Week 11

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

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

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

Theorem: We saw some examples where $\mathcal{U}$ convex:

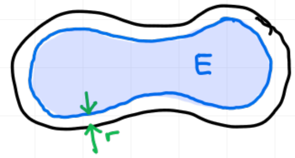
① $\mathcal{U}(z) = z^\alpha$, $\alpha > 1$ ② $\mathcal{U}(z) = z \log z$ ③ $\mathcal{U}(z) = -z^\alpha$ $1 \geq \alpha \geq 1 - \frac{1}{d}$

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

From Brunn-Minkowski to isoperimetric inequality:

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

I can take it as a definition $\leftarrow$ $\llcorner$ $\liminf_{r \downarrow 0} \dfrac{\mathcal{L}^d(E+B_r) - \mathcal{L}^d(E)}{r}$



If E C^2 set $\longrightarrow P(E)$ coincides with the surface area of ∂E.

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

Lower semicontinuity of functionals

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

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

Then, $\mathcal{U} = \begin{cases} \int \mathcal{U}(\rho) dx & \text{if } \mu = \rho dx \ll \mathcal{L}^d \\ +\infty & \text{otherwise} \end{cases}$ 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: $\frac{1}{2} W_2^2(\mu, \gamma_d) \leq \text{Ent}_{\gamma_d}(\mu)$ $\forall \mu \in \mathcal{P}_2(\mathbb{R}^d)$

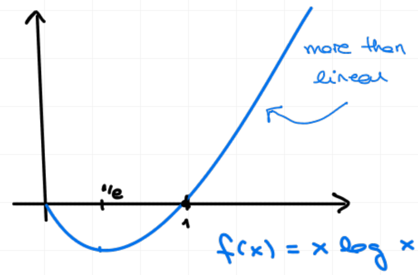
$\gamma_d = \frac{1}{(2\pi)^{d/2}} e^{-\frac{|x|^2}{2}} dx$ Gaussian

• 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 $(\mathbb{R}^d, |\cdot|_d, \gamma_d)$ and in $\mathbb{R}$ $\Longrightarrow$ it holds in $(\mathbb{R}^{d+1}, |\cdot|_{d+1}, \gamma_{d+1})$.

Proof:



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

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

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

$\longrightarrow$ it has a minimum at $\frac{1}{e}$.

By Jensen's inequality, $\int \rho \log\rho\, \lambda \geq \int \rho\lambda \log \overbrace{\int \rho \lambda}^{=1} = 0$

with equality $\Leftrightarrow$ ρ constant.

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

$$\int \rho \log \rho \, d\gamma_d = \int \frac{f}{\cancel{G_d}} \log \frac{f}{G_d} \cancel{G_d}\, dx$$

$$= \underbrace{\int f \log f\, dx}_{\text{geodesic. convex}} + \overbrace{\int f \frac{|x|^2}{2} dx}^{1\text{-convex, since } x \mapsto \frac{|x|^2}{2} \text{ 1-convex in } \mathbb{R}^d} - \overbrace{\log (2\pi)^{d/2} \overbrace{\int f\, dx}^{=1}}^{\text{constant}}$$

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

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

$\Longrightarrow$ $f(x) \geq \frac{\lambda}{2}|x|^2 + f(0)$. Indeed, $f(x) = f(0) + f'(0)x + \frac{x^2}{2} \overbrace{f''(\xi)}^{\geq \lambda}$

$\xi \in [0, x]$

More in general:

Lemma: (X, d) geodesic space, $\lambda > 0$, $\phi: X \longrightarrow \mathbb{R} \cup \{+\infty\}$ 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 $\phi = \mathrm{Ent}_{\gamma_d}(\cdot)$ 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, $t \in [0,1]$

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

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

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

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

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

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

- $\phi(x(t))$ is decreasing: $\frac{d}{dt}\phi(x(t)) = \nabla\phi(x(t)) \cdot \left(-\nabla\phi(x(t))\right) = -|\nabla\phi(x(t))|^2 \leq 0$

  → 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, $\dot{x}(t) = \lim \dfrac{x(t+h) - x(t)}{h} \in H$; if ϕ is $C^1(H)$,

$$\phi(x + \varepsilon v) = \phi(x) + \underbrace{d\phi(x)}\,\varepsilon v + o(\varepsilon)$$

linear functional, can be represented as $\langle \nabla\phi, \varepsilon v \rangle$

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: $\dfrac{x(t+\tau) - x(t)}{\tau} + \nabla\phi(x(t+\tau)) = 0$

→ without this term, this is called 'Explicit Euler Scheme'

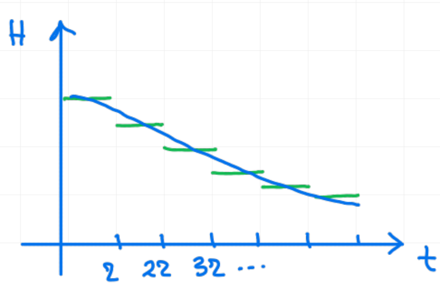
Rewrites as: $$\nabla\Bigg(\underbrace{\frac{\|x - x(t)\|^2}{2\tau} + \phi(x)}_{\psi(x)}\Bigg)(x(t+\tau)) = 0 \qquad (*)$$

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

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

where $\begin{cases} x_0^\tau = x_0 \\ x_{k+1}^\tau \text{ solves } \dfrac{x_{k+1}^\tau - x_k^\tau}{\tau} = -\nabla\phi(x_{k+1}^\tau) \end{cases}$



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

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

A <u>gradient flow of ϕ</u> is an absolutely continuous curve $x : [0,\infty) \to H$

s.t. $\begin{cases} \dot{x}(t) \in -\partial\phi(x(t)) \\ x(0) = x_0 \end{cases}$

↳ continuous curve, differentiable a.e. with $|\dot{x}| \in L^1_{loc}(0,+\infty)$ and $x(t) - x(s) = \int_s^t \dot{x}(z)\,dz$

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

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

$$\partial\psi_k^\tau \ni 0 \iff \frac{x_{k+1}^\tau - x_k^\tau}{\tau} + \partial\phi(x_{k+1}^\tau) \ni 0$$

$$\psi_k^\tau(x) = \frac{\|x - x_k^\tau\|^2}{2\tau} + \phi(x).$$

The equivalences follow from the fact that if f convex, $g \in C^1 \Rightarrow \partial(f+g)(x) = \partial f(x) + \nabla g(x)$.

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

<u>Lemma</u>: Given $\bar{x}$, let $\psi(x) = \dfrac{|x-\bar{x}|^2}{2\tau} + \phi(x)$ (convex lsc). Then, there $\exists$ a unique minimizer of ψ.

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

$$\phi(x) \geq \phi(x_0) + \langle p_0, x - x_0\rangle \geq -A|x| - B.$$

Hence, $\psi(x) = \dfrac{|x-\bar{x}|^2}{2\tau} + \phi(x) \geq \underbrace{\dfrac{|x-\bar{x}|^2}{2\tau} - A|x| - B}_{\to +\infty \text{ as } |x| \to +\infty}$

If $(x_j)_j$ is a minimizing sequence, $\psi(x_j) \to \min\psi$, $|x_j| \leq C \ \forall j$, up to a subseq. we have $x_j \to \tilde{x}$ (this is not true in Hilbert spaces! In reality is wrt weak convergence $x_j \rightharpoonup \tilde{x}$ 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: $\psi\left(\dfrac{x_1+x_2}{2}\right) < \dfrac{\psi(x_1)+\psi(x_2)}{2} = \min$ ↳ if x_1 & x_2 both min.

↑ strict convexity

$\square$