# Lectures on quantitative stability of optimal transport

Cyril Letrouit

May 14, 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

**1 General introduction** **2**

**2 Main results** **11**

Program of the lectures:

- May 14th: General introduction (Section 1), statement of the main results (Section 2).
- May 21st: The Kantorovich functional (Section **??**). This lecture is mostly based on a work by Alex Delalande and Quentin Mérigot [26], revisited in [47].
- May 28th: Gluing techniques (Section **??**), examples and counterexamples (Section **??**). This lecture is mostly based on my joint work with Quentin Mérigot [47].
- June 4th: Stability of maps (Section **??**), generalizations (Section **??**) and applications (Section **??**). 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 $\rho$ be a probability measure on a Polish (i.e., complete, separable metric) space $\mathcal{X}$ and $\mu$ 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 $\gamma$ of the space $\mathcal{P}(\mathcal{X} \times \mathcal{Y})$ of probability measures over $\mathcal{X} \times \mathcal{Y}$ whose marginals coincide with $\rho$ and $\mu$, 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 $\det(D^2 f) = g$ (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 $\mu_n$ to their limit $\mu$ (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 $\rho$, the Monge problem does not necessarily have a solution. If $\rho$ is a sum of Dirac masses but $\mu$ 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 $\rho$ 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 $\rho$ is absolutely continuous with respect to the Lebesgue measure. Then there exists between $\rho$ and $\mu$ a $\rho$-a.e. unique optimal transport map $T$ and a unique optimal transport plan $\gamma$, 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 $\nabla f = \nabla\phi$ $\rho$-a.e.*

If the support $\mathcal{X}$ of $\rho$ is the closure of a bounded connected open set, $\phi$ is uniquely determined on $\mathcal{X}$ up to additive constants. As a consequence of Brenier's theorem, for any convex function $\phi$ and any absolutely continuous $\rho \in \mathcal{P}_2(\mathbb{R}^d)$, the map $\nabla\phi$ is the optimal transport map from $\rho$ 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 $\rho$ and $\mu$. 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 $\rho$ and $(\mu_k)_{k\in\mathbb{N}}$ converge weakly to $\mu$. For each $k \in \mathbb{N}$, let $\gamma_k$ be an optimal transport plan between $\rho_k$ and $\mu_k$, 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 $\rho$ and $\mu$ is finite and, up to extraction of a subsequence, $\gamma_k$ converges weakly to some optimal transport plan $\gamma$ between $\rho$ and $\mu$.*

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 $\rho$ and consider stability with respect to the target measure only. The problem we are interested in reads

*If $\mu$ and $\nu$ 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 $\rho$ to $\mu$ (resp. $\rho$ to $\nu$) 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 $\rho$ and $\rho'$ 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, $\rho$ 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 $\rho$ consists of multiple connected components.

The source measure $\rho$ 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 $\mu$ 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, $(T_{\mu_n})$ is bounded in $L^2(\rho)$.

Let

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

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 $\eta > 0$,

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

where $\lambda$ denotes the Lebesgue measure on $\mathbb{R}^d$. The proof of (1.7) is provided in Appendix **??**. 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 $\phi_{\mu_n}$, using that $\rho(x) \geq \varepsilon$ on $K_\varepsilon$, and finally applying the Cauchy-Schwarz inequality. Since $(T_{\mu_n})$ 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 $\varepsilon$ 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 $\rho$. In particular, this implies (1.6).

From now on, we normalize $\phi_{\mu_n}$ in a way that $\phi_{\mu_n}(0) = 0$. By Arzelà-Ascoli, up to extraction of a subsequence omitted in the notation, $(\phi_{\mu_n})$ converges toward some $\phi$ uniformly over any $K_\varepsilon$. Of course, $\phi$ 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 $\phi$, $(\nabla \phi_{\mu_n})$ converges to $\nabla \phi$. Since $\phi$ is convex, it is differentiable almost everywhere, thus $T_{\mu_n}(x) \to T(x)$ for $\rho$-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 $(T_{\mu_n})$ 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 $(T_{\mu_n})$ in fact converges *strongly* to $T$ in $L^2(\rho)$.

Finally, let us observe that $T_\# \rho = \mu$ since $(T_{\mu_n})$ converges a.e. to $T$ and $(T_{\mu_n})$ 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 $\rho$ to $\mu$. We conclude that the full sequence $(T_{\mu_n})$ 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 $\mu, \nu$ 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 $\alpha$ as large as possible. The largest possible $\alpha$ is sometimes called the stability exponent (associated to $\rho$) 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 $\mu$ and $\nu$, 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 $\mu$ and $\nu$.

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 $\mu$ and $\nu$ 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 $\mu, \nu$. For instance, we will show that under some assumptions on $\rho$, for any compact set $\mathcal{Y} \subset \mathbb{R}^d$, there exist $C, \alpha > 0$ such that (1.10) holds for any $\mu, \nu$ 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 **??**. 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 **??**. 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 **??**). Finally, Kantorovich potentials have an economic interpretation which may help understand their meaning (see [64, 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 **??**) and the quadratic cost $c(x, y) = \frac{1}{2}\mathrm{dist}(x, y)^2$ on Riemannian manifolds (see Section **??**).

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 $\rho$ to $\mu$ but do not know exactly $\mu$ (due to some noise for instance) and have only access to some approximation $\nu$ of $\mu$, these inequalities tell us how close we may expect $T_\nu$ to be from $T_\mu$ (and $\phi_\nu$ from $\phi_\mu$), depending on some Wasserstein distance between $\mu$ and $\nu$.

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., $\rho, \mu, \nu$ are probability measures on $\mathbb{R}$, is particularly simple. Indeed, as soon as $\rho$ 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

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

is an optimal transport plan between $\mu$ and $\nu$. Indeed, (1.14) then follows immediately:

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

It is clear that $\gamma_{\mathrm{opt}}$ is an admissible transport plan between $\mu$ and $\nu$ 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 $\gamma$ is an optimal transport plan between two 1d probability measures, and $(x,y)$ and $(x',y')$ are in the support of $\gamma$ and $x < x'$, then necessarily $y \le y'$. This due to the convexity of the quadratic cost. Indeed, for any $x < x'$ and $y \le y'$, the inequality

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

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 $\rho$ to $\mu$, it is possible to give a completely explicit expression for $T_\mu$. Let us verify that

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

where

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

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

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

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

Now, since $T_\mu$ and $T_\nu$ are monotone, it is immediate to check that $\gamma_{\mathrm{opt}}$ is also monotone. But there is only one monotone admissible transport plan between $\mu$ and $\nu$, and thus $\gamma_{\mathrm{opt}}$ is optimal.

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

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

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

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 $\rho$ 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 $\rho$ 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 $\rho$ 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 $\mu, \nu$ 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 "$\rho$-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 $\mu$ and $\nu$ 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 $T_\mu, T_\nu$ 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 $\rho$ to $\mu$ and $\rho$ to $\nu$. Then $\widehat{\chi}$, which is the endpoint of the geodesic with this tangent vector, is an approximation of the midpoint between $\mu$ and $\nu$. It is also the midpoint of the so-called generalized Wasserstein geodesic (in the terminology of Ambrosio-Gigli-Savaré) between $\mu$ and $\nu$ 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 $\mu$ and $\nu$ coincides with the Wasserstein geodesic between $\mu$ and $\nu$.

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 $T_\mu$, before pushing forward $\rho$ 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 [63] and [64], the one by Santambrogio for "applied mathematicians" [58], the book by Peyré-Cuturi about computational aspects of optimal transport [56], and the very recent book by Chewi-Niles Weed-Rigollet about statistical optimal transport [20]. 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 [24]. Brenier presented his theorem in a short note [12] and gave details in an extended paper [13].

**§**1.2: The proof of Proposition 1.2 can be found in [64, Theorem 5.20]. Proposition 1.3 is a consequence of [13, Theorem 1.3] together with [64, Theorem 6.9], at least when $\mathcal{X}$ is smooth, bounded, and $\rho$ 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 [3]. 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 [58].

**§**1.4: Wasserstein geodesics are the main subject of the book by Ambrosio-Gigli-Savaré [2]. For a quick view on the subject, see [58, Chapter 5.4]. The interpretation of $W_2$ as a (pseudo) Riemannian manifold is due to Otto [55], who used it to study the long-time behavior of the porous medium equation. McCann introduced the concept of displacement interpolation in [50].

The paper [65] 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 [44] and to detect and visualize phenotypic differences between classes of cells [6].

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 [57], geometry processing [61], statistics [60] or machine learning [23]. The book chapters [56, Chapter 9.2], [20, 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 $\rho$ 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 $T_\mu$ from $\rho$ to $\mu$ 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 **??**, 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 $T_\mu$ is $K$-Lipschitz is very strong. First, it implies that the support of $\mu$ is connected. Second, to prove that $T_\mu$ is Lipschitz one has to invoke the regularity theory for optimal transport maps, which requires very strong assumptions on $\mu$. 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 $\rho$ 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 **??**), working under this kind of regularity assumptions on $T_\mu$. They have established stronger stability results (in terms of exponents) than what we present in these notes. For instance, under the assumption that $T_\mu$ 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 $T_\mu$ and $T_\mu^{-1}$. We shall explain a bit the proof techniques in Section **??**.

## 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 $\rho$ 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 $\mu$ and $\nu$. The discussion about the sharpness of these assumptions and the resulting stability inequalities is pretty long, and therefore we decided to devote Section **??** 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 $\rho$, while stability of optimal transport maps should hold under weaker (but still mysterious) assumptions. Notice that if the Poincaré inequality holds for $\rho$, 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 $\mu, \nu$ supported in $\mathcal{Y}$,*

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

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

$$\|T_\mu - T_\nu\|_{L^2(\rho)} \leq C W_1(\mu, \nu)^{1/9}. \tag{2.3}$$

Up to the logarithmic loss in (2.2), the inequality (2.2) is sharp, as discussed in Section **??**. 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 $\rho$ 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 $\rho$ be a probability density on a John domain $\mathcal{X} \subset \mathbb{R}^d$, and assume that $\rho$ 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 $\mu, \nu \in \mathcal{P}(\mathcal{Y})$,*

$$\|\phi_\mu - \phi_\nu\|_{L^2(\rho)} \leq C W_1(\mu, \nu)^{1/2}. \tag{2.4}$$

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

$$\|T_\mu - T_\nu\|_{L^2(\rho)} \leq C W_1(\mu, \nu)^{1/6}. \tag{2.5}$$

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 [26].

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 **??**.

### 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 $\rho$
- improving the stability exponents (1/9 in (2.3), 1/6 in (2.5)) for the source measures $\rho$ 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 $\rho$ in bounded domains. The assumption in Theorem 2.3 that $\rho$ 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 **??**.

(ii) Source measures $\rho$ 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 $\rho$ 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 **??**.

(iii) Source measures with disconnected support. If we replace the beginning of the statement of Theorem 2.3 by "Let $\rho$ 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 $\rho$ 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 [26], 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 **??**, 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 $\rho$ 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 $\rho$ as in Theorem 2.2, $\mathcal{X}$ has to be replaced by $\mathbb{R}^d$). The examples and counterexamples of Section **??** show an analogy, but not an equivalence, between the fact that the Poincaré inequality holds for $\rho$ 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 $\rho$-a.e. unique optimal transport map $T_{\mu_\theta}$ from $\rho$ to $\mu_\theta$ 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 $x_\theta$ and $x_{\theta+\pi}$. This cuts the disk into two (equal) halves, see Figure 1.

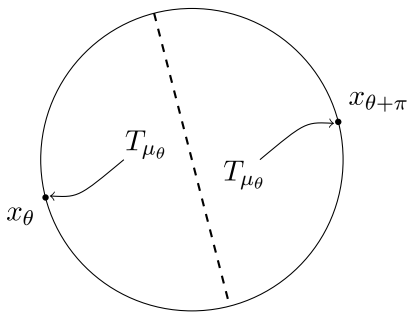


Figure 1: The optimal transport $T_{\mu_\theta}$ from $\rho$ to $\mu_\theta$.

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 $T_{\mu_0}$ and $T_{\mu_\theta}$ are at angular distance $\theta$, and $B$ is the set of points whose images under $T_{\mu_0}$ and $T_{\mu_\theta}$ 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 $\theta$ close enough to 0 and $p \geq 1$ arbitrary, the $W_p$ distance between $\mu_0$ and $\mu_\theta$ is obviously achieved by the map which sends $x_0$ to $x_\theta$ and $x_\pi$ 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 $\theta$ (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 **??**.

## 2.5 Bibliographical notes

§2.1: Theorem 2.1 is due to [31] and another proof has been given in [51]. The regularity theory of the Monge-Ampère equation and its link to regularity of optimal transport maps is explained in the survey [27]. The counterexample to the continuity of the optimal transport map is due to Caffarelli, see [16].

§2.2: Berman [7] was the first to obtain quantitative stability estimates without assuming regularity of the OT map. He derived dimension-dependent stability exponents for $\rho$ bounded above and below on a compact, convex domain, using complex geometry. Then, Delalande and Mérigot [26] improved his stability exponent, making it dimension-free, under the same assumptions on $\rho$. But more importantly, they introduced a robust proof technique based on the study of the Kantorovich functional, see Chapter **??**.

John domains were named in honor of F. John who introduced them in his work on elasticity [39]; Martio and Sarvas [49] 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 [51].

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

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

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

[6] 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.

[7] 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.

[8] 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.

[9] 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`

[10] 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.

[11] 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.

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

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

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

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

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

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

[18] 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.

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

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

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

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

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

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

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

[26] 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.

[27] 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.

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

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

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

[31] 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.

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

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

[34] 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`

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

[36] 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.

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

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

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

[40] 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.

[41] 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.

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

[43] Jun Kitagawa, Cyril Letrouit, and Quentin Mérigot. Stability of optimal transport maps on Riemannian manifolds.

[44] 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.

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

[46] 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.

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

[48] 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.

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

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

[51] 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.

[52] 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.

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

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

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

[56] 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.

[57] 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.

[58] 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.

[59] Daniel Spielman. Spectral and Algebraic Graph Theory. Ongoing draft available at

[60] 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.

[61] 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.

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

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

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

[65] 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.