Mobility functions, transport distances, and nonlinear diffusion

Giuseppe Savaré
http://www.imati.cnr.it/∼savare

Department of Mathematics, University of Pavia, Italy

Nonlinear Diffusions and Entropy Dissipation: From Geometry to Biology

BANFF, *May 10, 2010*

Jointly with *J. Dolbeault, B. Nazaret, J. Carrillo, S. Lisini, D. Slepcev*

Outline

Gradient flows and Wasserstein distance

$$\begin{cases} \partial_t u + \operatorname{div} \mathbf{w} = 0 & \textbf{(Continuity equation)} \\ \mathbf{w} = \mathbf{u}\,\mathbf{v} = -\mathbf{u}\,\mathrm{D}\psi & \textbf{(Flux structure)} \\ \psi = \dfrac{\delta \boldsymbol{\Phi}}{\delta \mathbf{u}} & \textbf{(Nonlinear variational condition)} \\ u(0,\cdot) = u_0 & u_0 \in L^1(\mathbb{R}^d),\ \ u_0 \geq 0. \end{cases}$$

Here Φ is an integral functional and $\dfrac{\delta \boldsymbol{\Phi}}{\delta \mathbf{u}}$ is its Euler-Lagrange first variation

$$\Phi(u) := \int \phi(x, u, \mathrm{D}u)\, dx, \qquad \frac{\delta \boldsymbol{\Phi}}{\delta \mathbf{u}} = \boldsymbol{\phi_{\mathbf{u}}(\mathbf{x}, \mathbf{u}, \mathrm{D}\mathbf{u}) - \operatorname{div} \phi_{\mathrm{D}\mathbf{u}}(\mathbf{x}, \mathbf{u}, \mathrm{D}\mathbf{u})}$$

[Jordan-Kinderlehrer-Otto '98 Otto '01] showed in many interesting cases that such kind of equations can be interpreted as

> the **"gradient flow"** of Φ with respect to the so called
> **"Wasserstein distance"**
> between probability density/measures.

Applications: existence and asymptotic behaviour of solutions, contracton properties, Logarithmic Sobolev Inequalities, approximation algorithms, ...
[Ambrosio-Gigli-S., Agueh, Brenier, Carrillo, Carlen, McCann, Gangbo, Giacomelli, Otto, Villani, Westdickenberg, ...]

Examples

Fokker-Planck equation:

$$\boxed{\partial_t u - \Delta u - \mathrm{div}(u\,\mathrm{D}V) = 0}$$

Is generated by the Relative Entropy functional w.r.t. $\gamma := \mathrm{e}^{-V}\,\mathrm{d}x$

$$\boldsymbol{\Phi}(u) = \int_{\mathbb{R}^d} u(\log u + V)\,\mathrm{d}x = \int_{\mathbb{R}^d} u(\log(u/\mathrm{e}^{-V}))\,\mathrm{d}x.$$

$$\frac{\boldsymbol{\delta\Phi}}{\boldsymbol{\delta u}} = \log u + 1 + V, \quad \mathbf{w} = -\mathbf{u}\mathrm{D}\Big(\frac{\boldsymbol{\delta\Phi}}{\boldsymbol{\delta u}}\Big) = -\mathbf{u}\Big(\frac{\mathrm{D}u}{u} + \mathrm{D}V\Big) = -\mathrm{D}u - u\,\mathrm{D}V.$$

Nonlinear diffusion:

$$\boxed{\partial_t u - \mathrm{div}\left(\mathrm{D}u^\beta\right) = 0}$$

Is generated by $\quad \boldsymbol{\Phi}(u) = \dfrac{1}{\beta - 1}\displaystyle\int_{\mathbb{R}^d} u^\beta\,\mathrm{d}x$

Thin film equation:

$$\boxed{\partial_t u + \mathrm{div}\left(u\mathrm{D}\Delta u\right) = 0}$$

Is generated by $\quad \boldsymbol{\Phi}(u) = \dfrac{1}{2}\displaystyle\int_{\mathbb{R}^d} |\mathrm{D}u|^2\,\mathrm{d}x$

DLSS/quantum drift-diffusion equation:

$$\boxed{\partial_t u + \mathrm{div}\left(u\mathrm{D}\frac{\Delta\sqrt{u}}{\sqrt{u}}\right) = 0}$$

Is generated by $\quad \boldsymbol{\Phi}(u) = \dfrac{1}{4}\displaystyle\int_{\mathbb{R}^d} \frac{|\mathrm{D}u|^2}{u}\,\mathrm{d}x = \int_{\mathbb{R}^d} |\mathrm{D}\sqrt{u}|^2\,\mathrm{d}x$

An formal motivation for the gradient flow structure

$$\partial_t u + \operatorname{div} \mathbf{w} = 0, \quad \mathbf{w} = \mathbf{u}\mathbf{v} = \boxed{-\mathbf{u}\mathbf{D}\psi} \quad \psi = \frac{\delta\boldsymbol{\Phi}}{\delta\mathbf{u}}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{\Phi}(\mathbf{u_t}) = \int_{\mathbb{R}^d} \partial_t u \frac{\delta\boldsymbol{\Phi}}{\delta\mathbf{u}} \,\mathrm{d}x = -\int_{\mathbb{R}^d} (\operatorname{div}\mathbf{w})\psi \,\mathrm{d}x = \int_{\mathbb{R}^d} \mathbf{w}\cdot\mathbf{D}\psi \,\mathrm{d}x$$
$$= \int_{\mathbb{R}^d} \mathbf{v}\cdot\mathbf{D}\psi\,\mathbf{u}\,\mathrm{d}x \geq -\Big(\int_{\mathbb{R}^d} \big|\mathbf{D}\psi\big|^2\mathbf{u}\,\mathrm{d}x\Big)^{1/2}\Big(\int_{\mathbb{R}^d} \big|\mathbf{v}\big|^2\mathbf{u}\,\mathrm{d}x\Big)^{1/2}$$

Ansatz: interpret

$$\Big(\int_{\mathbb{R}^d} \big|\mathbf{v}\big|^2\mathbf{u}\,\mathrm{d}x\Big)^{1/2}$$

as the “velocity” of the moving family u. If we want to decrease $\boldsymbol{\Phi}$ as fast as possible, we have to choose

$$\boxed{\mathbf{v} = -\mathbf{D}\psi}$$

Wasserstein distance: the Brenier dynamical approach

We interpret u as the density of a (probability) measure $\rho = u\,\mathrm{d}x$ and we consider a moving family $\rho_t = u_t\,\mathrm{d}x$ of probability measures, $t \in [0, T]$, satisfying the **continuity equation**

$$\partial_t \rho + \mathrm{div}\,(\rho\,\mathbf{v}) = 0; \quad \mathbf{v} \text{ is the velocity vector associated to } \rho$$

The **scalar velocity** at time t is given by

$$\mathcal{V}_{\mathbf{t}}[\rho_t] := \|\mathbf{v_t}\|_{L^2(\rho_t;\mathbb{R}^d)} = \left(\int_{\mathbb{R}^d} |\mathbf{v_t}(\mathbf{x})|^2\,\mathrm{d}\boldsymbol{\rho}_{\mathbf{t}}(\mathbf{x})\right)^{1/2}$$

The length of the curve ρ between t_0 and t_1

$$\mathcal{L}_{t_0}^{t_1}[\rho] := \int_{t_0}^{t_1} \mathcal{V}_t[\rho]\,\mathrm{d}t = \int_{t_0}^{t_1}\left(\int_{\mathbb{R}^d} |\mathbf{v_t}(\mathbf{x})|^2\,\mathrm{d}\boldsymbol{\rho}_{\mathbf{t}}(\mathbf{x})\right)^{1/2}\mathrm{d}t$$

Wasserstein distance W_2 between ρ_0 and ρ_1:

$$\mathrm{W}_2(\rho_0,\rho_1) := \min\left\{\mathcal{L}_0^T[\rho] : \rho|_{t=0} = \rho_0,\ \rho|_{t=T} = \rho_1\right\}.$$

"Useful" properties

If $\rho_i \in \mathscr{P}_2(\mathbb{R}^d)$, i.e.

$$\int_{\mathbb{R}^d} |x|^2 \, \mathrm{d}\rho_i(x) < +\infty$$

then $\mathrm{W}_2(\rho_0, \rho_1) < +\infty$. The "right" function space:

$$\mathscr{P}_2(\mathbb{R}^d) := \Big\{ \text{Borel probability measures } \rho \text{ with finite quadratic moment} \Big\}$$

- $\big(\mathscr{P}_2(\mathbb{R}^d), \mathrm{W}_2\big)$ is a **complete and separable** metric space.
- Convergence with respect to W_2 $\Leftrightarrow$ **Weak** convergence + convergence of the **quadratic moments.**
- **Lower semicontinuity** of $(\rho_0, \rho_1) \mapsto \mathrm{W}_2(\rho_0, \rho_1)$ with respect to weak/distributional convergence
- **Convexity** (but linear segments are not geodesics!)
- Existence of (constant speed, minimizing) **geodesics** connecting arbitrary measures $\rho_0, \rho_1 \in \mathscr{P}_2(\mathbb{R}^d)$: they are curves $\rho : s \in [0,1] \mapsto \rho_s$ s.t.

$$\mathrm{W}_2(\rho_0, \rho_1) = \mathfrak{L}_0^1[\rho], \quad \mathrm{W}_2(\rho_{s_0}, \rho_{s_1}) = |s_1 - s_0| \, \mathrm{W}_2(\rho_0, \rho_1), \quad 0 \le s_0 \le s_1 \le 1.$$

More general mobilities

In many applications one is interested in more general mobility functions to deal with equations of the type

$$\partial_t u - \operatorname{div}\left(\mathfrak{m}(\mathbf{u})\mathrm{D}\frac{\boldsymbol{\delta\Phi}}{\boldsymbol{\delta}\mathbf{u}}\right) = 0$$

which can be splitted as before

$$\begin{cases} \partial_t u + \operatorname{div}\mathbf{w} = 0 & \textbf{(Continuity equation)} \\ \mathbf{w} = \mathfrak{m}(\mathbf{u})\,\mathbf{v} = -\mathfrak{m}(\mathbf{u})\,\mathrm{D}\psi & \textbf{(Flux structure)} \\ \psi = \dfrac{\boldsymbol{\delta\Phi}}{\boldsymbol{\delta}\mathbf{u}} & \textbf{(Nonlinear variational condition)} \\ u(0,\cdot) = u_0 & u_0 \in L^1(\mathbb{R}^d),\ u_0 \geq 0. \end{cases}$$

$\mathfrak{m} : [0,+\infty) \to [0,+\infty)$ is a given **mobility function** associated to the equation.

Examples

Heat equation:

$$\partial_t u = \Delta u = \mathrm{div}\left(\mathbf{u}^{\boldsymbol{\alpha}} \mathbf{D}\boldsymbol{\phi}'_{\boldsymbol{\alpha}}(\mathbf{u})\right), \quad \boldsymbol{\phi}_{\boldsymbol{\alpha}}(u) = \frac{1}{(2-\alpha)(1-\alpha)} u^{2-\alpha}.$$

In this case $\mathfrak{m}(\mathbf{u}) = \mathbf{u}^{\boldsymbol{\alpha}}$, $\boldsymbol{\Phi}(u) := \int_{\mathbb{R}^d} \boldsymbol{\phi}_{\boldsymbol{\alpha}}(u)\,\mathrm{d}x = \frac{1}{(2-\alpha)(1-\alpha)} \int_{\mathbb{R}^d} u^{2-\alpha}\,\mathrm{d}x.$

Thin film (typically $\mathfrak{m}(\mathbf{u}) = \mathbf{u}^{\boldsymbol{\alpha}}$):

$$\partial_t u + \mathrm{div}\left(\mathfrak{m}(\mathbf{u})\,\mathbf{D}\boldsymbol{\Delta}\mathbf{u}\right) = 0, \quad \boldsymbol{\Phi}(u) := \frac{1}{2}\int_{\mathbb{R}^d} |\mathrm{D}u|^2\,\mathrm{d}x$$

Cahn-Hilliard:

$$\partial_t u + \mathrm{div}\left(\mathbf{u}(\mathbf{1}-\mathbf{u})\,\mathbf{D}\left(\boldsymbol{\Delta}\mathbf{u} - \mathbf{W}'(\mathbf{u})\right)\right) = 0 \quad \boldsymbol{\Phi}(u) := \int_{\mathbb{R}^d}\left(\frac{1}{2}|\mathrm{D}u|^2 + W(u)\right)\mathrm{d}x.$$

Chemotaxis with overcrowding prevention [Hillen-Painter '01]:

$$\partial_t u = \mathrm{div}\left(\mathrm{D}u + \mathfrak{m}(\mathbf{u})\mathrm{D}(W * u)\right) = \mathrm{div}\left(\mathfrak{m}(\mathbf{u})\mathbf{D}\left(\mathbf{F}'(\mathbf{u}) + \mathbf{W} * \mathbf{u}\right)\right)$$

$$\mathfrak{m}(\mathbf{u}) = \mathbf{u}(\mathbf{1}-\mathbf{u}), \quad \mathbf{F}(\mathbf{u}) = \mathbf{u}\log\mathbf{u} + (\mathbf{1}-\mathbf{u})\log(\mathbf{1}-\mathbf{u}),$$

$$\boldsymbol{\Phi}(u) = \int_{\mathbb{R}^d} \mathbf{F}(u)\,\mathrm{d}x + \iint_{\mathbb{R}^d\times\mathbb{R}^d} \mathbf{W}(x-y)u(x)u(y)\,\mathrm{d}x\mathrm{d}y$$

The modified gradient flow structure

$$\partial_t u + \operatorname{div} \mathbf{w} = 0, \quad \mathbf{w} = \mathfrak{m}(\mathbf{u})\mathbf{v} \boxed{= -\mathfrak{m}(\mathbf{u})\mathbf{D}\psi} \quad \psi = \frac{\delta \mathbf{\Phi}}{\delta \mathbf{u}}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{\Phi}(\mathbf{u_t}) = \int_{\mathbb{R}^d} \partial_t u \frac{\delta \mathbf{\Phi}}{\delta \mathbf{u}}\,\mathrm{d}x = -\int_{\mathbb{R}^d} (\operatorname{div}\mathbf{w})\psi\,\mathrm{d}x = \int_{\mathbb{R}^d} \mathbf{w}\cdot\mathbf{D}\psi\,\mathrm{d}x$$
$$= \int_{\mathbb{R}^d} \mathbf{v}\cdot\mathbf{D}\psi\,\mathfrak{m}(\mathbf{u})\,\mathrm{d}x \geq -\Big(\int_{\mathbb{R}^d} \big|\mathbf{D}\psi\big|^2\,\mathfrak{m}(\mathbf{u})\,\mathrm{d}x\Big)^{1/2}\Big(\int_{\mathbb{R}^d} \big|\mathbf{v}\big|^2\,\mathfrak{m}(\mathbf{u})\,\mathrm{d}x\Big)^{1/2}$$

As before, we can consider

$$\Big(\int_{\mathbb{R}^d} \big|\mathbf{v}\big|^2\,\mathfrak{m}(\mathbf{u})\,\mathrm{d}x\Big)^{1/2}$$

the “velocity” of the moving family u.
If we want to decrease $\mathbf{\Phi}$ as fast as possible, we have to choose

$$\boxed{\mathbf{v} = -\mathbf{D}\psi}$$

Weighted transport distances: the dynamical approach

We interpret u as the density of a (probability) measure $\rho = u\,\mathrm{d}x$ and we consider a moving family $\rho_t = u_t\,\mathrm{d}x$, $t \in [0, T]$, of nonnegative Borel measures, satisfying the **continuity equation**

$$\boxed{\partial_t \rho + \mathrm{div}\,\big(\mathfrak{m}(\mathbf{u})\,\mathbf{v}\big) = 0; \quad u \text{ is the Lebesgue density of } \rho}$$

The **scalar velocity** at time t is given by

$$\mathcal{V}_{\mathbf{t}}[\rho_t] := \|\mathbf{v_t}\|_{L^2(\mathfrak{m}(u_t);\mathbb{R}^d)} = \Big(\int_{\mathbb{R}^d} |\mathbf{v_t(x)}|^2\, \mathfrak{m}(\mathbf{u_t})\,\mathrm{d}x\Big)^{1/2}, \quad \rho_t = u_t \mathscr{L}^d,$$

The length of the curve ρ between t_0 and t_1

$$\mathfrak{L}_{t_0}^{t_1}[\rho] := \int_{t_0}^{t_1} \mathcal{V}_t[\rho]\,\mathrm{d}t = \int_{t_0}^{t_1} \Big(\int_{\mathbb{R}^d} |\mathbf{v_t(x)}|^2\, \mathfrak{m}(\mathbf{u_t})\,\mathrm{d}x\Big)^{1/2}\mathrm{d}t$$

Weighted transport distance $\mathrm{W}_{2,\mathfrak{m}}$ between ρ_0 and ρ_1:

$$\mathrm{W}_{2,\mathfrak{m}}(\rho_0, \rho_1) := \min\Big\{\mathfrak{L}_0^1[\rho] : \rho|_{t=0} = \rho_0,\ \rho|_{t=1} = \rho_1\Big\}.$$

Main problems

- How can we **make rigorous** this approach and when it is well posed, so that it defines a distance?
- Does the distance enjoys **similar/different properties** to the **Wasserstein** one?
- Could it be useful to study **evolution equations** and to get new geometric insights?
- Are there interesting **convexity properties** of the integral functionals and related **functional inequalities** ?

The variational problem

Problem

Given nonnegative densities $u_0, u_1 \in L^1_{\mathrm{loc}}(\mathbb{R}^d)$ find a minimizer of the action functional

$$\int_0^1 \int_{\mathbb{R}^d} |\mathbf{v_t}|^2 \mathfrak{m}(\mathbf{u_t}) \, \mathrm{d}x \, \mathrm{d}t \quad s.t. \quad \partial_t u + \mathrm{div}(\mathfrak{m}(\mathbf{u_t}) \mathbf{v_t}) = 0, \quad u\big|_{t=0,1} = u_{0,1}.$$

Direct method of the calculus of variations: fix the densities u_0, u_1 and take a minimizing sequence $(u^n_t, \mathbf{w}^n_t, \mathbf{v^n_t})$ with $\mathbf{w}^n_t = \mathfrak{m}(\mathbf{u^n_t})\mathbf{v^n_t}$, such that

$$\partial_t u^n_t + \mathrm{div}(\mathfrak{m}(\mathbf{u^n_t})\mathbf{v^n_t}) = 0, \quad u^n\big|_{t=0,1} = u_{0,1}, \quad \int_0^1 \int_{\mathbb{R}^d} |\mathbf{v^n_t}|^2 \mathfrak{m}(\mathbf{u^n_t}) \, \mathrm{d}x \, \mathrm{d}t \to \inf$$

Problem: sublevels of the minimizing functional are only **weakly*** relatively compact: we get **weak*** convergence of a suitable subsequence but the equation $\partial_t u_t + \mathrm{div}(\mathfrak{m}(\mathbf{u})\mathbf{v}) = 0$ is nonlinear in the couple $(u, \mathbf{v})$.

Basic idea: write everything in terms of $(u, \mathbf{w})$! Since $\mathbf{w} = \mathfrak{m}(\mathbf{u})\mathbf{v}$ we minimize

$$\mathcal{A}(u, \mathbf{w}) := \int_0^1 \int_{\mathbb{R}^d} A(u_t, \mathbf{w}_t) \, \mathrm{d}x \, \mathrm{d}t \quad \text{s.t.} \quad \partial_t u + \mathrm{div}\, \mathbf{w}_t = 0, \quad u\big|_{t=0,1} = u_{0,1}.$$

where $$\boxed{A(u, \mathbf{w}) := \frac{|\mathbf{w}|^2}{\mathfrak{m}(\mathbf{u})}}$$

Convexity (and l.s.c.) of the action requires a concave mobility

Lemma

The function

$$A : (u, \mathbf{w}) \in (0, +\infty) \times \mathbb{R}^d \to \frac{|\mathbf{w}|^2}{\mathfrak{m}(\mathbf{u})} \in [0, +\infty]$$

is **convex** *iff* $\mathfrak{m} : [0, +\infty) \to [0, \infty)$ *is* **concave**.

Two cases:

A) $\mathfrak{m} : [0, +\infty) \to [0, +\infty)$ is concave and nondecreasing.
Model example: $\mathfrak{m}(\mathbf{u}) = \mathbf{u}^{\alpha}$, $0 \le \alpha \le 1$. In this case $A(\lambda u, \lambda \mathbf{w})$ is superlinear as $\lambda \uparrow +\infty$, except when $\mathbf{w} = 0$.

B) $\mathfrak{m} : [0, M] \to [0, +\infty)$ is concave with $\mathfrak{m}(0) = \mathfrak{m}(M) = 0$.
Model example: $\mathfrak{m}(\mathbf{u}) = \mathbf{u}(\mathbf{M} - \mathbf{u})$. In this case $A(u, \mathbf{w}) = +\infty$ if $u > M$ and all the densities u are uniformly bounded.

A rigorous definition through convex functional of measures

To get weak* lower semicontinuity of $\mathcal{A}$, we extend it to

> *couples* $(\boldsymbol{\rho}, \boldsymbol{\nu})$ *where* $\boldsymbol{\rho} \in \mathcal{M}_{\rm loc}(\mathbb{R}^d)$ *is a* **nonnegative Radon measure** *and* $\boldsymbol{\nu} \in \mathcal{M}_{\rm loc}(\mathbb{R}^d; \mathbb{R}^d)$ *is a* **Radon vector measure.**

Moreover, *the function A is no more* 1−*homogeneous* in the couple $(\rho, \boldsymbol{\nu})$, so that **the definition of $\mathcal{A}$ also depends from a reference measure $\boldsymbol{\gamma} \in \mathcal{M}_{\rm loc}(\mathbb{R}^d)$** (usually the Lebesgue measure, but not necessarily).

Definition (The case of a sublinear mobility)

If $\rho \in \mathcal{M}_{\rm loc}(\mathbb{R}^d), \boldsymbol{\nu} \in \mathcal{M}_{\rm loc}(\mathbb{R}^d; \mathbb{R}^d)$ we set

$$\mathcal{A}(\rho, \boldsymbol{\nu} | \gamma) := \int_{\mathbb{R}^d} A\Big(\frac{\mathrm{d}\rho}{\mathrm{d}\gamma}, \frac{\mathrm{d}\boldsymbol{\nu}}{\mathrm{d}\gamma}\Big)\,\mathrm{d}\gamma$$

Given ρ_0, ρ_1 we have

$$\mathrm{W}^2_{2,\mathfrak{m},\gamma}(\rho_0, \rho_1) := \inf \Big\{ \int_0^1 \mathcal{A}(\rho_t, \boldsymbol{\nu}_t)\,\mathrm{d}t \quad \text{s.t.} \quad \partial_t \rho + \operatorname{div} \boldsymbol{\nu} = 0, \quad \rho_{|_{t=0,1}} = \rho_{0,1} \Big\}$$

We call $\mathcal{M}_{2,\mathfrak{m},\gamma}[\sigma]$ the collection of all measures at finite distance from σ.

A few remarks

Limiting cases:

$\mathbf{\mathfrak{m}(r) = r}$ $\leftrightarrow$ Wasserstein distance, $\mathrm{W}_{2,\mathfrak{m},\gamma} = \mathrm{W}_2$.
The definition depends only on $\mathrm{supp}(\gamma)$, the total mass of ρ_t is preserved along the continuity equation.

$\mathbf{\mathfrak{m}(r) \equiv 1}$ $\leftrightarrow$ Homogeneous $W^{-1,2}(\mathbb{R}^d)$ distance.

Homogeneous mobilities $\mathbf{\mathfrak{m}(r) = r^{\alpha},\ 0 < \alpha < 1}$: provides a family of distances "interpolating" between the L^2-Wasserstein and the $W^{-1,2}$-Sobolev one.

The role of γ.

1. Typically $\gamma = \mathscr{L}^{\mathbf{d}}$ (omitted in $\mathrm{W}_{2,\mathfrak{m}}$).
2. $\gamma = \mathscr{L}^{\mathbf{d}}|_{\Omega}$, Ω **open, bounded, and convex subset of** $\mathbb{R}^d$: equations in bounded domains with Neumann boundary condition.
3. $\gamma := \mathbf{e}^{-\mathbf{V}}\mathscr{L}^{\mathbf{d}}$ is a log-concave measure: Fokker-Planck equations (Beckner/convex Sobolev inequalities)
4. $\gamma := \mathscr{H}^{\mathbf{k}}|_{\mathbf{M}}$, M is k-dimensional Riemannian manifold embedded in $\mathbb{R}^d$: evolutions in Riemannian manifolds.

Simple topological properties of $\mathrm{W}_{2,\mathfrak{m},\gamma}$

Completeness: $\mathrm{W}_{2,\mathfrak{m},\gamma}$ is a (pseudo-)distance and $\mathcal{M}_{2,\mathfrak{m},\gamma}[\sigma]$ is a complete metric space.

Weak* topology: $\mathrm{W}_{2,\mathfrak{m},\gamma}$ is always stronger than the (local) weak*/distributional topology (convergence of the integral of $C^0_{\mathrm{c}}(\mathbb{R}^d)$ test functions)

Weak* relative compactness: $\mathrm{W}_{2,\mathfrak{m},\gamma}$-bounded sets are weakly* relatively compact.

Mass conservation and moment bounds

Let us consider for simplicity the case of

homogeneous and sublinear mobilities $\mathfrak{m}(\mathbf{u}) = \mathbf{u}^{\alpha}$, $0 < \alpha < 1$

Theorem

If

$$\int_{\mathbb{R}^d} (1+|x|)^{-r} \,\mathrm{d}\gamma(x) < +\infty \quad \textit{for some } r \leq r_\alpha := \frac{2}{1-\alpha},$$

then two elements ρ_0, ρ_1 at finite W-distance **have the same total mass.**

Moreover, if $r < r_\alpha$ then $\mathrm{W}_{2,\mathfrak{m},\gamma}$ is stronger than the Wasserstein distance W_p for $p = \frac{2}{2-\alpha}(1 - r/r_\alpha)$.

Particular cases:

If $\gamma(\mathbb{R}^d) < +\infty$ then mass is always preserved and $\mathrm{W}_{2,\mathfrak{m},\gamma}$ is stronger than W_1.

If γ has finite p-moment, then $\mathrm{W}_{2,\alpha,\gamma}$ is always stronger then W_p. In particular, if γ is a log-concave probability measure, then all the measures in $\mathcal{M}_{2,\mathfrak{m},\gamma}[\gamma]$ have finite moments of arbitrary orders.

If $\gamma = \mathscr{L}^{\mathbf{d}}$ then **mass is preserved iff** $\boxed{\alpha > 1 - 2/d}$.

A general semicontinuity result

Theorem (Dolbeault-Nazaret-S.)

Suppose that $\gamma^{\mathbf{n}} \rightharpoonup \gamma$, $\rho_i^n \rightharpoonup \rho_i$ *in* $\mathcal{M}_{\rm loc}(\mathbb{R}^d)$ *and* $\mathfrak{m}^{\mathbf{n}} \downarrow \mathfrak{m}$ *pointwise in* $[0, +\infty)$*. Then*

$$\liminf_{n \to +\infty} \mathrm{W}_{2,\mathfrak{m}^n,\gamma^n}(\rho_0^n, \rho_1^n) \geq \mathrm{W}_{2,\mathfrak{m},\gamma}(\rho_0, \rho_1).$$

Convexity, subadditivity, and rescaling

Theorem

The map $(\rho,\sigma)\mapsto \mathrm{W}^2_{2,\mathfrak{m},\gamma}(\rho,\sigma)$ *is convex, i.e. if*

$$\rho_\theta := (1-\theta)\rho_0+\theta\rho_1, \quad \sigma_\theta = (1-\theta)\sigma_0+\theta\sigma_1$$

then

$$\mathrm{W}^2_{2,\mathfrak{m},\gamma}(\rho_\theta,\sigma_\theta) \le (1-\theta)\mathrm{W}^2_{2,\mathfrak{m},\gamma}(\rho_0,\sigma_0)+\theta \mathrm{W}^2_{2,\mathfrak{m},\gamma}(\rho_1,\sigma_1)$$

Theorem

The map $(\rho,\sigma)\mapsto \mathrm{W}_{2,\mathfrak{m},\gamma}(\rho,\sigma)$ *is subadditive, i.e.*

$$\mathrm{W}_{2,\mathfrak{m},\gamma}(\rho_0+\rho_1,\sigma_0+\sigma_1) \le \mathrm{W}_{2,\mathfrak{m},\gamma}(\rho_0,\sigma_0)+\mathrm{W}_{2,\mathfrak{m},\gamma}(\rho_1,\sigma_1).$$

If $\mathbf{\mathfrak{m}(u) = u^{\alpha},\ 0<\alpha\le 1}$ then $\mathrm{W}^2_{2,\mathfrak{m},\gamma}$ is $(2-\alpha)$-homogeneous, i.e.

$$\mathrm{W}_{2,\mathfrak{m},\gamma}(\lambda\rho,\lambda\sigma) = \lambda^{2-\alpha}\mathrm{W}_{2,\mathfrak{m},\gamma}(\rho,\sigma) \quad \text{for every } \lambda>0.$$

Convolution

Theorem

If $\kappa \in C_c^\infty(\mathbb{R}^d)$ is a smooth, nonnegative, convolution kernel with $\int_{\mathbb{R}^d} \kappa(x)\,\mathrm{d}x = 1$, then

$$\mathrm{W}_{2,\mathrm{m},\gamma*\kappa}(\rho * \kappa, \sigma * \kappa) \leq \mathrm{W}_{2,\mathrm{m},\gamma}(\rho, \sigma)$$

Setting $\kappa_\varepsilon := \varepsilon^{-d}\kappa(\cdot/\varepsilon)$ we have

$$\lim_{\varepsilon \downarrow 0} \mathrm{W}_{2,\mathrm{m},\gamma*\kappa_\varepsilon}(\rho * \kappa_\varepsilon, \sigma * \kappa_\varepsilon) = \mathrm{W}_{2,\mathrm{m},\gamma}(\rho, \sigma)$$

Absolutely continuous curves and their metric velocity

A curve $(\rho_t)_{t\in[0,T]}$ is absolutely continuous in $\mathcal{M}_{2,\mathfrak{m},\gamma}$ if

$$\mathrm{W}_{2,\mathfrak{m},\gamma}(\rho_s,\rho_t) \le \int_s^t \delta(r),\mathrm{d}r \quad 0\le s\le t\le T, \quad \text{for some } \delta\in L^1(0,T).$$

Theorem

$(\rho_t)_{t\in[0,T]}$ is absolutely continuous iff there exists a Borel family of measures $(\boldsymbol{\nu}_t)_{t\in(0,T)}$ such that

$$\partial_t\rho_t + \operatorname{div}\boldsymbol{\nu}_t = 0, \quad \int_0^T A(\rho_t,\boldsymbol{\nu}_t|\gamma)^{1/2}\,\mathrm{d}t < +\infty.$$

In this case $\quad |\rho_t'| := \lim_{h\to 0}\dfrac{\mathrm{W}_{2,\mathfrak{m},\gamma}(\rho_t,\rho_{t+h})}{|h|} \le A(\rho_t,\boldsymbol{\nu}_t|\gamma)^{1/2}$ *a.e. in* $(0,T)$.

There exists a unique family $\tilde{\boldsymbol{\nu}}_t$ such that equality holds in the previous formula; it is characterized by

$$\tilde{\boldsymbol{\nu}}_t = \mathfrak{m}(\mathbf{u_t})\mathbf{v_t}\gamma, \quad u_t = \frac{\mathrm{d}\rho_t}{\mathrm{d}\gamma}, \quad \int_{\mathbb{R}^d}|\mathbf{v_t}|^2\,\mathfrak{m}(\mathbf{u_t})\,\mathrm{d}\gamma = A(\rho_t,\tilde{\boldsymbol{\nu}}_t) = |\rho_t'|^2$$

and $\quad \mathbf{v_t} \in \overline{\{\nabla\zeta : \zeta\in C_c^\infty(\mathbb{R}^d)\}}^{L^2(\mathbb{R}^d,\mathfrak{m}(u_t)\gamma)}$

Comparison with the Sobolev and the Wasserstein distance

Theorem

Suppose that $\gamma \in \mathcal{M}_{\rm loc}(\mathbb{R}^d)$ *is a bounded perturbation of a log-concave measure (e.g.* $\boldsymbol{\gamma = f e^{-V} dx}$ *where* f *is bounded and nonnegative,* V *is convex).*

1) *If* $\rho_i = u_i\gamma \in \mathcal{P}(\mathbb{R}^d)$ *have bounded densities* $u_i \le U < +\infty$ *then*

$$\mathrm{W}_{2,\mathrm{m},\gamma}(\rho_0,\rho_1) \le C_U \mathrm{W}_2(\rho_0,\rho_1).$$

2) *If* $\mathbf{m(r) = r^{\alpha}}$ *with* $\boldsymbol{\alpha > 1 - 2/d}$ *then*

$$\mathrm{W}_{2,\mathrm{m},\gamma}(\rho_0,\rho_1) \le C\mathrm{W}_2(\rho_0,\rho_1) \quad \textit{for every } \rho_0,\rho_1 \in \mathcal{P}(\mathbb{R}^d).$$

In particular, if $\boldsymbol{\gamma = \mathscr{L}^d|_\Omega}$*,* Ω *bounded and convex, then the distance* $\mathrm{W}_{2,\mathrm{m}}$ *between any couple of probability measures is finite. Its topology coincides with the weak** *convergence one.*

Theorem

Suppose that $\rho_i = u_i\gamma$ *with* $u_i \ge U > 0$*. Then*

$$\mathrm{W}_{2,\mathrm{m},\gamma}(\rho_0,\rho_1) \le C_U \|\rho_0 - \rho_1\|_{W^{-1,2}_{\gamma}(\mathbb{R}^d)}.$$

Summary

$\mathrm{W}_{2,\mathfrak{m},\gamma}$ enjoys similar properties to W_2 concerning

completeness, lower semicontinuity, convexity, convolution, rescaling, geodesic and absolutely continuous curves.

On the other hand

- The definition is not intrinsic but depends on a reference measure γ.
- Mass is not always preserved.
- Except for the case $\mathbf{m(r) = r^{\alpha}}$, $\alpha > 1 - 2/d$, in bounded domains, is not easy to characterize the set of measures at finite distance from, e.g., γ.
- Duality (à la Kantorovich) and optimal transport interpretation are missing.
- A precise characterization of geodesics ("as optimal transport") is also missing.

Energy functional and nonlinear diffusion

We consider now the case when $\gamma = \mathscr{L}^{d}|_{\Omega}$, where Ω is a bounded convex set, and we study behaviour of the integral functional

$$\mathscr{F}(\rho) := \int_{\Omega} F(u(x))\,\mathrm{d}x, \quad F : [0,+\infty) \to [0,+\infty).$$

We introduce the nonlinear function

$$P(u) := \int_0^u F''(z)\mathfrak{m}(z)\,\mathrm{d}z$$

and the associated "gradient flow"

$$\boxed{\partial_t u - \Delta P(u) = 0} \quad \text{in } \Omega \times (0,+\infty), \quad \partial_{\mathbf{n}} u = 0 \quad \text{on } \partial\Omega \times (0,+\infty).$$

Main examples:

Powers: $\mathfrak{m}(u) := u^{\alpha}$, $F(u) := \frac{1}{\beta-1}u^{\beta}$, $P(u) = \frac{\beta}{m}u^{m}$ where $\boxed{m = \alpha + \beta - 1}$

Heat equation, $P(u) = \beta\, u$: $\beta = 2 - \alpha$. In general

$$F''(r) = \frac{1}{\mathfrak{m}(r)}$$

Confining entropy and mobility: $\mathfrak{m}(u) = u(1-u)$,
$F(u) = u\log u + (1-u)\log(1-u)$, $P(u) = u$.

Displacement convexity in the Wasserstein case

A functional Φ is displacement convex in the space $\mathscr{P}_2(\mathbb{R}^d)$ if for every ρ_0, ρ_1 there exists a geodesic ρ_t connecting ρ_0 to ρ_1 such that

$$\Phi(\rho_t) \leq (1-t)\Phi(\rho_0) + t\Phi(\rho_1).$$

Theorem (McCann '97)

The energy functional $\mathscr{F}(\rho) := \int_\Omega F(u)\,\mathrm{d}x$, $\rho = u\mathrm{d}x$ *is displacement convex in* $\mathscr{P}_2(\mathbb{R}^d)$ *with respect to the Wasserstein distance if*

$$r \mapsto r^{-(1-1/d)}P(r) \quad \text{is nonnegative and non decreasing in } (0,+\infty).$$

In the power case $P(r) = r^m$ *is equivalent to* $\boxed{m \geq 1 - 1/d.}$

Theorem (Ambrosio-Gigli-S. '05, Daneri-S. '08, S.)

Φ *is displacement convex in* $\mathscr{P}_2(\mathbb{R}^d)$ *if and only if it generates a metric gradient flow: for every initial datum* $\rho_0 \in D(\Phi)$ *there exists a curve* ρ_t *solution of*

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\mathrm{W}_2^2(\rho_t, \sigma) \leq \Phi(\sigma) - \Phi(\rho_t). \tag{EVI}$$

In the case of $\mathscr{F}$ the curve $\rho_t = u_t\mathscr{L}^d$ solves the nonlinear diffusion equation

$$\partial_t u - \Delta P(u) = 0$$

Displacement convexity for weighted transport distances

McCann condition:

$$r \mapsto \frac{P(r)}{r^{1-1/d}} = \frac{P(r)}{\mathbf{m(r)}^{1-1/d}} \quad \text{is nonnegative and non decreasing in } (0,+\infty).$$

Theorem (Carrillo-Lisini-S.-Slepcev '09)

The functional $\mathscr{F}$ is displacement convex in $\mathcal{M}_{2,\mathrm{m}}(\Omega)$ with respect to the distance $\mathrm{W}_{2,\mathrm{m}}$ if

$$r \mapsto \frac{H(r)}{\mathbf{m(r)}^{\mathbf{1-1/d}}} \quad \textit{is nonnegative and non decreasing in } (0,+\infty),$$

where

$$H(r) := \int_0^r F''(z)\mathrm{m}(z)\mathrm{m}'(z)\,\mathrm{d}z. \quad H(r) = P(r) \quad \textit{if } \mathbf{m(r) = r}.$$

In the power case $P(r) = r^m$, the above condition is equivalent to

$$\boxed{m \geq 1 - \alpha/d} \ \textit{if } d > 1$$

.

Metric characterization of the gradient flow

Geodesic convexity is a consequence of the fact that the nonlinear diffusion equation is the metric gradient flow of $\mathscr{F}$ in $\mathcal{M}_{2,\mathfrak{m}}(\Omega)$.

Theorem

For every $\rho_0 \in D(\mathscr{F})$ there exists a unique solution ρ_t of the evolution variational inequality

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\mathrm{W}^2_{2,\mathfrak{m}}(\rho_t,\sigma) \le \mathscr{F}(\sigma) - \mathscr{F}(\rho_t) \quad \text{for every } \sigma \in \mathcal{M}_{2,\mathfrak{m}}(\Omega). \tag{EVI}$$

$\rho_t = u_t\mathscr{L}^d|_\Omega$ *is a (limit L^1-)solution of*

$$\partial_t u - \Delta P(u) = 0 \quad \text{in } \Omega \times (0,+\infty).$$

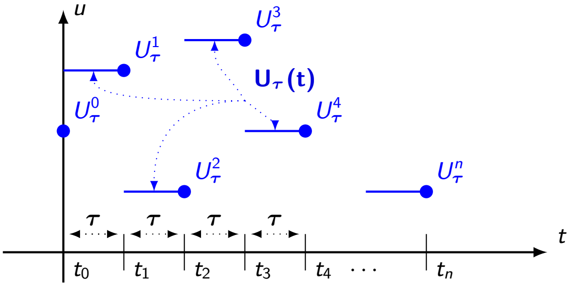
The map $\rho_0 \mapsto \rho_t$ does not expand the distance $\mathrm{W}_{2,\mathfrak{m}}$: if ρ_t, σ_t are two solutions then

$$\mathrm{W}_{2,\mathfrak{m}}(\rho_t,\sigma_t) \le \mathrm{W}_{2,\mathfrak{m}}(\rho_0,\sigma_0).$$

If $\mathbf{\mathfrak{m}(r) = r^\alpha}$ with $\boldsymbol{\alpha > 1 - 2/d}$ then the evolution semigroup can be extended to every Borel measure of finite mass in Ω.

Convergence of the JKO-De Giorgi's Minimizing movement scheme

- Choose a partition of $(0,+\infty)$ of **step size** $\tau > 0$



- Starting from $U_\tau^0 := \rho_0$ **find recursively** **minimizers** $U_\tau^n, \; n = 1, 2, \dots,$

$$\frac{U_\tau^n - U_\tau^{n-1}}{\tau} + \nabla\mathscr{F}(U_\tau^n) = 0 \quad \rightsquigarrow \quad \boxed{U_\tau^n \in \operatorname*{argmin}_V \frac{\mathrm{W}^2_{2,\mathrm{m}}(V, U_\tau^{n-1})}{2\tau} + \mathscr{F}(V)}$$

- $\mathbf{U}_\tau$ is the **piecewise constant** interpolant of $\{U_\tau^n\}_n$.
- We have the uniform Cauchy estimate [S.]

$$\sup_{[0,T]} W_{2,\mathrm{m}}(\mathbf{U}_\tau, \rho) \le C(T, \rho_0)\sqrt{\tau}$$

Log-Sobolev inequality and Wasserstein distance

Let $\gamma = e^{-V}\mathscr{L}^d \in \mathcal{P}(\mathbb{R}^d)$ be a **log-concave measure**, associated to a **uniformly convex** potential $V : \mathbb{R}^d \to \mathbb{R}^d$ satisfying

$$D^2 V(x) \geq \lambda I, \quad \lambda > 0.$$

γ satisfies the Logarithmic-Sobolev inequality

$$\boxed{\mathrm{Ent}(\rho|\gamma) = \int_{\mathbb{R}^d} u \log u \, d\gamma \leq \frac{2}{\lambda}\mathscr{I}(\rho|\gamma) = \frac{2}{\lambda}\int_{\mathbb{R}^d} \frac{|\nabla u|^2}{u}\, d\gamma}$$

[Otto-Villani '00] showed that this inequality can be interpreted as a sub-gradient bound for the relative entropy functional $\mathrm{Ent}(\cdot|\gamma)$, which satisfies the λ-convexity condition along a geodesic $(\rho_t)_{t\in[0,1]}$ in the Wasserstein space

$$\mathrm{Ent}(\rho_t|\gamma) \leq (1-t)\mathrm{Ent}(\rho_0|\gamma) + t\mathrm{Ent}(\rho_1|\gamma) - \frac{\lambda}{2} t(1-t) W_2^2(\rho_0, \rho_1)$$

It also related to Talagrand inequality

$$\boxed{\frac{\lambda}{2} W_2^2(\rho, \gamma) \leq \mathrm{Ent}(\rho|\gamma).}$$

They are modelled on the elementary inequalities

$$\frac{\lambda}{2}|x - x_{\min}|^2 \leq \psi(x) - \psi(x_{\min}) \leq \frac{2}{\lambda}|\nabla\psi(x)|^2$$

satisfied by a λ-convex function $\psi : \mathbb{R}^d \to \mathbb{R}$, $D^2\psi \geq \lambda I$.

Beckner/convex Sobolev inequalities

Theorem (Beckner '89)

Let $\rho = u\gamma \in \mathcal{P}(\mathbb{R}^d)$ *and let* $\beta \in (1,2]$. *If* $\mathscr{F}_\beta(\rho|\gamma) := \frac{1}{\beta-1}\int_{\mathbb{R}^d} u^\beta \,\mathrm{d}\gamma$ *we have*

$$\mathscr{F}_\beta(\rho|\gamma) - \mathscr{F}_\beta(\gamma|\gamma) \leq \frac{2}{\lambda}\mathscr{I}_\beta(\rho|\gamma), \quad \mathscr{I}_\beta(\rho|\gamma) := \frac{1}{\beta}\int_{\mathbb{R}^d} |\nabla u^{\beta/2}|^2 \,\mathrm{d}\gamma$$

These are particular cases of the **general convex Sobolev inequalities**. Let us suppose that $F : [0,+\infty) \to \mathbb{R}$ is a convex energy such that $1/F''$ is concave and let us set

$$\mathscr{I}(\rho|\gamma) := \int_{\mathbb{R}^d} \frac{1}{F''(u)} |\mathrm{D}u|^2 \,\mathrm{d}\gamma$$

Theorem (Bakry-Emery, Arnold-Markovich-Toscani-Unterreiter)

$$\mathscr{F}(\rho|\gamma) - \mathscr{F}(\gamma|\gamma) \leq \frac{2}{\lambda}\mathscr{I}(\rho|\gamma).$$

Convex Sobolev inequalities and weighted transport distances

If we introduce the *concave* mobility function $\mathfrak{m}(\mathbf{r}) := \dfrac{1}{F''(r)}$, the Convex Sobolev inequalities play the same role w.r.t. the weighted transport distances of the logarithmic Sobolev inequalities in the usual Wasserstein space.

Theorem (Dolbeault-Nazaret-S.)

The functional $\mathscr{F}(\cdot|\gamma)$ is λ-convex along any geodesic ρ_t of the weighted transport distance $\mathrm{W}_{2,\mathfrak{m},\gamma}$:

$$\mathscr{F}(\rho_t|\gamma) \leq (1-t)\mathscr{F}(\rho_0|\gamma) + t\mathscr{F}(\rho_1|\gamma) - \frac{\lambda}{2}t(1-t)\mathrm{W}^2_{2,\mathfrak{m},\gamma}(\rho_0,\rho_1)$$

It satisfies the weighted-Talagrand inequality

$$\frac{\lambda}{2}\mathrm{W}^2_{2,\mathfrak{m},\gamma}(\rho,\gamma) \leq \mathscr{F}(\rho|\gamma) - \mathscr{F}(\gamma|\gamma).$$

The solution $\rho_t = u_t\gamma$ of its gradient flow satisfies the Fokker-Planck equation

$$\partial_t u - \Delta u - \mathrm{div}(u\mathrm{D}V) = 0$$

and the metric EVI

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\mathrm{W}^2_{2,\mathfrak{m},\gamma}(\rho_t,\sigma) \leq \mathscr{F}(\sigma|\gamma) - \mathscr{F}(\rho_t|\gamma) - \frac{\lambda}{2}\mathrm{W}^2_{2,\mathfrak{m},\gamma}(\rho_t,\sigma). \qquad \text{(EVI)}$$

Open problems

- Find some further characterization of measures at finite $W_{2,\mathfrak{m},\gamma}$-distance.
- Develop a duality approach to the weighted distances and find a precise characterization of their geodesics. Curvature properties?
- Study the gradient flow of other integral functionals: potential and interaction energies do not behave well with respet to the weighted distances.
- What about non-concave mobilities?
- .....