A new Optimal Transport distance between nonnegative Radon measures

Léonard Monsaingeon CAMGSD-IST Lisbon
(with S. Kondratyev and D. Vorotnikov)

International Workshop on Calculus of Variations and its Applications
on the occasion of Luísa Mascarenha's 65th birthday
Costa de Caparica, December 2015

UT Austin | Portugal
International Collaboratory for Emerging Technologies, CoLab

FCT
Fundação para a Ciência e a Tecnologia
MINISTÉRIO DA CIÊNCIA, TECNOLOGIA E ENSINO SUPERIOR

CAMGSD

Static : Monge-Kantorovich problem and Wasserstein distance

$$\mathcal{W}_2^2(\rho_0,\rho_1) = \min_{\gamma\in\Gamma[\rho_0,\rho_1]} \iint_{\mathbb{R}^d\times\mathbb{R}^d} |x-y|^2 d\gamma(x,y)$$

Dynamical representation :

Benamou-Brenier formula '00

$$\mathcal{W}_2^2(\rho_0,\rho_1) = \min \int_0^1 \int_{\mathbb{R}^d} |\mathbf{v}_t|^2 \mathrm{d}\rho_t \mathrm{d}t$$

over curves $\rho\in\mathcal{C}_w([0,1];\mathcal{P}_2)$ and velocity fields $\mathbf{v}\in L^2(0,1;L^2(\mathrm{d}\rho_t))$ s.t.

$$\partial_t\rho_t + \mathrm{div}(\rho_t\mathbf{v}_t) = 0 \qquad \text{(C)}$$

Divergence form $\Rightarrow$ mass conservative, optimal $\mathbf{v}_t = \nabla p_t$

Definition/Theorem (KMV '15)

$\rho_0, \rho_1 \in \mathcal{M}_b^+$:

$$d^2(\rho_0, \rho_1) = \inf \int_0^1 \int_{\mathbb{R}^d} \left(|\nabla u_t|^2 + |u_t|^2\right) \mathrm{d}\rho_t \mathrm{d}t$$

over curves $\rho \in \mathcal{C}_w([0,1]; \mathcal{M}_b^+)$ and potentials $u \in L^2(0,1; H^1(\mathrm{d}\rho_t))$ s.t.

$$\partial_t \rho_t + \mathrm{div}(\rho_t \nabla u_t) = \rho_t u_t \qquad \text{(NCC)}$$

Rmks :

- Reaction term $\rightsquigarrow$ Minimum Principle, mass variations $|\rho_0| \neq |\rho_1|$
- Two strategies : displace or create/kill mass $\Rightarrow$ different geometry than $(\mathcal{P}_2, \mathcal{W}_2)$
- No finite moments or decay required

$$\partial_t \rho_t + \mathrm{div}(\rho_t \nabla u_t) = \rho_t u_t, \qquad d^2(\rho_0,\rho_1) = \inf \int_0^1 \|u_t\|^2_{H^1(\mathrm{d}\rho_t)} \mathrm{d}t$$

Theorem (KMV '15)

(i) $(\mathcal{M}^+, d)$ is a complete metric space

(ii) d is LSC w.r.t weak-$*$ CV of measures and metrizes the narrow topology

(iii) $d \leq \mathcal{W}_2$ on $\mathcal{P}_2$

(iv) compactly supported measures are dense : $d(\rho|_{B_R}, \rho) \to 0$ as $R \to \infty$

Theorem (KMV '15)

(i) Lipschitz curves $d(\rho_t,\rho_s) \leq L|t-s|$ can be represented by potentials $\|u_t\|_{H^1(\mathrm{d}\rho_t)} \leq L$, and vice-versa.

(ii) $(\mathcal{M}^+, d)$ is a geodesic space, i.e $d^2(\rho_0,\rho_1) = \int_0^1 \|u_t\|^2_{H^1(\mathrm{d}\rho_t)} \mathrm{d}t$ for some constant-speed curve $\|u_t\|_{H^1(\mathrm{d}\rho_t)} = cst$ with $d(\rho_t,\rho_s) = |t-s| d(\rho_0,\rho_1)$.

Rmk : for Wasserstein "narrow $\approx$ weak-$*$". More delicate here, no tightness

Geodesics $\partial_t \rho_t + \mathrm{div}(\rho_t \nabla u_t) = \rho_t u_t$, velocity $\zeta = \frac{d}{dt}\big|_{t=0} \rho_t \sim -\mathrm{div}(\rho \nabla u) + \rho u$.

Differential structure *à la Otto*, $\mathcal{M}_b^+ \approx$ Riemannian manifold

Tangent plane

$$T_\rho \mathcal{M}^+ = \{\zeta = -\mathrm{div}(\rho \nabla u) + \rho u : \quad u \in H^1(\mathrm{d}\rho)\} \qquad (\rho \neq 0 \text{ extremal})$$

and metrics

$$g_\rho(\zeta_1, \zeta_2) = \langle u_1, u_2 \rangle_{H^1(\mathrm{d}\rho)} = \int_{\mathbb{R}^d} (\nabla u_1 \cdot \nabla u_2 + u_1 u_2)\, \mathrm{d}\rho$$

Chain-rule $\frac{d}{dt}\big|_{t=0} \mathcal{F}(\rho_t) = g_\rho(\mathrm{grad}_d \mathcal{F}(\rho), \zeta)$

Gradients

$$\mathrm{grad}_d \mathcal{F}(\rho) = -\mathrm{div}\left(\rho \nabla \frac{\delta \mathcal{F}}{\delta \rho}\right) + \rho \frac{\delta \mathcal{F}}{\delta \rho}, \qquad \text{i-e } u = \frac{\delta \mathcal{F}}{\delta \rho}$$

Rmk : 2nd order calculus $g_\rho(\mathrm{Hess}_d \mathcal{F}(\rho) \cdot \zeta, \zeta) = \frac{d^2}{dt^2}\Big|_{geod} \mathcal{F}(\rho_t)$, more involved (generalized sticky particles)

Fitness-driven population dynamics [Cosner, *Th. Pop. Bio.*, '05]

$$\begin{cases} \partial_t \rho = -\operatorname{div}(\rho \nabla (m-\rho)) + \rho(m-\rho), & x \in \Omega \\ \rho \frac{\partial}{\partial \nu}(m-\rho) = 0, & x \in \partial\Omega \\ \rho|_{t=0} = \rho_0(x), & x \in \Omega. \end{cases}$$

Population density $\rho(t,x) \geq 0$, resources $m(x) \geq m_0 > 0$, fitness $m - \rho$.

Ignore BC's, $\operatorname{grad}_d \mathcal{F}(\rho) = -\operatorname{div}\left(\rho \nabla \frac{\delta \mathcal{F}}{\delta \rho}\right) + \rho \frac{\delta \mathcal{F}}{\delta \rho}$

$$\boxed{\mathcal{E}(\rho) = \frac{1}{2}\int_\Omega |\rho - m|^2 \mathrm{d}x \qquad \Rightarrow \qquad \frac{d\rho}{dt} = -\operatorname{grad}_d \mathcal{E}(\rho)}$$

Theorem (KMV '15)

There is $\gamma \equiv \gamma(\Omega, m, \rho_0)$ s.t. $\|\rho(t) - m\|_{L^2(\Omega)} \leq e^{-\gamma t}\|\rho_0 - m\|_{L^2(\Omega)}$

Least energy solution $\rho^* = m$, quantitative improvement on [Cosner-Winkler, '14]
No λ-displacement convexity : generalized Beckner inequality, stay away from 0

Thank you for your attention,

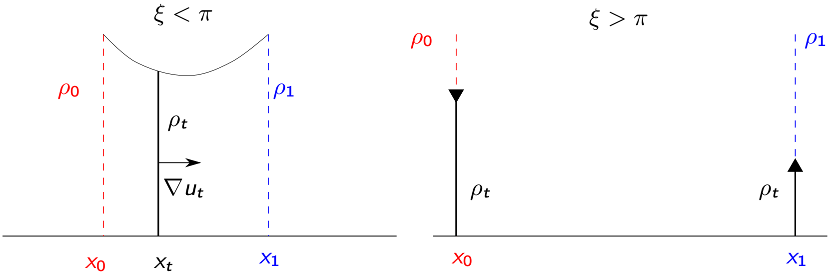
and

Happy birthday Luísa !

An illustrative example

One-point particles $\rho_0 = \delta_{x_0}, \rho_1 = \delta_{x_1}$ at distance $\xi = |x_1 - x_0|$

$$\begin{array}{ll} \xi < \pi : & \text{Wasserstein-like transport } \rho_t = \lambda(t)\delta_{x_t} \\ \xi > \pi : & \text{pure reaction (create/kill) } \rho_t = \kappa(t)\delta_{x_0} + \lambda(t)\delta_{x_1} \\ \xi = \pi : & \text{any combination ! } \rightsquigarrow \text{ non-uniqueness} \end{array}$$



$$\begin{array}{ll} \text{Short range :} & |\mathcal{W}_2(\rho_0,\rho_1) - d(\rho_0,\rho_1)| = o(\xi^2) \\ \text{Long range :} & d(\rho_0,\rho_1) \le d(\rho_0,0) + d(0,\rho_1) \le 4 \end{array}$$

$$\partial_t \rho = -\operatorname{div}(\rho \nabla (m-\rho)) + \rho(m-\rho) \qquad \leftrightarrow \qquad \frac{d\rho}{dt} = -\operatorname{grad}_d \mathcal{E}, \quad \mathcal{E}(\rho) = \frac{1}{2}\int_\Omega |\rho - m|^2$$

1. $\exists!$ and regularity of weak solutions [Cosner-Winkler, '14] : dissipation

$$\mathcal{D}(t) = -\frac{d}{dt}\mathcal{E}(\rho) = |\operatorname{grad}_d \mathcal{E}|^2 = \int_\Omega \left(|\nabla(\rho - m)|^2 + |\rho - m|^2\right) \mathrm{d}\rho$$

2. Entropy/Entropy production : No λ-convexity, generalized Beckner inequality

$$\Phi\left(\int_\Omega f\right)\int_\Omega |f-m|^2 \leq \int_\Omega \left(|\nabla(f-m)|^2 + |f-m|^2\right) f, \qquad \Phi(0) = 0 \text{ and } \Phi \uparrow$$

3. Comparison principle : $\int_\Omega \rho \geq \int_\Omega \min\{\rho_0, m\} = c_0 > 0$, $\gamma = \Phi(c_0)$

$$-\frac{d}{dt}\mathcal{E} = \mathcal{D} \geq 2\lambda\mathcal{E} \qquad \text{and} \qquad \mathcal{E}(t) \leq e^{-2\gamma t}\mathcal{E}(0)$$