Prékopa-Leindler inequality. It has been introduced by Delalande in [30] (see also his well-written PhD thesis [29]), and has been leveraged later, together with other tools, to prove quantitative stability of optimal transport for p-costs [61] and in Riemannian manifolds [49].

4 Gluing methods

In this chapter, we explain how to generalize the upper bound on the variance derived in Theorem 3.7 to more general measures ρ. The techniques presented here are very robust, and can handle non-degenerate densities on bounded domains (Theorem 2.3) as well as all families of source measures ρ listed in Section 2.3. We call "gluing methods" the tools developed in this chapter, since we use them to prove an upper bound on the variance of a function f with respect to a measure ρ defined on a union of convex sets by combining upper bounds on the variance of f with respect to the restriction of ρ to each of these convex sets. These gluing methods should not be confused with the procedure of gluing measures in optimal transport, which is something totally different.

4.1 Gluing arguments in a nutshell

This section remains at a panoramic level, while complete arguments are provided in the next section. We start with a toy example. Let Q_1, Q_2 be two cubes in $\mathbb{R}^d$ such that $Q_1 \cap Q_2 \neq \emptyset$, and let ρ be a probability density over $Q_1 \cup Q_2$, bounded above and below on $Q_1 \cup Q_2$. For $i = 1, 2$, let

$$\rho_{Q_i} = \frac{\rho_{|Q_i}}{\rho(Q_i)} \tag{4.1}$$

be the restriction of ρ to Q_i, normalized to be a probability measure. We show that

$$\mathrm{Var}_\rho(f) \leq C(\mathrm{Var}_{\rho_{Q_1}}(f) + \mathrm{Var}_{\rho_{Q_2}}(f)) \tag{4.2}$$

for some explicit constant C, roughly proportional to the quotient $\max(\rho(Q_1), \rho(Q_2))/\rho(Q_1 \cap Q_2)$. This is a quantitative version of the fact that if f is constant on Q_1 and constant on Q_2, then it is constant on $Q_1 \cup Q_2$, since $Q_1 \cap Q_2 \neq \emptyset$. The less the cubes overlap, the larger C has to be; and if $Q_1 \cap Q_2 = \emptyset$, then (4.2) becomes false.

We will prove a general version of (4.2) for a finite or infinite collection $\mathcal{F}$ of (well-chosen) cubes Q_i whose union is equal to the whole domain $\mathcal{X}$:

$$\mathrm{Var}_\rho(f) \leq C \sum_{Q_i \in \mathcal{F}} \rho(Q_i) \mathrm{Var}_{\rho_{Q_i}}(f) \tag{4.3}$$

for some $C < +\infty$ (and any f). For this inequality to be true, one has to make some assumptions on ρ; and to carefully design the family $\mathcal{F}$.

Once (4.3) is shown, it takes only a few lines to complete the proof of the stability of Kantorovich potentials, following similar arguments to what we did in Section 3.5. What we need to assume is:

(1) the variation of ρ on each cube is uniformly bounded over $\mathcal{F}$:

$$\sup_{Q_i \in \mathcal{F}} \frac{\sup_{Q_i} \rho}{\inf_{Q_i} \rho} \leq E < +\infty$$

(2) there exists $A > 0$ such that any $Q_i \in \mathcal{F}$ intersects at most A other cubes $Q_j \in \mathcal{F}$ (including itself).

So let us show how to deduce stability of Kantorovich potentials from (4.3). We first prove that for any $\psi_0, \psi_1 \in C^0(\mathcal{Y})$, there holds

$$\mathrm{Var}_\rho(\psi_1^* - \psi_0^*) \leq C' \langle \psi_0 - \psi_1 \mid (\nabla \psi_1^*)_{\#}\rho - (\nabla \psi_0^*)_{\#}\rho \rangle. \tag{4.4}$$

For this we apply Theorem 3.7 in each cube Q_i, to ρ_{Q_i} defined by (4.1) (with σ the normalized Lebesgue measure on Q_i). We get

$$\mathrm{Var}_{\rho_{Q_i}}(\psi_1^* - \psi_0^*) \leq eER_{\mathcal{Y}}\mathrm{diam}(\mathcal{X})\langle \psi_0 - \psi_1 \mid (\nabla\psi_1^*)_{\#}\rho_{Q_i} - (\nabla\psi_0^*)_{\#}\rho_{Q_i}\rangle$$

Combining with (4.3) we get

$$\mathrm{Var}_\rho(f) \leq CeER_{\mathcal{Y}}\mathrm{diam}(\mathcal{X}) \sum_{Q_i\in\mathcal{F}} \langle \psi_0 - \psi_1 \mid (\nabla\psi_1^*)_{\#}\rho_{|Q_i} - (\nabla\psi_0^*)_{\#}\rho_{|Q_i}\rangle. \tag{4.5}$$

We define a partition $\mathcal{F}'$ of $\mathcal{X}$ into convex sets as follows: $x, x' \in \mathcal{X}$ belong to the same element $P \in \mathcal{F}'$ if and only if they belong exactly to the same elements in $\mathcal{F}$. Each $P \in \mathcal{F}'$ is an intersection of at most A cubes according to (2), thus it is convex. Moreover,

$$\langle \psi_0 - \psi_1 \mid (\nabla\psi_1^*)_{\#}\rho_{|P} - (\nabla\psi_0^*)_{\#}\rho_{|P}\rangle \tag{4.6}$$

is non-negative for any $P \in \mathcal{F}'$ due to Theorem 3.7 (or more directly due to the convexity of $\mathcal{K}_{\rho_P}$ - we do not need strong convexity here). The sum in (4.5) may be written as a sum over $P \in \mathcal{F}'$ of the non-negative terms (4.6), each of them weighted by a coefficient between 1 and A. Recalling that the elements of $\mathcal{F}'$ form a partition of $\mathcal{X}$, we obtain (4.4) with $C' = ACeER_{\mathcal{Y}}\mathrm{diam}(\mathcal{X})$.

Finally, we apply (4.4) to $\psi_0 = \phi_\mu^*$ and $\psi_1 = \phi_\nu^*$, where the Legendre transform is computed as a supremum over $\mathcal{X}$, and ϕ_μ, ϕ_ν are the Kantorovich potentials from ρ to μ and ρ to ν. We get

$$\mathrm{Var}_\rho(\psi_1^* - \psi_0^*) \leq C'\langle \phi_\mu - \phi_\nu \mid \nu - \mu\rangle \leq C'\mathrm{diam}(\mathcal{Y})W_1(\mu, \nu)$$

where the last inequality comes from the Kantorovich-Rubinstein duality formula and the fact that ϕ_μ and ϕ_ν are diam($\mathcal{Y}$)-Lipschitz due to Proposition 3.1. This concludes the proof.

All in all, to establish the stability of Kantorovich potentials (2.4) (or stability of Kantorovich potentials in other situations, like those described in Section 2.3) we only need to show that (4.3) holds for some well-chosen family $\mathcal{F}$. To prove (4.3) we designed two strategies, which are complementary in terms of the families of probability measures ρ that they allow to handle:

- Strategy 1: an approach through spectral graph theory, which was chronologically the first we found. In this case, the constant C in (4.3) is related to the spectral gap of the Laplacian on a natural graph constructed from the family $\mathcal{F}$. This strategy is sufficient to prove most of our results, but sometimes in slightly weaker forms - for example with this approach we are able to prove Theorem 2.3 only for bounded, connected Lipschitz domains. In some specific cases, this strategy works whereas the second one fails, for instance for measures ρ which decay polynomially at infinity (see (2.8)).

- Strategy 2: an approach inspired by the proofs of Sobolev-Poincaré inequalities in the 1980's, where the elements of $\mathcal{F}$ are cubes, and we consider chains of cubes, called Boman chains, in which the variances are controlled. When applicable, this approach yields the results in their sharpest forms, and is exactly tailored to handle delicate cases like John domains in Theorem 2.3, and degenerate densities in bounded domains as in Section 2.3.

We describe these two strategies in more details in the next two sections.

4.2 Strategy 1: Gluing variances via spectral graph theory

Assume that ρ is a probability measure on a metric space $\mathcal{X}$, and $\mathcal{F} = \{Q_i\}_{i\in V}$ is a countable family of subsets of $\mathcal{X}$ such that

$$\bigcup_{i\in V} Q_i = \mathrm{supp}(\rho).$$

As in Section 4.1, we assume that there exists $A > 0$ such that any $Q_i \in \mathcal{F}$ intersects at most A other subsets $Q_j \in \mathcal{F}$ (including itself).

We construct a graph as follows: its vertices are given by the set V, and there is an edge between $i, j \in V$ if and only if $\rho(Q_i \cap Q_j) > 0$ (in which case we write $i \sim j$). Each vertex $i \in V$ is endowed with a weight $\delta_i = \rho(Q_i)$ and each edge (i, j) with a weigth $w_{ij} = \rho(Q_i \cap Q_j)$. We consider

$$\ell^2(V, \delta) = \Big\{ u : V \to \mathbb{R} \mid \sum_{i \in V} \delta_i u(i)^2 < +\infty \Big\}$$

and endow it with the scalar product $\langle u, v \rangle_\delta = \sum_{i \in V} \delta_i u(i) v(i)$ and the corresponding norm. We also consider the quadratic form $\mathcal{Q}$ with domain $\mathcal{D}$ given by

$$\mathcal{Q}(u) = \frac{1}{2} \sum_{i \sim j} w_{ij} (u(i) - u(j))^2, \qquad \mathcal{D} = \{ u \in \ell^2(V, \delta) \mid \mathcal{Q}(u) < +\infty \}.$$

Finally we define the Laplacian

$$Lu(i) = \frac{1}{\delta_i} \sum_{j \sim i} w_{ij} (u(i) - u(j))$$

which is positive and selfadjoint with respect to the scalar product $\langle \cdot, \cdot \rangle_\delta$. Due to (2) in Section 4.1, we know that for any $i \in V$,

$$\sum_{i \sim j} w_{ij} \leq A \delta_i, \tag{4.7}$$

hence L is a bounded operator. We also know that $\sum_{i \in V} \delta_i < +\infty$ since $\rho(\mathcal{X}) = 1$; hence the constant function $\mathbf{1}$ belongs to $\ell^2(V, \delta)$. Actually, it lies in the kernel of L, thus we define the spectral gap of L as

$$\lambda_2(L) = \inf \{ \mathcal{Q}(u) \mid \|u\|_\delta = 1, \langle u, \mathbf{1} \rangle = 0 \}.$$

Depending on the graph, $\lambda_2(L)$ may be positive or equal to 0. The next lemma is crucial and does not assume anything on ρ. It is useless if $\lambda_2(L) = 0$ (e.g., if the graph is not connected).

Lemma 4.1. *For any $i \in V$, let $\rho_{Q_i} = \frac{1}{\rho(Q_i)} \rho_{|Q_i}$. Let $f : \mathbb{R}^d \to \mathbb{R}$ be a continuous function. Then*

$$\mathrm{Var}_\rho(f) \leq A \left(1 + \frac{2A}{\lambda_2(L)} \right) \sum_{i \in V} \rho(Q_i) \mathrm{Var}_{\rho_{Q_i}}(f).$$

We describe the proof of Lemma 4.1 in broad lines. Let $m_i = \int_{Q_i} f \mathrm{d}\rho_{Q_i}$ be the mean of f over Q_i. Developing the variance $\mathrm{Var}_\rho(f)$ thanks to the identity $f(x) - f(y) = f(x) - m_i + m_i - m_j + m_j - f(y)$, it is not difficult to prove that

$$\mathrm{Var}_\rho(f) \leq S \sum_{i \in V} \delta_i \mathrm{Var}_{\rho_{Q_i}}(f) + \frac{1}{2} \sum_{i,j \in V} (m_i - m_j)^2 \delta_i \delta_j$$

where $S = \sum_{i \in V} \delta_i \leq A$. To upper bound the second term of the right-hand side we observe that $\langle m - \widetilde{m}, \mathbf{1} \rangle_\delta = 0$ where $\widetilde{m} = S^{-1} \sum_{i \in V} \delta_i m_i$ is "the mean of the means", hence by definition of the spectral gap,

$$\begin{aligned} \frac{1}{2} \sum_{i,j \in V} (m_i - m_j)^2 \delta_i \delta_j = S \|m - \widetilde{m}\|_\delta^2 &\leq \frac{S}{\lambda_2(L)} \langle m - \widetilde{m}, L(m - \widetilde{m}) \rangle_\delta \\ &= \frac{S}{2\lambda_2(L)} \sum_{i,j \in V} w_{ij} (m_i - m_j)^2. \end{aligned} \tag{4.8}$$

Then, let $m_{i\cap j} = \frac{1}{w_{ij}} \int_{Q_i \cap Q_j} f \,\mathrm{d}\rho$ be the mean of f over $Q_i \cap Q_j$. It is not difficult to prove that

$$(m_{i\cap j} - m_i)^2 \leq \frac{\rho(Q_i)}{w_{ij}} \mathrm{Var}_{\rho_i}(f)$$

and similarly for $(m_{i\cap j} - m_j)^2$. This allows to upper bound the right-hand side of (4.8) by a constant times $\sum_{i\in V} \rho(Q_i)\mathrm{Var}_{\rho_i}(f)$, which concludes the proof of Lemma 4.1.

In applications of the above lemma to concrete cases (with explicit ρ and $\mathcal{F}$), one needs to prove $\lambda_2(L) > 0$. For this, there is one extremely useful tool: the Cheeger inequality, which gives a lower bounds on the spectral gap of general Laplace operators. In the present context, due to the assumption (4.7), it takes the form

$$\lambda_2(L) \geq \frac{h^2}{2A} \tag{4.9}$$

where h is a constant defined as follows:

$$h = \inf_{\substack{U\subset V \\ 0<\mathrm{vol}(U)\leq \mathrm{vol}(V)/2}} \frac{|\partial U|}{\mathrm{vol}(U)}$$

with $\mathrm{vol}(U) = \sum_{i\in U} \delta_i$ and $|\partial U| = \sum_{i\in U, j\notin U} w_{ij}$.

With the Cheeger inequality at hand, we need to check in concrete applications whether $h > 0$ or not. And this depends strongly on ρ and on the construction of the family $\mathcal{F}$ (and thus on the corresponding graph). Also, is tempting to see the graph as a discretization of the domain $\mathcal{X}$, and to prove the Cheeger inequality on the graph as a consequence of a Cheeger inequality on the continuous domain $\mathcal{X}$, but it is not clear to me how to build a proof out of this intuition.

Let me mention one concrete example where we have been able to construct $\mathcal{F}$ and to check that $h > 0$. Consider $\rho(x) = (1 + |x|)^{-\beta}$ for some $\beta > d+2$; due to the radial symmetry of these distributions, each set Q is taken as the intersection of an annulus and an angular sector, see Figure 3. The associated graph is very simple, it is essentially the union of 2^d line graphs (see Figure 3), and the ratios $|\partial U|/\mathrm{vol}(U)$ can be lower bounded "by hand".

Similarly, if $\mathcal{X}$ is a bounded, connected Lipschitz domain, then taking for $\mathcal{F}$ the Whitney decomposition of the domain (with enlarged cubes, see below), it is possible to analyze the graph, which has some kind of hyperbolic structure.

4.3 Strategy 2: Gluing variances via Boman chains

In this section we explain the second gluing technique. This technique applies only to measures with bounded support in $\mathbb{R}^d$ (sometimes obtained from measures with full support, after a truncation argument). Roughly, it consists in fixing a central cube $Q_0 \in \mathcal{F}$ and estimating the variance in any cube $Q \in \mathcal{F}$ thanks to the variance in Q_0, via the construction of a chain of overlapping cubes going from Q_0 to Q.

Definition 4.2 (Boman chain condition)**.** *Let $A, B, C > 1$ with $B \in \mathbb{N}$. A probability measure ρ on an open set $\mathcal{X} \subset \mathbb{R}^d$ satisfies the Boman chain condition with parameters $A, B, C \in \mathbb{R}$ if there exists a covering $\mathcal{F}$ of $\mathcal{X}$ by open cubes $Q \in \mathcal{F}$ such that*

- *For any $x \in \mathbb{R}^d$,*

$$\sum_{Q\in\mathcal{F}} \chi_Q(x) \leq A\chi_{\mathcal{X}}(x). \tag{4.10}$$

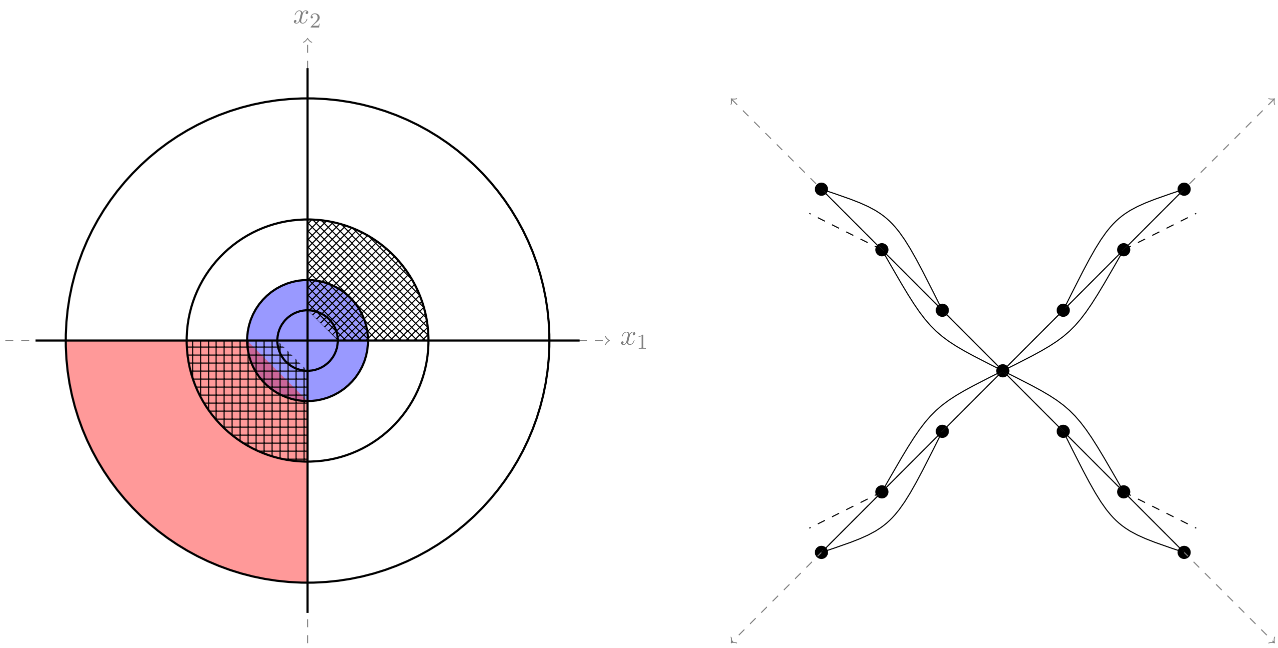

Figure 3: Here, ρ is radially symmetric, for instance $\rho(x) = (1 + |x|)^{-\beta}$. On the left, a few examples of sets Q (in red, in blue with a grid and with crosshatches). The central set Q in blue is different from the other ones, it covers the full unit disk, to make the graph on the right is connected. On the right, the associated graph in a neighborhood of its central point.

- *For some fixed cube Q_0 in $\mathcal{F}$, called the central cube, and for every $Q \in \mathcal{F}$, there exists a sequence $Q_0, Q_1, \ldots, Q_N = Q$ of distinct cubes from $\mathcal{F}$ such that for any $j \in \{0, \ldots, N\}$,*

$$Q \subset BQ_j \tag{4.11}$$

where BQ_j is the cube with same center as Q_j and sidelength multiplied by B.

- *Consecutive cubes of the above chain overlap quantitatively: for any $j \in \{0, \ldots, N-1\}$,*

$$\rho(Q_j \cap Q_{j+1}) \geq C^{-1} \max(\rho(Q_j), \rho(Q_{j+1})). \tag{4.12}$$

The condition (4.10) means that any point cannot belong to more than A cubes $Q \in \mathcal{F}$.

Remark 4.3. *Consecutive cubes of the above chain are comparable in size: for any $j \in \{0, \ldots, N-1\}$,*

$$C^{-1} \leq \frac{\rho(Q_j)}{\rho(Q_{j+1})} \leq C. \tag{4.13}$$

Indeed, $\frac{\rho(Q_j)}{\rho(Q_{j+1})} \geq \frac{\rho(Q_j \cap Q_{j+1})}{\rho(Q_{j+1})} \geq C^{-1}$ as a consequence of (4.12), and the reverse bound in (4.13) follows by the same argument.

Proposition 4.4. *If ρ is a probability measure on a John domain $\mathcal{X}$, with a density bounded above and below on $\mathcal{X}$, then ρ satisfies the Boman chain condition (for some A, B, C).*

The proof of this result is sketched in Appendix C. It relies on the Whitney decomposition of $\mathcal{X}$, a classical tool which allows to partition any open set into cubes whose sidelength is comparable to their distance to the boundary. The cubes of the Boman chain condition are obtained by enlarging each cube by a fixed given factor (e.g., each sidelength is multiplied by 6/5) to create some overlap between the cubes. See Figures 4 and 5 at the end of this chapter for illustrations of the Whitney decomposition and of the cubes of the Boman chain condition. Some kind of converse of Proposition 4.4 holds: if the characteristic function ρ of some bounded open set $\mathcal{X}$ satisfies the Boman chain condition, then $\mathcal{X}$ is a John domain.

Theorem 2.3 follows directly from the arguments explained below (4.3), together with Proposition 4.4 and the following result:

Proposition 4.5. *If ρ satisfies the Boman chain condition, there exists $C > 0$ such that for any f,*

$$\mathrm{Var}_\rho(f) \leq C \sum_{Q_i \in \mathcal{F}} \rho(Q_i) \mathrm{Var}_{\rho_{Q_i}}(f). \tag{4.14}$$

Proof. We set $f_Q = \frac{1}{\rho(Q)} \int_Q f \mathrm{d}\rho$ and $a_Q = (\mathrm{Var}_{\rho_Q}(f))^{1/2}$. Then

$$\begin{aligned} \mathrm{Var}_\rho(f) \leq \int_{\mathcal{X}} |f(x) - f_{Q_0}|^2 \mathrm{d}\rho(x) &\leq \sum_{Q \in \mathcal{F}} \int_Q |f(x) - f_{Q_0}|^2 \mathrm{d}\rho(x) \\ &\leq 2 \sum_{Q \in \mathcal{F}} \int_Q |f(x) - f_Q|^2 \mathrm{d}\rho(x) + \int_Q |f_Q - f_{Q_0}|^2 \mathrm{d}\rho(x) \qquad (4.15) \\ &= 2 \sum_{Q \in \mathcal{F}} \rho(Q) \mathrm{Var}_{\rho_Q}(f) + 2 \sum_{Q \in \mathcal{F}} \rho(Q) |f_Q - f_{Q_0}|^2. \end{aligned}$$

The first sum is bounded above by the right-hand side in (4.14), therefore we only need to upper bound the second sum. The triangle inequality yields

$$|f_Q - f_{Q_0}| \leq \sum_{j=0}^{N-1} |f_{Q_j} - f_{Q_{j+1}}| \tag{4.16}$$

We estimate each term in the sum separately:

$$\begin{aligned} |f_{Q_j} - f_{Q_{j+1}}|^2 &= \frac{1}{\rho(Q_j \cap Q_{j+1})} \int_{Q_j \cap Q_{j+1}} |f_{Q_j} - f_{Q_{j+1}}|^2 \\ &\leq \frac{2}{\rho(Q_j \cap Q_{j+1})} \left(\int_{Q_j \cap Q_{j+1}} |f_{Q_j} - f(x)|^2 \mathrm{d}\rho(x) + \int_{Q_j \cap Q_{j+1}} |f_{Q_{j+1}} - f(x)|^2 \mathrm{d}\rho(x) \right) \\ &\leq \frac{2}{\rho(Q_j \cap Q_{j+1})} \left(\int_{Q_j} |f_{Q_j} - f(x)|^2 \mathrm{d}\rho(x) + \int_{Q_{j+1}} |f_{Q_{j+1}} - f(x)|^2 \mathrm{d}\rho(x) \right) \\ &\leq 2C(a_{Q_j}^2 + a_{Q_{j+1}}^2). \end{aligned}$$

Taking the square root and plugging into (4.16), we obtain

$$|f_Q - f_{Q_0}| \leq (2C)^{1/2} \sum_{j=0}^{N-1} a_{Q_j} + a_{Q_{j+1}} \leq (8C)^{1/2} \sum_{Q \subset B\widetilde{Q}} a_{\widetilde{Q}}$$

where the sum $\sum_{Q \subset B\widetilde{Q}}$ means that we sum over all cubes $\widetilde{Q} \in \mathcal{F}$ such that $Q \subset B\widetilde{Q}$. By the Boman chain condition, Q_j and Q_{j+1} have this property, and notice that we use here the fact that the elements of the Boman chain in Definition 4.2 are distinct.

Therefore,

$$\rho(Q)|f_Q - f_{Q_0}|^2 \leq 8C\rho(Q)\Big(\sum_{Q\subset B\widetilde{Q}} a_{\widetilde{Q}}\Big)^2 = 8C\int_Q \Big(\sum_{Q\subset B\widetilde{Q}} a_{\widetilde{Q}}\Big)^2 \mathrm{d}\rho(x)$$
$$= 8C\int_Q \Big(\sum_{\widetilde{Q}\in\mathcal{F}} a_{\widetilde{Q}}\chi_{B\widetilde{Q}}(x)\Big)^2 \mathrm{d}\rho(x)$$

since for any $x \in Q$, the sum in the first line is equal to the sum in the second line.

Then we use an important lemma which says that

$$\Big\|\sum_{\widetilde{Q}\in\mathcal{F}} a_{\widetilde{Q}}\chi_{B\widetilde{Q}}\Big\|_{L^2(\rho)} \lesssim \Big\|\sum_{\widetilde{Q}\in\mathcal{F}} a_{\widetilde{Q}}\chi_{\widetilde{Q}}\Big\|_{L^2(\rho)}. \tag{4.17}$$

This lemma relies on some kind of Hardy-Littlewood maximal inequality (which merely requires a doubling property for ρ, and the hidden constant in (4.17) only depends on the doubling constant), see [54, Appendix A].

All in all,

$$\sum_{Q\in\mathcal{F}} \rho(Q)|f_Q - f_{Q_0}|^2 \lesssim \int_{\mathcal{X}} \Big(\sum_{\widetilde{Q}\in\mathcal{F}} a_{\widetilde{Q}}\chi_{B\widetilde{Q}}(x)\Big)^2 \mathrm{d}\rho(x) \lesssim \int_{\mathcal{X}} \Big(\sum_{\widetilde{Q}\in\mathcal{F}} a_{\widetilde{Q}}\chi_{\widetilde{Q}}(x)\Big)^2 \mathrm{d}\rho(x)$$
$$\lesssim \int_{\mathcal{X}} \sum_{\widetilde{Q}\in\mathcal{F}} a_{\widetilde{Q}}^2\chi_{\widetilde{Q}}(x)\mathrm{d}\rho(x) = \int_{\mathcal{X}} \sum_{\widetilde{Q}\in\mathcal{F}} \rho(\widetilde{Q})a_{\widetilde{Q}}^2$$

where in the third inequality we used the Cauchy-Schwarz inequality and the first condition in Definition 4.2. Plugging into (4.15) we get the result. □

Open question 4.6. *Is there an analog of Proposition 4.5 for the entropy*

$$\mathrm{Ent}_\rho(f) = \int_{\mathcal{X}} f\log f\mathrm{d}\rho - \Big(\int_{\mathcal{X}} f\mathrm{d}\rho\Big)\log\Big(\int_{\mathcal{X}} f\mathrm{d}\rho\Big)$$

instead of the variance?

4.4 Comments on John domains and relation to Sobolev-Poincaré inequalities

Readers familiar with the literature on Sobolev-Poincaré inequalities may have noticed some resemblance between our proof of Theorem 2.3 and the proof of Sobolev-Poincaré inequalities in John domains.

Maybe the easiest way to see a link between Poincaré-type inequalities and stability of Kantorovich potentials is to write that if the stability inequality for Kantorovich potentials (1.13) (with $p = 2$) holds, then

$$\|\phi_\mu - \phi_\nu\|_{L^2(\rho)} \leq CW_2(\mu,\nu)^\alpha \leq C\|\nabla\phi_\mu - \nabla\phi_\nu\|_{L^2(\rho)}^\alpha$$

according to (1.12). Hence some kind of Poincaré inequality with an exponent holds, but only for differences of convex functions (with uniformly bounded gradient). We are not aware of any literature on this kind of inequalities.

In another direction, it is possible to deduce from Proposition 4.5 the Poincaré inequality

$$\exists C_P > 0,\ \forall f\in C^1(\mathcal{X}), \qquad \mathrm{Var}_\rho(f) \leq C_P\int_{\mathcal{X}} |\nabla f|^2 d\rho \tag{4.18}$$

(but of course this requires ρ to satisfy the Boman chain condition, otherwise Proposition 4.5 does not apply). To prove (4.18), we first observe that there exists $C'_P > 0$ such that for any $Q \in \mathcal{F}$ there holds

$$\mathrm{Var}_{\rho_Q}(f) \le C'_P \int_Q |\nabla f|^2 d\rho_Q. \tag{4.19}$$

Indeed, since $\sup_Q \frac{M_{\rho_Q}}{m_{\rho_Q}} < +\infty$, it is sufficient to prove that (4.19) holds when ρ_Q is replaced by the normalized Lebesgue measure on Q, with a Q-independent constant C'_P. This latter fact is true because all Q are cubes, with uniformly bounded diameter. Summing (4.19) over $Q \in \mathcal{F}$ with weights $\rho(Q)$, and using (4.10) we get (4.18).

Let $f_{\mathcal{X}} = |\mathcal{X}|^{-1} \int_{\mathcal{X}} f \mathrm{d}x$. When $\mathcal{X} \subset \mathbb{R}^d$ is a John domain, the Sobolev-Poincaré inequality

$$\left(\int_{\mathcal{X}} |f - f_{\mathcal{X}}|^{pd/(d-p)} \mathrm{d}x \right)^{(d-p)/pd} \le C \left(\int_{\mathcal{X}} |\nabla f|^p \mathrm{d}x \right)^{1/p} \tag{4.20}$$

holds for $1 \le p < d$. This has been shown around 1985 by Bojarski, following ideas that he attributes to Boman, and our gluing techniques are adapted from this literature. It is also known that if $\mathcal{X} \subset \mathbb{R}^d$ is a domain of finite volume that satisfies a separation property, and $1 \le p < d$, then

$$\mathcal{X} \text{ satisfies (4.20)} \Leftrightarrow \mathcal{X} \text{ is a John domain.} \tag{4.21}$$

The separation property, which we do not discuss here, is automatically valid for simply connected planar domains. And without an additional assumption on $\mathcal{X}$ such as the separation property, the equivalence (4.21) is not true. Let us illustrate this on an example: take $\mathcal{X} = \mathbb{D} \setminus E$ where $\mathbb{D}$ is the unit disk and $E = \bigcup_{k=1}^{\infty} E_k$ where E_k consists of $k!$ equally spaced points on the circle $\{|x| = 1 - 2^{-k}\}$. Then $\mathcal{X}$ is not a John domain, but since E is of dimension 0, the Sobolev-Poincaré inequality (4.20) holds in $\mathcal{X}$ (it can be deduced by integration by parts from the Sobolev-Poincaré inequality in $\mathbb{D}$). This example may be transposed to the optimal transport setting with source measure ρ equal to the uniform density on $\mathcal{X}$. Then optimal transport maps and potentials coincide with those obtained when the source measure is equal to the uniform probability density on $\mathbb{D}$. And for the latter, stability follows from Theorem 2.3. Therefore we have exhibited a non-John domain for which optimal transport stability inequalities hold. When trying to prove a converse statement to Theorem 2.3, one should keep this example in mind.

Several families of examples of bounded connected domains which are not John domains have been considered in the literature. For instance, domains with an outward cusp, and the so-called room-and-passage domains. In Section 5.3 we show that in these examples, stability of Kantorovich potentials fails, even in a very weak sense. This shows the relevance of the John domain condition in Theorem 2.3, at least regarding stability of Kantorovich potentials.

4.5 Bibliographical notes

§4.1: A gluing argument for finite families $\mathcal{F}$ has been used in the work [20] in the slightly different context of stability of Wasserstein barycenters. This insight was the starting point of my collaboration with Quentin Mérigot [54], in which we worked out the general gluing methods presented here. These methods turned out to be useful to address stability of optimal transport in Riemannian manifolds too [49].

§4.2: Spectral graph theory is a classical topic, see for instance the books [25], [67] and the beautiful expository notes by Luca Trevisan [70]. The Cheeger inequality in finite graphs is of course covered in these references. For infinite graphs, the book chapter [48, Chapter 13.1] is particularly clear. I did not find any reference for the elementary Lemma 4.1, but it seems difficult to believe that no one ever used such arguments.

§4.3: Boman introduced in [11] the chains now known as Boman chains. His goal was to prove L^p estimates for solutions to some over-determined elliptic systems of PDEs in regions with irregular boundary. Bojarski discovered in [10] how to use these chains to prove Sobolev-Poincaré inequalities in John domains. We borrowed several computations from this very inspiring work. The proof of Proposition 4.4 is due to him and relies on the Whitney decomposition, for which we refer for instance to [37, Appendix J.1]. The converse fact that any bounded open subset of $\mathbb{R}^d$ supporting Boman chains is a John domain was proved in [17].

§4.4: A converse to Bojarski's result was proved in the paper "Sobolev-Poincaré implies John" [16]. We borrowed our discussion about the separation property from this paper. Since the 1980's, many authors have been using variants of the Boman chain condition. It turns out that for many results, a good framework is that of metric spaces endowed with a doubling measure, see the memoir [39] which develops the theory of Sobolev spaces and proves Poincaré-type inequalities in this setting.

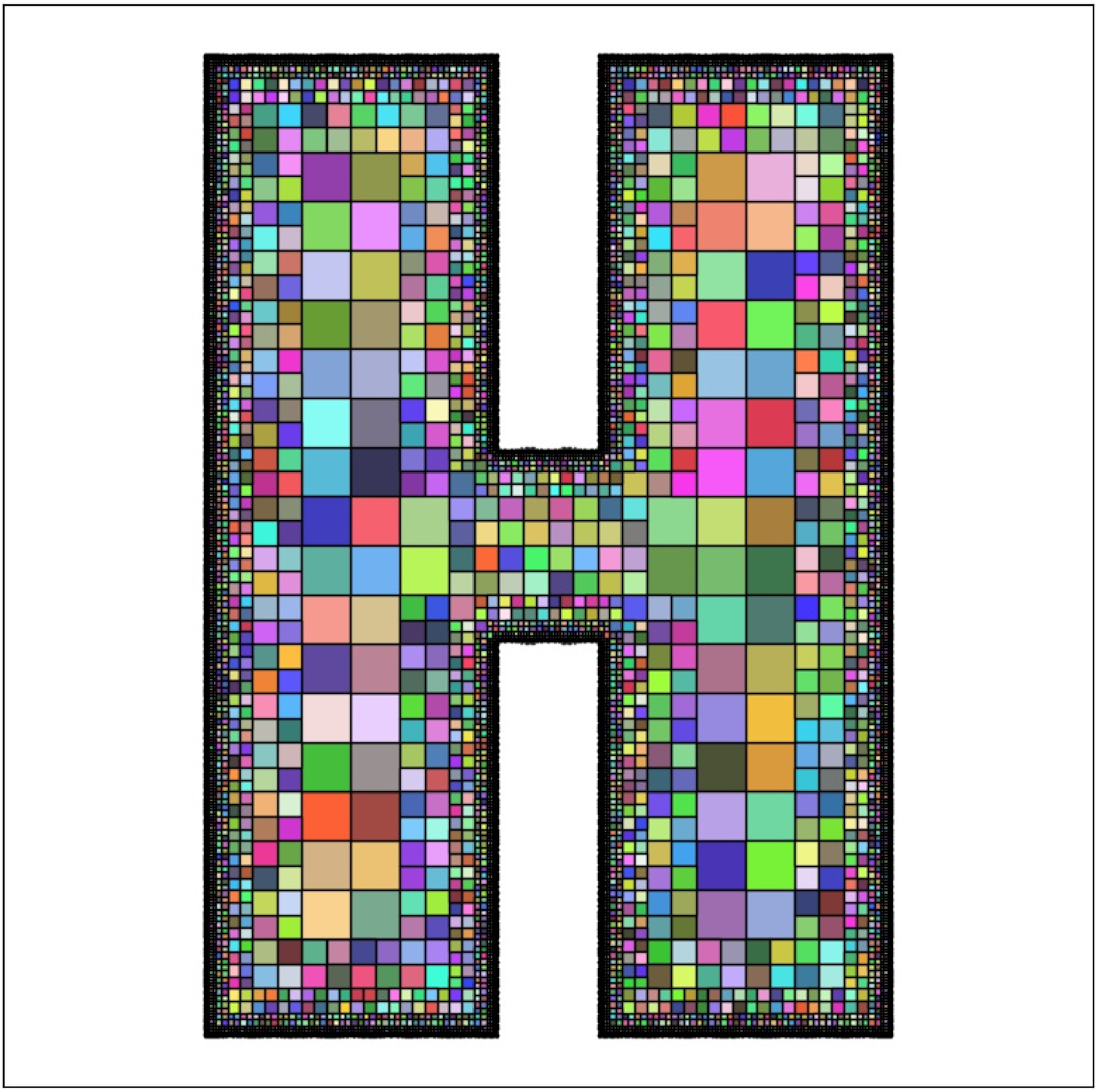

Figure 4: The Whitney decomposition of an H-shape in 2d. Each cube has a sidelength comparable to its distance to the boundary of the H. Courtesy of Quentin Mérigot.

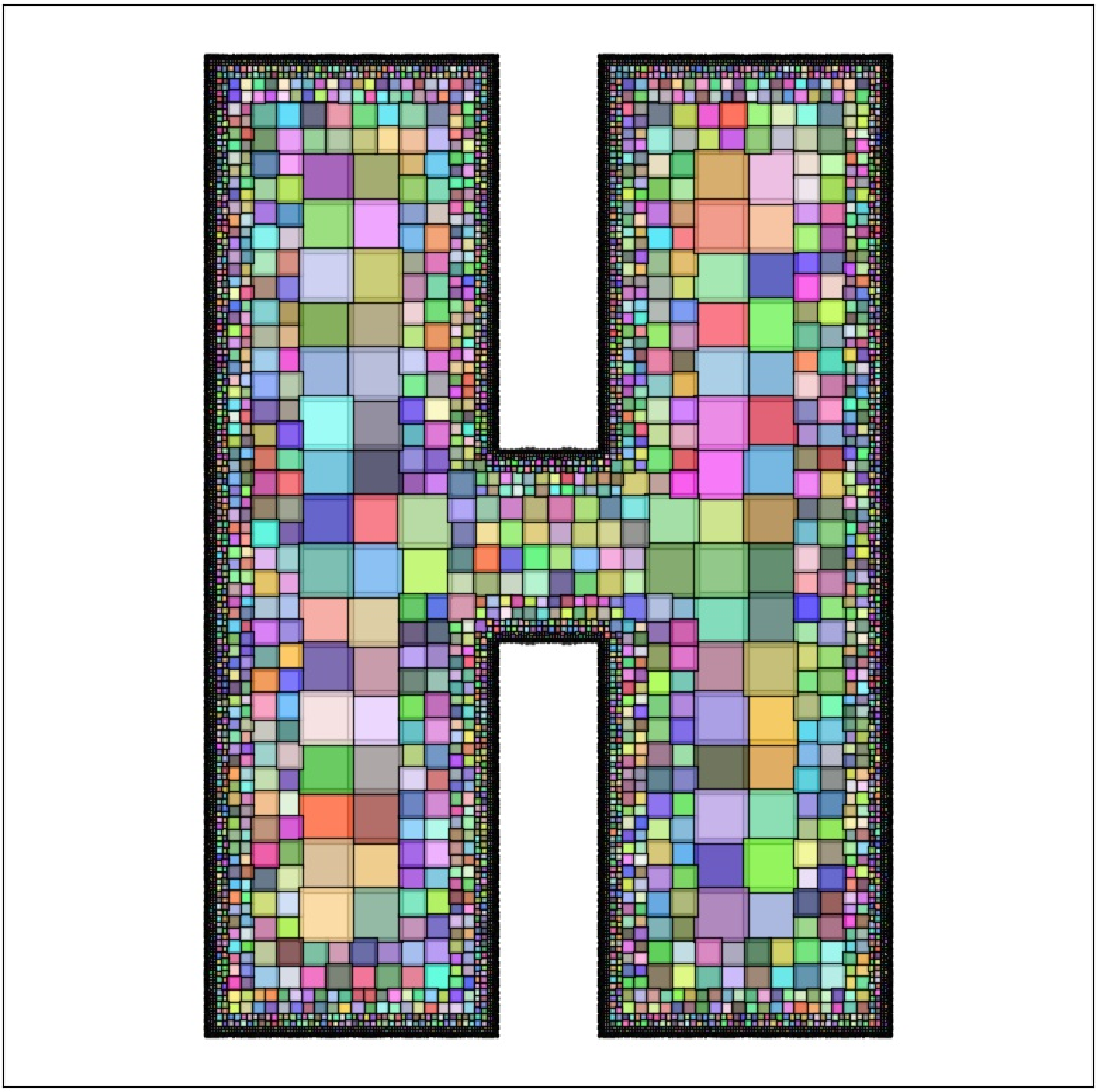

Figure 5: A Boman family for the uniform probability density on the H-shape, obtained by enlarging the sidelength of each cube of the Whitney decomposition by the same factor. This induces some overlap between the cubes. Courtesy of Quentin Mérigot.

5 Examples and counterexamples

In the previous chapters, we did not discuss the sharpness of our results. The only example we provided, in Section 2.4, showed that for optimal transport *maps*, without further assumptions on the target measures, the inequality

$$\|T_\mu - T_\nu\|_{L^2(\rho)} \le C W_1(\mu,\nu)^\alpha \tag{5.1}$$

fails for $\alpha > 1/2$: there exists no $C > 0$ such that for any μ, ν supported in $\mathcal{Y} = \mathbb{D}^2$, (5.1) holds.

In this section, we discuss the sharpness of our results regarding stability of *Kantorovich potentials*, by providing explicit computations on carefully chosen examples. This allows us to show some kind of sharpness in two respects:

- we show that the stability exponents $\alpha = 1/2$ for Kantorovich potentials in our main results is sharp in the Gaussian case (but the additional log-factor is probably not sharp), and becomes sharp in bounded domains as the dimension goes to $+\infty$.
- we show that in typical examples of domains $\mathcal{X}$ which are not John domains, no bound of the form (1.11) can hold when ρ has a density bounded above and below on $\mathcal{X}$. This shows that our "John domain" assumption in Theorem 2.3 is truly meaningful.

The main idea which guides the design of our examples is that to test stability, it seems much easier to choose convex potentials ϕ_1, ϕ_2 and to compute the Wasserstein distance between $(\nabla\phi_1)_\#\rho$ and $(\nabla\phi_2)_\#\rho$, than to choose two measures μ, ν and to try to compute the associated potentials ϕ_μ, ϕ_ν. Of course, the two points of view are in the end equivalent.

For the stability of optimal transport *maps*, unfortunately we have no good example to test the sharpness of stability exponents beside that of Section 2.4. In particular, we do not know the optimal exponents in (2.3) and (2.5).

5.1 Asymptotic sharpness of exponent in the ball

When ρ is the uniform density on the unit ball $B_d(0,1)$ of $\mathbb{R}^d$, Theorem 2.3 provides us with an inequality of the form

$$\|\phi_\mu - \phi_\nu\|_{L^2(\rho)} \le C_d W_1(\mu,\nu)^\alpha \tag{5.2}$$

with $\alpha = 1/2$. We show that the exponent $\alpha = 1/2$ is asymptotically sharp as $d \to +\infty$.

Denote by ω_d the Euclidean volume of the unit ball $B_d(0,1)$ of $\mathbb{R}^d$ and by σ_{d-1} the Euclidean area of the unit sphere $\mathbb{S}^{d-1} \subset \mathbb{R}^d$. Let

$$\rho_d(x) = \frac{1}{\omega_d}\mathbf{1}_{B_d(0,1)}$$

be the uniform probability density on the unit ball of $\mathbb{R}^d$. Consider for any $\varepsilon \in (0,1)$ the radial and convex functions

$$\phi_\varepsilon^{(1)}(x) = |x|, \qquad \phi_\varepsilon^{(2)}(x) = \max(|x|, \varepsilon).$$

Then

$$\int_{B_d(0,1)} \left(\phi_\varepsilon^{(2)} - \phi_\varepsilon^{(1)}\right) \mathrm{d}\rho = \frac{\sigma_{d-1}}{\omega_d}\int_0^\varepsilon r^{d-1}(\varepsilon - r)\mathrm{d}r = \frac{\varepsilon^{d+1}}{d+1}$$

and

$$\int_{B_d(0,1)} \left(\phi_\varepsilon^{(2)} - \phi_\varepsilon^{(1)}\right)^2 \mathrm{d}\rho = \frac{\sigma_{d-1}}{\omega_d}\int_0^\varepsilon r^{d-1}(\varepsilon - r)^2\mathrm{d}r = \frac{2\varepsilon^{d+2}}{(d+1)(d+2)}.$$

Hence,

$$\mathrm{Var}(\phi_\varepsilon^{(2)} - \phi_\varepsilon^{(1)})^{1/2} \sim C_d \varepsilon^{(d+2)/2} \tag{5.3}$$

as $\varepsilon \to 0$, with $C_d = (2/(d+1)(d+2))^{1/2}$. Finally, denoting by $\delta_{\mathbb{S}^{d-1}}$ the uniform probability measure on $\mathbb{S}^{d-1}$,

$$(\nabla \phi_\varepsilon^{(1)})_\# \rho_d = \delta_{\mathbb{S}^{d-1}}, \qquad (\nabla \phi_\varepsilon^{(2)})_\# \rho_d = (1-\varepsilon^d)\delta_{\mathbb{S}^{d-1}} + \varepsilon^d \delta_0$$

hence

$$W_1((\nabla \phi_\varepsilon^{(1)})_\# \rho_d, (\nabla \phi_\varepsilon^{(2)})_\# \rho_d) = \varepsilon^d.$$

We conclude that for any d,

$$\mathrm{Var}(\phi_\varepsilon^{(2)} - \phi_\varepsilon^{(1)})^{1/2} \underset{\varepsilon \to 0}{\sim} C_d W_1((\nabla \phi_\varepsilon^{(1)})_\# \rho_d, (\nabla \phi_\varepsilon^{(2)})_\# \rho_d)^{(d+2)/2d}$$

i.e., it is necessary for α to be $\leq (d+2)/2d$ in order for (5.2) to be true. This tends to $1/2$ as $d \to +\infty$.

5.2 (Almost) sharpness of exponent for Gaussians

It seems natural to test the sharpness of our exponents in the Gaussian case too. Let

$$\rho(x) = (2\pi)^{-\frac{d}{2}} e^{-\frac{|x|^2}{2}}$$

be the standard Gaussian. In this case, recall that the stability inequality for Kantorovich potentials is given by (2.2). Consider for any $r \in (0, +\infty)$ the radial and convex function

$$\phi_r(x) = (|x| - r)_+ - c_r \tag{5.4}$$

with c_r chosen in a way that $\int_{\mathbb{R}^d} \phi_r(x) d\rho(x) = 0$. Brenier's theorem guarantees that $\nabla \phi_r$ is the optimal transport map from ρ to $(\nabla \phi_r)_\# \rho$. For r and r' close enough (and r large enough), we compare $\|\phi_r - \phi_{r'}\|_{L^2(\rho)}$ to $W_1(\mu, \nu)$ where $\mu = (\nabla \phi_r)_\# \rho$ and $\nu = (\nabla \phi_{r'})_\# \rho$.

For r large, we set $r' = r + \frac{1}{r}$ and compute

$$\begin{aligned}
(2\pi)^{d/2}(c_r - c_{r'}) &= \int_r^{+\infty} (s-r) s^{d-1} e^{-s^2/2} ds - \int_{r'}^{+\infty} (s-r') s^{d-1} e^{-s^2/2} ds \\
&= (r'-r) \int_{r'}^{+\infty} s^{d-1} e^{-s^2/2} ds + \int_r^{r'} (s-r) s^{d-1} e^{-s^2/2} ds \\
&= O(r^{d-3} e^{-r^2/2})
\end{aligned}$$

and

$$\begin{aligned}
\|(|\cdot| - r)_+ - (|\cdot| - r')_+\|_{L^2(\rho)}^2 &= (r'-r)^2 \int_{r'}^{+\infty} s^{d-1} e^{-s^2/2} ds + \int_r^{r'} (s-r)^2 s^{d-1} e^{-s^2/2} ds \\
&= \Theta(r^{d-4} e^{-r^2/2})
\end{aligned}$$

where we write $f(r) = \Theta(g(r))$ if the quotient $f(r)/g(r)$ remains bounded above and below by positive constants as $r \to +\infty$. We deduce that

$$\|\phi_{r'} - \phi_r\|_{L^2(\rho)}^2 = \|(|\cdot| - r)_+ - (|\cdot| - r')_+\|_{L^2(\rho)}^2 - |c_r - c_{r'}|^2 = \Theta(r^{d-4} e^{-r^2/2}). \tag{5.5}$$

We then turn to the computation of $W_1((\nabla\phi_{r'})_{\#}\rho, (\nabla\phi_r)_{\#}\rho)$. We observe that

$$(\nabla\phi_r)_{\#}\rho = \rho(B(0,r))\delta_0 + (1-\rho(B(0,r))\sigma_{\mathbb{S}^{d-1}}$$

where $\sigma_{\mathbb{S}^{d-1}}$ is the uniform probability measure on $\mathbb{S}^{d-1}$. We have an analogous expression for $(\nabla\phi_{r'})_{\#}\rho$, and we deduce

$$W_1((\nabla\phi_{r'})_{\#}\rho, (\nabla\phi_r)_{\#}\rho) = \rho(B(0,r')) - \rho(B(0,r)) = \Theta(r^{d-2}e^{-r^2/2}). \tag{5.6}$$

It follows from (5.5) and (5.6) that

$$\|\phi_{r'} - \phi_r\|_{L^2(\rho)} = \Theta(W_1^{1/2}|\log W_1|^{-1})$$

where W_1 is a short notation for $W_1((\nabla\phi_{r'})_{\#}\rho, (\nabla\phi_r)_{\#}\rho)$. This shows that the exponent $1/2$ in (2.2) is sharp (but the log factor probably not).

We also observe that the above example does not prove the sharpness of the exponent of stability of optimal transport maps (2.3) (and indeed, we conjecture that the correct exponent is $1/2$ and not $1/6$).

Remark 5.1. *The above proof can be adapted to other contexts. For instance, when $\rho(x) = c_{\beta,d}(1+|x|)^{-\beta}$ ($\beta > d+2$), it is possible to derive sharp stability exponents for Kantorovich potentials using the same family of radial Kantorovich potentials (5.4). Also, when ρ blows up at the boundary of a ball or is the spherical uniform distribution (see "Degenerate densities ρ in bounded domains" in Section 2.3), this same family may be used to find upper bounds on the stability exponents for Kantorovich potentials.*

5.3 Strong instability for room-and-passage domains

We turn to another explicit computation, this time aimed at showing the relevance of the John domain condition in Theorem 2.3. For this, we consider domains that are considered in the literature as typical instances of non-John domains, and show that if ρ is bounded above and below on such domain, then stability of Kantorovich potentials cannot hold, even in a very weak sense.

We could seek for even stronger, and hope that the John domain condition is necessary and sufficient for Theorem 2.3 to hold. However, this cannot be true, as explained in Section 4.4. In analogy, John domains support Sobolev-Poincaré inequalities, but there exist non-John domains which also support Sobolev-Poincaré inequalities. To remedy this issue, it has been shown that a domain satisfying a certain separation property supports Sobolev-Poincaré inequalities if and only if it is a John domain. The proof of this fact is delicate, and it would be interesting to look for an analogous converse result to Theorem 2.3.

In this section we prove:

Theorem 5.2. *There exists a non-empty, bounded, path-connected domain $\mathcal{X} \subset \mathbb{R}^d$ such that for any probability density ρ bounded above and below on $\mathcal{X}$ the inequality*

$$\forall \mu,\nu \in \mathcal{P}(B(0,2)), \qquad \|\phi_\mu - \phi_\nu\|_{L^2(\rho)} \le C W_p(\mu,\nu)^\alpha \tag{5.7}$$

fails for any $C, \alpha > 0$ and $p \in [1,+\infty)$ (where ϕ_μ, ϕ_ν denote the Kantorovich potentials between ρ and μ and ρ and ν respectively).

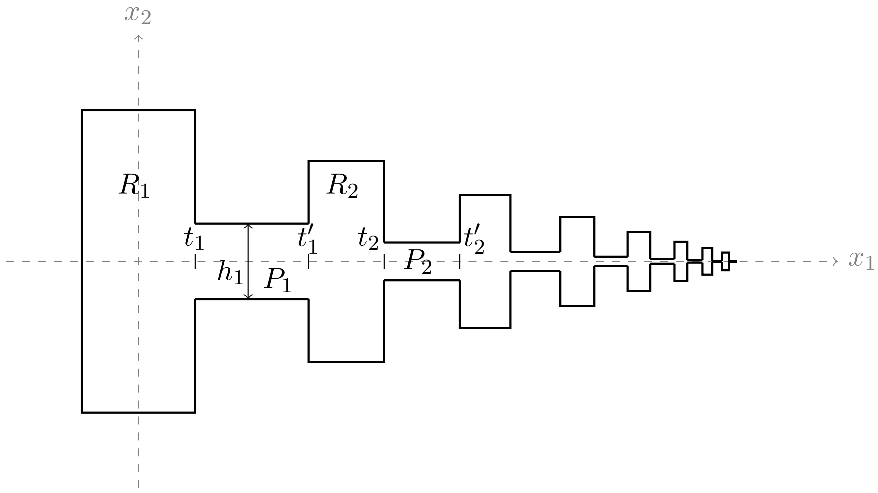

Figure 6: Room-and-passage domain

The counterexample $\mathcal{X}$ is a so-called "room-and-passage" domain, a typical example of a non-John domain. It is endowed with a probability density which is bounded above and below on the support, for instance the uniform density. We consider the case $d = 2$ for simplicity, but the computations may be modified to cover any dimension d.

As depicted on Figure 6, a room-and-passage domain in $\mathbb{R}^2$ is a connected and bounded set made of an infinite union of rectangles with variable lengths and widths. For simplicity, we assume that the axes of these rectangles are parallel to the coordinate axes. We call length of a rectangle the length of its side parallel to the x_1 axis, and width that of its side parallel to the x_2 axis. The rectangles are of two types, which alternate along the x_1-axis: the rooms R_n, $n \in \mathbb{N}$; and the passages P_n, $n \in \mathbb{N}$. The key assumption we make is that the passages have a width h_n which decreases very fast as n tends to $+\infty$, much faster than the other typical lengths of R_n and P_n. To start, we keep h_n free, as well as the other parameters of the rectangles, but we shall fix them later.

Proof of Theorem 5.2. We write $P_n = [t_n, t'_n] \times [-h_n/2, h_n/2]$, and set

$$\phi_n(x) = |x_1 - t_n|, \qquad \phi'_n(x) = |x_1 - t'_n| \tag{5.8}$$

for $x = (x_1, x_2) \in \mathbb{R}^2$. Since ϕ_n (resp. ϕ'_n) is convex, it differs from the Kantorovich potential from ρ to $(\nabla\phi_n)_\#\rho$ (resp. $(\nabla\phi'_n)_\#\rho$) only by a constant. Now, the idea is that $\nabla\phi_n$ and $\nabla\phi'_n$ coincide on $\mathcal{X} \setminus P_n$, and this set has ρ-volume almost 1, which makes $(\nabla\phi_n)_\#\rho$ and $(\nabla\phi'_n)_\#\rho$ extremely close in Wasserstein distance: their Wasserstein distance is proportional to $\rho(P_n)$ which is of order $h_n(t'_n - t_n)$. The quantity $\mathrm{Var}(\phi_n - \phi'_n)$ is much larger (but very small too!) since $|\phi_n - \phi'_n|$ is equal to $|t'_n - t_n|$ in the largest part of $\mathcal{X}$.

More precisely, both $\mu_n = (\nabla\phi_n)_\#\rho$ and $\mu'_n = (\nabla\phi'_n)_\#\rho$ are supported on $\{A, B\}$ where $A = (-1, 0)$ and $B = (1, 0)$, and the subset of points of $\mathcal{X}$ such that $\nabla\phi_n \neq \nabla\phi'_n$ is P_n. Since $\mathrm{dist}(A, B) = 2$, we get that for any $p \geq 1$,

$$W_p(\mu_n, \mu'_n) = 2\rho(P_n)^{1/p}. \tag{5.9}$$

We turn to the computation of $\mathrm{Var}_\rho(\phi_n - \phi'_n)$. For this, we observe that

$$\phi'_n(x) - \phi_n(x) = \begin{cases} t'_n - t_n & \text{if } x_1 \leq t_n \\ t_n - t'_n & \text{if } x_1 \geq t'_n \end{cases} \tag{5.10}$$

and

$$|\phi_n(x) - \phi_n'(x)| \le |t_n - t_n'| \text{ if } x \in P_n. \tag{5.11}$$

Therefore

$$\|\phi_n - \phi_n'\|_{L^2(\rho)}^2 \ge |t_n - t_n'|^2(1 - \rho(P_n)). \tag{5.12}$$

Then, we evaluate the mean of $\phi_n - \phi_n'$. We set

$$v_n = \rho(\{x \in \mathcal{X} \mid x_1 \le t_n\}) \qquad \text{and} \qquad w_n = \rho(\{x \in \mathcal{X} \mid x_1 \ge t_n'\}).$$

Then $v_n \to 1$ and $w_n \to 0$ as $n \to +\infty$, and for any $n \in \mathbb{N}^*$,

$$v_n + w_n + \rho(P_n) = 1.$$

Therefore, for n large enough, using (5.10) and (5.11),

$$\begin{aligned} 0 \le \int_{\mathcal{X}} (\phi_n' - \phi_n)\mathrm{d}\rho &\le (t_n' - t_n)v_n + (t_n - t_n')w_n + \rho(P_n)|t_n - t_n'| \\ &= (1 - 2w_n)(t_n' - t_n). \end{aligned} \tag{5.13}$$

We deduce that

$$\begin{aligned} \mathrm{Var}_\rho(\phi_n - \phi_n') &= \|\phi_n - \phi_n'\|_{L^2(\rho)}^2 - \left(\int_{\mathcal{X}} (\phi_n - \phi_n')\mathrm{d}\rho\right)^2 \\ &\ge |t_n - t_n'|^2(1 - \rho(P_n) - (1 - 2w_n)^2) \\ &= |t_n - t_n'|^2(4w_n - \rho(P_n) - 4w_n^2) \\ &\ge |t_n - t_n'|^2(\rho(R_{n+1}) - \rho(P_n)) \end{aligned} \tag{5.14}$$

since $w_n \to 0$ and $w_n \ge \rho(R_{n+1})$.

There remains to choose the parameters of the rooms R_n and the passages P_n. We choose h_n small enough compared to all other lengths, in particular $\rho(R_{n+1}) - \rho(P_n) \ge \frac{1}{2}\rho(R_{n+1}) \gtrsim \lambda(R_{n+1})$ where recall that ρ has density bounded below on $\mathcal{X}$. Then for any $\alpha > 0$ and $p \in [1, +\infty)$,

$$\frac{W_p(\mu_n, \mu_n')^\alpha}{\mathrm{Var}_\rho(\phi_n - \phi_n')} \lesssim \frac{\lambda(P_n)^{\alpha/p}}{|t_n - t_n'|^2\lambda(R_{n+1})} \lesssim \frac{|t_n - t_n'|^{(\alpha-2p)/p} h_n^{\alpha/p}}{\lambda(R_{n+1})}.$$

and we see that choosing h_n small enough compared to all other parameters, this quantity tends to 0 as $n \to +\infty$, for any $\alpha > 0$, $p \in [1, +\infty)$, which concludes the proof. □

Remark 5.3. *The above computations being essentially* 1*-dimensional, one may easily turn them into an example of a source measure ρ whose support is a segment of $\mathbb{R}$, and for which stability does not hold even in a very weak sense.*

Beside room-and-passage domains, domains with an outward cusp are another well-known category of non-John domains. By outward cusp, we mean that in some local coordinates, the equation defining the domain is $|y| \le f(x)$ for some $f : [0, +\infty) \to \mathbb{R}$ with $f'(0) = 0$. For instance $f(x) = x^s$ with $s > 1$, or $f(x) = e^{-1/x^2}$. The former are called (polynomial) s-cusps, the latter exponential cusps. It is not difficult to modify the above computations to show that domains with an exponential cusp could also be used to prove Theorem 5.2.

Regarding polynomial cusps, we need another definition. For $s \ge 1$, an s-John domain is a domain for which the condition (2.6) is replaced by

$$\mathrm{dist}(\gamma(t), \mathcal{X}^c) \ge \eta t^s$$

i.e., the same condition as John domains except that t is replaced by t^s in the right-hand side. It follows that s-John domains may have polynomial s-cusps.

Open question 5.4. *Does a stability inequality $\|\phi_\mu - \phi_\nu\|_{L^2(\rho)} \le CW_p(\mu,\nu)^\alpha$ hold in s-John domains, for some C, p, α which may depend on the domain (notably on s)?*

In s-John domains with large s, the exponent α necessarily has to be less good (i.e., larger) than 1/2. The reader can compute as an exercise an upper bound on the stability exponent in a domain containing an s-cusp, using similar sequences as in (5.8).

Let us hazard a final speculative comment: it seems to us that there is at least a formal resemblance between the example detailed in Section 5.3 and the Kannan-Lovasz-Simonovitz conjecture. This conjecture asserts that the Poincaré constant of log-concave measures can be checked on linear test functions. The analogy we see with our example is that we only need simple test functions (namely, distance functions to hyperplanes) to prove that optimal transport potentials are unstable. Therefore, it is tempting to formulate the following vague question:

Open question 5.5. *Is it true that for more general ρ's, some simple family of test functions is sufficient to guarantee stability/instability of optimal transport potentials?*

5.4 Bibliographical notes

The examples in Sections 5.1 and 5.2 come respectively from [49] and [54].

§5.3: The example presented in this section comes from [54]. Room-and-passage domains date back at least to the 1937 monograph by Courant and Hilbert [26, pp. 521-523], who used them to show that the embedding of $H^1(\mathcal{X})$ in $L^2(\mathcal{X})$ is not necessarily compact (see also [5]). Indeed, consider ϕ_n a function which is equal to a constant in the n-th room R_n and which drops linearly to 0 in the adjacent passages P_{n-1} and P_n, reaching the value 0 at the midpoint of each of these passages. Choosing the constant in a way that $\|\phi_n\|_{L^2(\mathcal{X})} = 1$, we obtain an orthonormal family of functions. If the passages are narrow enough (i.e., if h_n is small enough), then $\|\nabla\phi_n\|_{L^2(\mathcal{X})} \to 0$ as $n \to +\infty$. Since no subsequence of (ϕ_n) converges in $L^2(\mathcal{X})$, $H^1(\mathcal{X})$ is not compactly embedded in $L^2(\mathcal{X})$. This example is fundamentally related to the fact that the Poincaré inequality fails in $\mathcal{X}$, and that 0 is in the essential spectrum of the Neumann Laplacian on $\mathcal{X}$. In this direction, Hempel-Seco-Simon [42] used room-and-passage domains to provide examples of domains with prescribed essential spectrum. Finally, many papers have considered s-John domains, see for instance [38].

6 Stability of optimal transport maps: proof ideas

To establish stability of optimal transport maps, the main tool that we use is an inequality of the form

$$\|\nabla f - \nabla g\|_{L^2(\rho)} \leq C L^{2/3} \|f - g\|_{L^2(\rho)}^{1/3} \tag{6.1}$$

for f, g convex and L-Lipschitz on the support of ρ (of course (6.1) cannot hold without some assumption on f, g). Applying it to $f = \phi_\mu$, $g = \phi_\nu$ and using the stability of Kantorovich potentials, it immediately yields the stability of optimal transport maps in John domains (2.5). To get stability of optimal transport maps when the source measure ρ is log-concave, i.e., (2.3), one needs additional truncation arguments that we will not detail here.

The surprising inequality (6.1) crucially relies on the assumption that both ϕ_μ and ϕ_ν are convex. Its proof relies on two main ingredients:

- first, a one-dimensional version of the inequality (6.1);
- second, integral geometric techniques which allow to extend the one-dimensional inequality to higher dimensions.

Until now, this strategy has not led to optimal results, in the sense that the exponent $1/6$ in Theorem 2.3 does not match the upper bound $1/2$ on the exponent provided by the explicit example written in Section 2.4. Perhaps one would need a direct approach to stability of optimal transport maps, not going through stability of Kantorovich potentials, to get sharper exponents.

6.1 The 1d inequality

We start with a one-dimensional version of (6.1).

Proposition 6.1. *Let $I \subset \mathbb{R}$ be a compact segment. Let $u, v : I \to \mathbb{R}$ be two convex functions whose derivatives (defined a.e. on I) are uniformly bounded over I. Then*

$$\|u' - v'\|_{L^2(I)}^2 \leq 8(\|u'\|_{L^\infty(I)} + \|v'\|_{L^\infty(I)})^{4/3} \|u - v\|_{L^2(I)}^{2/3}. \tag{6.2}$$

This inequality looks like a Poincaré inequality, but in the wrong sense! It holds only because we are applying it to a difference of convex functions. Indeed, taking $I = [0,1]$, $u = 0$ and $v = \sin(nx)$ shows that (6.2) cannot hold without assuming something on u, v. One can get some intuition about (6.2) by drawing the graphs of u' and v', which are non-decreasing functions. Then u, v are obtained as areas under the curves and it may be seen that $|u' - v'|$ cannot be large on some quantitative fraction of I without having $|u - v|$ large at some point. Another remark is that (6.2) is invariant under affine transformations, hence it is sufficient to prove the result on $I = [0,1]$. Finally, the exponents in (6.2) are optimal, as may be seen by taking $u(x) = L|x - \frac{1}{2}|$ and $v = \max(u, \varepsilon)$.

We shall only streamline the proof of Proposition 6.1, and for $I = [0,1]$ (which we may assume thanks to a scaling argument). First integrating by parts,

$$\int_0^1 |u' - v'|^2 = \left[(u - v)(u' - v')\right]_0^1 - \int_0^1 (u - v)(u'' - v'')$$
$$\leq 2\|u - v\|_{L^\infty}(\|u'\|_{L^\infty} + \|v'\|_{L^\infty}) + \|u - v\|_{L^\infty}\left(\int_0^1 |u''| + \int_0^1 |v''|\right)$$

But since u is convex,

$$\int_0^1 |u''| = \int_0^1 u'' = u'(1) - u'(0) \leq 2\|u'\|_{L^\infty},$$

and similarly for v, thus we conclude that

$$\int_0^1 |u' - v'|^2 \leq 4\|u - v\|_{L^\infty}(\|u'\|_{L^\infty} + \|v'\|_{L^\infty}). \tag{6.3}$$

The second step is to bound the L^∞ norm of $f = u - v$ with its L^2-norm using that the Lipschitz constant of $f = u - v$ is less than $\|u'\|_{L^\infty} + \|v'\|_{L^\infty}$. This second step does not use the fact that f is the difference of two convex functions: considering the worst case scenario where f is piecewise affine, equal to 0 except around the maximum of $\|f\|_{L^\infty}$ where it looks like a "tent", we get

$$\|f\|^2_{L^2([0,1])} \geq \frac{1}{4} \min\left(\frac{\|f\|_{L^\infty}}{2\|f\|_{\mathrm{Lip}}}, 1\right) \|f\|^2_{L^\infty}.$$

And a little more work combined with (6.3) allows to conclude.

6.2 Higher dimension: an integral-geometric argument

We prove the following generalization of Proposition 6.1 to higher dimensions

Proposition 6.2. *Let $L > 0$ and let K be a compact subset of $\mathbb{R}^d$ whose boundary has finite $(d-1)$-dimensional measure. Then there exists $C > 0$ such that for any $u, v : K \to \mathbb{R}$ convex on any segment included in K and L-Lipschitz,*

$$\|\nabla u - \nabla v\|_{L^2(K)} \leq C\|u - v\|^{1/3}_{L^2(K)}.$$

Then the stability of maps (2.5) follows immediately by combining (2.4) with Proposition 6.2 for $K = \overline{\mathcal{X}}$. To prove Proposition 6.2, one possibility is to rely on integral-geometric techniques, i.e., expressing a multidimensional integral in terms of integrals over lines (or geodesics, in Riemannian geometry). We start from the formula

$$\int_{\mathbb{R}^d} f(x)^2 \mathrm{d}x = \int_{e^\perp} \int_{\mathbb{R}} f(y + te)^2 \mathrm{d}t \mathrm{d}y.$$

valid for any $e \in \mathbb{S}^{d-1}$, where $e^\perp$ denotes the hyperplane (through the origin) perpendicular to the unit vector e. Applying this to $f(x) = \langle F(x), e\rangle$, and then integrating over $e \in \mathbb{S}^{d-1}$, we get

$$\int_{\mathbb{S}^{d-1}} \int_{\mathbb{R}^d} \langle F(x), e\rangle^2 \mathrm{d}x \mathrm{d}\sigma(e) = \int_{\mathbb{S}^{d-1}} \int_{e^\perp} \int_{\mathbb{R}} \langle F(y + te), e\rangle^2 \mathrm{d}t \mathrm{d}y \mathrm{d}\sigma(e)$$

where σ is the uniform probability measure on $\mathbb{S}^{d-1}$. We observe that the LHS is equal to $C_d\|F\|^2_{L^2(\mathbb{R}^d)}$ for some $C_d > 0$ depending only on d. We apply this to F given by $\nabla u - \nabla v$ inside K, and extended by 0 outside K. We get

$$\|\nabla u - \nabla v\|^2_{L^2(K)} = C_d^{-1} \int_{\mathbb{S}^{d-1}} \int_{e^\perp} \|u'_{\ell_e^y} - v'_{\ell_e^y}\|^2_{L^2(\ell_e^y \cap K)} \mathrm{d}y \mathrm{d}\sigma(e) \tag{6.4}$$

where ℓ_e^y denotes the oriented line $y + e\mathbb{R}$ and $u_{\ell_e^y} = u|_{\ell_e^y \cap K}$, $v_{\ell_e^y} = v|_{\ell_e^y \cap K}$. The set $\ell_e^y \cap K$ may be decomposed as a finite union of intervals $I^i_{\ell_e^y}$, $i = 1, \ldots, n_{\ell_e^y}$ in which we can apply Proposition 6.1 (the 1d inequality, using that $\|u'\|_{L^\infty}, \|v'\|_{L^\infty} \leq L$). We get

$$\begin{aligned}\|u'_{\ell_e^y} - v'_{\ell_e^y}\|^2_{L^2(\ell_e^y \cap K)} = \sum_{i=1}^{n_{\ell_e^y}} \|u'_{\ell_e^y} - v'_{\ell_e^y}\|^2_{L^2(I^i_{\ell_e^y})} &\leq 8(2L)^{4/3} \sum_{i=1}^{n_{\ell_e^y}} \|u_{\ell_e^y} - v_{\ell_e^y}\|^{2/3}_{L^2(I^i_{\ell_e^y})} \\ &\leq 8(2L)^{4/3} n_{\ell_e^y}^{2/3} \|u_{\ell_e^y} - v_{\ell_e^y}\|^{2/3}_{L^2(\ell_e^y \cap K)}\end{aligned}$$

where the last inequality comes from Jensen's inequality. Combining this with (6.4) and then using Hölder's inequality, we get

$$\begin{aligned}\|\nabla u-\nabla v\|_{L^2(K)}^2 &\lesssim \int_{\mathbb{S}^{d-1}}\int_{e^\perp} n_{\ell_e^y}^{2/3}\|u_{\ell_e^y}-v_{\ell_e^y}\|_{L^2(\ell_e^y\cap K)}^{2/3}\mathrm{d}y\mathrm{d}\sigma(e)\\ &\lesssim \left(\int_{\mathbb{S}^{d-1}}\int_{e^\perp} n_{\ell_e^y}\mathrm{d}y\mathrm{d}\sigma(e)\right)^{2/3}\left(\int_{\mathbb{S}^{d-1}}\int_{e^\perp}\|u_{\ell_e^y}-v_{\ell_e^y}\|_{L^2(\ell_e^y\cap K)}^{2}\mathrm{d}y\mathrm{d}\sigma(e)\right)^{1/3}.\end{aligned}$$

The second parenthesis is equal to $C_d\|u-v\|_{L^2(K)}^2$ due to the same argument which led to (6.4). Regarding the first parenthesis, we observe that $n_{\ell_e^y}\le \#(\ell_e^y\cap\partial K)$ and then we use the Cauchy-Crofton formula, which asserts that

$$\int_{\mathbb{S}^{d-1}}\int_{e^\perp}\#(\ell_e^y\cap\partial K)\mathrm{d}y\mathrm{d}\sigma(e)=\mathscr{H}^{d-1}(\partial K)<+\infty \tag{6.5}$$

where $\mathscr{H}^{d-1}$ denotes the $(d-1)$-dimensional Hausdorff measure. This concludes the proof of Proposition 6.2.

Here, we should warn the reader that for the Cauchy-Crofton formula (6.5) to hold, one actually needs to assume that the boundary ∂K is rectifiable. If this is not assumed, then it is still true that the left-hand side in (6.5) is finite when $\mathscr{H}^{d-1}(\partial K)<+\infty$ (but the equality in (6.5) does not necessarily hold). To prove this and also to extend Theorem 2.3 to Riemannian manifolds, we devised a more robust argument based on the definition of the integral-geometric measure, with the help of Antoine Julia and Federer's book [34]. This outer measure, defined following Caratheodory's construction, compares easily (almost by definition) with the $(d-1)$-dimensional Hausdorff measure and may be shown to count the number of intersections of short geodesic curves with $\partial\mathcal{X}$.

6.3 Where do we lose sharpness of the exponents?

If we summarize our proofs, we have proved stability of optimal transport maps thanks to a chain of inequalities of the form

$$\|T_\mu-T_\nu\|_{L^2(\rho)}^6\lesssim\|\phi_\mu-\phi_\nu\|_{L^2(\rho)}^2\lesssim\langle\psi_\nu-\psi_\mu,\mu-\nu\rangle\lesssim W_1(\mu,\nu).$$

As seen in Section 5.1 the two inequalities on the right become asymptotically sharp as $d\to+\infty$. The inequality on the left is sharp in any dimension. However, the example proving this sharpness is not the same as in Section 5.1: take $\rho=\mathbf{1}_{[0,1]^d}$ and

$$\phi_\varepsilon^{(1)}(x)=|x_1|,\qquad \phi_\varepsilon^{(2)}(x)=\max(\phi_\varepsilon^{(1)},\varepsilon)$$

for $x=(x_1,\dots,x_d)\in\mathbb{R}^d$. Then $\|\nabla\phi_\varepsilon^{(1)}-\nabla\phi_\varepsilon^{(2)}\|_{L^2(\rho)}^2=\varepsilon$, whereas $\mathrm{Var}(\phi_\varepsilon^{(1)}-\phi_\varepsilon^{(2)})\approx\varepsilon^3$, showing the sharpness of the exponents in the left hand side inequality.

In other words, the cases of equality in this chain of inequalities do not match, and we have no example where this chain of inequalities is indeed saturated. Thus, one of the main open problems that remains to be solved is the following:

Open question 6.3. *Obtain sharp exponents for the stability of optimal transport maps (in* (2.3) *and* (2.5) *for instance).*

6.4 Bibliographical notes

§6.1: The one-dimensional inequality given in Proposition 6.1 was proved in [31]. This is a refinement of Theorem 3.5 in [21], in which the upper bound involved the uniform distance $\|u - v\|_{L^\infty}$.

§6.2: Integral-geometric techniques are explained in Federer's book [34]. As already mentioned, there is an alternative path to prove Proposition 6.2, which relies on the Caratheodory construction, see for instance [34, Chapter 2.10]. This alternative strategy was used in [49] to extend Theorem 2.3 to Riemannian manifolds.

References

[1] Martial Agueh, and Guillaume Carlier. "Barycenters in the Wasserstein space." SIAM Journal on Mathematical Analysis 43.2 (2011): 904-924.

[2] Luigi Ambrosio, Nicola Gigli, and Giuseppe Savaré. Gradient flows: in metric spaces and in the space of probability measures. Springer Science & Business Media, 2008.

[3] Luigi Ambrosio, Federico Glaudo, and Dario Trevisan. "On the optimal map in the 2-dimensional random matching problem." Discrete & Continuous Dynamical Systems: Series A 39.12 (2019).

[4] Alexandr Andoni, Assaf Naor, and Ofer Neiman. "Snowflake universality of Wasserstein spaces." Annales Scientifiques de l'Ecole Normale Superieure. Vol. 51. No. 3. 2018.

[5] Charles J. Amick. "Some remarks on Rellich's theorem and the Poincaré inequality." Journal of the London Mathematical Society 2.1 (1978): 81-93.

[6] Sivaraman Balakrishnan, and Tudor Manole. "Stability Bounds for Smooth Optimal Transport Maps and their Statistical Implications." arXiv preprint arXiv:2502.12326 (2025).

[7] Saurav Basu, Soheil Kolouri, and Gustavo K. Rohde. "Detecting and visualizing cell phenotype differences from microscopy images using transport-based morphometry." Proceedings of the National Academy of Sciences 111.9 (2014): 3448-3453.

[8] Robert J. Berman. "Convergence rates for discretized Monge–Ampère equations and quantitative stability of optimal transport." Foundations of Computational Mathematics 21.4 (2021): 1099-1140.

[9] Sergei Bobkov and Michel Ledoux. "From Brunn-Minkowski to Brascamp-Lieb and to logarithmic Sobolev inequalities. " Geometric and Functional Analysis, Vol. 10 (2000) 1028-1052.

[10] Bogdan Bojarski. "Remarks on Sobolev imbedding inequalities." Complex Analysis Joensuu 1987: Proceedings of the XIIIth Rolf Nevanlinna-Colloquium, held in Joensuu, Finland, Aug. 10–13, 1987. Berlin, Heidelberg: Springer Berlin Heidelberg, 2006.

[11] Jan Boman. "L^p-estimates for very strongly elliptic systems". Department of Mathematics, University of Stockholm, Sweden, 1982. Available at https://mathoverflow.net/questions/320172/famous-but-unavailable-paper-of-jan-boman

[12] Nicolas Bonneel, Julien Rabin, Gabriel Peyré, Hanspeter Pfister. "Sliced and Radon Wasserstein barycenters of measures." Journal of Mathematical Imaging and Vision 51 (2015): 22-45.

[13] Herm Jan Brascamp, and Elliott H. Lieb. "On extensions of the Brunn-Minkowski and Prékopa-Leindler theorems, including inequalities for log concave functions, and with an application to the diffusion equation." Journal of Functional Analysis 22.4 (1976): 366-389.

[14] Yann Brenier. "Décomposition polaire et réarrangement monotone des champs de vecteurs." CR Acad. Sci. Paris Sér. I Math. 305 (1987): 805-808.

[15] Yann Brenier. "Polar factorization and monotone rearrangement of vector-valued functions." Communications on Pure and Applied Mathematics 44.4 (1991): 375-417.

[16] Stephen M. Buckley, and Pekka Koskela. “Sobolev-Poincaré implies John.” Mathematical Research Letters 2.5 (1995): 577-593.

[17] Stephen M. Buckley, Pekka Koskela, and Guozhen Lu. “Boman equals John.” Proceedings of the XVI Rolf Nevanlinna Colloquium. W. de Gruyter, 1996.

[18] Luis A. Caffarelli. “The regularity of mappings with a convex potential.” Journal of the American Mathematical Society 5.1 (1992): 99-104.

[19] Luis A. Caffarelli. “Monotonicity properties of optimal transportation and the FKG and related inequalities.” Communications in Mathematical Physics 214 (2000): 547-563.

[20] Guillaume Carlier, Alex Delalande and Quentin Merigot. “Quantitative stability of barycenters in the Wasserstein space.” Probability Theory and Related Fields 188.3 (2024): 1257-1286.

[21] Frédéric Chazal, David Cohen-Steiner, and Quentin Mérigot. “Boundary measures for geometric inference”. Foundations of Computational Mathematics, 10(2):221–240, 2010.

[22] Victor Chernozhukov, Alfred Galichon, Marc Hallin, and Marc Henry. “Monge-Kantorovich depth, quantiles, ranks and signs.” Annals of Statistics 45.1 (2017): 223-256.

[23] Sinho Chewi, Jonathan Niles-Weed, and Philippe Rigollet. “Statistical optimal transport.” arXiv preprint arXiv:2407.18163 (2024).

[24] Lénaïc Chizat, Alex Delalande, Tomas Vaškevičius. “Sharper exponential convergence rates for Sinkhorn’s algorithm in continuous settings.” Arxiv preprint arXiv:2407.01202 (2024).

[25] Fan RK Chung. Spectral graph theory. Vol. 92. American Mathematical Soc., 1997.

[26] Richard Courant, and David Hilbert. Methoden der mathematischen Physik, Vol. 2, Springer, Berlin, 1937.

[27] Marco Cuturi. “Sinkhorn distances: Lightspeed computation of optimal transport.” Advances in neural information processing systems 26 (2013).

[28] Marco Cuturi, and Arnaud Doucet. ”Fast computation of Wasserstein barycenters.” International conference on machine learning. PMLR, 2014.

[29] Alex Delalande. Quantitative Stability in Quadratic Optimal Transport. Diss. Université Paris-Saclay, 2022.

[30] Alex Delalande. “Nearly tight convergence bounds for semi-discrete entropic optimal transport.” International Conference on Artificial Intelligence and Statistics. PMLR, 2022.

[31] Alex Delalande and Quentin Merigot. “Quantitative stability of optimal transport maps under variations of the target measure.” Duke Mathematical Journal 172.17 (2023): 3321-3357.

[32] Guido De Philippis, and Alessio Figalli. “The Monge–Ampère equation and its link to optimal transportation.” Bulletin of the American Mathematical Society 51.4 (2014): 527-580.

[33] Vincent Divol, Jonathan Niles-Weed, and Aram-Alexandre Pooladian. “Tight stability bounds for entropic Brenier maps.” arXiv preprint arXiv:2404.02855 (2024).

[34] Herbert Federer. Geometric measure theory. Springer-Verlag Berlin Heidelberg New York, 1969

[35] Anatole Gallouët, Quentin Mérigot, and Boris Thibert. "Strong c-concavity and stability in optimal transport." arXiv preprint arXiv:2207.11042 (2022).

[36] Nicola Gigli. "On Hölder continuity-in-time of the optimal transport map towards measures along a curve." Proceedings of the Edinburgh Mathematical Society 54.2 (2011): 401-409.

[37] Loukas Grafakos. Classical Fourier Analysis. Graduate texts in mathematics Vol. 249. Springer, 2008.

[38] Piotr Hajłasz, and Pekka Koskela. "Isoperimetric inequalities and imbedding theorems in irregular domains." Journal of the London Mathematical Society 58.2 (1998): 425-450.

[39] Piotr Hajłasz, and Pekka Koskela. Sobolev met Poincaré. Vol. 688. American Mathematical Soc., 2000.

[40] Marc Hallin. "On distribution and quantile functions, ranks and signs in $\mathbb{R}^d$: a measure transportation approach." Preprint 2017, available at https://ideas.repec.org/p/eca/wpaper/2013-258262.html

[41] Bernard Helffer, and Johannes Sjöstrand. "On the correlation for Kac-like models in the convex case." Journal of statistical physics 74 (1994): 349-409.

[42] Rainer Hempel, Luis A. Seco, and Barry Simon. "The essential spectrum of Neumann Laplacians on some bounded singular domains." Journal of Functional Analysis 102.2 (1991): 448-483.

[43] Jean-Baptiste Hiriart-Urruty and Claude Lemaréchal. Fundamentals of convex analysis'. Springer, 2004.

[44] Jan-Christian Hütter, and Philippe Rigollet. "Minimax estimation of smooth optimal transport maps." The Annals of Statistics, 49(2) (2021): 1166-1194.

[45] Fritz John. "Rotation and strain." Communications on Pure and Applied Mathematics 14.3 (1961): 391-413.

[46] Jon Johnsen. "On the spectral properties of Witten-Laplacians, their range projections and Brascamp-Lieb's inequality." Integral Equations and Operator Theory 36.3 (2000): 288-324.

[47] Leonid Kantorovich. "On the translocation of masses." Dokl. Akad. Nauk. USSR 37 (1942), 199–201. English translation in Management science 5.1 (1958): 1-4.

[48] Matthias Keller, Daniel Lenz, and Radoslaw K. Wojciechowski. Graphs and discrete Dirichlet spaces. Vol. 358. Berlin, Germany: Springer, 2021.

[49] Jun Kitagawa, Cyril Letrouit, and Quentin Mérigot. "Stability of optimal transport maps on Riemannian manifolds." ArXiv preprint arXiv:2504.05412.

[50] Boaz Klartag. Regularity through convexity in high dimensions. Lecture notes available at https://www.weizmann.ac.il/math/klartag/sites/math.klartag/files/uploads/klartag-notes-by-brazitikos.pdf

[51] Soheil Kolouri, and Gustavo K. Rohde. “Transport-based single frame super resolution of very low resolution face images.” Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2015.

[52] Thibaut Le Gouic, and Jean-Michel Loubes. “Existence and consistency of Wasserstein barycenters.” Probability Theory and Related Fields 168 (2017): 901-917.

[53] Dorian Le Peutrec. “On Witten Laplacians and Brascamp–Lieb’s inequality on manifolds with boundary.” Integral Equations and Operator Theory 87.3 (2017): 411-434.

[54] Cyril Letrouit, and Quentin Mérigot. “Gluing methods for quantitative stability of optimal transport maps.” arXiv preprint arXiv:2411.04908 (2024).

[55] Wenbo Li, and Ricardo H. Nochetto. “Quantitative stability and error estimates for optimal transport plans.” IMA Journal of Numerical Analysis 41.3 (2021): 1941-1965.

[56] Olli Martio, and Jukka Sarvas. “Injectivity theorems in plane and space.” Annales Fennici Mathematici 4.2 (1979): 383-401.

[57] Robert J. McCann. “A convexity principle for interacting gases.” Advances in mathematics 128.1 (1997): 153-179.

[58] Robert J. McCann. ”Polar factorization of maps on Riemannian manifolds.” Geometric & Functional Analysis 11.3 (2001): 589-608.

[59] Quentin Mérigot, Alex Delalande, and Frederic Chazal. “Quantitative stability of optimal transport maps and linearization of the 2-Wasserstein space.” International Conference on Artificial Intelligence and Statistics. PMLR, 2020.

[60] Gaspard Monge. “Mémoire sur la théorie des déblais et des remblais.” In Histoire de l’Académie Royale des Sciences de Paris (1781), pp. 666–704.

[61] Octave Mischler and Dario Trevisan. “Quantitative stability in optimal transport for general power costs.” arXiv preprint arXiv:2407.19337 (2024).

[62] Raimo Näkki, and Jussi Väisälä. “John disks.” Expositiones Math. 9 (1991), 3–43.

[63] Felix Otto. “The geometry of dissipative evolution equations: the porous medium equation.” Comm. Partial Differential Equations 26 (2001): 101-174.

[64] Gabriel Peyré, and Marco Cuturi. “Computational optimal transport: With applications to data science.” Foundations and Trends in Machine Learning 11.5-6 (2019): 355-607.

[65] Julien Rabin, Gabriel Peyré, Julie Delon and Marc Bernot. “Wasserstein barycenter and its application to texture mixing.” International conference on scale space and variational methods in computer vision. Berlin, Heidelberg: Springer Berlin Heidelberg, 2011.

[66] Filippo Santambrogio. Optimal Transport for Applied Mathematicians: Calculus of Variations, PDEs, and Modeling. Progress in Nonlinear Differential Equations and Their Applications, Vol. 87. Birkhäuser, 2015.

[67] Daniel Spielman. Spectral and Algebraic Graph Theory. Ongoing draft available at http://cs-www.cs.yale.edu/homes/spielman/sagt/

[68] Sanvesh Srivastava, Cheng Li, and David B. Dunson. "Scalable Bayes via barycenter in Wasserstein space." Journal of Machine Learning Research 19.8 (2018): 1-35.

[69] Justin Solomon, Fernando de Goes, Gabriel Peyré, Marco Cuturi, Adrian Butscher, Andy Nguyen, Tao Du, Leonidas Guibas. "Convolutional wasserstein distances: Efficient optimal transportation on geometric domains." ACM Transactions on Graphics (ToG) 34.4 (2015): 1-11.

[70] Luca Trevisan. Lecture notes on graph partitioning, expanders and spectral methods. Available at https://lucatrevisan.github.io/books/expanders-2016.pdf

[71] Cédric Villani. Topics in optimal transportation. Graduate studies in mathematics, Vol. 58. American Mathematical Soc., 2021.

[72] Cédric Villani. Optimal transport: old and new. Grundlehren der mathematischen Wissenschaften, Vol. 338. Berlin: Springer, 2009.

[73] Wei Wang, Dejan Slepčev, Saurav Basu, John A. Ozolek, and Gustavo K. Rohde. "A linear optimal transportation framework for quantifying and visualizing variations in sets of images." International journal of computer vision 101 (2013): 254-269.