Metric measure spaces with Riemannian Ricci curvature bounded from below
Lecture II

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

Dipartimento di Matematica, Università di Pavia

Outline

1. **Curves, upper gradient and slopes in metric spaces**

2. **Cheeger energy, Sobolev spaces $W^{1,2}(\boldsymbol{X}, \mathsf{d}, \mathfrak{m})$, nonlinear Laplacian and metric-heat flow.**

3. **Test plans, weak upper gradients**

4. **The Wasserstein gradient flow of the Entropy functional**

Absolutely continuous curves and upper gradients

$(\boldsymbol{X}, \mathsf{d})$ is a complete and separable metric space.

Absolutely continuous curves with finite p-energy $\mathrm{AC}^p([a,b];\boldsymbol{X})$: curves $\mathsf{x} : [a,b] \to \boldsymbol{X}$ such that

$$\mathsf{d}(\mathsf{x}(s),\mathsf{x}(r)) \le \int_r^s m(t)\,\mathrm{d}t \quad a \le r \le s \le b, \quad \text{for some } m \in L^p(a,b), \qquad (\star)$$

$\mathrm{Lip}([a,b];\boldsymbol{X}) = \mathrm{AC}^\infty([a,b];\boldsymbol{X})$.

Metric derivative:

$$\boxed{|\dot{\mathsf{x}}|(t) := \lim_{h\to 0} \frac{\mathsf{d}(\mathsf{x}(t+h),\mathsf{x}(t))}{|h|}}$$

$|\dot{\mathsf{x}}| \in L^p(a,b)$ and provides the minimal function such that $(\star)$ holds.

Geodesics: $\mathsf{x} : [0,1] \to \boldsymbol{X}$ with

$$\mathsf{d}(\mathsf{x}(s),\mathsf{x}(t)) = |t-s|\mathsf{d}(x(0),\mathsf{x}(1)), \quad |\dot{\mathsf{x}}|(t) \equiv \mathsf{d}(\mathsf{x}(0),\mathsf{x}(1)).$$

Integrals:

$$\int_{\mathsf{x}} f := \int_a^b f(\mathsf{x}(t))\,|\dot{\mathsf{x}}|(t)\,\mathrm{d}t, \qquad \int_{\partial\mathsf{x}} f = f(\mathsf{x}(b)) - f(\mathsf{x}(a))$$

Upper gradients

Let $f : \boldsymbol{X} \to \mathbb{R}$ and $\mathcal{C}$ be a collection of absolutely continuous curves (invariant by restrictions).

A Borel function $g : \boldsymbol{X} \to [0, +\infty]$ is an **upper gradient for f on $\mathcal{C}$** if

$$\boxed{\left|f(\mathsf{x}(b)) - f(\mathsf{x}(a))\right| \le \int_{\mathsf{x}} g = \int_a^b g(\mathsf{x}(t))|\dot{\mathsf{x}}|(t)\,\mathrm{d}t} \quad \forall \mathsf{x} \in \mathcal{C} \quad \text{defined in } [a,b]$$

Equivalently, whenever $\mathsf{x} \in \mathcal{C}$ with $g \circ \mathsf{x} \in L^1(a,b)$

$$f \circ \mathsf{x} \in \mathrm{AC}([a,b]), \quad \left|\frac{\mathrm{d}}{\mathrm{d}t}(f \circ \mathsf{x})\right| \le (g \circ \mathsf{x})\,|\dot{\mathsf{x}}| \quad \mathscr{L}^1\text{-a.e.}$$

When $\mathcal{C}$ contains all the absolutely continuous curves we just say that g is an upper gradient.

Gradient flows

Suppose that g is an upper gradient for $f : \boldsymbol{X} \to (-\infty, +\infty]$ on $\mathcal{C}$.

A curve $\mathsf{x} : [0,\infty) \to D(f)$ in $\mathcal{C}$ is a **metric gradient flow for f w.r.t. g** if $g \circ \mathsf{x} \in L^2(0,\infty)$ and

$$\boxed{-\frac{\mathrm{d}}{\mathrm{d}t} f(\mathsf{x}(t)) = g^2(\mathsf{x}(t)) = |\dot{\mathsf{x}}|^2(t)} \quad \text{a.e. in } (0,\infty)$$

Equivalent **dissipation inequality**:

$$f(\mathsf{x}(t)) + \int_0^t \Big(\frac{1}{2}|\dot{\mathsf{x}}|^2(r) + \frac{1}{2} g^2(\mathsf{x}(r))\Big)\,\mathrm{d}r \le f(\mathsf{x}(0)) \quad \text{for every } t > 0.$$

The definition is purely metric and we will apply it to various spaces, as $L^2(\boldsymbol{X}, \mathfrak{m})$ or $\mathscr{P}_2(\boldsymbol{X})$ endowed with the Wasserstein distance W_2.

In the case $f = f_t(x)$ is time-dependent

$$f_t(\mathsf{x}(t)) + \int_0^t \Big(\frac{1}{2}|\dot{\mathsf{x}}|^2(r) + \frac{1}{2} g_t^2(\mathsf{x}(r))\Big)\,\mathrm{d}r \le f_0(\mathsf{x}(0)) + \int_0^t \partial_r f_r(\mathsf{x}(r))\,\mathrm{d}r$$

for every $t > 0$.

Slopes

If $f : \boldsymbol{X} \to (-\infty, +\infty]$ and $f(x)$ is finite

$$|\mathrm{D}^{\pm} f|(x) := \limsup_{y \to x} \frac{\big(f(y) - f(x)\big)_{\pm}}{\mathsf{d}(x, y)}, \quad |\mathrm{D} f|(x) := \limsup_{y \to x} \frac{\big|f(y) - f(x)\big|}{\mathsf{d}(x, y)}$$

$$|\mathrm{D} f| = \max\Big(|\mathrm{D}^{-} f|, |\mathrm{D}^{+} f|\Big).$$

If f is **Lipschitz**, then $|\mathrm{D}^{\pm} f|$ are upper gradients.

f is geodesically K-convex if every $x_0, x_1 \in D(f)$ can be connected by a geodesic x along which

$$f(\mathsf{x}(t)) \leq (1-t) f(\mathsf{x}(0)) + t f(\mathsf{x}(1)) - \frac{K}{2} t(1-t) \mathsf{d}^2(\mathsf{x}(0), \mathsf{x}(1))$$

If f is geodesically K-convex then $|\mathrm{D}^{-} f|$ is an upper gradient.

Cheeger energy

Let $\mathfrak{m} \in \mathscr{P}(\boldsymbol{X})$. If $f \in \mathrm{Lip}_b(\boldsymbol{X})$ we consider the integral functional

$$f \mapsto \frac{1}{2}\int |\mathrm{D}f|^2 \, \mathrm{d}\mathfrak{m}$$

Cheeger energy: L^2-relaxation

$$\mathsf{Ch}(f) := \inf \Big\{ \liminf_{n\to\infty} \frac{1}{2}\int |\mathrm{D}f_n|^2 \, \mathrm{d}\mathfrak{m} : f_n \in \mathrm{Lip}_b(\boldsymbol{X}), \quad f_n \xrightarrow{L^2} f \Big\}$$

Take an **optimal sequence** $f_n \in \mathrm{Lip}_b(\boldsymbol{X})$ with

$$f_n \xrightarrow{L^2} f, \quad \frac{1}{2}\int |\mathrm{D}f_n|^2 \, \mathrm{d}\mathfrak{m} \longrightarrow \mathsf{Ch}(f)$$

Integral representation by the minimal relaxed slope

- $|\mathrm{D}f_n| \xrightarrow{L^2} |\mathrm{D}f|_w$, $\mathsf{Ch}(f) = \frac{1}{2}\int |\mathrm{D}f|_w^2 \, \mathrm{d}\mathfrak{m}$.
- If $h_n \in \mathrm{Lip}_b(\boldsymbol{X})$, $h_n \xrightarrow{L^2} f$ and $|\mathrm{D}h_n| \rightharpoonup G$ in L^2, then $G \geq |\mathrm{D}f|_w$.

In particular $|\mathrm{D}f|_w$ is unique.

Properties of the Cheeger energy and of the relaxed slope

Let $f, g \in D(\mathsf{Ch})$.

- $|\mathrm{D}f|_w = 0$ $\mathfrak{m}$-a.e. on $f^{-1}(N)$ if $\mathscr{L}^1(N) = 0$.
- **Locality**: $|\mathrm{D}f|_w = |\mathrm{D}g|_w$ $\mathfrak{m}$-a.e. on $\{f - g = c\}$ for all constant $c \in \mathbb{R}$.
- **Chain rule:** $|\mathrm{D}\phi(f)|_w = |\phi'(f)|\,|\mathrm{D}f|_w$ for any $\phi \in \mathrm{Lip}(\mathbb{R})$.
- $|\mathrm{D}(\alpha f + \beta g)|_w \le \alpha|\mathrm{D}f|_w + \beta|\mathrm{D}g|_w$ whenever $\alpha, \beta \ge 0$.
- **Leibnitz inequality:** $|\mathrm{D}(fg)|_w \le |f|\,|\mathrm{D}g|_w + |g|\,|\mathrm{D}f|_w$ if f, g are bounded.
- **Ch is a convex, 2-homogeneous, l.s.c. functional.**

Sobolev space $W^{1,2}(\boldsymbol{X}, \mathsf{d}, \mathfrak{m})$: $f \in L^2(\boldsymbol{X}, \mathfrak{m})$ with $\mathsf{Ch}(f) < \infty$ and norm

$$\|f\|^2_{W^{1,2}} := \int \Big(|f|^2 + |\mathrm{D}f|^2_w\Big)\,\mathrm{d}\mathfrak{m}.$$

!! Ch can be non-quadratic in general !!

Fisher information:

$$\mathsf{F}(f) := \int_{f>0} \frac{|\mathrm{D}f|^2_w}{f}\,\mathrm{d}\mathfrak{m} = 4\int |\mathrm{D}\sqrt{f}|^2_w\,\mathrm{d}\mathfrak{m} = 8\mathsf{Ch}(\sqrt{f}).$$

(Nonlinear) Laplacian

Given $f \in D(\mathsf{Ch})$ consider the (possibly empty) set $\partial\mathsf{Ch}(f) \subset L^2(\boldsymbol{X}, \mathfrak{m})$ defined by

$$\xi \in L^2(\boldsymbol{X}, \mathfrak{m}) : \quad \boxed{\int \xi(g - f)\,\mathrm{d}\mathfrak{m} \le \mathsf{Ch}(g) - \mathsf{Ch}(f)} \quad \forall\, g \in D(\mathsf{Ch}).$$

If $\partial\mathsf{Ch}(f)$ is non empty, it is **closed and convex:**
we denote by **-Δf its element of minimal L^2-norm.**

- **Rough integration by parts**

$$\Big|\int f\,\Delta g\,\mathrm{d}\mathfrak{m}\Big| \le \int |\mathrm{D}f|_w\,|\mathrm{D}g|_w\,\mathrm{d}\mathfrak{m}$$

- **Integral chain rule**: $\phi \in \mathrm{Lip}(\mathbb{R})$, $\phi' \ge 0$

$$-\int \Delta f\,\phi(f)\,\mathrm{d}\mathfrak{m} = \int |\mathrm{D}f|^2\phi'(f)\,\mathrm{d}\mathfrak{m}$$

- **Monotonicity,** $\phi' \ge 0$.

$$-\int (\Delta f - \Delta g)\phi(f - g)\,\mathrm{d}\mathfrak{m} \ge 0$$

- $\Delta(\lambda f) = \lambda\Delta(f)$, $\Delta(f + c) = \Delta f$.

Nonlinear heat flow

Generation results for gradient flows of convex l.s.c. functionals in Hilbert spaces [Brezis '70]:

For every $f \in L^2(\boldsymbol{X}, \mathfrak{m})$ there exists a unique locally lipschitz curve $f_t = \mathsf{P}_t f$ with

$$\boxed{\frac{\mathrm{d}}{\mathrm{d}t_+} f_t = \Delta f_t} \quad \text{for every } t > 0.$$

- $(\mathsf{P}_t)_{t\geq 0}$ is a **semigroup of contractions** in every $L^p(\boldsymbol{X}, \mathfrak{m})$: $\|\mathsf{P}_t f - \mathsf{P}_t g\|_{L^p} \leq \|f - g\|_{L^p}$.
- **Regularization effect:** $\|\Delta f_t\|_{L^2} \leq t^{-1}\|f\|_{L^2}$
- P_t is **order preserving:** $f \leq g \ \Rightarrow\ \mathsf{P}_t f \leq \mathsf{P}_t g$.
- P_t is **mass preserving** and $\mathsf{P}_t c \equiv c$.
- **Entropy dissipation:**

$$\frac{\mathrm{d}}{\mathrm{d}t} \int f_t \log f_t \, \mathrm{d}\mathfrak{m} = -\mathsf{F}(f_t).$$

Test plans and weak upper gradients.

Test plan: a dynamic plan (i.e. a probability measure $\boldsymbol{\pi} \in \mathscr{P}(\mathrm{C}([0,1];\boldsymbol{X}))$ on the path space) such that

- $\boldsymbol{\pi}$ is concentrated on $\mathrm{AC}([0,1];\boldsymbol{X})$
- $(\mathrm{e}_t)_\sharp \boldsymbol{\pi} \le C\mathfrak{m}$, i.e. $\boldsymbol{\pi}\Big[\mathsf{x} : \mathsf{x}(t) \in B\Big] \le C\mathfrak{m}(B)$.

$\Lambda \subset \mathrm{AC}([0,1];\boldsymbol{X})$ is **negligible** if $\boldsymbol{\pi}(\Lambda) = 0$ for every test plan $\boldsymbol{\pi}$.

Weak upper gradient for $f : \boldsymbol{X} \to \mathbb{R}$: a $\mathfrak{m}$-measurable function $G : \boldsymbol{X} \to [0,\infty]$ satisfying

$$\boxed{\Big|\int_{\partial \mathsf{x}} f\Big| \le \int_{\mathsf{x}} G < \infty} \quad \text{for a.e. } \mathsf{x} \in \mathrm{AC}([0,1];\boldsymbol{X}).$$

- Weak upper gradient are **invariant w.r.t. modification of G and f in $\mathfrak{m}$-negligible sets.**
- If f has a weak upper gradient, then **f is Sobolev along a.e. curve**,

$$\Big|\frac{\mathrm{d}}{\mathrm{d}t} f \circ \mathsf{x}\Big| \le G \circ \mathsf{x}\,|\dot{\mathsf{x}}| \quad \text{a.e. in } (0,1), \ \text{for a.e. } \mathsf{x} \in \mathrm{AC}([0,1];\boldsymbol{X}).$$

- If $f \in W^{1,2}(\boldsymbol{X}, \mathsf{d}, \mathfrak{m})$ then $|\mathrm{D}f|_w$ **is a weak upper gradient for f.**

Absolutely continuous curves of measures

Let $\mu \in \mathrm{AC}^2([0,1];\mathscr{P}_2(\boldsymbol{X}))$.

Representation theorem [Lisini]

There exists a dynamic plan $\boldsymbol{\pi}$ such that

- $\boldsymbol{\pi}$ respresents μ_t: $\mu_t = (\mathsf{e}_t)_\sharp \boldsymbol{\pi}$ for every $t \in [0,1]$, i.e.

$$\boxed{\int \varphi(\mathsf{x}(t))\,\mathrm{d}\boldsymbol{\pi}(\mathsf{x}) = \int \varphi\,\mathrm{d}\mu_t}$$

- $\boldsymbol{\pi}$ is concentrated on $\mathrm{AC}^2([0,1];\boldsymbol{X})$ and

$$\int \Big(\int_0^1 |\dot{\mathsf{x}}|^2\,\mathrm{d}t\Big)\,\mathrm{d}\boldsymbol{\pi}(\mathsf{x}) = \int_0^1 |\dot{\mu}_t|^2\,\mathrm{d}t < \infty$$

- $$\boxed{|\dot{\mu}_t|^2 = \int |\dot{\mathsf{x}}|^2(t)\,\mathrm{d}\boldsymbol{\pi}(\mathsf{x})} \quad \text{for a.e. } t \in (0,1).$$

Fisher information is an upper gradient of the Entropy

Suppose

$$\mu \in \mathrm{AC}^2([0,1]; \mathscr{P}_2(\boldsymbol{X})) \text{ with } \mu_t = \varrho_t \mathfrak{m},\ \|\varrho_t\|_\infty \le C,\ \int_0^1 \mathsf{F}(\varrho_t)\,\mathrm{d}t < \infty.$$

Entropy-Fisher dissipation formula:

The map

$$t \mapsto \mathrm{Ent}_{\mathfrak{m}}(\mu_t) = \int \varrho_t \log \varrho_t \,\mathrm{d}\mathfrak{m}$$

is absolutely continuous and

$$\Big|\frac{\mathrm{d}}{\mathrm{d}t}\mathrm{Ent}_{\mathfrak{m}}(\mu_t)\Big| \le \sqrt{\mathsf{F}(\varrho_t)}\,|\dot\mu_t|$$

The Fisher information is a **Wasserstein upper gradient** for the Entropy on the class of curves with uniformly bounded densities.

Wasserstein Gradient flow of the entropy:

$$\mathrm{Ent}_{\mathfrak{m}}(\mu_t) + \frac{1}{2}\int_0^t \Big(|\dot\mu_t|^2 + \mathsf{F}(\varrho_t)\Big)\,\mathrm{d}r \le \mathrm{Ent}_{\mathfrak{m}}(\mu_0). \qquad \text{(EDI)}$$

The Heat flow concides with the Wasserstein gradient flow of the entropy

Assume

$$\mu_0 = \varrho_0 \mathfrak{m} \in \mathscr{P}_2(\boldsymbol{X}), \quad \text{with} \quad \varrho_0 \in L^\infty(\boldsymbol{X}; \mathfrak{m}).$$

Theorem

Setting $\varrho_t = \mathsf{P}_t \varrho_0$ and $\mu_t = \varrho_t \mathfrak{m}$, we have

$$\mathrm{Ent}_{\mathfrak{m}}(\mu_t) + \frac{1}{2}\int_0^t \Big(|\dot\mu_t|^2 + \mathsf{F}(\varrho_t)\Big)\,\mathrm{d}r \le \mathrm{Ent}_{\mathfrak{m}}(\mu_0). \qquad \text{(EDI)}$$

μ_t is a Wasserstein gradient flow of the Entropy.

*μ_t is the **unique solution of** (EDI) in the class of absolutely continuous curves with uniformly bounded densities.*

[Jordan-Kinderleherer-Otto, Otto, AGS, Ambrosio-S.-Zambotti, Erbar, Villani, Gigli, Gigli-Kuwada-Ohta, AGS]

Proof

Euristics in $\mathbb{R}^n$.

Basic tools:

- Dual Kantorovich characterization of the Wasserstein distance
- Precise pointwise solution of the Hamilton-Jacobi equation given by the Hopf-Lax formula
- Kuwada Lemma:

$$\boxed{|\dot\mu_t|^2 \le \mathsf{F}(\varrho_t).}$$

Applications to the structure of Sobolev space $W^{1,2}(\boldsymbol{X}, \mathsf{d}, \mathfrak{m})$
[Cheeger, Shanmugalingam, Koskela-MacManus]

Euristics

In $\mathbb{R}^n$: $\mu_t = \varrho_t \mathscr{L}^n$

$$\partial_t \mu_t = \Delta \mu_t = \mathrm{div}\Big(\mu_t \nabla \log \varrho_t\Big)$$

The measures μ_t are evolving transported by the vector field $v_t := -\nabla \log \varrho_t$.

Lisini's representation: **$\boldsymbol{\pi}$ is concentrated on characteristic curves** solving

$$\dot{\mathsf{x}}(t) = -\nabla \log \varrho_t(\mathsf{x}(t)) = -\frac{\nabla \varrho_t}{\varrho_t}(\mathsf{x}(t))$$

thus

$$-\frac{\mathrm{d}}{\mathrm{d}t}\log(\varrho_t(\mathsf{x}(t)) + \big(\partial_t \log \varrho_t\big)(\mathsf{x}(t)) = \nabla \log \varrho_t(\mathsf{x}(t)) \cdot \dot{\mathsf{x}}(t) = \frac{1}{2}\frac{|\nabla \varrho_t(\mathsf{x}(t))|^2}{\varrho_t(\mathsf{x}(t))} + \frac{1}{2}|\dot{\mathsf{x}}|^2(t)$$

Integrating w.r.t. $\boldsymbol{\pi}$

$$-\partial_t \int \log(\varrho_t(\mathsf{x}(t))\,\mathrm{d}\boldsymbol{\pi} + \int \varrho_t^{-1}\partial_t \varrho_t(\mathsf{x}(t))\,\mathrm{d}\boldsymbol{\pi} = -\partial_t \int \varrho_t \log \varrho_t \,\mathrm{d}\mathbf{m} + \int \Delta \varrho_t \,\mathrm{d}\mathbf{m}$$

$$= -\frac{\mathrm{d}}{\mathrm{d}t}\mathrm{Ent}_{\mathbf{m}}(\mu_t) = \int \Big(\frac{1}{2}\frac{|\nabla \varrho_t(\mathsf{x}(t))|^2}{\varrho_t(\mathsf{x}(t)} + \frac{1}{2}|\dot{\mathsf{x}}(t)|^2\Big)\,\mathrm{d}\boldsymbol{\pi} = \frac{1}{2}\mathsf{F}(\mu_t) + \frac{1}{2}|\dot{\mu}_t|^2$$

Dual Kantorovich characterizations of the Wasserstein distance

$$W_1(\mu_0,\mu_1) = \min\Big\{ \int \mathsf{d}(x_0,x_1)\,\mathrm{d}\boldsymbol{\mu} : \boldsymbol{\mu} \text{ coupling for } \mu_0,\mu_1 \Big\}$$

Dual characterization:

$$W_1(\mu_0,\mu_1) = \sup\Big\{ \int \phi\,\mathrm{d}\mu_1 - \int \phi\,\mathrm{d}\mu_0 : \phi(x_1) - \phi_0(x_0) \le \mathsf{d}(x_0,x_1) \Big\}$$

$$W_2^2(\mu_0,\mu_1) = \min\Big\{ \int \mathsf{d}^2(x_0,x_1)\,\mathrm{d}\boldsymbol{\mu} : \boldsymbol{\mu} \text{ coupling for } \mu_0,\mu_1 \Big\}$$

Dual characterization

$$\frac{1}{2}W_2^2(\mu_0,\mu_1) = \boxed{\sup\Big\{ \int \mathsf{Q}_1\phi\,\mathrm{d}\mu_1 - \int \phi\,\mathrm{d}\mu_0 : \phi \in \mathrm{Lip}_b(\boldsymbol{X}) \Big\}}$$

where

$$\boxed{\mathsf{Q}_t\phi(x) := \inf_y \frac{1}{2t}\mathsf{d}^2(x,y) + \phi(y).}$$

The Hopf-Lax semigroup

Assume that $\boldsymbol{X}$ is compact. Let $\phi \in \mathrm{Lip}(\boldsymbol{X})$ and

$$\boxed{\mathsf{Q}_t\phi(x) := \inf_y \frac{1}{2t}\mathsf{d}^2(x,y) + \phi(y).}$$

Then the map $t \mapsto \mathsf{Q}_t\phi$ is **Lipschitz** from $[0,\infty)$ to $\mathrm{C}(\boldsymbol{X})$,
$\mathsf{Q}_t\phi$ is **Lipschitz** for every $t \geq 0$
for every $x \in \boldsymbol{X}$

$$\boxed{\frac{\mathrm{d}}{\mathrm{d}t}\mathsf{Q}_t\phi + \frac{1}{2}|\mathrm{D}\mathsf{Q}_t\phi|^2 \leq 0} \tag{HJ}$$

for every $t > 0$ with at most countably many exceptions.
If moreover $(\boldsymbol{X}, \mathsf{d})$ is a geodesic space, then equality holds in (HJ) for every $t > 0$ with at most countably many exceptions.