Quantitative Stability in Optimal Transport: From Classical Potentials to Stable Transport Maps

Dario Trevisan (University of Pisa)

July 7, 2026

- "Random" mailing list $\longrightarrow$ Probability / Statistics math
- cvgmt.sns.it $\longrightarrow$ Calculus of variations

Lectures 01–02 recap

1. **Dual potentials as prices.** No-arbitrage condition:

$$\boxed{u(x) \leq v(y) + c(x,y)}$$

2. **Hard c-transforms.**

$$Q_c v(x) = \inf_y \{v(y) + c(x,y)\}, \qquad \widehat{Q}_c u(y) = \sup_x \{u(x) - c(x,y)\}.$$

3. **Softmin and entropy.**

$$Q_c^\varepsilon v = -\varepsilon \log \int e^{-(v+c)/\varepsilon} \, d\mu.$$

4. **Zero-temperature limit.**

$$Q_c^\varepsilon \longrightarrow Q_c, \qquad \text{Gibbs kernels} \longrightarrow \text{contact sets}.$$

5. **Rockafellar theorem.**

$$c\text{-cyclical monotonicity} \iff \text{existence of contact potentials}.$$

Lecture 03 roadmap

In this lecture we reinterpret quadratic optimal transport as nonlinear potential theory.

1. **Brenier theorem.**
2. **From transport maps to (Monge-Ampére, Poisson) PDEs.**
3. **PDE intermezzo.**
4. **Dynamic OT.**
5. **PDE upper and lower bonds.**

Differentiation on the contact set

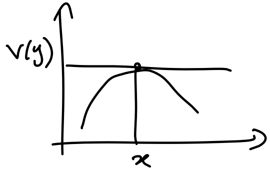
Specialize to $c(x, y) = |x - y|^p/p$, $p > 1$.

$X = Y = \mathbb{R}^d \quad d \geq 1 \quad p > 1$

The dual feasibility reads

$$\boxed{u(x) \leq v(y) + \frac{1}{p}|x - y|^p} \quad \text{for every } x, y.$$

On the contact set $(x, y) \in \Gamma$



Fix y,

$$u(x) = v(y) + \frac{1}{p}|x - y|^p.$$

$(x,y) \in \Gamma$

Assume that u is differentiable at x. Differentiating w.r.t x,

$$\nabla u(x) = |x - y|^{p-2}(x - y) = (x - y)^{p-1}.$$

$z \in \mathbb{R}^d \longmapsto z^{p-1} = |z|^{p-2} z$

Inverting we obtain, $\boxed{y = x - (\nabla u(x))^{1/(p-1)}}$. For $p = 2$ it is particularly simple.

Brenier's theorem

Let

$$\rho, \mu \in \mathcal{P}_2(\mathbb{R}^d),$$

$\int |x|^2 d\rho(x) + \int |y|^2 d\mu(y) < +\infty$

and assume that ρ is absolutely continuous with respect to Lebesgue measure.

Then there exists a convex function

$$\phi : \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$$

such that

- Rademacher thm
- $u(x)$ locally Lipschitz

$$T = \nabla \phi$$

is the unique optimal transport map from ρ to μ for the cost

$$c(x, y) = \frac{1}{2}|x - y|^2 \quad \text{or equivalently} \quad c(x, y) = -x \cdot y$$

The pushforward equation (Monge–Ampére)

$\boxed{T_\# \rho = \mu}$

The condition that the coupling is induced by a map T means

$$\int h(T(x))\rho(x)\,dx = \int h(y)\mu(y)\,dy$$

$y = T(x)$

for all test functions h.
If T is sufficiently smooth and invertible, the change of variable gives

$$\boxed{\rho(x) = \mu(T(x)) \det DT(x).}$$

For quadratic transport cost, Brenier

$$\boxed{T(x) = x - \nabla u(x),} \quad DT(x) = I - D^2 u(x).$$

Therefore

$$\boxed{\rho(x) = \mu(x - \nabla u(x)) \det(I - D^2 u(x)).}$$

$T(x) = \nabla \phi$

Standard Monge–Ampère

With Brenier's potential

$$\phi(x) = \frac{1}{2}|x|^2 - u(x), \quad \Rightarrow \boxed{T = \nabla\phi.}$$

The change-of-variables formula becomes

$$\rho(x) = \mu(\nabla\phi(x)) \det D^2\phi(x).$$

Equivalently,

$$\boxed{\det D^2\phi(x) = \frac{\rho(x)}{\mu(\nabla\phi(x))}.}$$

Brenier turns optimal transport into nonlinear PDE.

(Formal) linearization of Monge–Ampère

close to uniform

Assume ρ and μ are close, so that $T(x) = x - \nabla u(x) \approx x$. Then

$$\rho(x) = \mu(x - \nabla u(x)) \det(I - D^2 u(x))$$

for small $\nabla u, \nabla^2 u \approx 0$

$\det(I+A) =$
$= 1 + \mathrm{Tr}(A) + \ldots$

$$\det(I - D^2 u) \approx 1 - \Delta u \quad \text{and} \quad \mu(x - \nabla u(x)) = \mu(x) \approx 1$$

Formally,

$$\rho - \mu \approx -\mu \Delta u \approx \operatorname{div}(\rho \nabla u).$$

$\mu \approx \rho \approx 1$

$$\boxed{\text{Quadratic OT linearizes to a (weighted) Poisson equation.}}$$

Same argument for p-cost gives p′-Poisson equation.

$\Delta u = \operatorname{div} \nabla u$
$\Delta_{p'}(u) = \operatorname{div}\left((\nabla u)^{p'-1}\right)$

PDE intermezzo: flux formulation

$\boxed{\Delta u = \rho - \mu}$

Let $\Omega \subset \mathbb{R}^d$ be a domain with smooth boundary, and let

$$f (= \rho - \mu) \in L^1(\Omega), \qquad \int_\Omega f \, dx = 0.$$

For $p > 1$, consider the primal flux problem

$$\inf \left\{ \frac{1}{p} \int_\Omega |J|^p \, dx : -\operatorname{div} J = f, \quad J \cdot n = 0 \text{ on } \partial\Omega \right\}.$$

The constraint says that the flux J transports positive and negative parts of f.

$\boxed{\text{Claim: this is the linearized analogue of the primal transport problem.}}$

PDE intermezzo: dual formulation

Introduce a Lagrange multiplier φ for

$$-\operatorname{div} J = f\,.$$

Formally,

$$\begin{aligned}\mathcal{L}(J,\varphi) &= \frac{1}{p}\int_\Omega |J|^p\,dx + \int_\Omega \varphi(-\operatorname{div} J - f)\,dx \\ &= \frac{1}{p}\int_\Omega |J|^p\,dx + \int_\Omega \nabla\varphi\cdot J\,dx - \int_\Omega \varphi f\,dx\,.\end{aligned}$$

Minimizing in J and then changing $\varphi \mapsto -\varphi$ gives

$$\sup_\varphi \left\{ \int_\Omega \varphi f\,dx - \frac{1}{p'}\int_\Omega |\nabla\varphi|^{p'}\,dx \right\}.$$

PDE intermezzo: the p'-Poisson equation

The dual functional is

$$\sup_{\varphi} \left\{ \int_\Omega \varphi f \, dx - \frac{1}{p'} \int_\Omega |\nabla \varphi|^{p'} \, dx \right\}.$$

The Euler–Lagrange equation is

$$-\operatorname{div}\left(|\nabla \varphi|^{p'-2} \nabla \varphi \right) = f.$$

The optimal flux is

$$J = |\nabla \varphi|^{p'-2} \nabla \varphi,$$

up to the sign convention chosen in the constraint.

$$\boxed{\text{The } p\text{-flux problem is dual to a } p'\text{-Poisson equation.}}$$

Back to the Kantorovich semi-dual

Assume $X = Y$. The semi-dual reads

$$\mathcal{D}_\mu(v) = \int \mathrm{Q}_c v \, d\rho - \int v \, d\mu.$$

Add and subtract $\int v \, d\rho$:

$$\mathcal{D}_\mu(v) = \int (\mathrm{Q}_c v - v) \, d\rho - \int v \, d(\mu - \rho).$$
$$= \int v d(\rho - \mu) - \int (\mathcal{H}_c v) d\rho$$

$$\boxed{\mathcal{H}_c v = v - Q_c v}$$
$$\approx \frac{|\nabla\varphi|^{p'}}{p'}$$

This looks like a PDE dual functional:

$$\int_\Omega \varphi f \, dx - \frac{1}{p'} \int_\Omega |\nabla \varphi|^{p'} \, dx$$

forcing term $-$ energy term

From Q_t to a differential equation

$\boxed{p = p' = 2}$

Recall the time-dependent Hopf–Lax transform (quadratic cost for simplicity)

$$Q_t v(x) = v_t = = \boxed{\inf_y \left\{ v(y) + \frac{|x-y|^2}{2t} \right\}}.$$

We already saw the semigroup inequality (equality in Euclidean case)

$$Q_{s+t} v = Q_s (Q_t v.)$$

Can we describe the evolution $t \mapsto Q_t v$ by a PDE?

It should be the idempotent analogue of the heat equation $\boxed{\partial_t v_t = \frac{1}{2} \Delta v_t.}$

The infinitesimal generator

Hopf-Lax formula

$$v_t(x) := Q_t v(x) = \inf_{y} \left\{ v(y) + \frac{|x-y|^2}{2t} \right\}. \qquad y_t(x) \in \operatorname{argmin}_y \{ \quad \}$$

By the envelope theorem,

$$\boxed{\nabla v_t(x) = \frac{x - y_t(x)}{t}}, \quad \text{and} \quad \boxed{\partial_t v_t(x) = -\frac{|x - y_t(x)|^2}{2t^2}}.$$

Therefore

$$\partial_t v_t(x) = -\frac{1}{2}\left|\frac{x - y_t(x)}{t}\right|^2 = -\frac{1}{2}|\nabla v_t(x)|^2.$$

Hopf–Lax formula is (the) solution operator for Hamilton–Jacobi.

Economic meaning of the Hamilton–Jacobi semigroup

Recall

$$\mathrm{Q}_t v(x) = \inf_y \left\{ v(y) + \frac{|x-y|^2}{2t} \right\}.$$

Interpretation:

$$Q_t(x) = \text{cheapest value of replacing good } x \text{ within time } t.$$

As t increases, replacement becomes easier, so values relax downward.

$$\partial_t v_t + \frac{1}{2}|\nabla v_t|^2 = 0.$$

$$\boxed{\text{Values decrease more where local gain opportunities are large.}}$$

Continuous-time dual viewpoint

The static semi-dual $c(x,y) = \frac{|x-y|^2}{2}$

$$\frac{1}{2}W_2^2(\rho,\mu) = sup_v \int Q_1 v \, d\rho - \int v \, d\mu = \sup_v \int v d(\rho-\mu) + \int (Q_1 v - v) d\rho$$

can be replaced by a continuous-time dual:

$$\frac{1}{2}W_2^2(\rho,\mu) = \sup_{v_t} \left\{ \int v_1 \, d\rho - \int v_0 \, d\mu \right\} = \boxed{\sup_{v_t} \left\{ \int v_0 d(\rho-\mu) - \frac{1}{2}\int\int_0^1 |\nabla v_t|^2 d\rho \right\}}$$

subject to

$$\partial_t v_t + \frac{1}{2}|\nabla v_t|^2 \leq 0.$$

The HJ inequality is the infinitesimal no arbitrage constraint.

General power costs

For

$$c(x,y) = \frac{1}{p}|x-y|^p, \qquad p' = \frac{p}{p-1},$$

the time-dependent transform is

$$\mathrm{Q}_t v(x) = \inf_y \left\{ v(y) + \frac{|x-y|^p}{p\, t^{p-1}} \right\}.$$

The infinitesimal generator becomes $\frac{1}{p'}|\nabla v_t|^{p'}$, hence the continuous-time dual constraint is

$$\partial_t v_t + \frac{1}{p'}|\nabla v_t|^{p'} \leq 0.$$

The p-transport cost leads to the p'-Hamiltonian.

Theorem: Benamou–Brenier formula

Let $p > 1$, $p' = p/(p-1)$, and let $\rho, \mu \in \mathcal{P}_p(\mathbb{R}^d)$. Then

$$\frac{1}{p} W_p^p(\rho, \mu) = \inf_{\rho_t, J_t} \int_0^1 \int_{\mathbb{R}^d} \frac{1}{p} \frac{|J_t(x)|^p}{\rho_t(x)^{p-1}} \, dx \, dt,$$

where the infimum is taken over narrowly continuous curves

$$t \mapsto \rho_t \in \mathcal{P}_p(\mathbb{R}^d)$$

and fluxes J_t satisfying the continuity equation

$$\partial_t \rho_t + \operatorname{div} J_t = 0$$

in the sense of distributions, with boundary conditions $\rho_0 = \mu$, $\rho_1 = \rho$
Equivalently, writing $J_t = \rho_t b_t$,

$$\frac{1}{p} W_p^p(\rho_0, \rho_1) = \inf_{\rho_t, b_t} \int_0^1 \int_{\mathbb{R}^d} \frac{1}{p} |b_t(x)|^p \, d\rho_t(x) \, dt.$$

References: Benamou–Brenier, *A computational fluid mechanics solution to the Monge–Kantorovich mass transfer problem*, Numer. Math. 2000; Ambrosio–Gigli–Savaré, *Gradient Flows*, Ch. 8; Villani, *Topics in Optimal Transportation*, Ch. 7; Santambrogio, *Optimal Transport for Applied Mathematicians*, Ch. 5.

Rigorous PDE estimates: linear upper bounds

Let $p > 1$, and let $\rho, \mu,$ be probability (densities) on a bounded domain $\Omega \subset \mathbb{R}^d$, with $\rho \geq m > 0$. Let ϕ solve the Neumann Poisson problem

$$-\Delta\phi = \rho - \mu \quad \text{in } \Omega, \qquad \nabla\phi \cdot n = 0 \quad \text{on } \partial\Omega.$$

Then one has upper bounds of the form

$$W_p^p(\rho_{\cancel{0}}\, dx, \cancel{\rho_1}\,\mu\, \cancel{dx}) \leq C_{\Omega,p}\, m^{1-p} \boxed{\int_\Omega |\nabla\phi|^p\, dx}.$$

For $p = 2$, this gives the familiar estimate

$$W_2^2(\rho_0\, dx, \rho_1\, dx) \lesssim m^{-1} \|\rho_1 - \rho_0\|^2_{\dot{H}^{-1}}.$$

References: Peyré, *Comparison between W_2 distance and H^{-1} norm*; Ledoux, *On optimal matching of Gaussian samples*; Ambrosio–Stra–Trevisan, *A PDE approach to a 2-dimensional matching problem*; see also Ambrosio–Vitillaro–Trevisan, Prop. 2.1.

Rigorous PDE estimates: nonlinear two-sided bounds

Let $p > 1$, $p' = p/(p-1)$, and let ρ_0, ρ_1 be probability densities on the flat torus $\mathbb{T}^d$. Let ϕ solve the p'-Poisson equation

$$-\operatorname{div}\left(|\nabla\phi|^{p'-2}\nabla\phi\right) = \rho_1 - \rho_0, \qquad \int_{\mathbb{T}^d} \phi\, dm = 0.$$

periodic b.c.

Assume that both densities are close to 1:

$$c := 2 \max_{i=0,1} \|\rho_i - 1\|_{L^\infty(\mathbb{T}^d)}$$

is sufficiently small. Then there exist errors

$$\delta(c,p),\ \bar{\delta}(c,p) \to 0 \qquad \text{as } c \to 0,$$

$\mathbb{T}^d \to [0,1]^d$

such that

$$(1-\delta(c,p)) \int_{\mathbb{T}^d} |\nabla\phi|^{p'}\, dm \le W_p^p(\rho_0 m, \rho_1 m) \le (1+\bar{\delta}(c,p)) \int_{\mathbb{T}^d} |\nabla\phi|^{p'}\, dm.$$

Reference: Ambrosio–Vitillaro–Trevisan, *Sharp PDE estimates for random two-dimensional bipartite matching with power cost function*, Thm. 1.2. For domains with boundary: Ambrosio–Tassoni–Trevisan, in preparation.

Proof of the upper bound: quadratic case

$p = 2$

$$\frac{1}{2} W_2^2(\rho, \mu) = \sup_{\partial_t v_t + \frac{1}{2}|\nabla v_t|^2 \le 0} \left\{ \int v_1 \, d\mu - \int v_0 \, d\rho \right\}.$$

We use the Dacorogna-Moser interpolation: set $\rho_t = (1-t)\rho + t\mu$. Let ϕ solve

$$-\Delta\phi = \rho - \mu, \qquad \nabla\phi \cdot n = 0.$$

Then, with flux $J := \nabla\phi$,

$$\partial_t \rho_t = \mu - \rho, \qquad \operatorname{div} J = \Delta\phi = \mu - \rho, \quad \text{hence} \quad \partial_t\rho_t - \operatorname{div} J = 0.$$

continuity equation

For any Hamilton–Jacobi subsolution v_t,

$$\int v_1 \, d\mu - \int v_0 \, d\rho = \int_0^1 \frac{d}{dt}\left(\int v_t \rho_t \, dx \right) dt$$

$$= \int_0^1 \int \partial_t v_t \, \rho_t + v_t \, \partial_t \rho_t \, dx \, dt$$

$$\le \int_0^1 \int \overset{-\frac{1}{2}|\nabla v_t|^2}{\cancel{\partial_t v_t}} \, \rho_t - \nabla v_t \cdot \frac{J}{\sqrt{\rho_t}} \, dx \, dt . \sqrt{\rho_t} \qquad + - \int \frac{|J|^2}{\rho_t}$$

Conclusion: the H^{-1} upper bound

Using

$$\partial_t v_t + \frac{1}{2}|\nabla v_t|^2 \leq 0,$$

we get

$$\int v_1 \, d\mu - \int v_0 \, d\rho \leq \int_0^1 \int \left[-\frac{1}{2}\rho_t |\nabla v_t|^2 - \nabla v_t \cdot J \right] dx \, dt$$

$$\leq \int_0^1 \int \frac{|J|^2}{2\rho_t} \, dx \, dt.$$

Assume (for simplicity) $\rho \geq m$ and also $\mu \geq m$. Then $\rho_t \geq m$, and therefore

$$\int v_1 \, d\mu - \int v_0 \, d\rho \leq \frac{1}{2m} \int |\nabla \phi|^2 \, dx.$$

Taking the supremum over v_t,

$$\frac{1}{2} W_2^2(\rho, \mu) \leq \frac{1}{2m} \int |\nabla \phi|^2 \, dx.$$

Bridge to Lecture 04: stability

$\boxed{\text{OT potentials are nonlocal analogues of nonlinear elliptic potentials.}}$

Quadratic cost (Brenier)

$$T(x) = x - \nabla u(x) \quad \Rightarrow \quad -\Delta\phi \approx \rho - \mu.$$

Power costs:

$$T(x) = x - (\nabla u(x))^{1/(p-1)} \quad \Rightarrow \quad -\operatorname{div}\left(|\nabla\varphi|^{p'-2}\nabla\varphi\right) = \rho - \mu.$$

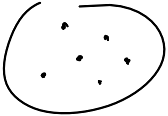
Lecture 04 question:

$\boxed{\text{Can PDE stability estimates predict stability of OT maps?}}$

References for Lecture 03

- Y. Brenier, *Polar factorization and monotone rearrangement of vector-valued functions*, Comm. Pure Appl. Math., 1991.
- L. Ambrosio, N. Gigli, G. Savaré, *Gradient Flows in Metric Spaces and in the Space of Probability Measures*, Ch. 8.
- C. Villani, *Topics in Optimal Transportation*, Ch. 7.
- F. Santambrogio, *Optimal Transport for Applied Mathematicians*, Ch. 5.
- L. C. Evans, *Partial Differential Equations*, sections on variational methods and Hamilton–Jacobi equations.
- L. Ambrosio, F. Vitillaro, D. Trevisan, *Sharp PDE estimates for random two-dimensional bipartite matching with power cost function*, 2024.

$\rho \geq m$ $\qquad \mu$



$$W_2^2(\rho,\mu) \le \left(W_2\left(\rho,\frac{\mu+\rho}{2}\right) + W_2\left(\mu,\frac{\mu+\rho}{2}\right)\right)^2$$

$$(a+b)^2 \le a^2+b^2+2ab$$

$$\le (1+\varepsilon)a^2 + C(\varepsilon)b^2$$

$$\le \left(W_2\left(\rho,\frac{\mu+\rho}{2}\right) + W_2\left(\frac{\mu}{2},\frac{\rho}{2}\right)\right)^2$$

$$\le (1+\varepsilon)\, W_2^2\left(\frac{\mu}{2},\frac{\rho}{2}\right) + C(\varepsilon)\, W_2^2\left(\rho,\frac{\mu+\rho}{2}\right)$$

$$\le \frac{1+\varepsilon}{2}\, W_2^2(\mu,\rho) + C(\varepsilon)\, W_2^2\left(\rho,\frac{\mu+\rho}{2}\right)$$

$$W_2^2(\rho,\mu) \le \tilde{C}(\varepsilon)\, W_2^2\left(\rho,\frac{\mu+\rho}{2}\right)$$