Calculus and heat flow in metric measure spaces and spaces with Riemannian curvature bounds from below

L. Ambrosio

Scuola Normale Superiore, Pisa
http://cvgmt.sns.it

Outline

SCUOLA NORMALE SUPERIORE

Overview

A.-Gigli-Savaré: *Calculus and heat flows in metric measure spaces and applications to spaces with Ricci bounds from below.*
http://cvgmt.sns.it, submitted.
(some results and proofs)

A.-Gigli-Savaré: *Riemannian Ricci curvature bounds in metric measure spaces.*
In preparation.
(just statements, no proofs)

Some by now "classical" results

Let us consider in $\mathbb{R}^n$ the heat equation ($u_t(x) = u(t,x)$)

$$\partial_t u_t = \Delta u_t$$

Classically it can be viewed as the gradient flow of the energy

$$\mathrm{Dir}(u) := \frac{1}{2}\int_{\mathbb{R}^n} |\nabla u|^2\,dx \qquad (+\infty \text{ if } u \notin H^1(\mathbb{R}^n))$$

in the Hilbert space $H = L^2(\mathbb{R}^n)$.

Formally, $t \mapsto u_t$ solves the ODE $u' = -\nabla\mathrm{Dir}(u)$ in H because

$$\mathrm{Dir} \text{ "differentiable" at } u \qquad \Longleftrightarrow \qquad -\Delta u \in L^2,\ \nabla\mathrm{Dir}(u) = -\Delta u$$

In 1998, Jordan-Kinderlehrer-Otto proved that the same equation arises as gradient flow of the *entropy* functional

$$\mathrm{Ent}(\rho\mathscr{L}^n) := \int_{\mathbb{R}^n} \rho \log \rho \, dx \qquad (+\infty \text{ if } \mu \text{ is not a.c. w.r.t. } \mathscr{L}^n)$$

in the space $\mathscr{P}_2(\mathbb{R}^n)$ of probability measures with finite quadratic moments, with respect to Wasserstein distance W_2.

$$W_2^2(\mu,\nu) := \min\left\{\int_{\mathbb{R}^n\times\mathbb{R}^n} |x-y|^2 \, d\gamma(x,y) : \ (\pi_1)_\sharp\gamma = \mu, \ (\pi_2)_\sharp\gamma = \nu\right\}.$$

Push forward notation. $f : X \to Y$ Borel induces a map $f_\sharp : \mathscr{P}(X) \to \mathscr{P}(Y)$:

$$f_\sharp\mu(B) := \mu\bigl(f^{-1}(B)\bigr) \qquad \forall B \in \mathscr{B}(X).$$

Proofs of this equivalence

1. By the so-called Otto calculus (formal);

2. Prove that the implicit time discretization scheme (Euler scheme), traditionally used for the approximation of gradient flows, when done with energy $\mathrm{Ent}(\mu)$ and distance W_2, does converge to the heat equation.

3. Give a meaning to what "gradient flow of Ent w.r.t. W_2 means", and check that solutions of this gradient flow are solutions to the heat equation. Then, apply uniqueness for $\partial_t u_t = \Delta u_t$.

The last strategy is more abstract, but still uses the differentiable structure of $\mathbf{R}^n$. The question is to understand deeper reasons for this equivalence, in particular on which structural properties of the space it depends (Riemannian manifolds, Finsler spaces, Wiener spaces, sub-Riemannian spaces, etc.)

SCUOLA NORMALE SUPERIORE

Metric measure spaces

Let us consider a *metric measure space* $(X, d, \mathfrak{m})$, with $\mathfrak{m} \in \mathscr{P}(X)$. In this framework it is still possible to define a "Dirichlet energy", that we call Cheeger functional:

$$Ch(f) := \inf\left\{\liminf_{n\to\infty}\int_X |\nabla f_n|^2\,d\mathfrak{m}:\ f_n \in \mathrm{Lip}(X),\ \int_X |f_n - f|^2\,d\mathfrak{m} \to 0\right\},$$

where

$$|\nabla g|(x) := \limsup_{y\to x}\frac{|g(y)-g(x)|}{d(y,x)}$$

is the *slope* (also called local Lipschitz constant).
Also, one can consider the so-called *relative entropy functional* $\mathrm{Ent}_{\boldsymbol{m}} : \mathscr{P}(X) \to [0, +\infty]$

$$\mathrm{Ent}_{\boldsymbol{m}}(\rho\mathfrak{m}) := \int_X \rho\log\rho\,d\mathfrak{m} \qquad (+\infty \text{ if } \mu \text{ is not a.c. w.r.t. } \mathfrak{m}).$$

SCUOLA NORMALE SUPERIORE

The basic result is that the equivalence between L^2-gradient flow of Ch and W_2-gradient flow of $\mathrm{Ent}_{\boldsymbol{m}}$ *always* holds, if the latter is properly understood. But, without additional assumptions on the space, both objects can be trivial.

Example. Let $X = [0,1]$, d the Euclidean distance, $\mathfrak{m} = \sum_{n\geq 1} 2^{-n}\delta_{q_n}$, where $\{q_n\}_{n\geq 1}$ is an enumeration of $[0,1]\cap\mathbb{Q}$. Let $A_n \supset \mathbb{Q}\cap X$ be open sets with $\mathscr{L}^1(A_n) \to 0$ and

$$\chi_n(t) := \int_0^t \bigl(1 - \chi_{A_n}(s)\bigr)\,ds \qquad t \in [0,1].$$

Then $f \circ \chi_n \to f$ in $L^2(X,\mathfrak{m})$ for all $f \in \mathrm{Lip}(X)$ and $f \circ \chi_n$ is locally constant in $\mathbb{Q}\cap X$ hence

$$\mathrm{C}h(f) = 0 \qquad \forall f \in \mathrm{Lip}(X).$$

It follows that $\mathrm{C}h \equiv 0$ in $L^2(X,\mathfrak{m})$.

Identification of weak gradients

A closely related question, relevant in particular for the second paper, is the identification of weak gradients. The first one, that we call *relaxed* gradient $|\nabla f|_*$, is the object that provides integral representation to Ch:

$$\mathrm{C}h(f) = \frac{1}{2}\int_X |\nabla f|_*^2 \, d\mathfrak{m} \qquad \forall f \in D(\mathrm{C}h).$$

It has all the natural properties (locality, chain rules, etc.) a weak gradient should have.

This gradient is useful when doing "vertical" variations $\epsilon \mapsto f + \epsilon g$ (i.e. in the *dependent* variable).

SCUOLA NORMALE SUPERIORE

Identification of weak gradients

But, when computing variations of the entropy, the "horizontal" variations $\epsilon \to f(\gamma_\epsilon)$ (i.e. in the *independent* variable) are necessary. These are related to another weak gradient $|\nabla f|_w$, defined as follows.

We require the so-called upper gradient property

$$|f(\gamma_1) - f(\gamma_0)| \leq \int_\gamma G$$

along "almost all" curves γ in $AC^2([0,1];X)$ and then we define $|\nabla f|_w$ as the element with smallest $L^2(X,\mathfrak{m})$ norm.

The remarkable fact is that these two gradients *always* coincide (and, of course, maybe both trivial without extra assumptions). The proof of this identification uses ideas from optimal transportation, as lifting of solutions to the heat flow to probability measures in $AC^2\big([0,1];X\big)$ and the energy dissipation rate of $\mathrm{Ent}_{\boldsymbol{m}}$.

SCUOLA NORMALE SUPERIORE

Why gradients are not trivial in Lott-Sturm-Villani spaces

In these spaces one imposed convexity of W_2 geodesics of $\mathrm{Ent}_{\boldsymbol{m}}$ (the so-called $CD(0,\infty)$ condition) or of functionals

$$\rho\mathfrak{m} \mapsto -\int_X \rho^{1-1/N}\,d\mathfrak{m}$$

(the $CD(0,N)$ condition).

In this case the gradient flow of $\mathrm{Ent}_{\boldsymbol{m}}$ is not trivial, and since it coincides with the heat flow, also this is not trivial.

The energy dissipation rate is

$$\begin{aligned}\frac{d}{dt}\int_X \rho_t \log\rho_t\,d\mathfrak{m} &= \int_X \log\rho_t\Delta\rho_t\,d\mathfrak{m} = -\int_{\{\rho_t>0\}}\frac{|\nabla\rho_t|^2}{\rho_t}\,d\mathfrak{m}\\ &= -4\int_X |\nabla\sqrt{\rho_t}|^2\,d\mathfrak{m}.\end{aligned}$$

SCUOLA NORMALE SUPERIORE

Standing assumptions (for the lectures).

$$(X, d) \text{ compact metric space, } \mathfrak{m} \in \mathscr{P}(X)$$

Prerequisites.

Basic facts of Optimal Transportation and Measure Theory

References. Villani's monographs ('03, '09), [A.-Gigli-Savaré] '08, A.-Gigli user's guide '11.

The Hamilton-Jacobi semigroup

$$Q_t f(x) := \inf_{y \in X} f(y) + \frac{1}{2t} d^2(x,y) \qquad \text{(Hopf-Lax formula)}$$

Theorem. *f bounded, lower semicontinuous. It holds:*
(1) $Q_t f(x) \uparrow f(x)$ as $t \downarrow 0$;
(2) $Q_t(Q_s f(x)) \geq Q_{t+s} f(x)$, with equality if (X,d) is geodesic;
(3) $\frac{d^+}{dt} Q_t f(x) + \frac{1}{2} |\nabla Q_t f(x)|^2 \leq 0$;
(4) $Q_t f(x)$ restricted to $(\epsilon, \infty) \times X$ is Lipschitz for all $\epsilon > 0$.

SCUOLA NORMALE SUPERIORE

Sketch of proof

(1) It follows by the lower semicontinuity of f, which ensures also that minimizers do exist (exercise)
(2) It follows by

$$
\begin{aligned}
& \inf_y \left(\inf_z f(z) + \frac{1}{2s} d^2(z,y) \right) + \frac{1}{2t} d^2(x,y) \\
= \; & \inf_z \inf_y \left(\frac{1}{2s} d^2(z,y) + \frac{1}{2t} d^2(x,y) \right) + f(z) \\
\geq \; & \inf_z \frac{1}{2(s+t)} d^2(x,z) + f(z)
\end{aligned}
$$

noticing that the last inequality is an equality in geodesic spaces.
In order prove (3), we set

$$
\begin{cases}
D_f^+(x,t) := \max\{d(x,y) : \; y \text{ minimizer}\} \\
D_f^-(x,t) := \min\{d(x,y) : \; y \text{ minimizer}\}.
\end{cases}
$$

Sketch of proof

Since limit of minimizers is a minimizer, D^+ is upper semicontinuous, while D^- is lower semicontinuous. In addition $D_f^+(x,t) \geq D_f^-(x,t) \geq D_f^+(x,s)$ if $0 < s < t$.

We prove first that $\frac{d^\pm}{dt} Q_t f(x) = -[D_f^\pm(x,t)]^2/(2t^2)$.

Choosing x_t at maximum distance and x_s at minimum distance yields

$$\begin{aligned} Q_s f(x) - Q_t f(x) &\leq \frac{1}{2s} d^2(x,x_t) + f(x_t) - f(x_t) - \frac{1}{2t} d^2(x,x_t) \\ &= \frac{(D_f^+(x,s))^2}{2}\left(\frac{1}{s} - \frac{1}{t}\right) \end{aligned}$$

$$\begin{aligned} Q_s f(x) - Q_t f(x) &\geq \frac{1}{2s} d^2(x,x_s) + f(x_s) - f(x_s) - \frac{1}{2t} d^2(x,x_s) \\ &= \frac{(D_f^-(x,s))^2}{2}\left(\frac{1}{s} - \frac{1}{t}\right) \end{aligned}$$

Sketch of proof

To conclude, suffices to show that $|\nabla Q_t f| \le D_f^+(x,t)/t$. The same trick used before, now for variations in space, yields:

$$\begin{aligned} Q_t f(x) - Q_t f(y) &\le \frac{1}{2t} d^2(x,z) + f(z) - f(z) - \frac{1}{2t} d^2(z,y) \\ &\le d(x,y)\Big(\frac{D^-(y,t)}{t} + \frac{d(x,y)}{2t}\Big) \end{aligned}$$

and we can use the upper semicontinuity of D_f^+ to conclude.
If y is kept fixed and we let $x \to y$ we obtain the sharper inequality

$$|\nabla^+ Q_t f|(y)| \le \frac{D_f^-(y,t)}{t},$$

where the *ascending slope* $|\nabla^+ f|$ is defined by

$$|\nabla^+ f|(y) := \limsup_{x \to y} \frac{[f(x) - f(y)]^+}{d(x,y)}.$$

Hamilton-Jacobi and optimal transportation

Why the Hopf-Lax formula and the Hamilton-Jacobi equation are relevant in the theory of optimal transport?

***c*-transform.** Given a cost function $c : X \times Y \to \mathbb{R}$ the c-transforms are $\varphi^c : Y \to \mathbb{R} \cup \{-\infty\}$, $\psi^c : X \to \mathbb{R} \cup \{-\infty\}$ are defined by

$$\varphi^c(y) := \inf_{x \in X} c(x,y) - \varphi(x), \qquad \psi^c(x) := \inf_{y \in Y} c(x,y) - \psi(y)$$

Notice the analogy with convex analysis: $\psi^c = (-\psi)^*$ if X is Hilbert and $c(x,y) = \langle x, y \rangle$. The relation with the HL formula is also obvious:

$$\psi^c = Q_1(-\psi).$$

Then, we say that $\varphi : X \to \mathbb{R} \cup \{-\infty\}$ is *c-concave* if $\varphi = \psi^c$ for some $\psi : Y \to \mathbb{R} \cup \{-\infty\}$. As in convex analysis, $\varphi \mapsto \varphi^c$ is an involution in the class of c-concave functions: $(\varphi^c)^c = \varphi$.

Hamilton-Jacobi and optimal transportation

Definition. *We say that a c-concave function* $\varphi : X \to \mathbb{R} \cup \{-\infty\}$ *is a Kantorovich potential relative to* (μ, ν) *if it satisfies*

$$(*) \qquad \varphi(x) + \varphi^c(y) = c(x, y) \qquad \text{for } \gamma\text{-a.e. } (x, y)$$

for any optimal plan γ *from* μ *to* ν*.*

Corollary. *If* φ *is a Kantorovich potential from* μ *to* ν *it holds:*

$$|\nabla^+ \varphi|(x) \le d(x, y) \qquad \text{for } \gamma\text{-a.e. } (x, y).$$

In particular $\int |\nabla^+ \varphi|^2 \, d\mu \le W_2^2(\mu, \nu)$*.*

Proof. Since $\varphi = (\varphi^c)^c$ we may write $\varphi = Q_1(-\varphi^c)$.
On the other hand, the optimality condition (*) gives

$$D^-_{-\varphi^c}(x, 1) \le d(x, y) \qquad \gamma\text{-a.e. in } X \times X.$$

The classical Brenier theorem and an example

In the Euclidean case $c(x, y) = |x - y|^2/2$, if φ is differentiable at x and

$$\varphi(x) + \varphi^c(y) = \frac{1}{2}|x - y|^2$$

one can differentiate at x and obtain $\nabla\varphi(x) + (x - y) = 0$, which tells us that y is uniquely determined by x and

$$|\nabla\varphi|(x) = d(x, y).$$

Example. $X = [0, 1]$, $\mu = \delta_0$, $\mu_t = t^{-1}\chi_{[0,t]}\mathscr{L}^1$. In this case

$$\varphi(x) = \frac{x^2}{2} - x, \qquad \int |\nabla^+\varphi|^2 \, d\mu_0 = 0$$

while

$$W_2^2(\mu_0, \mu_1) = \frac{1}{3}.$$

Optimal transport and Kantorovich potentials in geodesic spaces

If (X, d) is Polish and geodesic we may formulate the optimal transport problem in terms of *geodesic plans* i.e. probability measures $\boldsymbol{\pi}$ concentrated in the Polish space $\mathrm{Geo}(X)$ of constant speed geodesics:

$$(*) \qquad \min\left\{\int d^2(\gamma_0,\gamma_1)\,d\boldsymbol{\pi}(\gamma):\ (e_0)_\sharp\boldsymbol{\pi}=\mu,\ (e_1)_\sharp\boldsymbol{\pi}=\nu\right\}.$$

Here $e_t : C([0,1];X) \to X$ are the evaluation maps, namely $e_t(\gamma) = \gamma_t$.

The relation with the classical optimal plans γ of Kantorovich theory is that if $\boldsymbol{\pi}$ is a minimizer in (*), then $(e_0, e_1)_\sharp\boldsymbol{\pi}$ is an optimal plan, and that any optimal γ admits a (possibly nonunique) "lifting" $\boldsymbol{\pi}$.

Theorem. *Any constant speed geodesic μ_t can be represented as $(e_t)_\sharp\boldsymbol{\pi}$ for a suitable optimal geodesic plan $\boldsymbol{\pi}$. Conversely, any optimal geodesic plan $\boldsymbol{\pi}$ induces a constant speed geodesic $(e_t)_\sharp\boldsymbol{\pi}$.*

Optimal transport and Kantorovich potentials in geodesic spaces

Theorem. *Let μ_t, $t \in [0,1]$ be a constant speed geodesic and let φ be a Kantorovich potential relative to μ_0, μ_1. Then, for all $t \in (0,1]$, $\varphi_t := Q_t(-\varphi^c)$ is a Kantorovich potential, relative to the scaled cost c/t, from μ_{1-t} to μ_1.*

Sketch of proof. It is obvious that $\varphi_t + \varphi \leq c_t$. The key implication is

$$\varphi(\gamma_0) + \varphi^c(\gamma_1) = c(\gamma_0, \gamma_1) \quad \text{implies} \quad \varphi_t(\gamma_{1-t}) + \varphi^c(\gamma_1) = c_t(\gamma_0, \gamma_1)$$

hence, if $\boldsymbol{\pi}$ is an optimal geodesic plan, $\varphi + \varphi^c = c$ γ_1-a.e. implies $\varphi_t + \varphi^c = c/t$ γ_t-a.e., where

$$\gamma_t := (\gamma_{1-t}, \gamma_1)_\sharp \boldsymbol{\pi}$$

is an optimal plan from μ_{1-t} to μ_1.

Cheeger's energy and relaxed slopes

$$\mathsf{C}h(f) := \frac{1}{2}\inf\left\{\liminf_{h\to\infty}\int|\nabla f_h|^2\,d\mathfrak{m}:\ f_h\in \mathrm{Lip}(X),\ \int_X|f_h-f|^2\,d\mathfrak{m}\to 0\right\}$$

By construction $\mathsf{C}h$ is lower semicontinuous, and it is easily seen to be convex. Can we provide an integral representation to it?
Relaxed slope: $G\in L^2(X,\mathfrak{m})$ is a relaxed slope of f if G bounds from above a function in

$$\left\{\text{weak } L^2 \text{ limit points of } |\nabla f_n|,\ f_n\in\mathrm{Lip}(X),\ \|f_n-f\|_2\to 0\right\}$$

or equivalently in

$$\left\{\textit{strong } L^2 \text{ limit points of c.c. of } |\nabla f_n|,\ f_n\in\mathrm{Lip}(X),\ \|f_n-f\|_2\to 0\right\}$$

Cheeger's energy and relaxed slopes

We call *minimal relaxed slope* and denote by $|\nabla f|_*$ the function with smallest $L^2(X, \mathfrak{m})$ norm among relaxed slopes.

Theorem. *Let* $f \in D(\mathrm{C}h)$. *Then:*
(1) $\mathrm{C}h(f) = \frac{1}{2}\int |\nabla f|_*^2 \, d\mathfrak{m}$;
(2) if G_1, G_2 *are relaxed slopes, so is* $\min\{G_1, G_2\}$;
(3) $|\nabla f|_* \leq G$ $\mathfrak{m}$*-a.e. for any relaxed slope* G;
(4) $g = f$ $\mathfrak{m}$*-a.e. on a Borel set* B *implies* $|\nabla f|_* = |\nabla g|_*$ $\mathfrak{m}$*-a.e. on* B.

Calculus rules. *If* $N \subset \mathbb{R}$ *is Lebesgue negligible, then* $|\nabla f|_* = 0$ *a.e. in* $f^{-1}(N)$.

$$|\nabla \phi(f)|_* \leq |\phi'(f)||\nabla f|_* \qquad \textit{with equality if } \phi' \geq 0.$$

Sketch of proof

(1) Any weak limit point of $|\nabla f_n|$ yields a relaxed slope, hence $\mathsf{C}h(f) \geq \frac{1}{2}\int |\nabla f|_*^2\, d\mathfrak{m}$. Writing $G \leq |\nabla f|_*$ as the strong limit of convex combinations of $|\nabla f_n|$ we have

$$\int |\nabla f|_*^2\, d\mathfrak{m} \geq \int G^2\, d\mathfrak{m} \geq \liminf_n \int |\nabla f_n|^2\, d\mathfrak{m} \geq 2\mathsf{C}h(f).$$

(2) By approximation, suffices to show that $\chi_{X\setminus B}G_1 + \chi_B G_2$ is a relaxed slope if B is closed. Set $\rho(x) = \mathrm{dist}(x,B)$, $\chi_r(x) = \min\{1, r^{-1}\rho\}$, so that $\chi_r \uparrow \chi_{X\setminus B}$ as $r \downarrow 0$, and pass to the limit in

$$|\nabla(\chi_r f_{n,1} + (1-\chi_r)f_{n,2})| \leq \chi_r|\nabla f_{n,1}| + (1-\chi_r)|\nabla f_{n,2}| + \mathrm{Lip}(\chi_r)|f_{n,1} - f_{n,2}|.$$

(3) Just take $\tilde{G} := \min\{|\nabla f|_*, G\}$. Its L^2 norm would be strictly smaller than $\||\nabla f|_*\|_2$, were the set $\{|\nabla f|_* > G\}$ with positive $\mathfrak{m}$-measure.

Heat flow and Laplacian

Let's start with some reminders on the classical theory of gradient flows of convex and l.s.c. functionals $F : H \to \mathbb{R} \cup \{+\infty\}$ in a Hilbert space H.

Subdifferential ∂F. It is the multivalued map defined by

$$\partial F(x) := \{p \in H : \ F(x) + \langle p, y - x \rangle \leq F(y) \ \ \forall y \in H\}$$

for all $x \in D(F) := \{F < \infty\}$. The *gradient* $\nabla F(x)$ is the element with minimal norm in $\partial F(x)$.

Gradient flow. It is a locally absolutely continuous map $x : (0, \infty) \to H$ satisfying

$$-x'(t) \in \partial F\big(x(t)\big) \qquad \text{for a.e. } t > 0.$$

In addition, we say that $x(t)$ starts from $\bar{x}$ if $\lim_{t \downarrow 0} x(t) = \bar{x}$.

Heat flow and Laplacian

Theorem. (Existence and uniqueness) *For all $\bar{x} \in \overline{D(F)}$ there exists a unique gradient flow starting from $\bar{x}$ and the induced semigroup*

$$S_t : [0,\infty) \times \overline{D(F)} \to \overline{D(F)}$$

is contractive. In addition, we have the regularizing effects:

(1) *$S_t\bar{x} \in D(\partial F) \subset D(F)$ for all $t > 0$ and*

$$F(S_t\bar{x}) \leq \inf_{v \in D(F)} F(v) + \frac{1}{2t} d^2(v,\bar{x});$$

(2) *$\frac{d^+}{dt} S_t\bar{x} = -\nabla F(S_t\bar{x})$ for all $t > 0$;*

(3) *$t \mapsto |\nabla F|^2(S_t\bar{x})$ is nonincreasing, so that $S_t\bar{x}$ is Lipschitz in (ϵ,∞) for all $\epsilon > 0$;*

(4) *$\frac{d^+}{dt} F(S_t\bar{x}) = -|\nabla F|^2(S_t\bar{x}) = -|\frac{d^+}{dt} S_t\bar{x}|^2$ for all $t > 0$.*

SCUOLA NORMALE SUPERIORE

Heat flow and Laplacian

According to these results, we may choose $H = L^2(X, \mathfrak{m})$ and $F = \mathrm{C}h$ and define

$$-\Delta f := \text{the element with minimal } L^2\text{-norm of } \partial \mathrm{C}h(f)$$

so that (by the density of $D(\mathrm{C}h) \supset \mathrm{Lip}(X)$ in $L^2(X, \mathfrak{m})$) we obtain a L^2 *heat flow* $\mathbf{h}_t f$ solving

$$\frac{d}{dt}\mathbf{h}_t \bar{f} = \Delta \mathbf{h}_t \bar{f}$$

starting from any initial condition $\bar{f} \in L^2(X, \mathfrak{m})$.

Remarks. (1) $\Delta = \Delta_{d,\mathfrak{m}}$. Even in the classical situations, $\Delta f = \mathrm{div}(\nabla f)$, where ∇f depends on the metric (to associate a vector ∇f to df) while div depends on the volume form $\mathfrak{m}$, via the adjoint formula

$$\int g \, \mathrm{div} F \, d\mathfrak{m} = -\int \langle \nabla g, F \rangle \, d\mathfrak{m}.$$

Heat flow and Laplacian

(2) Δ need not to be linear in this context! Take $X = \mathbb{R}^2$ with the L^∞ norm, to get

$$\mathsf{C}h(f) = \int \left(\left| \frac{\partial f}{\partial x} \right| + \left| \frac{\partial f}{\partial y} \right| \right)^2 dxdy.$$

Nowithstanding this potential lack of linearity, a reasonable calculus can be developed:

$$-\int g \Delta f \, d\mathfrak{m} \le \int |\nabla f|_* |\nabla g|_* \, d\mathfrak{m},$$

$$-\int \phi(f) \Delta f \, d\mathfrak{m} = \int \phi'(f) |\nabla f|_*^2 \, d\mathfrak{m}.$$

The first one follows by

$$\mathsf{C}h(f) - \epsilon \int g \Delta f \, d\mathfrak{m} \le \int |\nabla (f + \epsilon g)|_*^2 \, d\mathfrak{m}$$

noticing that $|\nabla (f + \epsilon g)|_* \le |\nabla f|_* + \epsilon |\nabla g|_*$ for $\epsilon > 0$.

Heat flow and Laplacian

Properties of the heat flow. *(1) Homogeneity:* $\mathbf{h}_t(\lambda f) = \lambda \mathbf{h}_t f \ \forall \lambda \in \mathbb{R}$*;*
(2) Comparison principle: if $f \le g$*, then* $\mathbf{h}_t f \le \mathbf{h}_t g$ *for all* $t \ge 0$*;*
(3) Energy dissipation: if $f_t : X \to J$ *and* $\mathrm{e} : J \to \mathbb{R}$ *is convex and locally* $C^{1,1}$*, then*

$$\int (\mathbf{h}_t f)\, d\mathfrak{m} = \int \mathrm{e}(f)\, d\mathfrak{m} + \int_0^t \int \mathrm{e}''(\mathbf{h}_s f) |\nabla \mathbf{h}_s f|_*^2 \, d\mathfrak{m} ds.$$

(4) Mass preservation: $\int \mathbf{h}_t f \, d\mathfrak{m} = \int f \, d\mathfrak{m}$ *for all* $t \ge 0$.

Strictly speaking, (3) does not cover the most interesting case, the case of the entropy $\mathrm{e}(z) = z \log z$ when $\mathbf{h}_t f \ge 0$:

$$\int \mathbf{h}_t f \log \mathbf{h}_t f \, d\mathfrak{m} = \int f \log f \, d\mathfrak{m} + \int_0^t \int_{\{\mathbf{h}_s f > 0\}} \frac{|\nabla \mathbf{h}_s f|_*^2}{\mathbf{h}_s f} \, d\mathfrak{m} ds.$$

It can be recovered by the approximation $f \mapsto \max\{f, \epsilon\}$.

SCUOLA NORMALE SUPERIORE

Absolutely continuous functions and metric speed

A curve $\gamma : [0,1] \to X$ is said to be absolutely continuous if

$$(*) \qquad d(\gamma_t, \gamma_s) \le \int_t^s f(r)\, dr \qquad \forall [t,s] \subset [0,1]$$

for some $f \in L^1(0,1)$.

If γ is absolutely continuous, the *metric speed* $|\dot\gamma| : [0,1] \to [0,\infty]$ is defined by

$$|\dot\gamma_t| := \lim_{h \to 0} \frac{d(\gamma_{t+h}, \gamma_t)}{|h|}.$$

It is possible to prove that the limit exists for a.e. t, that $|\dot\gamma| \in L^1(0,1)$, and that it is the minimal L^1 function for which the bound (*) holds.

Kuwada lemma

Lemma. *Let* $f_0 \in L^2(X, \mathfrak{m})$ *a probability density,* $f_t = \mathbf{h}_t f_0$*. Then the curve* $\mu_t := f_t \mathfrak{m}$ *is absolutely continuous in* $\mathscr{P}(X)$ *and*

$$|\dot{\mu}_t|^2 \le \int_{\{f_t>0\}} \frac{|\nabla f_t|_*^2}{f_t} \, d\mathfrak{m} \qquad \text{for a.e. } t > 0.$$

It is convenient to introduce the Fisher information functional, defined $\{\rho : \sqrt{\rho} \in D(\mathsf{C}h)\}$, as follows:

$$F(\rho) := 4 \int |\nabla \sqrt{\rho}|_*^2 \, d\mathfrak{m} = \int_{\{\rho>0\}} \frac{|\nabla \rho|_*^2}{\rho} \, d\mathfrak{m}$$

(the last equality follows by chain rule).
We prove an integral version of the lemma, namely

$$W_2^2(\mu_t, \mu_s) \le \ell \int_t^s F(f_r) \, dr$$

with $0 \le s < t < \infty$ and $\ell := (s - t)$.

Proof of Kuwada lemma

By Kantorovich's duality formula, suffices to show

$$\int -\varphi \, d\mu_t + \int Q_1\varphi \, d\mu_s \leq \frac{\ell}{2} \int_t^s F(f_r) \, dr,$$

where φ runs in the class of bounded continuous functions. Replacing φ by $Q_\epsilon\varphi$ and letting $\epsilon \downarrow 0$ we can assume that $Q_t\varphi$ is Lipschitz in $[0,1] \times X$.

Now we set $g(r) := \int Q_r\varphi \, d\mu_{t+\ell r}$ and we write the inequality as

$$\int_0^1 g'(r) \, dr \leq \frac{\ell}{2} \int_t^s F(f_r) \, dr.$$

SCUOLA NORMALE SUPERIORE

Evaluation of $g'(r)$

Using the HJ subsolution property of $Q_r\varphi$ and the "integration by parts" we get

$$\begin{aligned} g'(r) &= \int (\frac{d}{dr}Q_r\varphi) f_{t+\ell r}\,d\mathfrak{m} + \ell \int Q_r\varphi \Delta f_{t+\ell r}\,d\mathfrak{m} \\ &\leq -\frac{1}{2}\int |\nabla Q_r\varphi|_*^2 f_{t+\ell r}\,d\mathfrak{m} + \ell \int |\nabla Q_r\varphi| \sqrt{f_{t+\ell r}} \frac{|\nabla f_{t+\ell r}|_*}{\sqrt{f_{t+\ell r}}}\,d\mathfrak{m}. \end{aligned}$$

Eventually the Young inequality gives

$$g'(r) \leq \frac{\ell^2}{2} F(f_{t+\ell r}) = \frac{\ell}{2}\frac{d}{dr}F(f_{t+\ell r})$$

and an integration in $(0, 1)$ with respect to r gives the result.

SCUOLA NORMALE SUPERIORE

Gradient flow of $\mathrm{Ent}_{\mathfrak{m}}$

Since the ambient space $\mathscr{P}(X)$ is not linear (at least if we take the viewpoint of optimal transportation), what does it mean?

Key idea. (De Giorgi) Encode the system $x'(t) = -\nabla F(x(t))$ in a single differential inequality, by looking at the rate of energy dissipation:

$$\text{(DG)} \qquad \frac{d}{dt} F(x(t)) \leq -\frac{1}{2}|\nabla F|^2(x(t)) - \frac{1}{2}|x'(t)|^2.$$

Indeed, in a sufficiently smooth setting, along *any* curve $y(t)$, we have

$$\begin{aligned} \frac{d}{dt} F(y(t)) &= \langle \nabla F(y(t)), y'(t) \rangle \\ &\geq -|\nabla F(y(t))||y'(t)| \quad (= \text{iff } -y'(t) \text{ is parallel to } \nabla F(y(t))) \\ &\geq -\frac{1}{2}|\nabla F|^2(y(t)) - \frac{1}{2}|y'(t)|^2 \quad (= \text{iff } |\nabla F|(y(t)) = |y'(t)|). \end{aligned}$$

All terms in (DG) make sense in a metric space (X, d): $|x'|$ can be replaced metric derivative and $|\nabla F|$ by the *descending slope* $|\nabla^- F|$, so that the speed is 0 at minimum points.

(EDE) and (EDI) flows in metric spaces

$$\text{(EDI)} \quad F(x(t)) + \int_0^t \frac{1}{2}|x_r'|^2 + \frac{1}{2}|\nabla^- F|^2(x(r))\,dr \le F(x(0)) \quad \forall t \ge 0.$$

$$\text{(EDE)} \quad F(x(t)) + \int_0^t \frac{1}{2}|x_r'|^2 + \frac{1}{2}|\nabla^- F|^2(x(r))\,dr = F(x(0)) \quad \forall t \ge 0.$$

Lemma. *If F is convex, l.s.c in a geodesic metric space, we have the upper gradient property*

$$F(x(0)) \le F(x(t)) + \int_0^t |x_r'||\nabla^- F|(x(r))\,dr.$$

As a consequence, (EDE) and (EDI) are equivalent for F, $|x_t'| = |\nabla^- F|(x(t))$ for a.e. $t > 0$, $t \mapsto F(x(t))$ is locally a.c. in $(0,\infty)$, with derivative $-|\nabla^- F|^2(x(t))$.

Lott-Sturm-Villani spaces

In these spaces (I consider only the case $N = \infty$, $K = 0$) one requires convexity along geodesics, namely for all $\mu_0,\ \mu_1 \in D(\mathrm{Ent}_{\mathfrak{m}})$ there *exists* a constant speed geodesic μ_t satisfying

$$\mathrm{Ent}_{\mathfrak{m}}(\mu_t) \le (1-t)\mathrm{Ent}_{\mathfrak{m}}(\mu_0) + t\mathrm{Ent}_{\mathfrak{m}}(\mu_1).$$

Consequences of convexity:

• Duality formula for the slope.

$$|\nabla^{-}\mathrm{Ent}_{\mathfrak{m}}|(\mu) = \sup_{\nu\neq\mu} \frac{[\mathrm{Ent}_{\mathfrak{m}}(\mu) - \mathrm{Ent}_{\mathfrak{m}}(\nu)]^{+}}{W_2(\mu,\nu)}.$$

It implies, among other things, that $\mu \mapsto |\nabla^{-}\mathrm{Ent}_{\mathfrak{m}}|(\mu)$ is l.s.c.

• Upper gradient property. The previous formula for the slope implies a *one-sided and local* Lipschitz estimate

$$\mathrm{Ent}_{\mathfrak{m}}(\mu_t) - \mathrm{Ent}_{\mathfrak{m}}(\mu_t) \le |\nabla^{-}\mathrm{Ent}|(\mu_t) W_2(\mu_t,\mu_s).$$

This, together with the l.s.c. of $\mathrm{Ent}_{\mathfrak{m}}$, can be used to show that $t \mapsto \mathrm{Ent}_{\mathfrak{m}}(\mu_t)$ is absolutely continuous and the upper gradient property.

Fisher bounds slope

Proposition. *In a LSV space, assume that* $\rho \in L^1(X, \mathfrak{m})$ *is a probability density with* $\sqrt{\rho} \in D(\mathsf{C}h)$*. Then*

$$|\nabla^- \mathrm{Ent}_{\mathfrak{m}}|^2(\rho\mathfrak{m}) \le \int_{\{\rho>0\}} \frac{|\nabla \rho|_*^2}{\rho}\, d\mathfrak{m} \left(= 4 \int |\nabla \sqrt{\rho}|_*^2\, d\mathfrak{m} \right).$$

It is precisely this inequality that prevents, in *LSV* spaces, triviality of the theory!

Proof. By approximation (recall that $\mathsf{C}h$ is defined by approximation with Lipschitz functions and that $|\nabla^- \mathrm{Ent}_{\mathfrak{m}}|$ is l.s.c.) we can assume that $\sqrt{\rho} \in \mathrm{Lip}(Z)$.

By truncation, also that $c^{-1} \ge \sqrt{\rho} \ge c > 0$, so that $\log \rho \in \mathrm{Lip}(X)$.

Let us consider another density η and an optimal plan $\boldsymbol{\pi}_\eta$ from ρ to η.

SCUOLA NORMALE SUPERIORE

Fisher bounds slope

Then

$$\begin{aligned}\mathrm{Ent}_{\mathfrak{m}}(\rho\mathfrak{m}) - \mathrm{Ent}_{\mathfrak{m}}(\eta\mathfrak{m}) &\leq \int \log\rho(\rho-\eta)\,d\mathfrak{m} = \int \log\rho(x) - \log\rho(y)\,d\boldsymbol{\pi}_\eta \\ &\leq \int \big(|\nabla^- \log\rho|(x) + \omega_x(y)\big)\,d(x,y)\,d\boldsymbol{\pi}_\eta(x,y) \\ &\leq W_2(\eta\mathfrak{m},\rho\mathfrak{m})\left(\int (|\nabla^- \log\rho|(x) + \omega_x(y))^2\,d\boldsymbol{\pi}_\eta\right)^{1/2}\end{aligned}$$

where $\omega_x(y)$ is a uniformly bounded modulus of continuity with $\omega_x(x) = 0$.

Dividing both sides by $W_2(\eta\mathfrak{m},\rho\mathfrak{m})$ and letting $\eta\mathfrak{m} \to \rho\mathfrak{m}$ gives the result, by the convergence of $\boldsymbol{\pi}_\eta$ to the identity plan, concentrated on the diagonal.

Identification of gradient flows

Traditional strategy.

$$\begin{cases} \{\text{ g.f. of } \mathrm{Ent}_{\mathfrak{m}}\} \subset \{\text{g.f. of } \mathrm{Dir}\ \} \\ \\ \text{Uniqueness of g.f. of } \mathrm{Dir} \end{cases} \Longrightarrow \ =$$

New strategy.

$$\begin{cases} \{\text{ g.f. of } \mathrm{Dir}\} \subset \{\text{g.f. of } \mathrm{Ent}_{\mathfrak{m}}\ \} \\ \\ \text{Uniqueness of g.f. of } \mathrm{Ent}_{\mathfrak{m}} \end{cases} \Longrightarrow \ =$$

This is possible thanks to the recent uniqueness result of Gigli, a surprising result because Otha-Sturm show that there is no contractivity of W_2 in LSV spaces.

SCUOLA NORMALE SUPERIORE

Proof of identification of gradient flows

We want to show that any L^2 heat flow $f_t := \mathbf{h}_t f_0$ (with f_0 probability density) is a W_2-gradient flow with $\mu_t := f_t\mathfrak{m}$, i.e.

$$\int f_t \log f_t \, d\mathfrak{m} + \int_0^t \frac{1}{2}|\dot{\mu}_r|^2 + \frac{1}{2}|\nabla^-\mathrm{Ent}_\mathfrak{m}|^2(\mu_r)\, dr \le \int f_0 \log f_0 \, d\mathfrak{m}.$$

Indeed, Kuwada lemma and Fisher bounds slope give for a.e. t

$$|\dot{\mu}_t|^2 \le \int_{\{f_t>0\}} \frac{|\nabla f_t|_*^2}{f_t}\, d\mathfrak{m}, \qquad |\nabla^-\mathrm{Ent}_\mathfrak{m}|^2(f_t\mathfrak{m}) \le \int_{\{f_t>0\}} \frac{|\nabla f_t|_*^2}{f_t}\, d\mathfrak{m}$$

while the Hilbertian energy dissipation gives

$$\frac{d}{dt}\int f_t \log f_t = -\int_{\{f_t>0\}} \frac{|\nabla f_t|_*^2}{f_t}\, d\mathfrak{m} \qquad \text{for a.e. } t > 0.$$

Proof of identification of gradient flows

A byproduct of the proof above is that all inequalities should be equalities (so that L^2 gradient flows provide (*EDE*) solutions, even though Gigli's result applies to the larger class of (*EDI*) solutions), so that the energy dissipation rates are equal a.e.:

$$|\nabla^- \mathrm{Ent}_{\mathfrak{m}}|^2(f_t\mathfrak{m}) = \int_{\{f_t>0\}} \frac{|\nabla f_t|_*^2}{f_t}\, d\mathfrak{m} \qquad \text{for a.e. } t>0.$$

By letting $t \downarrow 0$, this can be used to show that Fisher *coincides* with slope:

$$|\nabla^- \mathrm{Ent}_{\mathfrak{m}}|^2(f\mathfrak{m}) = \int_{\{f>0\}} \frac{|\nabla f|_*^2}{f}\, d\mathfrak{m}.$$

Metric Sobolev spaces

Let's start from the Euclidean case. We discuss only the case $W^{1,2}$, although all $W^{1,p}$ spaces $1 \leq p \leq \infty$ and even the BV spaces could be treated.

$$W^{1,2}(\mathbb{R}^n) := \left\{ u \in L^2(\mathbb{R}^n) : \frac{\partial u}{\partial x_i} \in L^2(\mathbb{R}^n),\ 1 \leq i \leq n \right\}$$

$$H^{1,2}(\mathbb{R}^n) := \left\{ \text{completion of } C^\infty \cap W^{1,2} \text{ for the } W^{1,2} \text{ norm} \right\}$$

"$H = W$" theorem by Meyers-Serrin in 1960, even in any open domain.

Levi's definition, 1906, $n = 2$. $u \in BL^{1,2}(\mathbb{R}^2)$ if:

(a) for a.e. x, $u(x, \cdot)$ is a.c. in $\mathbb{R}$ and, for a.e. y, $u(\cdot, y)$ is a.c. in $\mathbb{R}$.

(b) $\int |\frac{\partial u}{\partial x}|^2 + |\frac{\partial u}{\partial y}|^2 dxdy < \infty$.

SCUOLA NORMALE SUPERIORE

Metric Sobolev spaces

Theorem. *$BL^{1,2}(\mathbb{R}^n) \subset W^{1,2}(\mathbb{R}^n)$. In addition, any $u \in W^{1,2}(\mathbb{R}^n)$ has a version (for instance $\limsup_\epsilon u * \rho_\epsilon$) in $BL^{1,2}(\mathbb{R}^2)$.*

In metric spaces, the W definition roughly corresponds to the Cheeger's energy (with $\mathrm{Lip}(X)$ playing the role of C^∞), while Levi's definition corresponds to the Shanmugalingam's notion of Newtonian space $N^{1,2}(X, d, \mathfrak{m})$.

Definition. Given $\Gamma \subset AC([0,1]; X)$ we define

$$\mathrm{Mod}_2(\Gamma) := \inf \left\{ \left(\int g^2 \, d\mathfrak{m} \right)^{1/2} \right\},$$

where the infimum runs among all Borel functions $g : X \to [0, \infty]$ satisfying $\int_\gamma g \geq 1$ for all $\gamma \in \Gamma$.

Metric Sobolev spaces

We define $N^{1,2}(X,d,\mathfrak{m})$ by

$$\left\{f : X \to \mathbb{R} :\ \left|\int_{\partial\gamma} f\right| \le \int_\gamma G \text{ for } \mathrm{Mod}_2\text{-a.e. } \gamma,\ \int G^2\,d\mathfrak{m} < \infty\right\}$$

and $|\nabla f|_S$ as the G with smallest L^2 norm.

Absolute continuity lemma. *Any $f \in N^{1,2}(X,d,\mathfrak{m})$ is a.c. along Mod_2-a.e. curve γ.*

Proof. Let Γ be the set of curves γ where the u.g. property with $|\nabla f|_S$ does not hold,

$$\Gamma_1 := \{\gamma :\ \gamma \supset \gamma' \in \Gamma\}$$

$$\Gamma_2 := \left\{\gamma :\ \int_\gamma |\nabla f|_S = \infty\right\}$$

Γ_1 is Mod_2-negligible because Γ is, Γ_2 is Mod_2-negiglible by the inequality

$$\mathrm{Mod}_2\left(\Gamma_2 \cap \{\int_\gamma |\nabla f|_S = \infty\}\right) \le \frac{1}{n^2}\int |\nabla f|_S^2\,d\mathfrak{m} \to 0.$$

SCUOLA NORMALE SUPERIORE

Metric Sobolev spaces

If $\gamma \notin (\Gamma_1 \cup \Gamma_2)$ we have

$$\left|\int_{\partial\gamma'} f\right| \leq \int_{\gamma'} |\nabla f|_S \leq \int_{\gamma} |\nabla f|_S < \infty \qquad \forall \gamma' \subset \gamma$$

which yields immediately the absolute continuity property of $t \mapsto f(\gamma(t))$ for Mod_2-a.e. γ.
The gradient $|\nabla f|_S$ has pointwise minimality properties analogous to $|\nabla f|_*$, in particular if G satisfies the weak upper gradient property

$$\left|\int_{\partial\gamma} f\right| \leq \int_{\gamma} G \qquad \text{for } \mathrm{Mod}_2\text{-a.e. curve } \gamma$$

then $|\nabla f|_S \leq G$ $\mathfrak{m}$-a.e. in X.

Identification of weak gradients

Are the gradients $|\nabla f|_*$, $|\nabla f|_S$ equal? While the first gradient is relevant in connection with the L^2 heat flow and the "vertical" derivative, the second one is relevant in connection with the derivative of Ent_m and the "horizontal" derivative.

If we assume doubling and Poincaré then we can approximate any $f \in N^{1,2}(X, d, \mathfrak{m})$ (Semmes, Cheeger) in the strong norm and even in the Lusin sense by Lipschitz maps. This leads to the equality of gradients.

With "optimal transportation tools" we can provide the equivalence of gradients without doubling & Poincaré. This requires an approximation by Lipschitz functions f_n in the *weak* topology, namely

$$\limsup_{n\to\infty} \int |\nabla f_n|^2 \, d\mathfrak{m} \leq \int |\nabla f|_S^2 \, d\mathfrak{m}, \qquad \int |f_n - f|^2 \, d\mathfrak{m} \to 0.$$

Notice that as soon as we know that the Sobolev spaces are Hilbertian, this yields approximation also in the strong topology (while without doubling & Poincaré the Lusin approximation seems to be out of reach).

Two auxiliary results

Lemma. *If* $\eta \in \mathscr{P}(C([0,1];X))$ *concentrated on* $AC^2([0,1];X)$ *has bounded time marginals, then*

$$[\eta(\Gamma)]^2 \leq C(\eta)\left(\int\int_0^1 |\dot\gamma_s|^2\,ds\,d\eta(\gamma\right)\mathrm{Mod}_2(\Gamma) \qquad \forall\Gamma\subset AC^2([0,1];X).$$

In particular $\mathrm{Mod}_2(\Gamma)=0$ *implies* $\eta(\Gamma)=0$*.*

Proof. If g is admissible for Γ we have

$$[\eta(\Gamma)]^2 \leq \left(\int_0^1\int g(\gamma_s)|\dot\gamma_s|\,ds\,d\eta(\gamma)\right)^2.$$

Then, it suffices to apply Hölder and to minimize w.r.t. g.

Superposition principle. ([AGS], Lisini, Calc. Var. & PDE) *Let* $(\mu_t)_{t\in[0,T]}\subset\mathscr{P}(X)$ *be absolutely continuous with* L^2*-integrable metric speed. Then there exists* $\eta\in\mathscr{P}(C([0,1];X))$ *concentrated on* $AC^2([0,1];X)$ *and satisfying*

(1) $\mu_t=(e_t)_\sharp\eta$ *for all* $t\in[0,T]$*;*

(2) $|\dot\mu_t|^2=\int|\dot\gamma_t|^2\,d\eta(\gamma)$ *for a.e.* $t\in(0,T)$*.*

Proof of the equivalence

Lemma. (Stability of weak upper gradients) *If* $f_n \to f$ *in* $L^2(X,\mathfrak{m})$, $G_n \to G$ *weakly in* $L^2(X,\mathfrak{m})$ *and*

$$\left|\int_{\partial\gamma} f_n\right| \leq \int_\gamma G_n \qquad \text{for } \mathrm{Mod}_2\text{-a.e. curve } \gamma,$$

then there is a version $\tilde f$ *of* f *satisfying*

$$\left|\int_{\partial\gamma} \tilde f\right| \leq \int_\gamma G \qquad \text{for } \mathrm{Mod}_2\text{-a.e. curve } \gamma.$$

Using this lemma with f_n equal to the optimal sequence in the definition of Ch and $G_n = |\nabla f_n|$, weakly convergent to $|\nabla f|_*$, we obtain

$$|\nabla f|_S \leq G = |\nabla f|_* \qquad \mathfrak{m}\text{-a.e. in } X.$$

The proof of the converse inequality is constructive: we need Lipschitz functions f_n satisfying $f_n \to f$ in $L^2(X,\mathfrak{m})$ and

$$\limsup_{n\to\infty} \int |\nabla f_n|_*^2 \, d\mathfrak{m} \leq \int |\nabla f|_S^2 \, d\mathfrak{m}.$$

SCUOLA NORMALE SUPERIORE

Proof of the equivalence

Suffices to find $f_n \in D(Ch)$ satisfying $\limsup_n \int |\nabla f_n|_*^2 \, d\mathfrak{m} \le \int |\nabla f|_S^2 \, d\mathfrak{m}$.
By a truncation argument, $0 < c \le f \le c^{-1} < \infty$ and by homogeneity $\int f^2 \, d\mathfrak{m} = 1$.
We set $k = f^2$, $k_t = \mathbf{h}_t f^2$, $\mu_t = k_t \mathfrak{m} \in \mathscr{P}(X)$. Then

$$
\begin{aligned}
& \int k \log k - k_t \log k_t \, d\mathfrak{m} \\
\le \ & \int \log k \rho (k - k_t) \, d\mathfrak{m} = \int \log k(\gamma_0) - \log k(\gamma_t) \, d\eta(\gamma) \\
\le \ & \int \int_0^t |\nabla \log k|_S(\gamma_s) |\dot\gamma_s| ds d\eta(\gamma) \\
\le \ & \left(\int_0^t \int |\nabla \log k|_S^2 (\gamma_s) d\eta ds \right)^{1/2} \left(\int_0^t \int |\dot\gamma_s|^2 \, d\eta(\gamma) ds \right)^{1/2} \\
\le \ & \frac{1}{2} \int_0^t \int |\nabla \log k|_S^2 h_s d\mathfrak{m} ds + \frac{1}{2} \int |\dot\mu_s|^2 \, ds.
\end{aligned}
$$

Proof of the approximation

By the Kuwada lemma we get

$$\int k\log k - k_t\log k_t\,d\mathfrak{m}$$
$$\leq \frac{1}{2}\int_0^t\int |\nabla\log k|_S^2 h_s d\mathfrak{m}ds + \frac{1}{2}\int_0^t\int_{\{k_s>0\}}\frac{|\nabla k_s|_*^2}{k_s}\,d\mathfrak{m}ds.$$

The entropy dissipation formula then gives

$$\int_0^t\int_{\{k_s>0\}}\frac{|\nabla k_s|_*^2}{k_s}\,d\mathfrak{m}ds \leq \int_0^t\int |\nabla\log k|_S^2 k_s d\mathfrak{m}ds,$$

so that the identity $|\nabla\log k|_S = |\nabla k|_S/k = 2|\nabla f|_S/f$ we get

$$\frac{4}{t}\int_0^t \mathsf{C}h(\sqrt{k_s})\,ds \leq \frac{4}{t}\int_0^t\int\frac{|\nabla f|_S^2}{f^2}k_s\,d\mathfrak{m}ds.$$

Letting $t\downarrow 0$ and using the w^*-convergence in $L^\infty(X,\mathfrak{m})$ of k_s to $k=f^2$ gives the result.

Strong $CD(0,\infty)$ condition and *EVI* gradient flows

For simplicity I will talk about the case $\mathrm{Ric} \geq 0$ only ($K = 0$).
Definition. We say that $(X, d, \mathfrak{m})$ is a *strong* $CD(0,\infty)$ space if for all $\mu_0,\ \mu_1 \in \mathscr{P}(X)$ with finite entropy there exists an optimal geodesic plan $\boldsymbol{\pi}$ between them satisfying

$$\mathrm{Ent}_{\mathfrak{m}}\big((e_t)_\sharp(h\boldsymbol{\pi})\big) \leq (1-t)\mathrm{Ent}_{\mathfrak{m}}\big((e_0)_\sharp(h\boldsymbol{\pi})\big)+t\mathrm{Ent}_{\mathfrak{m}}\big((e_1)_\sharp(h\boldsymbol{\pi})\big) \quad \forall t \in [0,1],$$

whenever $h \in C_b(\mathrm{Geo}(X))$, $h \geq 0$, $\int h\, d\boldsymbol{\pi} = 1$.
Definition. In a metric space (E, d), a locally absolutely continuous curve $u(t)$ is an (*EVI*) solution to the gradient flow of $F : X \to \mathbb{R}\cup\{+\infty\}$ if for all $v \in D(F)$ it holds

$$\frac{d}{dt}\frac{1}{2}d^2(u(t),v) + F(u(t)) \leq F(v) \qquad \text{for a.e. } t \in (0,\infty).$$

This formulation of gradient flows is equivalent in Hilbert spaces, but in general *stronger* than the one based on energy dissipation.

Riemannian Ricci lower bounds

As shown by Ohta-Sturm, all Minkowski spaces ($\mathbb{R}^n$ endowed with the Lebesgue measure and any norm $\|\cdot\|$) satisfy the $CD(0,\infty)$ condition.
Question. Is there a more restrictive notion, still stable and (strongly) consistent with the Riemannian case, that rules out Minkowsky spaces? The answer is yes.
Definition. [AGS] We say that $(X, d, \mathfrak{m})$ has *Riemannian Ricci curvature* bounded from below if one of the following equivalent conditions hold:

(i) $(X, d, \mathfrak{m})$ is a strong $CD(K, \infty$ space and the L^2 heat flow $\mathbf{h}_t$ is linear;

(ii) $(X, d, \mathfrak{m})$ is a strong $CD(K, \infty$ space and the W_2 heat flow $\mathbf{H}_t$ is additive (i.e. convex and concave) on $\mathscr{P}(X)$;

(iii) $\mathbf{H}_t\mu$ is a gradient flow in the *EVI* sense.

One of the main results is indeed the proof of this equivalence.

Properties of $RCD(K,\infty)$ spaces

• Strong consistency with the Riemannian case and stability under Gromov-Hausdorff limits.

• Tensorization: if $(X, d_X, \mathfrak{m}_X)$ and $(Y, d_Y, \mathfrak{m}_Y)$ are $RCD(0,\infty)$, so is $(X\times Y, \sqrt{d_X^2+d_Y^2}, \mathfrak{m}_X\times\mathfrak{m}_Y)$.

• The heat flow $\mathbf{h}_t$ is L^2-selfadjoint, leaves $\mathrm{Lip}(X)$ invariant and regularizes $L^\infty(X,\mathfrak{m})$ to $C_b(X)$. We have also the Bakry-Emery estimate

$$|\nabla(\mathbf{h}_t f)|_*^2 \le \mathbf{h}_t|\nabla f|_*^2 \qquad \mathfrak{m}\text{-a.e. in } X.$$

• Compatibility with the theory of Dirichlet forms: the Dirichlet form

$$\mathcal{E}(u,v) := \frac{Ch(u+v)-Ch(u-v)}{4}$$

associated to the quadratic form Ch induces a distance $d_{\mathcal{E}}$ equal to d and a local energy measure $[u]$ equal to $|\nabla u|_*^2\mathfrak{m}$.

Properties of $RCD(K,\infty)$ spaces

In particular the theory of Dirichlet forms can be applied to obtain a unique (in law) *Brownian motion* in $(X,d,\mathfrak{m})$, i.e. a Markov process $\boldsymbol{X}_t$ with continuous sample paths satisfying

$$\mathbf{P}(\boldsymbol{X}_t|\boldsymbol{X}_0=x)=\mathbf{H}_t\delta_x \qquad \forall x\in X,\ t\geq 0.$$