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

Dario Trevisan (University of Pisa)

July 6, 2026

Entry points into optimal transport

There are already excellent introductions to the standard theory.

- C. Villani, *Topics in Optimal Transportation*; *Optimal Transport: Old and New*.
- F. Santambrogio, *Optimal Transport for Applied Mathematicians*.
- L. Ambrosio, N. Gigli, G. Savaré, *Gradient Flows*.
- A. Figalli, F. Glaudo, *An Invitation to Optimal Transport, Wasserstein Distances, and Gradient Flows*.
- L. Ambrosio, E. Brué, D. Semola, *Lectures on Optimal Transport*.
- G. Peyré, M. Cuturi, *Computational Optimal Transport*.

$\boxed{\text{These lectures do not try to replace these references.}}$

Course philosophy

The standard route often begins with the (Monge-)Kantorovich primal problem:

$$\text{transport plans} \quad \longrightarrow \quad \text{duality} \quad \longrightarrow \quad \text{Kantorovich potentials.}$$

We will try to reverse it.

$$\boxed{\text{What is the potential behind optimal transport?}}$$

Lectures 01–02 roadmap

In the first two lectures we build the dual viewpoint step by step.

1. **Dual** potentials as prices with a no-arbitrage condition.
2. Hard c-**transforms.**
3. Softmin **entropic regularization** and zero-temperature limit
4. **Wasserstein distance** of order p.
5. Cyclical monotonicity and **Rockafellar theorem.**

Economic starting point: replacement costs

Suppose that a unit of mass at $x \in X$ can be replaced by a unit of mass at $y \in Y$.

$$v(y) = \text{price of using target point } y,$$

$$c(x, y) = \text{cost of replacing } x \text{ by } y.$$

Then the total replacement cost of x through y is

$$v(y) + c(x, y).$$

Admissibile valuation

No arbitrage

> A source value $u(x)$ is admissible only if it is no larger than every possible replacement cost.

The admissibility condition is therefore

$$u(x) \leq v(y) + c(x, y) \qquad \forall x \in X, \ y \in Y.$$

Equivalently,

$$u(x) - v(y) \leq c(x, y).$$

If this failed for some pair (x, y), then

$$u(x) > v(y) + c(x, y),$$

so x would be valued more than the cost of replacing it by y.

Arbitrage-free valuations

$$u(x) \leq v(y) + c(x, y) \qquad \forall x, y.$$

$$\begin{aligned} u(x) &= \text{value assigned to source point } x, \\ v(y) &= \text{price assigned to target point } y, \\ v(y) + c(x, y) &= \text{cost of replacing } x \text{ by } y. \end{aligned}$$

Kantorovich potentials are arbitrage-free valuations.

Dual gain

Let $\rho \in \mathcal{P}(X)$, $\mu \in \mathcal{P}(Y)$,

$$\rho = \text{distribution of items to be sold / valued},$$

$$\mu = \text{distribution of items to be bought / used as replacements}.$$

For admissible prices

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

the dual gain is

$$\boxed{\int_X u\, d\rho - \int_Y v\, d\mu.}$$

The dual problem

The Kantorovich (dual) cost is

$$W_c(\rho, \mu) := \sup_{u(x) \leq v(y) + c(x,y)} \left\{ \int_X u \, d\rho - \int_Y v \, d\mu \right\}.$$

$\boxed{\text{Maximize value discrepancy (gain) under the no-arbitrage constraint.}}$

Sign convention

Standard notation:

$$\boxed{\varphi(x) + \psi(y) \le c(x,y).}$$

Our notation:

$$u(x) - v(y) \le c(x,y), \qquad \text{equivalently} \qquad u(x) \le v(y) + c(x,y).$$

$$u = \varphi, \qquad v = -\psi.$$

This convention is adapted to the c-transform operator

$$\boxed{Qv(x) = \inf_y \{v(y) + c(x,y)\},}$$

the cheapest replacement value induced by v.

The primal problem   Monge - Kantorovich

Couplings:

$$\Pi(\rho, \mu) = \{\pi \in \mathcal{P}(X \times Y) : (p_X)_\# \pi = \rho,\ (p_Y)_\# \pi = \mu\}.$$

Primal problem:

$$W_c(\rho, \mu) \overset{?}{=} \inf_{\pi \in \Pi(\rho,\mu)} \int_{X \times Y} c(x, y)\, d\pi(x, y).$$

$$\text{primal} = \text{feasible matchings},$$
$$\text{dual} = \text{feasible prices}.$$

Weak duality

$$W_c(\varrho,\mu) = \sup_{u \oplus v \le c} \int_X u \, d\varrho - \int_Y v \, d\mu$$

$\underline{\forall\ \pi \in \Pi(\varrho,\mu)}$

$$\int_X u \, d\varrho = \int_{X \times Y} u(x) \, d\pi(x,y) \qquad \int_Y v \, d\mu = \int_{X \times Y} v(y) \, d\pi(x,y)$$

$$= \sup_{u \oplus v \le c} \int_{X \times Y} \underbrace{(u(x) - v(y))}_{\le c(x,y)} \, d\pi(x,y)$$

$$\le \cancel{\sup_{u \oplus v \le c}} \int_{X \times Y} c(x,y) \, d\pi(x,y) \leftarrow \inf_{\pi}$$

Theorem: Kantorovich duality

Let X, Y be Polish spaces, let

$$\rho \in \mathcal{P}(X), \qquad \mu \in \mathcal{P}(Y),$$

and let

$$c : X \times Y \to (-\infty, +\infty]$$

be lower semicontinuous and bounded from below. Then,

$$\inf_{\pi \in \Pi(\rho,\mu)} \int_{X \times Y} c(x,y)\, d\pi(x,y) = \sup_{u(x) \leq v(y) + c(x,y)} \left\{ \int_X u\, d\rho - \int_Y v\, d\mu \right\},$$

where the supremum is taken over integrable admissible potentials.

Moreover, if the primal value is finite, an optimal plan exists.

References: Villani, *Topics in Optimal Transportation*, Thm. 1.3;
Ambrosio–Gigli–Savaré, *Gradient Flows*, Ch. 6;
Santambrogio, *Optimal Transport for Applied Mathematicians*, Ch. 1.

Formal derivation: introduce multipliers

$\pi(A \times Y) = \rho(A) \quad \forall A$

$$\boxed{\int_{X\times Y} u(x)\, d\pi(x,y) = \int_X u(x)\, d\rho(x)} \quad \forall u$$

Primal:

$$\inf_{\pi \geq 0} \left\{ \int c \, d\pi : (\mathrm{pr}_X)_\# \pi = \rho, \ (\mathrm{pr}_Y)_\# \pi = \mu \right\}.$$

Lagrange multipliers for the two marginal constraints:

$$u : X \to \mathbb{R}, \qquad v : Y \to \mathbb{R}.$$

$$\mathcal{L}(\pi, u, v) = \int c \, d\pi + \underbrace{\int u \, d\rho - \int u(x) \, d\pi}_{} - \underbrace{\int v \, d\mu + \int v(y) \, d\pi}_{}.$$

Formal derivation

$$\mathcal{L}(\pi;(u,v)) = \int c\,d\pi - \int u(x)\,d\pi + \int u(x)\,\rho(x) + \int v(y)\,d\pi - \int v(y)\,d\mu(y)$$

$$\sup_{(u,v)} \mathcal{L}(\pi;(u,v)) = \int c\,d\pi + \sup_u \Big(\underbrace{- \int u(x)\,d\pi + \int u\,d\rho}_{+\infty} \Big) + \sup_v \Big(\underbrace{\qquad\qquad}_{+\infty} \Big)$$

unless $(P_X)_\# \pi = \rho$ unless $(P_Y)_\# \pi = \mu$

$$\inf_\pi \sup_{(u,v)} \mathcal{L}(\pi;(u,v)) = \inf_{\pi \in \Pi(\rho,\mu)} \int c\,d\pi$$

$$\sup_{(u,v)} \inf_\pi \mathcal{L}(\pi;(u,v)) = \inf_{\pi \geq 0} \int_{X \times Y} \underbrace{(c - u(x) + v(y))}_{-\infty \text{ unless } \boxed{u(x) \leq v(y) + c}}\,d\pi + \left(\int u\,d\rho - \int v\,d\mu \right)$$

$\boxed{u(x) \leq v(y) + c} \to 0$

Formal duality

The min–max principle (formally) gives

$$
\begin{aligned}
\inf_{\pi\in\Pi(\rho,\mu)} \int c\, d\pi &= \inf_{\pi\geq 0} \sup_{u,v} \mathcal{L}(\pi, u, v) \\
&= \sup_{u,v} \inf_{\pi\geq 0} \mathcal{L}(\pi, u, v) \\
&= \sup_{u(x)\leq v(y)+c(x,y)} \left\{ \int u\, d\rho - \int v\, d\mu \right\}.
\end{aligned}
$$

$=$

The (hard) c-transform

Given v, admissibility says

$$u(x) \leq v(y) + c(x,y) \qquad \forall y \in Y.$$

Therefore

$$u(x) \leq \inf_{y \in Y} \{v(y) + c(x,y)\}.$$

Define the c-transform of v

$$\boxed{\mathrm{Q}_c v(x) := \inf_{y \in Y} \{v(y) + c(x,y)\}.}$$

write $\mathrm{Q} = \mathrm{Q}_c$ if c is clear. Standard notation ($v = -\psi$)

$$\mathrm{Q}_c v \longleftrightarrow \psi^c.$$

Semi-dual formulation

The two-potential problem can be reduced one potential replacing u with $\mathrm{Q}v$.

Kantorovich functional:

$$\mathcal{D}_\mu(v) := \int_X \mathrm{Q}v \, d\rho - \int_Y v \, d\mu.$$

We make the problem less symmetric but reduce to a single potential v.

Eliminating either potential

Given v, the best admissible u is

$$u = \mathrm{Q}v, \qquad \mathrm{Q}v(x) := \inf_{y \in Y} \{v(y) + c(x,y)\}.$$

Given u, admissibility

$$u(x) \le v(y) + c(x,y)$$

is equivalent to

$$v(y) \ge u(x) - c(x,y) \qquad \forall x.$$

So the smallest admissible v is

$$\boxed{\widehat{\mathrm{Q}}u(y) := \sup_{x \in X} \{u(x) - c(x,y)\}.}$$

$$\boxed{\text{We can eliminate either variable: } u = \mathrm{Q}v \text{ or } v = \widehat{\mathrm{Q}}u.}$$

The two transforms encode the same constraint

$$\mathrm{Q}v(x) := \inf_{y \in Y}\{v(y) + c(x,y)\}, \quad \widehat{\mathrm{Q}}u(y) := \sup_{x \in X}\{u(x) - c(x,y)\}.$$

The following are equivalent:

$$u(x) \le v(y) + c(x,y) \qquad \forall x, y,$$

$$u \le \mathrm{Q}v,$$

$$\widehat{\mathrm{Q}}u \le v.$$

General properties of Q and $\widehat{\mathrm{Q}}$

$$v \le w \quad \Longrightarrow \quad \mathrm{Q}v \le \mathrm{Q}w, \qquad u \le z \quad \Longrightarrow \quad \widehat{\mathrm{Q}}u \le \widehat{\mathrm{Q}}z.$$

$$\mathrm{Q}(v + a) = \mathrm{Q}v + a, \qquad \widehat{\mathrm{Q}}(u + a) = \widehat{\mathrm{Q}}u + a. \qquad a \in \mathbb{R}$$

$$\|\mathrm{Q}v - \mathrm{Q}w\|_\infty \le \|v - w\|_\infty, \qquad \|\widehat{\mathrm{Q}}u - \widehat{\mathrm{Q}}z\|_\infty \le \|u - z\|_\infty.$$

$$u \le \mathrm{Q}\widehat{\mathrm{Q}}u, \qquad \widehat{\mathrm{Q}}\mathrm{Q}v \le v.$$

$$\mathrm{Q}\widehat{\mathrm{Q}}\mathrm{Q} = \mathrm{Q}, \qquad \widehat{\mathrm{Q}}\mathrm{Q}\widehat{\mathrm{Q}} = \widehat{\mathrm{Q}}.$$

Soft c-transform / Entropic c-transform

The hard transform is

$$Q_c v(x) = \inf_{y \in Y} \{v(y) + c(x,y)\}.$$

Its soft version is for $\varepsilon > 0$ (temperature):

$$\boxed{Q_c^\varepsilon v(x) := -\varepsilon \log \int_Y \exp\left(-\frac{v(y) + c(x,y)}{\varepsilon}\right) \, d\mu(y).}$$

Interpretation: soft replacement value of x. Notice that μ explicitly appears!

Gibbs variational principle

For a function $F : Y \to \mathbb{R}$,

$$-\varepsilon \log \int_Y e^{-F/\varepsilon}\, d\mu = \inf_{\gamma \in \mathcal{P}(Y)} \left\{ \int_Y F\, d\gamma + \varepsilon \mathrm{Ent}(\gamma \mid \mu) \right\}.$$

Apply this with

$$F_x(y) = v(y) + c(x, y).$$

Then

$$\mathrm{Q}_c^\varepsilon v(x) = \inf_{\gamma \in \mathcal{P}(Y)} \left\{ \int_Y \Big(v(y) + c(x, y) \Big)\, d\gamma(y) + \varepsilon \mathrm{Ent}(\gamma \mid \mu) \right\}.$$

As $\varepsilon \downarrow 0$,

$$\boxed{\mathrm{Q}_c^\varepsilon v(x) \longrightarrow \mathrm{Q}_c v(x).}$$

Gibbs kernel

The minimizer in the Gibbs variational principle is Gibbs distribution $\frac{e^{-F/\varepsilon}}{Z}$

$$d\gamma_x^\varepsilon(y) = \frac{\exp\left(-(v(y)+c(x,y))/\varepsilon\right)}{\int_Y \exp\left(-(v(z)+c(x,z))/\varepsilon\right)\, d\mu(z)}\, d\mu(y).$$

Equivalently,

$$d\gamma_x^\varepsilon(y) = e^{\overbrace{(Q_c^\varepsilon v(x)-v(y)-c(x,y))/\varepsilon}}\, d\mu(y).$$

The soft adjoint transform

Hard adjoint transform:

$$\widehat{\mathrm{Q}}_c u(y) = \sup_{x \in X} \{u(x) - c(x,y)\}.$$

Its soft version is

$$\boxed{\widehat{\mathrm{Q}}_c^{\varepsilon} u(y) := \varepsilon \log \int_X \exp\left(\frac{u(x) - c(x,y)}{\varepsilon}\right) \, d\rho(x).}$$

For the soft transforms the exact identity is lost:

$$\mathrm{Q}_c^{\varepsilon} \widehat{\mathrm{Q}}_c^{\varepsilon} \mathrm{Q}_c^{\varepsilon} \neq \mathrm{Q}_c^{\varepsilon} \qquad \text{in general.}$$

Instead of one exact projection, one iterates the two soft transforms: this is the Sinkhorn mechanism.

Maslov dequantization

The soft minimum is based on Laplace principle

$$-\varepsilon \log \int e^{-F/\varepsilon} \longrightarrow \inf F \qquad \text{as } \varepsilon \downarrow 0.$$

This is often called the *Maslov dequantization*.

The usual operations are transformed as

$$+ \quad \rightsquigarrow \quad \min, \qquad \times \quad \rightsquigarrow \quad +.$$

$\boxed{\text{Classical probability becomes idempotent, or tropical, probability.}}$

(From the soft-max one can obtain a max, $+$ idempotent probability)

From heat to Hopf–Lax

Let P_t^ε be the heat semigroup at diffusivity ε:

$$P_t^\varepsilon f(x) = \int_{\mathbb{R}^d} f(y) \frac{1}{(2\pi\varepsilon t)^{d/2}} e^{-|x-y|^2/(2\varepsilon t)}\, dy.$$

Apply it to

$$f = e^{-v/\varepsilon}.$$

Then

$$-\varepsilon \log P_t^\varepsilon(e^{-v/\varepsilon})(x) = -\varepsilon \log \int \exp\left(-\frac{v(y) + |x-y|^2/(2t)}{\varepsilon}\right) \frac{dy}{(2\pi\varepsilon t)^{d/2}}.$$

$$\boxed{\text{This is a soft version of the Hopf–Lax operator.}}$$

By the Laplace-(Varadhan) principle,

$$-\varepsilon \log P_t^\varepsilon(e^{-v/\varepsilon})(x) \longrightarrow \inf_y \left\{ v(y) + \frac{|x-y|^2}{2t} \right\} = Q_t v(x)$$

with $c(x,y) = |x-y|^2/2t$.

$$\boxed{\text{Hopf–Lax is the tropical shadow of the heat semigroup.}}$$

Why this is useful? we can guess many properties, e.g.

$$P_{t+s} = P_t P_s \quad \Rightarrow \quad Q_{t+s} = Q_t Q_s.$$

Wasserstein distance of order p

Let $(X = Y, d)$ be a (bounded) metric space and $p \geq 1$. For the cost

$$c(x, y) = d(x, y)^p,$$

the dual formulation gives

$$W_p^p(\rho, \mu) := \sup_{u(x) \leq v(y) + d(x,y)^p} \left\{ \int_X u \, d\rho - \int_X v \, d\mu \right\}.$$

By Kantorovich duality, this is equal to the usual primal value

$$\inf_{\pi \in \Pi(\rho,\mu)} \int_{X \times X} d(x, y)^p \, d\pi(x, y).$$

$W_p(\rho, \mu)$

In this lecture, we use the dual formula as the definition and the primal formula as its representation.

A dual proof of the triangle inequality : $W_p(\rho,\mu) \le W_p(\rho,\eta) + W_p(\eta,\mu)$

Fix $p > 1$. For $t > 0$, define the rescaled cost

$$c_t(x,y) := \frac{d(x,y)^p}{t^{p-1}}. \quad \boxed{p=2}$$

Correspondingly, define

$$Q_t v(x) := \inf_y \left\{ v(y) + \frac{d(x,y)^p}{t^{p-1}} \right\}.$$

$$\boxed{p-1 = \frac{p}{p'} \qquad p' = \frac{p}{p-1}}$$

Then

$$\boxed{W_{c_t}(\rho,\mu) = t^{1-p} W_p^p(\rho,\mu).}$$

The parameter t plays the role of time in the Hopf–Lax formula.

The key semigroup (inequality) $Q_{(s+t)}V \le Q_t(Q_s V)$

$$Q_{s+t}V(x) = \inf_y \left\{ V(y) + \frac{d(x,y)^p}{(s+t)^{p-1}} \right\}$$

$$\frac{d(x,y)}{(t+s)^{1/p'}} \le \frac{d(x,z)}{t^{1/p'}} t^{\frac{1}{p'}} + \frac{d(z,y)}{s^{1/p'}} s^{1/p'} \le \left(\frac{d(x,z)^p}{t^{p/p'}} + \frac{d(z,y)^p}{s^{p/p'}} \right)^{1/p} \frac{(t+s)^{1/p'}}{(t+s)^{1/p'}}$$

$$\boxed{\frac{d(x,y)^p}{(t+s)^{p-1}} \le \frac{d(x,z)^p}{t^{p-1}} + \frac{d(z,y)^p}{s^{p-1}}} \quad \forall s,t>0$$

<u>Lemma</u> $\frac{A^p}{(t+s)^{p-1}} \le \frac{B^p}{t^{p-1}} + \frac{C^p}{s^{p-1}} \quad \forall s,t>0 \implies A \le B+C$

$$
\begin{aligned}
Q_{s+t} V(x) &= \inf_{y} \left\{ V(y) + \frac{d(x,y)^p}{(s+t)^{p-1}} \right\} \\
&\le \inf_{z} \inf_{y} \left\{ V(y) + \boxed{\frac{d(x,z)^p}{s^{p-1}}} + \frac{d(z,y)^p}{t^{p-1}} \right\} \\
&= \inf_{z} \Bigg\{ \underbrace{\inf_{y} \left\{ V(y) + \frac{d(z,y)^p}{t^{p-1}} \right\}}_{} + \frac{d(x,z)^p}{s^{p-1}} \Bigg\} \\
&= \inf_{z} \left\{ Q_t V(z) + \frac{d(x,z)^p}{s^{p-1}} \right\} = Q_s \big(Q_t V \big)(x)
\end{aligned}
$$

Dual triangle inequality

$$W_{\frac{d^p}{(s+t)^{p-1}}}(\rho,\mu) = \boxed{\frac{1}{(s+t)^{p-1}} \underbrace{W_p^p(\rho,\mu)}_{A^p}}$$

$$= \sup_v \left\{ \int Q_{(s+t)} v \, d\rho - \int v \, d\mu \right\} \le \sup_v \left\{ \int Q_t(Q_s v) \, d\rho - \int v \, d\mu \right\}$$

$$- + \int Q_s v \, d\eta$$

$$= \sup_v \left[\int Q_t(Q_s v) \, d\rho - \int Q_s v \, d\eta \right] + \left[\int Q_s v \, d\eta - \int v \, d\mu \right]$$

$$\forall s,t \geq 0$$

$$\le W_{\frac{d^p}{t^{p-1}}}(\rho,\eta) + \sup_v \left[\int Q_s v \, d\eta - \int v \, d\mu \right]$$

$$= W_{\frac{d^p}{t^{p-1}}}(\rho,\eta) + W_{\frac{d^p}{s^{p-1}}}(\eta,\mu) = \boxed{\frac{1}{t^{p-1}} \overbrace{W_p^p(\rho,\eta)}^{B^p} + \frac{1}{s^{p-1}} \overbrace{W_p^p(\eta,\mu)}^{C^p}}$$

Taking the supremum

Taking the supremum over v, we obtain

$$(s+t)^{1-p} W_p^p(\rho,\mu) \leq t^{1-p} W_p^p(\rho,\eta) + s^{1-p} W_p^p(\eta,\mu).$$

Optimize in s, t to conclude.

Remark $p=1$ $\quad Q_d v(x) = \inf_y \{ v(y) + d(x,y) \}$

$Q_d v = v$ iff v is Lip $\quad \mathrm{Lip}(v) = 1 \quad |v(x) - v(y)| \leq d(x,y)$

$$W_1(\rho,\mu) = \sup_{\mathrm{Lip}(v) \leq 1} \left\{ \int v \, d\rho - \int v \, d\mu \right\}$$

Exercises: quadratic cost and Gaussian examples

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

1. **Translations.** Let $\mu = (\tau_m)_{\#}\rho$, where

$$\tau_m(x) = x + m.$$

   Show directly from the dual definition that τ_m is optimal.

2. **Dilations.** Let $T_a(x) = ax$, $a > 0$. Show that T_a is optimal from

$$\rho = \mathcal{N}(0, \sigma_0^2) \quad \text{to} \quad \mu = \mathcal{N}(0, \sigma_1^2), \qquad a = \frac{\sigma_1}{\sigma_0}.$$

   Is T_a optimal for arbitrary ρ?

3. **Gaussian to Gaussian.** If

$$\rho = \mathcal{N}(m_0, \Sigma_0), \qquad \mu = \mathcal{N}(m_1, \Sigma_1),$$

   show that the optimal potentials are quadratic and the optimal map is affine.

Quadratic cost turns Gaussian OT into completing squares.

Hints for the quadratic exercises

Translation. Try affine potentials

$$u(x) = -m \cdot x, \qquad v(y) = -m \cdot y + \frac{1}{2}|m|^2.$$

Check

$$\frac{1}{2}|x-y|^2 - \big(u(x) - v(y)\big) = \frac{1}{2}|y - x - m|^2.$$

Dilation in one dimension. Try

$$u(x) = \frac{1-a}{2}x^2, \qquad v(y) = \frac{1-a}{2a}y^2.$$

Check

$$\frac{1}{2}(x-y)^2 - \big(u(x) - v(y)\big) = \frac{1}{2a}(y - ax)^2.$$

General Gaussians. Try affine $T(x) = m_1 + A(x - m_0)$ and quadratic potentials.

Contact set and c-subdifferentials

Recall

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

Assume

$$u = Qv, \qquad v = \widehat{Q}u,$$

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

Define the subdifferentials

$$\partial_c u(x) := \Big\{ y \in Y : u(x) = v(y) + c(x,y) \Big\},$$

$$\hat{\partial}_c v(y) := \Big\{ x \in X : u(x) = v(y) + c(x,y) \Big\}.$$

Then the contact set is defined as

$$\Gamma(u,v) = \Big\{ (x,y) : y \in \partial_c u(x) \Big\} = \Big\{ (x,y) : x \in \hat{\partial}_c v(y) \Big\}.$$

Complementary slackness

Let $\pi \in \Pi(\rho, \mu)$ and let (u, v) be admissible:

$$c(x, y) - u(x) + v(y) \geq 0.$$

If π is primal optimal and (u, v) is dual optimal, then

$$\int_{X \times Y} \Big(c(x, y) - u(x) + v(y) \Big) \, d\pi(x, y) = 0.$$

Since the integrand is nonnegative,

$$c(x, y) - u(x) + v(y) = 0 \qquad \pi\text{-a.e.}$$

Equivalently,

$$u(x) = v(y) + c(x, y) \qquad \pi\text{-a.e.}$$

$$\boxed{\text{Optimal transport plans are concentrated on the contact set.}}$$

Definition: c-cyclical monotonicity

A set $\Gamma \subset X \times Y$ is c-cyclically monotone if for every finite family

$$(x_i, y_i) \in \Gamma, \qquad i = 1, \ldots, N,$$

with $y_{N+1} = y_1$,

$$\boxed{\sum_{i=1}^{N} c(x_i, y_i)} \le \boxed{\sum_{i=1}^{N} c(x_i, y_{i+1})}.$$

Contact implies c-cyclical monotonicity

For $(x_i, y_i) \in \Gamma$, contact gives

$$\boxed{u(x_i) - v(y_i) = c(x_i, y_i).} \leftarrow$$

Admissibility gives

$$\boxed{u(x_i) - v(y_{i+1}) \leq c(x_i, y_{i+1}).}$$

$$\sum_{i=1}^{N} c(x_i, y_i) = \sum_{i=1}^{N} u(x_i) - \sum_{i=1}^{N} v(y_i) = \sum_{i=1}^{N} u(x_i) - \sum_{i=1}^{N} v(y_{i+1})$$

$$= \sum_{i=1}^{N} u(x_i) - v(y_{i+1}) \leq \sum_{i=1}^{N} c(x_i, y_{i+1})$$

Example: monotonicity

Let

$$c(x,y) = \frac{1}{2}|x-y|^2, \quad x, y \in \mathbb{R}^d.$$

For two contact pairs,

$$(x_1, y_1), (x_2, y_2) \in \Gamma,$$

cyclical monotonicity gives

$$|x_1 - y_1|^2 + |x_2 - y_2|^2 \leq |x_1 - y_2|^2 + |x_2 - y_1|^2.$$

$$- x_1 \cdot y_1 - x_2 \cdot y_2 \leq - x_1 \cdot y_2 - x_2 \cdot y_1$$

$$x_1 (y_2 - y_1) - x_2 (y_2 - y_1) \leq 0$$

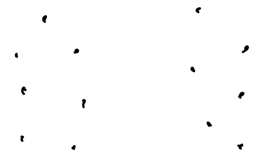
$$(x_1 - x_2) \cdot (y_1 - y_2) \geq 0$$

$\rightsquigarrow$ $y = y(x)$ is increasing

Theorem: optimal plans are c-cyclically monotone

Let X, Y be Polish spaces and let

$$c : X \times Y \to (-\infty, +\infty]$$



be lower semicontinuous and bounded from below. Let

$$\pi \in \Pi(\rho, \mu)$$

be an optimal transport plan.

Then π is concentrated on a c-cyclically monotone set: there exists $\Gamma \subset X \times Y$ such that

$$\pi(\Gamma) = 1,$$

and for every $N \geq 1$, every $(x_i, y_i)_{i=1}^{N} \subset \Gamma$, and every (cyclic) permutation $\sigma \in S_N$,

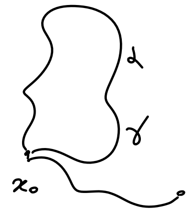
$$\sum_{i=1}^{N} c(x_i, y_i) \leq \sum_{i=1}^{N} c(x_i, y_{\sigma(i)}).$$

References: Villani, *Topics in Optimal Transportation*, Thm. 2.3; Ambrosio–Gigli–Savaré, *Gradient Flows*, Ch. 6; Santambrogio, *Optimal Transport for Applied Mathematicians*, Ch. 1.

Closed one-forms versus cyclic work

Smooth Poincare lemma:

$$d\alpha = 0 \qquad \Longrightarrow \qquad \alpha = d\phi.$$



Work along closed curves:

$$\int_\gamma \alpha = 0.$$

Nonsmooth convex analogue (Rockafellar's theorem):

$$\Gamma \text{ cyclically monotone} \qquad \Longrightarrow \qquad \boxed{\Gamma \subset \partial\phi.} \qquad \phi \text{ convex}$$

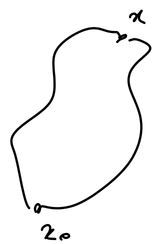
w.r.t. $-x \cdot y = c(x,y)$

Cyclical monotonicity is nonsmooth closedness.

Classical Poincare construction

If $d\alpha = 0$ on a simply connected domain, define

$$\phi(x) = \phi(x_0) + \int_{x_0}^{x} \alpha.$$



Closedness gives path independence.

closedness $\Longrightarrow$ path-independent work $\Longrightarrow$ potential.

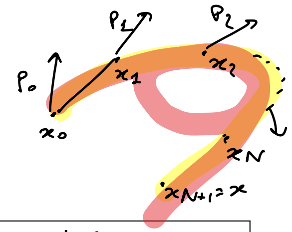
Rockafellar construction

Let $\Gamma \subset \mathbb{R}^d \times \mathbb{R}^d$ be cyclically monotone (w.r.t $c(x,p) = -x \cdot p$).
Fix $(x_0, p_0) \in \Gamma$. Define

$$\phi(x) = \sup_{\text{chains}} \overbrace{\sum_i p_i \cdot (x_{i+1} - x_i)}^{\text{"work along the chain"}}.$$

Here

$$(x_i, p_i) \in \Gamma, \qquad x_{N+1} = x.$$



$\boxed{\text{Line integral } \int \alpha \text{ becomes discrete work along chains.}}$

Cyclic monotonicity is the closedness condition

For a cyclically monotone set, $\sum_{i=0}^{N} p_i \cdot (x_{i+1} - x_i) \leq 0$ whenever $x_{N+1} = x_0$.

$\Rightarrow$ closed discrete loops have no positive work.

This prevents the chain formula

$$\phi(x) = \sup_{\text{chains ending at } x} \sum_i p_i \cdot (x_{i+1} - x_i)$$

from being made larger by inserting artificial cycles.

$$\boxed{\text{Cyclical monotonicity is the no-positive-work condition.}}$$

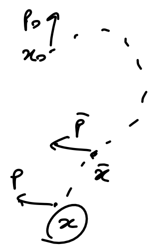
The potential has the right subgradients

Define

$$\phi(x) := \sup_{\text{chains ending at } x} \sum_i p_i \cdot (x_{i+1} - x_i).$$

Take a point $(\bar{x}, \bar{p}) \in \Gamma$. Any chain ending at $\bar{x}$ can be extended to end at x by adding the step

$$(\bar{x}, \bar{p}) : \qquad \bar{p} \cdot (x - \bar{x}).$$



Therefore

$$\phi(x) \geq \phi(\bar{x}) + \bar{p} \cdot (x - \bar{x}).$$

This means

$$\boxed{\bar{p} \in \partial\phi(\bar{x}).}$$

Theorem: c-Rockafellar

Let X, Y be Polish spaces and let

$$c : X \times Y \to [0, +\infty]$$

be lower semicontinuous.
If

$$\Gamma \subset X \times Y$$

is c-cyclically monotone and $c < +\infty$ on Γ, then there exist admissible potentials u, v such that

$$u(x) \leq v(y) + c(x, y) \qquad \forall x, y,$$

and

$$u(x) = v(y) + c(x, y) \qquad \forall (x, y) \in \Gamma.$$

In other words,

$$\Gamma \subset \big\{ (x, y) : u(x) = v(y) + c(x, y) \big\}.$$

Conversely, every subset of the contact set of admissible potentials is c-cyclically monotone.
Reference: Villani, *Topics in Optimal Transportation*, Thm. 2.3.

The c-Rockafellar construction

Given $\Gamma \subset X \times Y$, c-cyclically monotone. Fix a base point $(x_0, y_0) \in \Gamma$.

$$\boxed{\text{With our sign convention, the chain formula is an infimum.}}$$

Define u by a infimum over chains:

$$u(x) := \inf_{\text{chains}} \left\{ \sum_{i=0}^{N} \Big[c\big(\overbrace{x_{i+1}, y_i) - c(x_i,}^{\text{work from } x_i \text{ to } x_{i+1} \rightsquigarrow -p_i(x_{i+1}-x_i)} y_i\big) \Big] \right\},$$

where $(x_i, y_i) \in \Gamma$, $x_{N+1} = x$.

Then define $v = \widehat{Q}u = \underline{\sup_x \{u(x) - c(x, y)\}}.$

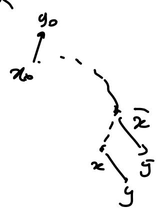
Why the c-construction works

By construction

$$u(x) \le v(y) + c(x, y) \qquad \forall x, y. \quad \text{OK}$$

Now take $(\bar{x}, \bar{y}) \in \Gamma$. The chain construction gives

$$\hat{Q}u(\bar{y})\ \sup_{x} \{ u(x) - c(x, \bar{y}) \} \le u(\bar{x}) - c(\bar{x}, \bar{y}) \qquad \forall x$$



Taking the supremum over x,

$$v(\bar{y}) = \sup_{x} \{ u(x) - c(x, \bar{y}) \} \le u(\bar{x}) - c(\bar{x}, \bar{y}).$$

Together with admissibility,

$$u(\bar{x}) \le v(\bar{y}) + c(\bar{x}, \bar{y}),$$

we obtain

$$u(\bar{x}) = v(\bar{y}) + c(\bar{x}, \bar{y}).$$

Duality from Rockafellar

From optimality:

$$\operatorname{supp} \pi \text{ is } c\text{-cyclically monotone}.$$

By c-Rockafellar, there exist u, v such that

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

and equality holds on $\operatorname{supp} \pi$.
Therefore

$$\int c \, d\pi = \int (u(x) - v(y)) \, d\pi = \int u \, d\rho - \int v \, d\mu.$$

Main route:

optimality $\Rightarrow$ c-cyclical monotonicity $\Rightarrow$ potential $\Rightarrow$ dual equality.

Lectures 01–02 recap

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

$$u(x) \le v(y) + c(x,y)$$

2. **Hard c-transforms.**

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

3. **Softmin and entropy.**

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

4. **Zero-temperature limit.**

$$\mathrm{Q}_c^\varepsilon \longrightarrow \mathrm{Q}_c, \qquad \text{Gibbs kernels} \longrightarrow \text{contact sets.}$$

5. **Rockafellar theorem.**

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