

Week 11

Def: Let X be a geodesic metric space, $F: X \rightarrow \mathbb{R} \cup \{+\infty\}$, $\lambda \in \mathbb{R}$. F is λ -geodesically convex if $\forall x, y \in X$, $\exists \eta: [0,1] \rightarrow X$ constant speed geodesic between x and y such that: $F \circ \eta: [0,1] \rightarrow \mathbb{R}$ is λ -convex, i.e.,

$$F(\eta(t)) \leq (1-t)F(x) + tF(y) - \frac{\lambda}{2} t(1-t) d^2(x,y) \quad \forall t \in [0,1].$$

Theorem: $V: \mathbb{R}^d \rightarrow \mathbb{R}^+ \cup \{+\infty\}$ is λ -convex $\Leftrightarrow V: \mathcal{P}_2(X) \rightarrow \mathbb{R}^+ \cup \{+\infty\}$ is λ -convex

Theorem: We saw some examples where U convex:

① $U(z) = z^\alpha, \alpha > 1$ ② $U(z) = z \log z$ ③ $U(z) = -z^\alpha, 1 \geq \alpha \geq 1 - \frac{1}{d}$

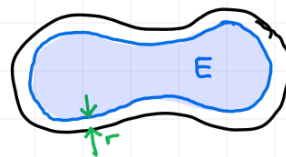
Theorem (Brunn-Minkowski): $A, B \subseteq \mathbb{R}^d$ compact. $\lambda^d(A+B)^{1/d} \geq \lambda^d(A)^{1/d} + \lambda^d(B)^{1/d}$

From Brunn-Minkowski to isoperimetric inequality:

Theorem: $P(E) \geq P(B_1)$ $\forall E \in \mathcal{C}^2$ set with $\lambda^d(E) = \lambda^d(B_1)$.

I can take it as a definition

$$\leftarrow \liminf_{r \downarrow 0} \frac{\lambda^d(E+B_r) - \lambda^d(E)}{r}$$



If $E \in \mathcal{C}^2$ set $\rightarrow P(E)$ coincides with the surface area of ∂E .

Proof: $P(E) \geq \liminf_{r \rightarrow 0} \frac{[\lambda^d(E)^{1/d} + r \lambda^d(B_1)^{1/d}]^d - \lambda^d(E)}{r} = d \lambda^d(E)^{\frac{d-1}{d}} \frac{1}{r} \cdot \lambda^d(B_1)^{1/d} = d \lambda^d(B_1)^{1/d} = P(B_1).$

Lower semicontinuity of functionals (see exercises for $\int V d\mu$ & $\int W(x,y) d\mu(x) d\mu(y)$)

Theorem: Let U with $\frac{U(z)}{z} \rightarrow +\infty$ as $z \rightarrow +\infty$ and (HP1).

Then, $U = \begin{cases} \int U(\rho) dx & \text{if } \mu = \rho dx \ll \lambda^d \\ +\infty & \text{otherwise} \end{cases}$ is l.s.c. in $\mathcal{P}_2(\mathbb{R}^d)$ wrt narrow convergence (and hence also W_2 -convergence).

Proof: Not done (see books Villani or Santambrogio).

Talagrand inequality (about relative entropy wrt a Gaussian)

Given $\mu, \lambda \in \mathcal{P}(X)$, we define the relative entropy of μ wrt λ : $\rightarrow$ Radon-Nikodym derivative

$$\text{Ent}_\lambda(\mu) := \begin{cases} \int \rho \log \rho d\lambda & \text{if } \mu \ll \lambda, \mu = \rho \lambda \\ +\infty & \text{if } \mu \not\ll \lambda. \end{cases}$$

Theorem: $\frac{1}{2} W_2^2(\mu, \gamma_d) \leq \text{Ent}_{\gamma_d}(\mu) \quad \forall \mu \in \mathcal{P}_2(\mathbb{R}^d)$

$\frac{1}{(2\pi)^{d/2}} e^{-\frac{|x|^2}{2}}$ Gaussian

- particularities: constant independent of the dimension

- Originally: • $d=1$, use explicitly the OT map (monotone rearrangement).

↳ the inequality is stable under tensorization: the ineq. holds in $(\mathbb{R}^d, 1 \cdot 1_d, \gamma_d)$ and in $\mathbb{R} \Rightarrow$ it holds in $(\mathbb{R}^{d+1}, 1 \cdot 1_{d+1}, \gamma_{d+1})$.

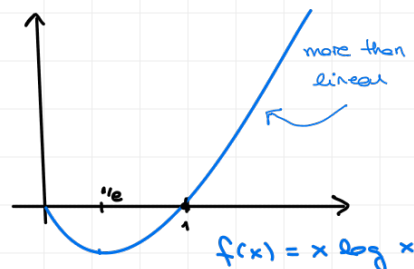
Proof:

① $\text{Ent}_\lambda(\mu) \geq 0$ with equality $\Leftrightarrow \mu = \lambda$.

$$s \mapsto s \log s \text{ convex, } (s \log s)' = 1 + \log s$$

$$(s \log s)'' = \frac{1}{s} > 0 \text{ convex}$$

→ it has a minimum at $\frac{1}{e}$.



By Jensen's inequality, $\int p \log p \lambda \geq \int p \lambda \log \int p \lambda = 0$

with equality $\Leftrightarrow p$ constant.

② $\mu = p d\gamma_d = f dx$ where $f = p G_d$, $\gamma_d = G_d dx = \frac{1}{(2\pi)^{d/2}} e^{-\frac{|x|^2}{2}} dx$

$$\int p \log p d\gamma_d = \int \frac{f}{G_d} \log \frac{f}{G_d} G_d dx$$

1-convex, since $x \mapsto \frac{|x|^2}{2}$ 1-convex in $\mathbb{R}^d$

$$= \underbrace{\int f \log f dx}_{\text{geodesic. convex}} + \underbrace{\int f \frac{|x|^2}{2} dx - \log(2\pi)^{d/2} \int f dx}_{\text{constant}}$$

→ $\text{Ent}_{\gamma_d}(\cdot)$ is 1-convex in $\mathcal{P}_2^a(\mathbb{R}^d)$.

③ Remark in $\mathbb{R}$: $f: \mathbb{R} \rightarrow \mathbb{R}$ C^2 , λ -convex, $\lambda > 0$, minimum at 0 $\stackrel{!}{\geq} \lambda$

$\Rightarrow f(x) \geq \frac{\lambda}{2} |x|^2 + f(0)$. Indeed, $f(x) = f(0) + f'(0)x + \frac{x^2}{2} f''(\xi)$ $\xi \in [0, x]$

More in general:

Lemma: (X, d) geodesic space, $\lambda > 0$, $\phi: X \rightarrow \mathbb{R} \cup \{+\infty\}$ is λ -convex with minimum in $x_{\min}$. Then: $\phi(x) - \phi(x_{\min}) \geq \frac{\lambda}{2} d^2(x, x_{\min})$.

With this lemma, with $\phi = \text{Ent}_{\gamma_d}(\cdot)$ in $(\mathcal{P}_2^a(\mathbb{R}^d), W_2)$, we conclude.

Proof of lemma: Let $\gamma(t)$ be the geodesic between $x_{\min}$ and x , $t \in [0, 1]$

$$\phi(x_{\min}) \leq \phi(\gamma(t)) \leq \underbrace{(1-t)\phi(x_{\min}) + t\phi(x)}_{\text{Jensen}} - \frac{\lambda}{2} t(1-t) d^2(x, x_{\min})$$

→ $\cancel{\phi(x_{\min})} \leq \cancel{\phi(x)} + \frac{\lambda}{2} \cancel{t(1-t)} d^2(x, x_{\min})$, let $t \downarrow 0$. □

Gradient flows in $\mathbb{R}^d$ and in Hilbert spaces

Let $H = \mathbb{R}^d$ or a Hilbert space. Let $\phi: H \rightarrow \mathbb{R}$ $\mathcal{C}^1$.

The gradient flow of ϕ , $x: [0, +\infty) \rightarrow H$, starting at x_0 is:

$$(GF) \begin{cases} \dot{x}(t) = -\nabla \phi(x(t)) \\ x(0) = x_0 \end{cases}$$

- $\phi(x(t))$ is decreasing: $\frac{d}{dt} \phi(x(t)) = \nabla \phi(x(t)) \cdot (-\nabla \phi(x(t))) = -|\nabla \phi(x(t))|^2 \leq 0$
→ if it is constant, we are in a stationary point.

In particular, if ϕ convex has a unique minimizer, we expect that $x(t)$ converges towards it.

In a Hilbert space, $\dot{x}(t) = \lim_{h \rightarrow 0} \frac{x(t+h) - x(t)}{h} \in H$; if ϕ is $\mathcal{C}^1(H)$,

$$\phi(x + \varepsilon v) = \phi(x) + \underbrace{d\phi(x)}_{\text{linear functional, can be represented as } \langle \nabla \phi, v \rangle} \varepsilon + o(\varepsilon)$$

If $\nabla \phi$ is Lipschitz, Cauchy-Lipschitz or Picard-Lindelöf theorems

guarantee the $\exists!$ of solution from any initial datum → they require too much regularity

The implicit Euler scheme

Given $\tau > 0$, we discretize the derivative: $\frac{x(t+\tau) - x(t)}{\tau} + \nabla \phi(x(t+\tau)) = 0$

Rewrites as:
$$\nabla \left(\underbrace{\frac{\|x - x(t)\|^2}{2\tau} + \phi(x)}_{\psi(x)} \right)(x(t+\tau)) = 0 \quad (*)$$

If ϕ convex → ψ is convex and a solution to $(*)$ is a global minimizer of ψ .

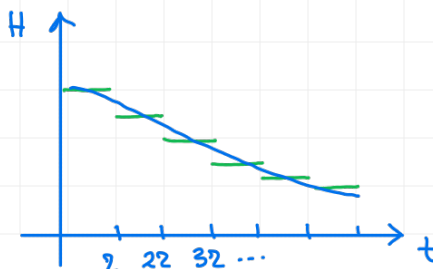
↪ We now consider a minimizing movement scheme from x_0 : the curve

$x: [0, \infty) \rightarrow H$, x piecewise constant on intervals of length τ ,

$$x(t) = x_k^\tau \quad \text{if } t \in [k\tau, (k+1)\tau),$$

where $\begin{cases} x_0^\tau = x_0 \\ x_{k+1}^\tau \text{ solves } \end{cases}$

$$x_{k+1}^\tau \text{ solves } \frac{x_{k+1}^\tau - x_k^\tau}{\tau} = -\nabla \phi(x_{k+1}^\tau)$$



Gradient flow of convex, lsc function ϕ (we drop regularity!)

Let $\phi: H \rightarrow \mathbb{R} \cup \{+\infty\}$ convex lsc.

A gradient flow of ϕ is an absolutely continuous curve $x: [0, \infty) \rightarrow H$

s.t.
$$\begin{cases} \dot{x}(t) \in -\partial\phi(x(t)) \\ x(0) = x_0 \end{cases} \quad \begin{array}{l} \hookrightarrow \text{continuous curve, differentiable a.e. with} \\ |\dot{x}| \in L^1_{loc}(0, \infty) \text{ and } x(t) - x(s) = \int_s^t \dot{x}(z) dz \end{array}$$

By analogy we can construct discrete solutions with implicit Euler schemes:

we set:
$$\begin{cases} x_0^2 = x_0 \\ x_{k+1}^2 \text{ obtained from } x_k^2 \text{ as a solution to } \frac{x_{k+1}^2 - x_k^2}{2} \in -\partial\phi(x_{k+1}^2) \end{cases}$$

$$\begin{aligned} \partial\psi_k^2 \ni 0 &\Leftrightarrow \frac{x_{k+1}^2 - x_k^2}{2} + \partial\phi(x_{k+1}^2) \ni 0 \\ \rightarrow \psi_k^2(x) &= \frac{\|x - x_k^2\|^2}{22} + \phi(x). \end{aligned}$$

The equivalences follow from the fact that if f convex, $g \in C' \Rightarrow \partial(f+g)(x) = \partial f(x) + \nabla g(x)$.

We can now define a curve as before. There are deep theorems which allow to let $2 \downarrow 0$ to find $x^2(t)$ (piecewise constant) converge to a gradient flow.

Lemma: Given $\bar{x}$, let $\psi(x) = \frac{|x - \bar{x}|^2}{22} + \phi(x)$. Then, there $\exists$ a unique minimizer of ψ . $\hookrightarrow$ convex lsc

Proof in $\mathbb{R}^d$: Fix $p_0 \in H$ s.t. $p_0 \in \partial\phi(x_0)$:

$$\phi(x) \geq \phi(x_0) + \langle p_0, x - x_0 \rangle \geq -A|x| - B. \quad \rightarrow +\infty \text{ as } |x| \rightarrow +\infty$$

$$\text{Hence, } \psi(x) = \frac{|x - \bar{x}|^2}{22} + \phi(x) \geq \frac{|x - \bar{x}|^2}{22} - A|x| - B$$

If $(x_j)_j$ is a minimizing sequence, $\psi(x_j) \rightarrow \min \psi$, $|x_j| \leq C \forall j$, up to a subseq. we have $x_j \rightarrow \tilde{x}$ (this is not true in Hilbert spaces! In reality is wrt weak convergence $x_j \rightarrow \tilde{x}$ by Banach-Alaoglu).

By lsc of ψ , $\min \psi = \lim_j \psi(x_j) \geq \psi(\tilde{x})$. Hence $\tilde{x}$ is a min.

Uniqueness from strict convexity: $\psi\left(\frac{x_1 + x_2}{2}\right) < \frac{\psi(x_1) + \psi(x_2)}{2} = \min$
 $\uparrow$ strict convexity $\hookrightarrow$ if $x_1 \neq x_2$ both min. □