Stability of Schrödinger potentials and application to PDEs

Maxime Laborde
in collaboration with G. Carlier and L. Chizat

Université Paris Cité

CANUM, May 31, 2024

Motivation: Dynamical urban planning model

Solve

$$\begin{cases} \partial_t \rho_1 - \Delta \rho_1^p - \mathrm{div}(\rho_1 \nabla (V_1 + \varphi)) = 0, \\ \partial_t \rho_2 - \Delta \rho_2^q - \mathrm{div}(\rho_2 \nabla (V_2 + \psi)) = 0, \\ \rho_1(0,\cdot) = \rho_{1,0},\ \rho_2(0,\cdot) = \rho_{2,0}, \end{cases}$$

on a compact, convex subset Ω of $\mathbb{R}^n$ with no-flux boundary conditions, where $p, q \geqslant 1$, V_1 and V_2 are smooth potentials and φ and , ψ are:

- Potentials of Kantorovich: Optimal transport problem
- Potentials of Schrödinger: Entropic regularization of optimal transport problem

Plan

Table of contents

Wasserstein distance

- Kantorovich relaxation (1942): given a source $\rho_1 \in \mathcal{P}(\Omega)$ and a target $\rho_2 \in \mathcal{P}(\Omega)$

$$W_2^2(\rho_1, \rho_2) = \min \left\{ \iint_{\Omega \times \Omega} |y - x|^2 \, d\gamma(x, y) \, : \, \gamma \in \Pi(\rho_1, \rho_2) \right\},$$

  where

$$\Pi(\rho_1, \rho_2) = \left\{ \gamma \in \mathcal{P}(\Omega \times \Omega) \, : \, \pi_{1\#}\gamma = \rho_1 \text{ and } \pi_{2\#}\gamma = \rho_2 \right\}.$$

- Wasserstein distance: W_2 define a metric on $\mathcal{P}(\Omega)$ and an optimal γ is called optimal transport plan between ρ_1 and ρ_2.

Dual formulation and Brenier's theorem

Dual formulation:

$$W_2^2(\rho_1, \rho_2) = \max\left\{ \int_\Omega \varphi_1(x)\, d\rho_1(x) + \int_\Omega \varphi_2(x)\, d\rho_2(x) \;:\; \varphi_1(x) + \varphi_2(y) \leqslant |x-y|^2 \right\}.$$

Solutions (φ_1, φ_2) are called Kantorovich potential.

Theorem (Brenier 1989)

If $\rho_1 \ll \mathcal{L}_{|\Omega}$, then Kantorovich's problem admits a unique solution γ induced by a map T, i.e. $\gamma = (Id, T)_{\#}\rho_1$.
Moreover, T is the gradient of a convex function and satisfies $T = Id - \nabla\varphi$ where φ is a Kantorovich potential.

Entropic regularization

- Regularized optimal transport problem:

$$W_{c,\epsilon}(\rho_1,\rho_2) := \inf_{\gamma\in\Pi(\rho_1,\rho_2)} \left\{ \iint_{\Omega\times\Omega} c(x,y)\, d\gamma(x,y) + \epsilon \iint_{\Omega\times\Omega} \gamma(\log(\gamma)-1) \right\}.$$

Can be rewritten as

$$W_{c,\epsilon}(\rho_1,\rho_2) = \epsilon \inf_{\gamma\in\Pi(\rho_1,\rho_2)} \mathcal{H}(\gamma|G_\epsilon),$$

where $G_\epsilon := e^{-\frac{c}{\epsilon}}$ and $\mathcal{H}$ is the relative entropy defined by

$$\mathcal{H}(\gamma|\mu) := \begin{cases} \int_{\Omega\times\Omega}(\log\left(\frac{d\gamma}{d\mu}\right)-1)d\gamma & \text{if } \gamma \ll \mu \\ +\infty & \text{otherwise.} \end{cases}$$

Entropic regularization

- Regularized optimal transport problem:

$$W_{c,\epsilon}(\rho_1,\rho_2) := \inf_{\gamma\in\Pi(\rho_1,\rho_2)} \left\{ \iint_{\Omega\times\Omega} c(x,y)\, d\gamma(x,y) + \epsilon \iint_{\Omega\times\Omega} \gamma(\log(\gamma)-1) \right\}.$$

Can be rewritten as

$$W_{c,\epsilon}(\rho_1,\rho_2) = \epsilon \inf_{\gamma\in\Pi(\rho_1,\rho_2)} \mathcal{H}(\gamma|G_\epsilon),$$

where $G_\epsilon := e^{-\frac{c}{\epsilon}}$ and $\mathcal{H}$ is the relative entropy defined by

$$\mathcal{H}(\gamma|\mu) := \begin{cases} \int_{\Omega\times\Omega}(\log\left(\frac{d\gamma}{d\mu}\right)-1)d\gamma & \text{if } \gamma \ll \mu \\ +\infty & \text{otherwise.} \end{cases}$$

- Change of reference measure: Define

$$E(\rho_1,\rho_2) := \inf_{\gamma\in\Pi(\rho_1,\rho_2)} \left\{ \iint_{\Omega\times\Omega} c(x,y)\, d\gamma(x,y) + \epsilon\mathcal{H}(\gamma|\rho_1\otimes\rho_2) \right\}$$

Remark: $W_{c,\epsilon}(\rho_1,\rho_2) = E(\rho_1,\rho_2) + \epsilon\mathcal{H}(\rho_1) + \epsilon\mathcal{H}(\rho_2)$.

Dual problem and Schrödinger system

- Dual problem

$$E(\rho_1, \rho_2) = \max_{\phi_1, \phi_2} \int_\Omega \phi_1 \, d\rho_1 + \int_\Omega \phi_2 \, d\rho_2 - \int_{\Omega^2} e^{(\phi_1(x)+\phi_2(y))/\epsilon} G_\epsilon(x, y) \, d\rho_1(x) d\rho_2(y).$$

Dual problem and Schrödinger system

- Dual problem

$$
\begin{aligned}
&E(\rho_1, \rho_2) \\
&= \max_{\phi_1,\phi_2} \int_\Omega \phi_1 \, d\rho_1 + \int_\Omega \phi_2 \, d\rho_2 - \int_{\Omega^2} e^{(\phi_1(x)+\phi_2(y))/\epsilon} G_\epsilon(x,y) \, d\rho_1(x) d\rho_2(y).
\end{aligned}
$$

- Schrödinger system: Dual solutions (ϕ_1, ϕ_2) satisfies $\rho_1 \otimes \rho_2$ a.e.

$$
\begin{cases}
\phi_1(x) = -\epsilon \log\left(\int_\Omega e^{\frac{\phi_2(y)}{\epsilon}} G_\epsilon(x,y) \, d\rho_2(y) \right) \\
\phi_2(y) = -\epsilon \log\left(\int_\Omega e^{\frac{\phi_1(x)}{\epsilon}} G_\epsilon(x,y) \, d\rho_1(x) \right)
\end{cases}
$$

Dual problem and Schrödinger system

- Dual problem

$$E(\rho_1, \rho_2) \\ = \max_{\phi_1, \phi_2} \int_\Omega \phi_1 \, d\rho_1 + \int_\Omega \phi_2 \, d\rho_2 - \int_{\Omega^2} e^{(\phi_1(x)+\phi_2(y))/\epsilon} G_\epsilon(x,y) \, d\rho_1(x) d\rho_2(y).$$

- Schrödinger system: Dual solutions (ϕ_1, ϕ_2) satisfies $\rho_1 \otimes \rho_2$ a.e.

$$\begin{cases} \phi_1(x) = -\epsilon \log \left(\int_\Omega e^{\frac{\phi_2(y)}{\epsilon}} G_\epsilon(x,y) \, d\rho_2(y) \right) \\ \phi_2(y) = -\epsilon \log \left(\int_\Omega e^{\frac{\phi_1(x)}{\epsilon}} G_\epsilon(x,y) \, d\rho_1(x) \right) \end{cases}$$

- Regularity: same as c and unique in $\tilde{\mathcal{C}}^k := \mathcal{C}^k \times \mathcal{C}^k / \sim$ where

$$(\phi_1, \phi_2) \sim (\psi_1, \psi_2) \quad \Leftrightarrow \quad \exists \kappa \in \mathbb{R} \text{ such that } \phi_1 = \psi_1 + \kappa \text{ and } \phi_2 = \psi_2 - \kappa$$

Table of contents

Lipschitz stability of the Schrödinger map

The Schrödinger map:

$$S : (\rho_1, \rho_2) \longmapsto (\phi_1, \phi_2)$$

Theorem (Carlier-Chizat-L., '22)

For $k \in \mathbb{N}^*$, assume that $c \in \mathcal{C}^{k+1}(\Omega)$. The Schrödinger map $S : \mathcal{P}(\Omega)^2 \to \tilde{\mathcal{C}}^k$ is Lipschitz continuous, i.e. there exists $C > 0$ such that, for all $(\rho_1, \rho_2), (\mu_1, \mu_2) \in \mathcal{P}(\Omega)^4$,

$$\|S(\rho_1, \rho_2) - S(\mu_1, \mu_2)\|_{\tilde{\mathcal{C}}^k} \leq C(W_2^2(\rho_1, \mu_1) + W_2^2(\rho_2, \mu_2))^{1/2}.$$

Remark: can be extended to multimarignal case

Idea of proof

Denote $\boldsymbol{\rho} = (\rho_1, \rho_2)$ and $\boldsymbol{\phi} = (\phi_1, \phi_2)$

- Rewrite the Schrödinger system as

$$F(\boldsymbol{\phi}, \boldsymbol{\rho}) = 0$$

- For any optimal transport plan $\boldsymbol{\gamma} \in \Pi(\boldsymbol{\rho}, \boldsymbol{\mu})$, consider the interpolation

$$\boldsymbol{\rho}_t = ((1-t)\pi_1 + t\pi_2)_{\#}\boldsymbol{\gamma}$$

- Apply the implicit function theorem to

$$G : \begin{array}{ll} \tilde{\mathcal{C}}^k \times [0,1] & \longrightarrow \tilde{\mathcal{C}}^k \\ (\boldsymbol{\phi}, t) & \longmapsto F(\boldsymbol{\phi}, \boldsymbol{\rho}_t) \end{array}$$

Displacement smoothness

Let $\rho_0, \rho_1 \in \mathcal{P}(\Omega)$ and γ an optimal transport plan between ρ_0 and ρ_1. Define the Wasserstein geodesic $t \mapsto \rho_t$ by

$$\rho^t := ((1-t)\pi_1 + t\pi_2)_{\#}\gamma.$$

A functional $\mathcal{E} : (\rho_1, \dots, \rho_I) \mapsto \mathcal{E}(\rho_1, \dots, \rho_I)$ is said λ-geodesically convex if

$$t \in [0,1] \mapsto \mathcal{E}(\rho_1^t, \dots, \rho_I^t) \text{ is } \lambda\text{-convex.}$$

Corollary (Carlier-Chizat-L., '22)

If $c \in \mathcal{C}^2$, then there exists $\lambda > 0$ such that E and $-E$ are $(-\lambda)$-geodesically convex.

Table of contents

Optimal transport and labour market [Carlier-Ekeland '04]

- ρ_1, ρ_2 are the densities of inhabitants and firms in a city $\Omega \subset \mathbb{R}^2$,
- Commuting cost from x to y given by $c(x,y)$, e.g. $c(x,y) = |x-y|^2$,
- Where to work? Optimization problem over commuting cost and salary, $\psi(y)$,

$$\varphi(x) = \inf_{y\in\Omega} \{c(x,y) - \psi(y)\}.$$

- Construction of a transport map:

$$T(x) = \operatorname{argmin} \{c(x,y) - \psi(y)\}$$

  Then the equilibrium constraint reads $T_{\#}\rho_1 = \rho_2$.

- This problem is equivalent to solve the optimal transport problem

$$\inf_{\rho_2 = T_{\#}\rho_1} \int_\Omega c(T(x), x)\, d\rho_1(x),$$

  and φ and ψ are simply the Kantorovich potential associated to the dual problem.

Static urban planning model [Buttazzo-Santambrogio '05]

Optimal distribution in a city Ω: Existence and characterization of minimizers of

$$(\rho_1, \rho_2) \mapsto W_2^2(\rho_1, \rho_2) + \mathcal{F}(\rho_1) + \mathcal{G}(\rho_2),$$

where

- $\mathcal{F}$ represents a congestion effect for the inhabitants, for example

$$\mathcal{F}(\rho) = \int_\Omega F(\rho(x))\,dx = \int_\Omega \frac{F(\rho(x))}{\rho(x)}\rho(x),$$

  where F is convex and superlinear. $\frac{F(\rho)}{\rho}$ is unhappiness of a citizen living at a place with density ρ.

- $\mathcal{G}$ represents a concentration effect for the firms, for example

$$\mathcal{G}(\rho) := \iint_{\Omega\times\Omega} |x-y|^2\,d\rho(x)d\rho(y).$$

Dynamical urban planning model

[L. '20]

Dynamics of

$$\mathcal{E}(\rho) = W_2^2(\rho_1, \rho_2) + \mathcal{F}(\rho_1) + \mathcal{G}(\rho_2)$$

- Wasserstein gradient flow of $\mathcal{E}$: Formally,

$$\begin{cases} \partial_t \rho_1 - \text{div}(\rho_1 \nabla F'(\rho_1)) - \text{div}(\rho_1 \nabla \varphi) = 0, \\ \partial_t \rho_2 - \text{div}(\rho_2 \nabla G(|x-y|) * \rho_2) - \text{div}(\rho_2 \nabla \psi) = 0, \end{cases} \tag{1}$$

  where $(\varphi(t), \psi(t))$ is a pair of Kantorovich potential of $W_2(\rho_1(t), \rho_2(t))$, t-a.e.

- W_2^2 is not geodesically convex in general!

- Existence of weak solutions: JKO scheme

Dynamic with noise

Dynamics of

$$\mathcal{E}_\epsilon(\rho) = W^2_{2,\epsilon}(\rho_1, \rho_2) + \mathcal{F}(\rho_1) + \mathcal{G}(\rho_2)$$

- Wasserstein gradient flowof $\mathcal{E}_\epsilon$ satisfies

$$\begin{cases} \partial_t \rho_1 - \mathrm{div}(\rho_1 \nabla F'(\rho_1)) - \mathrm{div}(\rho_1 \nabla \phi_1) - \epsilon \Delta \rho_1 = 0, \\ \partial_t \rho_2 - \mathrm{div}(\rho_2 \nabla G(|x-y|) * \rho_2) - \mathrm{div}(\rho_2 \nabla \phi_2) - \epsilon \Delta \rho_2 = 0, \\ \rho_1(0, \cdot) = \rho_{1,0},\ \rho_2(0, \cdot) = \rho_{2,0}, \end{cases}$$

  where (ϕ_1, ϕ_2) are Schrödinger potentials.

- $\mathcal{E}_\epsilon$ is geodesically convex.

- Existence and uniqueness of Wasserstein gradient flow of $\mathcal{W}_{c,\epsilon}$

Asymptotic convergence

Proposition (Carlier-Chizat-L., '22)

Assume that $\mathcal{H}(\rho_i^0) < +\infty$ for every i, then $\boldsymbol{\rho}_t$, the WGF of $W_{c,\epsilon}$, converges at an exponential rate to the equilibrium $\boldsymbol{\rho}^*$, defined by

$$\rho_i^*(x) = \frac{\int_\Omega e^{-c(x,y)/\epsilon}\mathrm{d}y}{\int_{\Omega^2} e^{-c/\epsilon}}$$

i.e. there exists $\kappa > 0$ independent of $\boldsymbol{\rho}^0$ such that

$$W_{c,\epsilon}(\boldsymbol{\rho}_t) - W_{c,\epsilon}(\boldsymbol{\rho}^*) \leq e^{-\kappa t}(W_{c,\epsilon}(\boldsymbol{\rho}^0) - W_{c,\epsilon}(\boldsymbol{\rho}_*)).$$

Remark: E is not λ-geodesically convex with $\lambda > 0$

Thank you for your attention!