PART III : GRADIENT FLOWS IN $\mathcal{P}_2(H)$

Titolo nota 06/04/2008

- Convexity properties of $\frac{1}{2} W_2^2(\cdot, \sigma)$
- Convex functionals in $\mathcal{P}_2(H)$
- Examples of Gradient flows

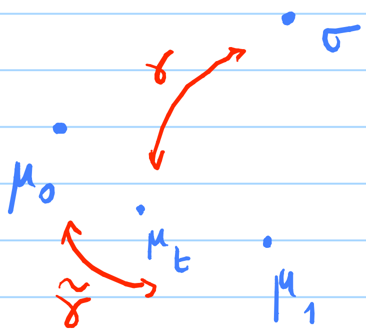
$\mathcal{P}_2(H)$ is a Positively Curved metric space, i.e.

$$\frac{1}{2} W_2^2(\mu_t, \sigma) \geq \frac{1}{2}(1-t)\, W_2^2(\mu_0, \sigma) + \frac{t}{2}\, W_2^2(\mu_1, \sigma) - \frac{t(1-t)}{2}\, W_2^2(\mu_0, \mu_1)$$

for all constant speed geodesics $\mu_t : [0,1] \longrightarrow \mathcal{P}_2(H)$.

PROOF. We use the following [illegible]: $\forall t \in [0,1]$ $\forall \tilde{\gamma} \in \Gamma_o(\mu_0,\mu_1)$ $\forall \gamma \in \Gamma_o(\mu_t,\sigma)$ there exists $\eta \in \mathcal{P}_2(\underset{\mu_0}{H} \times \underset{\mu_1}{H} \times \underset{\sigma}{H})$ with

$$(\pi^1,\pi^2)_\# \eta = \tilde{\gamma}, \qquad ((1-t)\pi^1 + t\pi^2, \pi^3)_\# \eta = \gamma$$



We use also the Hilbertian identity

$$\frac{1}{2}\left|(1-t)x_1 + t x_2) - x_3\right|^2 = \frac{(1-t)}{2}|x_1-x_3|^2 + \frac{t}{2}|x_2-x_3|^2 - \frac{t(1-t)}{2}|x_1-x_2|^2$$

to get:

$$W_2^2(\mu_t,\sigma) = \int |x_1 - x_2|^2 \, d\tilde{\gamma} = \int |(1-t)x_1 + t x_2 - x_3|^2 \, d\eta =$$

$$= \int (1-t)|x_1-x_3|^2 \, d\eta + \int t\,|x_2-x_3|^2 \, d\eta \; - \; t(1-t)\int |x_1-x_2|^2 \, d\eta$$

$$\geq \qquad\qquad \geq \qquad\qquad =$$

$$(1-t)\,W_2^2(\mu_0,\sigma) \quad + \quad t\,W_2^2(\mu_1,\sigma) \; - \; t(1-t)\,W_2^2(\mu_0,\mu_1) \qquad \blacksquare$$

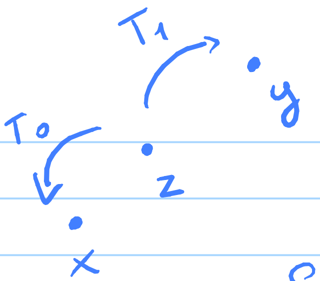
To prove the lemma, remember that $\Gamma_o(\mu_t,\mu_0) = \{(Id \times T_0)_\# \mu_t\}$

$\Gamma_o(\mu_t,\mu_1) = \{(Id \times T_1)_\# \mu_t\}$

injective!

$$(x,y) \in \operatorname{supp} \tilde{\gamma} \longrightarrow z = (1-t)x + ty \longrightarrow (x,y) =: (T_0(z), T_1(z))$$

$z \to (x,y)$ is well-defined on supp μ_t by cyclical monotonicity of supp $\tilde{\gamma}$



(**) $\quad (T_0, T_1)_\# \mu_t = \gamma$

(*) $\quad (1-t)\, T_0(z) + t\, T_1(z) = z$ on $\operatorname{supp} \mu_t$.

Set now $\alpha(z, x_3) := (T_0(z), T_1(z), x_3)$ and

$$\eta := \alpha_\#(\gamma)$$

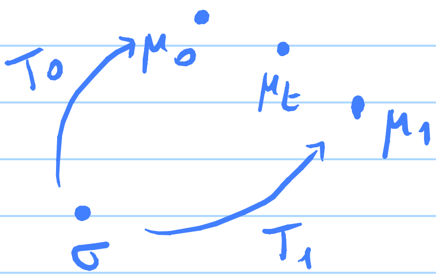
$$((1-t)\pi^1 + t\pi^2, \pi^3)_\# \eta = ((1-t)T_0 + tT_1, x_3)_\# \gamma = \gamma \quad \text{by } (*)$$

$$(\pi^1, \pi^2)_\# \eta = (T_0, T_1)_\# \gamma = (T_0, T_1)_\# \mu_t = \gamma \quad \text{by } (**)$$

∎

Since $\mathcal{P}_2(H)$ is (PC), we have to find different curves along which $\frac{1}{2} W_2^2(\cdot, \sigma)$ is 1-convex

(INTERPOLATING CURVE WITH BASE σ)



When $H = \mathbb{R}^n$ and $\sigma << \mathcal{L}^n$, we set

$$\mu_t := \left((1-t)\, T_0 + t\, T_1 \right)_{\#} \sigma$$

T_0 optimal from σ to μ_0, T_1 optimal from σ to μ_1

In general, given $\gamma_0 \in \Gamma_0(\sigma, \mu_0)$, $\gamma_1 \in \Gamma_0(\sigma, \mu_1)$, find $\eta \in \mathcal{P}(H \times H \times H)$ with $(\pi^1, \pi^2)_{\#} \eta = \gamma_0$, $(\pi^1, \pi^3)_{\#} \eta = \gamma_1$ and set

$$\mu_t := \left((1-t)\, x_2 + t\, x_3 \right)_{\#} \sigma \qquad \left[\eta = \int (\gamma_0)_{x_1}(dx_2) \times (\gamma_1)_{x_1}(dx_3)\, d\sigma(x_1) \right]$$

[redacted] $\frac{1}{2} W_2^2(\cdot, \sigma)$ is 1-convex along the interpolating curves with base σ.

PROOF Since $(x_1, (1-t)x_2 + t x_3)_\# \eta \in \Gamma(\sigma, \mu_t)$, we have

$$W_2^2(\mu_t, \sigma) \le \int |(1-t)x_2 + t x_3 - x_1|^2 \, d\eta$$

$$\begin{array}{ccccc} = (1-t) \displaystyle\int |x_2 - x_1|^2 \, d\eta & + & t \displaystyle\int |x_3 - x_1|^2 \, d\eta & - & t(1-t) \displaystyle\int |x_1 - x_2|^2 \, d\eta \\ \shortparallel & & \shortparallel & & \leq \\ (1-t) W_2^2(\sigma, \mu_0) & + & t W_2^2(\sigma, \mu_1) & - & t(1-t) W_2^2(\mu_0, \mu_1). \end{array}$$

CONVEX FUNCTIONALS IN $\mathcal{P}_2(H)$

- POTENTIAL ENERGY $\quad V : H \longrightarrow \mathbb{R} \cup \{+\infty\} \quad (\inf V > -\infty)$

$$\Phi(\mu) = \int V \, d\mu$$

(1) V λ-convex, $\lambda \geq 0$ $\Longrightarrow$ Φ λ-convex along all interpolating curves

(2) V λ-convex $\Longleftrightarrow$ Φ is λ-convex along constant speed geodesics

Proof of (1):

$$\Phi(\mu_t) = \int V \, d\mu_t = \int V((1-t)x_2 + t x_3) \, d\gamma(x_1, x_2, x_3)$$

$$\leq (1-t) \int V(x_2)\, d\eta + t \int V(x_3)\, d\eta - \frac{\lambda}{2} t(1-t) \int |x_2 - x_3|^2 d\eta$$

$$\leq (1-t)\, \Phi(\mu_0) + t\, \Phi(\mu_1) - \frac{\lambda}{2} t(1-t)\, W_2^2(\mu_0, \mu_1).$$

The proof of (2) is analogous.

- INTERACTION ENERGY $\quad W : H^k \longrightarrow \mathbb{R} \cup \{+\infty\} \quad (\inf W > -\infty)$

$$\Phi(\mu) := \int W(x_1, \ldots, x_k)\, d\overbrace{\mu \times \cdots \times \mu}^{k \text{ times}}$$

W λ-convex in H^k, $\lambda \geq 0$ $\Longrightarrow$ Φ $k\lambda$-convex

W λ-convex in H^k $\Longrightarrow$ Φ $k\lambda$-convex along geodesics

$(K=2)$: $\int W((1-t)x_1+ty_1, (1-t)x_2+ty_2)\, d\gamma(x_1,y_1)\, d\gamma(x_2,y_2) \leq$

$(1-t)\int W(x_1,x_2)\, d\gamma(x_1,y_1)\, d\gamma(x_2,y_2) + t\int W(y_1,y_2)\, d\gamma(x_1,y_1)\, d\gamma(x_2,y_2)$

$- \frac{\lambda}{2} t(1-t) \int |(y_1,y_2)-(x_1,x_2)|^2\, d\gamma(x_1,y_1)\, d\gamma(x_2,y_2)$

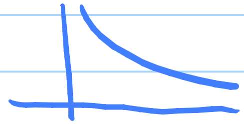
$\qquad\qquad \rotatebox{90}{=}\; \lambda t(1-t)\, W_2^2(\mu_0,\mu_1)$

- (INTERNAL ENERGY)

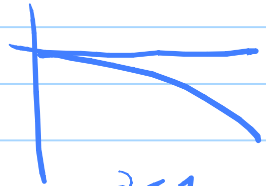
$\Phi(\rho) := \int_{\mathbb{R}^n} U(\rho)\, dx$ with $U(0)=0$ and

$t \to t^n \, U(t^{-n})$ convex $\searrow$ in $(0,+\infty)$ (*)

<u>Examples</u> • z^a, $a \geq 1$ $\qquad t^{n(1-a)}$



• $\frac{1}{a-1} z^a$ $\qquad a \geq 1 - \frac{1}{n}$ $\qquad \frac{1}{1-a} t^{(1-a)n}$



$a < 1.$

• $\frac{1}{a-1}(z^a - z) \longrightarrow z \ln z$

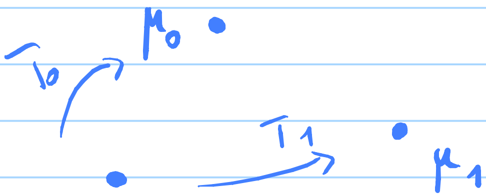
THEOREM 3. If (*) holds, Φ is convex along all interpolating curves.

PROOF. • Explicit formula for the interpolated densities

• Concavity of

$$A \longrightarrow (\det A)^{1/n}$$

in $\{A \in \mathrm{Sym}^{n \times n} \mid A \geq 0\}$ (see CAFFARELLI-CABRÉ)



$\sigma \ll \mathcal{L}^n$

$\sigma = \rho \mathcal{L}^n$

$T_0 = \nabla \varphi_0$ optimal from σ to μ_0

$T_1 = \nabla \varphi_1$ optimal from σ to μ_1

$T_t = ((1-t) T_0 + t T_1)$, $\mu_t = (T_t)_\# (\rho \mathcal{L}^n)$

$$\mu_t = \frac{\rho}{\det \nabla T_t} \circ \left(T_t \Big|_{\Sigma^c} \right)^{-1} \mathcal{L}^n \qquad (**)$$

$$\Phi(\mu_t) \underset{\text{"}y = T_t(x)\text{"}}{=} \int \det \nabla T_t \; U\left(\frac{\rho}{\det \nabla T_t} \right) dx$$

$A \longrightarrow \det A \; U\left(\frac{\rho(x)}{\det A}\right)$ is convex, being the composition of the concave map $A \to (\det A)^{1/n}$ and the convex ↘ map $z \longrightarrow z^n U(\rho(x)/z^n)$.

Let $\Gamma(x)$ be a monotone operator. Then:

(1) for $\mathcal{L}^n$-a.e. x $\quad \Gamma(x) = \{p(x)\}$ is single-valued;

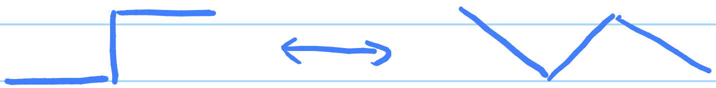
(2) for $\mathcal{L}^n$-a.e. x $\quad \exists\, A(x)$, $n \times n$ matrix, with

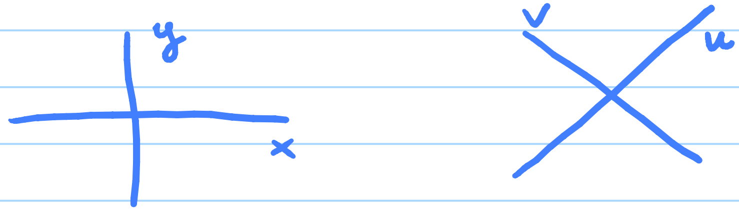
$$\lim_{\substack{y \to x \\ q \in \Gamma(y)}} \frac{|q - p(x) - A(x)(y-x)|}{|y-x|} = 0 .$$

If $\Gamma(x) = \partial\varphi(x)$, $A(x) \in \mathrm{Sym}^{n \times n}$ for a.e. x $\quad \left(A = D^2\varphi / \mathcal{L}^n \right)$

PROOF Monotone operators $\overset{\text{(MINTY)}}{\iff}$ 1-Lipschitz maps

ALEXANDROV $\iff$ RADEMACHER





$$\begin{cases} u = (y+x)/\sqrt{2} \\ v = (y-x)/\sqrt{2} \end{cases}$$

See A. - ALBERTI, '88

All these examples can be combined

$$\Phi(\mu) = \int U(\rho) + V\rho \, dx + \int W \rho(x) \rho(y) \, dx dy$$

$$\mu = \rho \mathcal{L}^n$$

and we get, by the general theory, existence, uniqueness, stability, error estimates, ...

Since $\mathcal{P}_2(\mathbb{R}^n)$ has a differentiable structure, we can go back to a differentiable description of all these results.

DEFINITION Let $F : \mathcal{P}_2(H) \to \mathbb{R} \cup \{+\infty\}$ be convex along geodesics.

$$\partial F(\mu) = \Big\{ \xi \in L^2(\mu; H) \;\Big|\; F(\nu) \geq F(\mu) + \int \langle \xi(x), y - x \rangle \, d\gamma \quad \forall \nu, \ \forall \gamma \in \Gamma_o(\mu, \nu) \Big\}$$

When γ is induced by T, we have

$$F(\nu) \geq F(\mu) + \int \langle \xi(x), T(x) - x \rangle \, d\mu(x)$$

DEFINITION μ_t is a (GF) if:

(1) $\frac{d}{dt} \mu_t + \nabla \cdot (v_t \mu_t) = 0$, with $\|v_t\|_{L^2(\mu_t)} \in L^2_{loc}(0, +\infty)$

(2) $-v_t \in \partial F(\mu_t)$ for a.e. t.

THEOREM 4. The (GF), (EVI) and (EDI) formulations are equivalent.

The proof uses, as in the Hilbertian case, the formula for the derivative of $\frac{1}{2}W_2^2(\cdot,\sigma)$ along a.c. curves:

$$\frac{d}{dt}\frac{1}{2}W_2^2(\mu_t,\sigma) = -\int \langle v_t(x), y-x\rangle \, d\gamma \qquad \forall \gamma \in \Gamma_o(\mu_t,\sigma)$$

where v_t is the tangent velocity field to μ_t.

ENTROPY AND RELATIVE ENTROPY

$$F(\rho) = \int \rho \ln \rho + \rho V \, dx \qquad \mu = \rho \mathcal{L}^n$$

Then $\nabla F(\rho) = \dfrac{\nabla \rho}{\rho} + V$, so that

$$\frac{d}{dt}\rho_t + \operatorname{div}\big(-\underbrace{\nabla F(\rho_t)}_{= v_t}\, \rho_t\big) = 0 \qquad \text{becomes}$$

$$\frac{d}{dt}\rho_t = \operatorname{div}\left(\nabla \rho_t + \rho_t \nabla V\right) \qquad (\text{FP equation})$$

Proof of $\nabla F(\rho) = \frac{\nabla \rho}{\rho} + \nabla V$

$(T = \nabla \varphi)$

$$\frac{d}{dt} F\left((Id + tT)_{\#}\rho \right) \Big|_{t=0} = \int \left\langle \frac{\nabla \rho}{\rho} + \nabla V, T \right\rangle \rho \, dx$$

$$\|$$

$$\int \rho_t \ln \rho_t + \int V \rho_t = \int \rho \ln \frac{\rho}{\det(I + t\nabla T)} + \int V \rho_t$$

$$= \int \rho \ln \rho + \int V \rho + t \left(- \int \rho \, \Delta T + \int \langle \nabla V, \nabla T \rangle \rho \right) + o(t)$$

A more intrinsic representation of F

Set $\gamma = e^{-V}\mathcal{L}^n$, so that $\mu = \rho\,\mathcal{L}^n = u\,\gamma$, $\quad u = \rho\, e^{V}$

Then $$\int \rho \ln \rho + \rho V \, dx = \int u \ln u \, d\gamma =: \mathcal{H}(\mu|\gamma)$$

relative entropy w.r.t. γ

Once we take γ as reference measure, the whole picture works equally well in ∞ dimensions.

The convexity condition on V becomes

$$\ln \gamma\big((1-t)A + tB\big) \geq (1-t) \ln \gamma(A) + t \ln \gamma(B)$$

γ log-concave $\iff$ $\mathcal{H}(\cdot \mid \gamma)$ is convex along geodesics

$\iff$ $\mathcal{H}(\cdot \mid \gamma)$ is convex along all interpolating curves

$\Rightarrow$ (EVI) solutions to $\mathcal{H}(\cdot \mid \gamma)$ exist and are stable.

Again, we can find a differential description, and connect this with the traditional viewpoint based on Dirichlet forms:

$$\mathcal{E}_\gamma(u,v) = \int \langle \nabla u, \nabla v \rangle \, d\gamma$$

$P_t u = L^2$ semigroup generated by $\mathcal{E}_\gamma$

$$\frac{d}{dt} \langle P_t u, v \rangle = \mathcal{E}_\gamma(P_t u, v) \quad \forall v$$

$S_t \mu =$ WASSERSTEIN semigroup generated by $\mathcal{H}(\cdot \mid \gamma)$

$$P_t u(x) = \int u \, d S_t \delta_x$$

(See A. - SAVARÈ - ZAMBOTTI ....)