The geodesic problem and Gromov-Hausdorff convergence

4.1 Metric derivative and geodesics in metric spaces

Let (E,d) be a metric space. Given a curve $\gamma:[a,b]\to E$, $\Gamma:=\gamma([a,b])$

$Var_{a'}^{b'}(\gamma)=\sup\{\sum_{i=1}^{n-1} d(\gamma(t_i),\gamma(t_{i+1}))\mid a'\le t_1<\dots<t_n\le b'\}$ for every pair a', b' with $a\le a'\le b'\le b$. (length)

We say that γ is rectifiable if $Var(\gamma)<\infty$.

Rmk 4.1.1 If $\gamma:[a,b]\to E$ is a curve of finite variation and $\ell(t)=Var_a^t(\gamma)$, then

$$d(\gamma(t+h),\gamma(t))\le \ell(t+h)-\ell(t)\qquad a\le t\le t+h\le b$$

Therefore, if $\ell(t)$ is cts, then γ is also cts. If ℓ is Lipschitz cts, so is γ.

$Var(\gamma)<+\infty$ does not imply that γ is cts

For $x,y\in E$, the problem of the geodesics joining x and y has two possible formulations.

$$\min\{Var(\gamma)\mid \gamma\in Lip([a,b],E),\ \gamma(a)=x,\ \gamma(b)=y\}$$

Def^n 4.1.2 (Metric derivative) Given a curve $\gamma:[a,b]\to E$, we define the metric derivative of γ at the point $t\in(a,b)$ as the limit $\lim_{h\to0}\frac{d(\gamma(t+h),\gamma(t))}{|h|}$ whenever it exists, we denote it by $|\dot\gamma|(t)$.

Eg. 4.1.3 $E=\mathbb{R}^n$. If γ is diff'ble at the pt t, then $|\dot\gamma|(t)$ is the euclidean norm of the derivative of γ at the pt t

Eg. 4.1.4 If $E=\mathbb{R}^n$ with $\|\cdot\|$, d is induced by $\|\cdot\|$, then $|\dot\gamma|(t_0)=\left\|\frac{d\gamma}{dt}(t_0)\right\|$ at any pt t_0 where γ is diff'ble.

Thm 4.1.6. For each Lipschitz curve $\gamma:[a,b]\to E$ the metric derivative exists at $\mathcal{L}^1$-almost every point in $[a,b]$. Moreover, it holds that $Var(\gamma)=\int_a^b|\dot\gamma|(t)\,dt$.

Proof: Since γ is a cts function, $T=\gamma([a,b])$ is a cpt metric space, separable. $\{x_n\}$ be a seq. dense in T.

$\varphi_n(t):=d(\gamma(t),x_n)\in Lip([a,b])$

By the Rademacher thm, $\dot\varphi_n(t)$ exists at $\mathcal{L}^1$-a.e. $[a,b]$. Defining $m(t):=\sup_n|\dot\varphi_n(t)|$

We will prove that $|\dot\gamma|(t)=m(t)$ for $\mathcal{L}^1$-a.e. t.

Observe that $x\mapsto d(x,x_n)$ is 1-Lipschitz. $\forall n$ we have

$$\lim_{h\to0}\frac{d(\gamma(t+h),\gamma(t))}{|h|}\ge\lim_{h\to0}\frac{|\varphi_n(t+h)-\varphi_n(t)|}{|h|}=|\dot\varphi_n(t)|\quad\text{holds for } \mathcal{L}^1\text{a.e. } t\in[a,b].$$

taking the supremum wrt. n we obtain $\lim_{h\to0}\frac{d(\gamma(t+h),\gamma(t))}{|h|}\ge m(t)$ for a.e. $t\in[a,b]$

On the other hand, $d(\gamma(t),\gamma(s))=\sup_n|d(\gamma(t),x_n)-d(\gamma(s),x_n)|\le\sup_n\int_s^t|\dot\varphi_n(\tau)|d\tau\le\int_s^t m(\tau)\,d\tau$

$Lip(\varphi_n)\le Lip(\gamma)$ (the Lip constant). Hence $|m(t)|\le Lip(\gamma)$ and m is integrable over $[a,b]$. If $t\in(a,b)$ is a Lebesgue pt for m, $\lim_{h\to0}\frac{d(\gamma(t+h),\gamma(t))}{|h|}\le\lim_{h\to0}\left|\frac1h\int_t^{t+h}m(\tau)d\tau\right|=m(t)$.

Since a.e. $t\in[a,b]$ is a Lebesgue pt

Since $m(t)=|\dot\gamma|(t)$ $\sum_{i=1}^{n-1}d(\gamma(t_{i+1}),\gamma(t_i))\le\sum_{i=1}^{n-1}\int_{t_i}^{t_{i+1}}|\dot\gamma|(\tau)d\tau\le\int_a^b|\dot\gamma|(\tau)\,d\tau\quad\forall\, a\le t_1<\dots<t_n\le b$

Take the sup over partitions. $Var(\gamma)\le\int_a^b|\dot\gamma|(t)dt$

choose $\varepsilon>0$. $h=\frac{b-a}{n}$, $t_i=a+ih$, $n\ge2$ s.t. $h\le\varepsilon$

$$\frac1h\int_a^{b-\varepsilon}d(\gamma(t+h),\gamma(t))\,dt\le\frac1h\int_0^h\sum_{i=0}^{n-2}d(\gamma(\tau+t_{i+1}),\gamma(\tau+t_i))\,d\tau\le\frac1h\int_0^h Var(\gamma)\,d\tau=Var(\gamma).$$

$$\int_a^{b-\varepsilon}|\dot\gamma|(t)\,dt=\int_a^{b-\varepsilon}\lim_{h\to0}\frac{d(\gamma(t+h),\gamma(t))}{h}\,dt\le\lim_{h\to0}\frac1h\int_a^{b-\varepsilon}d(\gamma(t+h),\gamma(t))\,dt\le Var(\gamma)$$

$\Rightarrow\int_a^b|\dot\gamma|(t)\,dt\le Var(\gamma)$.

4.2 Reparametrization.

Thm 4.2.1 Suppose that $\gamma \in Lip([a,b],E)$, $L = Var(\gamma)$. Then there exists a Lipschitz curve $\tilde{\gamma}:[0,L] \to E$ s.t. $|\dot{\tilde{\gamma}}| = 1$ a.e. on $[0,L]$ and $\tilde{\gamma}([0,L]) = \gamma([a,b])$

Proof: WLOG assume $a=0$. Let $\ell(t) = Var_0^t(\gamma)$. Define $h(t) = \inf\{s \in [0,b] : \ell(s) = t\}$ $t \in [0,L]$.
$h(0)=0$, h is non decreasing, $h(L) \le b$. Moreover, $\ell(h(t)) = t$ $\gamma(s) = \gamma(h(\ell(s)))$ for $t \in [0,L]$ and $s \in [0,b]$.

Observe that $h(\ell(s)) \le s$ and
$$d(\gamma(h(\ell(s))),\gamma(s)) \le Var_{h(\ell(s))}^{s}(\gamma) = \ell(s) - \ell(h(\ell(s))) = \ell(s) - \ell(s) = 0.$$

Define $\tilde{\gamma}(t) := \gamma(h(t))$ $t \in [0,L]$. $Var_0^t(\tilde{\gamma}) = t$ $\forall t \in [0,L]$.

If $0 \le t_1 \le \cdots \le t_n \le t$, letting $s_i = h(t_i)$.
$$\sum_{i=1}^{n-1} d(\tilde{\gamma}(t_{i+1}), \tilde{\gamma}(t_i)) = \sum_{i=1}^{n-1} d(\gamma(s_{i+1}), \gamma(s_i)) \le Var_0^{s_n}(\gamma) = \ell(s_n) = t_n \le t.$$

On the other hand, for $t \in [0,L]$ and $\varepsilon > 0$, $\exists$ $0 = s_1 \le \cdots \le s_n = h(t)$ s.t.
$$\ell(h(t)) \le \varepsilon + \sum_{i=1}^{n-1} d(\gamma(s_{i+1}), \gamma(s_i))$$
$$\Rightarrow \quad t \le \varepsilon + \sum_{i=1}^{n-1} d(\gamma(h(\ell(s_{i+1}))), \gamma(h(\ell(s_i))))$$
$$= \varepsilon + \sum_{i=1}^{n-1} d(\tilde{\gamma}(\ell(s_{i+1})), \tilde{\gamma}(\ell(s_i)))$$
$$\le \varepsilon + Var_{\ell(s_1)}^{\ell(s_n)}(\tilde{\gamma})$$
$$= \varepsilon + Var_0^{\ell(h(t))}(\tilde{\gamma})$$
$$= \varepsilon + Var_0^{t}(\tilde{\gamma}) \qquad \tilde{\gamma} \text{ is 1-Lipschitz.}$$

Finally, $t = Var_0^t(\tilde{\gamma}) = \int_0^t |\dot{\tilde{\gamma}}|(r)\,dr$ $\forall t \in [0,1]$.

differentiating w.r.t. t, $|\dot{\tilde{\gamma}}|(t) = 1$ for a.e. t.

Def^{n}: We say γ is rectifiable if $\ell(\gamma) := \sup \sum_{i=0}^{N-1} d(\gamma(t_{i+1}), \gamma(t_i)) < \infty$

where sup is over all finite collections $a = t_0 < t_1 < \cdots < t_N = b$

Thm. If γ is rectifiable then there exists a curve $\alpha : [0, \ell(\gamma)] \to X$

s.t. i) $\forall s \in [0, \ell(\gamma)]$, $\ell(\alpha|_{[0,s]}) = s$

ii) In particular, α is 1-Lipschitz α is called the arclength parametrization.

$$d(\alpha(s), \alpha(s')) \leq \ell(\gamma|_{[s,s']}) = |s - s'|$$

ii) $\gamma(s) = \alpha(h(s))$ for some increasing $h : [a,b] \to [0, \ell(\gamma)]$

Ex. Take $f : [0,1] \to [0,1]$ be the Cantor-Vitali function. Let $\gamma(t) = (t, f(t))$
$[0,1] \to \mathbb{R}^2 = X$

γ is rectifiable, but it is not Lipschitz.

Ex Find the right h - compute it!

Def^{n}: We say $u : X \to \mathbb{R}$ is abs cts on a rectifiable curve γ if

$$u \circ \alpha : [0, \ell(\gamma)] \to \mathbb{R}.$$

is abs cts where α is the arclength parametrization of γ.

Ex: Let $u : \mathbb{R}^n \to \mathbb{R}$ be C^1, $\gamma : [a,b] \to \mathbb{R}^n$ C^1.

$u \circ \gamma : [a,b] \to \mathbb{R}$ is C^1 $(u \circ \gamma)'(t) = \nabla u(\gamma(t)) \cdot \gamma'(t)$.

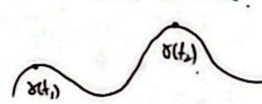
It follows that $(u\circ\gamma)(b) - (u\circ\gamma)(a) = \int_a^b \underbrace{\langle \nabla u(\gamma(t)), \gamma'(t) \rangle}_{\text{not on } X.} dt$

$$|u(\gamma(t_2)) - u(\gamma(t_1))| \leq \int_{t_1}^{t_2} |\nabla u(\gamma(t))| |\gamma'(t)| dt.$$

$|\nabla u| : X \to \mathbb{R}$.

If γ is arclength parametrized, then,

$$|u(\gamma(t_2)) - u(\gamma(t_1))| \leq \int_{t_1}^{t_2} \underbrace{|\nabla u(\gamma(t))|}_{\text{object on } X = \mathbb{R}^n} dt.$$



Recall: $f : [a,b] \to \mathbb{R}$ is abs cts iff there is $g \in L^1([a,b])$ s.t. $f(x) = f(a) + \int_a^x g(t) dt$.

Lemma: Suppose there exists some function $\rho \geq 0$ on $[a,b]$ s.t. $\forall t_1, t_2$, $\rho \in L^1([a,b])$.

$$|u(\gamma(t_2)) - u(\gamma(t_1))| \leq \int_{t_1}^{t_2} \rho(t) dt.$$

then $u \circ \gamma$ is absolutely continuous, i.e. u is AC on γ.

Proof: Any collection of disjoint (t_j, t_{j+1}) satisfy $* = \sum_j |u(\gamma(t_{j+1})) - u(\gamma(t_j))| \leq \sum_j \int_{t_j}^{t_{j+1}} \rho(t) dt$

$$= \int_{\bigsqcup_j (t_j, t_{j+1})} \rho(t) dt.$$

By AC of $\int$, $\forall \varepsilon > 0$, $\exists \delta > 0$ s.t. $|A| < \delta \Rightarrow \int_A |\rho(t)| dt < \varepsilon$.

So if $\sum_j |t_{j+1} - t_j| < \delta \Rightarrow * < \varepsilon$.

For $u : X \to \mathbb{R}$, we will seek $\rho \geq 0$ s.t. $|u(\gamma(t_2)) - u(\gamma(t_1))| \leq \int_{t_1}^{t_2} \rho(s) ds$

If such ρ belongs to $L^1([a,b])$ then u will be AC on γ.

In $X = \mathbb{R}^n$ with $u \in C^1(\mathbb{R}^n)$, $|\nabla u(\gamma(t))| = \rho(t)$ is the right ρ.

Benamou - Brenier Formulation.

Let's consider a density $\rho(t,x)$ $x \in \mathbb{R}^n$ $0<t<1$

We want $\rho(0,x) = f(x)$ $\rho(1,x) = g(x)$

Introduce velocity field $v(t,x)$.

Describe a mass-preserving flow via the continuity eq^{n}.

$\rho_t + \nabla \cdot (v\rho_t) = 0$.

Introduce Lagrangian coordinates $X(t,x)$ to "follow" the particle through the flow.

$$\begin{cases} \frac{\partial X}{\partial t}(t,x) = v(t, X(t,x)) \\ X(0,x) = x \end{cases}$$

Mass in E at $t=0$ = Mass in $X(t,E)$ at time t.

$\Rightarrow \int_E f(x)\,dx = \int_{X(t,E)} \rho(t,x)\,dx.$ $\forall\, 0<t\leq 1$.

This is the statement $X(t,\cdot)_{\#} f = \rho(t,\cdot)$.

and

$X(1,\cdot)_{\#} f = g$

$$\begin{aligned} W_2^2(f,g) &\leq \int_{\mathbb{R}^n} f(x)\,|X(1,x) - x|^2\,dx. \\ &= \int_{\mathbb{R}^n} f(x)\,|X(1,x) - X(0,x)|^2\,dx \\ &= \int_{\mathbb{R}^n} f(x)\,\left|\int_0^1 \frac{\partial X}{\partial t}(t,x)\,dt\right|^2 dx \\ &\leq \int_{\mathbb{R}^n}\int_0^1 f(x)\,\left|\frac{\partial X}{\partial t}(t,x)\right|^2 dt\,dx \\ &= \int_0^1 \int_{\mathbb{R}^n} f(x)\,|v(t,X(t,x))|^2\,dx\,dt. \end{aligned}$$

Fix t Define $S(x) = X(t,x)$.

$S_{\#} f = \rho(t,x)$

$\varphi(y) = |v(t,y)|^2$

$$\begin{aligned} &\int_{\mathbb{R}^n} f(x)\,|v(t,X(t,x))|^2\,dx \\ &= \int_{\mathbb{R}^n} f(x)\,\varphi(S(x))\,dx. \\ &= \int_{\mathbb{R}^n} \rho(t,y)\,\varphi(y)\,dy. \text{ (push forward)} \\ &= \int_{\mathbb{R}^n} \rho(t,y)\,|v(t,y)|^2\,dy. \end{aligned}$$

$W_2^2(f,g) \le \int_0^1\int_{\mathbb{R}^n} \rho(t,x)\,|v(t,x)|^2\,dx\,dt$

$\forall \rho, v$ satisfying $\begin{cases} \rho_t + \nabla\cdot(\rho v) = 0 \\ \rho(0,x) = f(x) \\ \rho(1,x) = g(x) \end{cases}$

Can we achieve equality?

Need $X(1,x) = T(x)$ the optimal map from f to g.

Need $\left|\int_0^1 \frac{\partial X}{\partial t}(t,x)\,dt\right|^2 = \int_0^1 \left|\frac{\partial X}{\partial t}(t,x)\right|^2 dt$

This holds if $\frac{\partial X}{\partial t}(t,x) = u(x)$ is constant in time.

Recall: $X(0,x) = x$.

$\Rightarrow$ $X(t,x) = x + u(x)t$.

$X(1,x) = x + u(x) = T(x)$.

$\Rightarrow$ $u(x) = T(x) - x$.

$X(t,x) = x + (T(x)-x)t = x(1-t) + tT(x)$.

Our velocity field should satisfy $v(t, X(t,x)) = u(x) = T(x) - x := (T-I)(x)$.

Let $y = X(t,x) = [(1-t)I + tT](x)$.

$\Rightarrow v(t,y) = (T-I)[(1-t)I + tT]^{-1}(y)$

achieves the optimum. if (f,g) are nice, else do "smoothed" approximations.

$W_2^2(f,g) = \inf_{\rho,v} \int_0^1\int_{\mathbb{R}^n} \rho(t,x)\,|v(t,x)|^2\,dx\,dt$ s.t. $\begin{cases} \rho_t + \nabla\cdot(\rho v) = 0 \\ \rho(0,x) = f(x) \\ \rho(1,x) = g(x). \end{cases}$ BB formulation.

E.g. $\min f(x)$ s.t. $Ax = b$. $L(x,\lambda) = f(x) + \lambda(Ax-b)$.

$\min_x \left\{ f(x) + \sup_\lambda \lambda(Ax-b) \right\}$. seek saddle pts of this.

Let $\varphi(t,x)$ be the Lagrange multiplier.

$\int_0^1\int_{\mathbb{R}^n} (\rho_t + \nabla\cdot(\rho v))\,\varphi\,dx\,dt$

IBP
$= \int_{\mathbb{R}^n} \big(g(x)\varphi(x,1) - f(x)\varphi(x,0)\big)\,dx - \int_0^1\int_{\mathbb{R}^n} (\rho_t\varphi_t + \rho v\cdot\nabla\varphi)\,dx\,dt$.

Lagrangian $L = \int_0^1\int_{\mathbb{R}^n} \left(\frac{\rho|v|^2}{2} - \rho\varphi_t - \rho v\cdot\nabla\varphi\right) dx\,dt + \underbrace{\int_{\mathbb{R}^n} \big(g(x)\varphi(x,1) - f(x)\varphi(x,0)\big)\,dx}_{G(\varphi)}$.

Letting $m = \rho v$ $\quad L(\varphi,\rho,m) = \int_0^1\int_{\mathbb{R}^n}\left(\frac{|m|^2}{2\rho} - \rho\varphi_t - m\cdot\nabla\varphi\right)dx\,dt + G(\varphi)$

Want to solve $\inf_{\rho,m} \sup_\varphi L(\varphi,\rho,m)$.

Let $h(\rho,m) = \frac{|m|^2}{2\rho}$ $\quad$ Claim: $h(\rho,m) = \sup_{(a,b)\in K} (a\rho + b\cdot m)$ to express the quadratic in terms of linear

$K = \{(a,b) \in \mathbb{R}\times\mathbb{R}^n : a + \frac{|b|^2}{2} \le 0\}$.

Proof Let (a^*,b^*) be the sup. Since $\rho > 0$, a^* as big as possible. $\Rightarrow a^* = -\frac{|b^*|^2}{2}$.

$\sup_{(a,b)\in K}(a\rho + b\cdot m) = \sup_{b\in\mathbb{R}^n}\left(-\frac{\rho|b|^2}{2} + b\cdot m\right)$. $\Rightarrow$ set gradient to 0. $-\rho b + m = 0$. $\Rightarrow b^* = \frac{m}{\rho}$.

$\sup_{(a,b)\in K}(a\rho + b\cdot m) = -\frac{|m|^2}{2\rho} + \frac{|m|^2}{\rho} = \frac{|m|^2}{2\rho} = h(\rho,m)$.

Saddle pt problem is $\inf_{\rho, m} \sup_{\varphi, (a,b) \in K} \int_0^1 \int_{\mathbb{R}^n} (a\rho + b \cdot m - \rho\varphi_t - m \cdot \nabla\varphi) \, dx dt + G(\varphi)$.

$$= \inf_{\rho, m} \sup_{\varphi, (a,b) \in K} \int_0^1 \int_{\mathbb{R}^n} [(a - \varphi_t)\rho + (b - \nabla\varphi) \cdot m] \, dx dt + G(\varphi).$$

Let $r = \begin{pmatrix} \rho \\ m \end{pmatrix}$. $\quad \xi = \begin{pmatrix} a \\ b \end{pmatrix}$ $\quad \nabla_{t,x}\varphi = \begin{pmatrix} \varphi_t \\ \nabla\varphi \end{pmatrix}$.

$\Rightarrow \inf_r \sup_{\varphi, \xi \in K} \langle \xi - \nabla_{t,x}\varphi, r \rangle + G(\varphi)$.

Let $\tau > 0$ be small to regularize the max

$\inf_r \sup_{\varphi, \xi \in K} \langle \xi - \nabla_{t,x}\varphi, r \rangle + G(\varphi) - \frac{\tau}{2} \| \xi - \nabla_{t,x}\varphi \|^2$.

optimize those 3 unknows ~~independently~~ independently

Given r_k, ξ_k, $\quad \varphi_{k+1} = \arg\max_\varphi -\langle \nabla_{t,x}\varphi, r \rangle + G(\varphi) - \frac{\tau}{2} \| \xi - \nabla_{t,x}\varphi \|^2$.

quadratic, unconstrained, ez.

Computing 1st variation $\Rightarrow$ Poisson equation with Neumann BC (homogeneous in space & non-homogeneous in time)

$\Delta_{t,x} \varphi_{k+1} = \nabla \cdot (\xi_k - \frac{r_k}{\tau})$.

Give r_k, φ_{k+1}, optimize for ξ_{k+1}.

$\xi_{k+1} = \arg\max_{\xi \in K} \langle \xi, r \rangle - \frac{\tau}{2} |\xi - \nabla_{t,x}\varphi|^2$.

Do this pointwise (quadratic)

Given φ_{k+1}, ξ_{k+1}, we do gradient descent in r.

$r_{k+1} = r_k - \tau(\xi_{k+1} - \nabla_{t,x}\varphi_{k+1})$. Iterate.

JKO flows

Let $F:\mathbb{R}^n \to \mathbb{R}$ be convex we want to minimize $F(x)$.

Gradient flow : $\begin{cases} x'(t) = -\nabla F(x(t)) \\ x(0) = x_0 . \end{cases}$

Discretize in time via Backward Euler.

$$\frac{X^{n+1}-X^n}{\tau} = -\nabla F(x^{n+1})$$

$$\Rightarrow \frac{X^{n+1}-X^n}{\tau} + \nabla F(x^{n+1}) = 0 .$$

$$\Rightarrow \nabla\left(\frac{|x-x^n|^2}{2\tau} + F(x)\right)\Big|_{x=x_{n+1}} = 0$$

$$\Rightarrow x^{n+1} \in \operatorname{argmin}\left\{\frac{|x-x^n|^2}{2\tau} + F(x)\right\} .$$

Can define a scheme like this on a metric space (X,d).

Let $F: X \to \mathbb{R}$ be lsc and bounded below.

· Define $X_\tau^{n+1} \in \operatorname{argmin}\left\{F(x) + \frac{d(x,x_\tau^n)^2}{2\tau}\right\}$.

Interpolate to all t $\quad X_\tau(t) = X_\tau^n$ if $t \in ((n-1)\tau, n\tau]$

Study limit as $\tau \to 0$.

Consider $F: P(\Omega) \to \mathbb{R}$ $d = W_2$. Ω is compact, F lsc bounded below.

We previously used the continuity eqt $\rho_t + \nabla\cdot(\rho_t v) = 0$ to "flow" densities.

Goal: Find velocity field v s.t. this flow agrees with $\lim_{\tau\to 0} X_\tau(t)$.

Investigate optimality condition in the JKO scheme.

We need to compute the 1st variation

We need to perturb $\rho \in P(\Omega)$ to $\rho + \varepsilon\chi$ · Need $\rho+\varepsilon\chi \in P(\Omega)$ st. $F(\rho+\varepsilon\chi)$ is well-defined

Restrict to χ st. $\sigma = \rho + \varepsilon\chi \in P(\Omega)$ $\forall$ small $\varepsilon > 0$.

$\Rightarrow \rho + \varepsilon\chi = \rho + \varepsilon(\sigma - \rho)$.

$\quad = \rho(1-\varepsilon) + \varepsilon\sigma \in P(\Omega)$ as long as $\rho, \sigma \in P(\Omega)$.

$\forall \sigma \in P(\Omega) \cap L_c^\infty(\Omega)$.

The first variation of F, $\frac{\delta F}{\delta\rho}(\rho)$ is s.t. $\frac{d}{d\varepsilon} F(\rho+\varepsilon\chi)\Big|_{\varepsilon=0} = \int \frac{\delta F}{\delta\rho}(\rho)\,\chi(x)\,dx$. $\quad \forall \chi = \sigma - \rho$, $\sigma \in P(\Omega) \cap L_c^\infty$

$$\int \left(\frac{\delta F}{\delta\rho} + c\right)\chi(x)\,dx = \int \frac{\delta F}{\delta\rho}\,\chi(x)\,dx + c\underbrace{\int \chi(x)\,dx}_{=0} .$$

The 1st variation is defined uniquely only up to additive constants.

Come back to $G(\rho) = F(\rho) + \frac{W_2^2(\rho, \rho_\tau^n)}{2\tau}$.

We need $\frac{\delta G}{\delta\rho}(\rho) = \frac{\delta F}{\delta\rho}(\rho) + \frac{1}{2\tau}\frac{\delta W_2^2}{\delta\rho}(\rho, \rho_\tau^n)$.

Use the dual formulation: $W_2^2(f,g) = 2\inf_{\pi\in\Pi(f,g)} \int \frac{|x-y|^2}{2} d\pi(x,y)$.

$$= 2 \max_{u,v} \left\{ \int uf\,dx + \int vg\,dy \;\middle|\; u(x)+v(y) \le \tfrac{1}{2}|x-y|^2 \right\}.$$

$$= 2 \max_{u} \left\{ \int uf\,dx + \int u^c g\,dy \right\}.$$

$$\frac{d}{d\varepsilon} W_2^2(f+\varepsilon\chi, g)\Big|_{\varepsilon=0} = 2\frac{d}{d\varepsilon} \max_{u} \left\{ \int u(f+\varepsilon\chi)dx + \int u^c g\,dy \right\}\Big|_{\varepsilon=0}$$

$$= 2\int u^*\chi\,dx$$ where u^* achieves the max.

potential associate with the cost $\frac{1}{2}|x-y|^2$.

When we do OT, the optimal map is $T(x) = x - \nabla u^*(x)$.

$$= x - (\nabla h)^{-1}(\nabla u^*(x))$$ where $h(z) = \frac{1}{2}|z|^2$.

$$\Rightarrow \frac{\delta W_2^2}{\delta\rho}(\rho, \rho_\tau^n) = 2u^*$$

$T(x) = x - \nabla u^*(x)$ is the optimal map from ρ to ρ_τ^n.

The JKO scheme is $\rho_\tau^{n+1} = \operatorname{argmin} \left\{ F(\rho) + \frac{W_2^2(\rho,\rho_\tau^n)}{2\tau} \right\}$

$$= \operatorname{argmin} G(\rho).$$

$$\Rightarrow \frac{\delta G}{\delta\rho}(\rho_\tau^{n+1}) + C = 0.$$

$$\Rightarrow \frac{\delta F}{\delta\rho}(\rho_\tau^{n+1}) + \frac{u^*}{\tau} = \text{constant}.$$

$$\Rightarrow 0 = \nabla\left(\frac{\delta F}{\delta\rho}\right) + \frac{\nabla u^*}{\tau}$$

$$= \nabla\left(\frac{\delta F}{\delta\rho}\right) + \frac{x - T(x)}{\tau}$$

$$\Rightarrow \underbrace{\frac{T(x)-x}{\tau}}_{\text{velocity!}} = \nabla\left(\frac{\delta F}{\delta\rho}\right).$$

of a flow from ρ_τ^{n+1} to ρ_τ^n.

The flow we want should have velocity $V(x) := -\left(\frac{T(x)-x}{\tau}\right)$

$$= -\nabla\left(\frac{\delta F}{\delta\rho}\right)$$

This is the velocity associated with our time discrete scheme.

If everything works out as $\tau \to 0$, we expect our JKO scheme to limit to this flow

$$\rho_t + \nabla\cdot(\rho v) = 0 \quad \text{or} \quad \rho_t - \nabla\cdot\left(\rho \nabla\frac{\delta F}{\delta\rho}\right) = 0.$$

This is the PDE associated with gradient flows of F in the W_2 metric.

Eg. $F(\rho) = \int \rho\log\rho\,dx$. We want a flow that maximizes entropy.

$$\frac{d}{d\varepsilon}F(\rho+\varepsilon\chi)\Big|_{\varepsilon=0} = \frac{d}{d\varepsilon}\int(\rho+\varepsilon\chi)\log(\rho+\varepsilon\chi)\,dx\Big|_{\varepsilon=0} = \int(\chi\log\rho + \chi)\,dx. \quad \Rightarrow \frac{\delta F}{\delta\rho} = \log\rho + 1.$$

$\nabla\left(\frac{\delta F}{\delta\rho}\right) = \nabla(\log\rho+1) = \frac{1}{\rho}\nabla\rho$. $\Rightarrow$ The GF is $0 = \rho_t - \nabla\cdot(\rho\cdot\frac{1}{\rho}\nabla\rho)$

$$= \rho_t - \nabla\cdot(\nabla\rho)$$

$$= \rho_t - \Delta\rho.$$

$\Rightarrow \rho_t = \Delta\rho$ (heat equation).

Eg. $F(\rho) = \int \rho \log \rho \, dx + \int V(x) \rho \, dx$

$\Rightarrow \rho_t - \Delta\rho - \nabla\cdot(\rho\nabla V) = 0$ Fokker - Planck.

Eg. $F(\rho) = \frac{1}{m-1} \int \rho^m \, dx$

$\Rightarrow \rho_t - \Delta(\rho^n) = 0$. porous medium.

Eg. $F(\rho) = \int \rho \log \rho - \frac{1}{2} \int |\nabla u_\rho|^2$ where $-\Delta u_\rho = \rho$.

$\Rightarrow \begin{cases} \rho_t + \nabla\cdot(\rho\nabla u) - \Delta\rho = 0 \\ -\Delta u = \rho \end{cases}$ (Keller - Segel, chemotaxis).

E.

$F(\rho) = \frac{1}{2} \iint W(x-y) \, d\rho(x) \, d\rho(y)$

$\Rightarrow \rho_t - \nabla\cdot(\rho((\nabla W) * \rho)) = 0$ (aggregation model)

Scanned with CamScanner

Otto's calculus

By BB formula $\cdot\ W_2^2(\bar\rho_0,\bar\rho_1) = \inf_{\rho_t, v_t} \left\{ \int_0^1 \left(\int_\Omega |v_t|^2 \rho_t \, dx \right) dt : \partial_t \rho + \operatorname{div}(v_t \rho_t) = 0 ,\ v_t \cdot \overset{\text{normal}}{\nu}\big|_{\partial\Omega} = 0 ,\ \rho_0 = \bar\rho_0 ,\ \rho_1 = \bar\rho_1 \right\}$

$$= \inf_{\rho_t} \left\{ \inf_{v_t} \int_0^1 \int_\Omega |v_t|^2 \rho_t \, dx\, dt : \partial_t \rho + \operatorname{div}(v_t\rho_t) = 0 ,\ v_t \cdot n|_{\partial\Omega} = 0 ,\ \rho_0 = \bar\rho_0 ,\ \rho_1 = \bar\rho_1 \right\}$$

$$= \inf_{\rho_t} \left\{ \int_0^1 \underbrace{\inf_{v_t} \left\{ \int_\Omega |v_t|^2 \rho_t \, dx \,\middle|\, \operatorname{div}(v_t\rho_t) = -\partial_t \rho_t ,\ v_t \cdot n|_{\partial\Omega} = 0 \right\}}_{\|\partial_t \rho_t\|^2_{\rho_t}} dt \,\middle|\, \rho_0 = \bar\rho_0 ,\ \rho_1 = \bar\rho_1 \right\}$$

In analogy with the Riemannian distance

$$\|\partial_t \rho_t\|^2_{\rho_t} := \inf_{v_t} \left\{ \int_\Omega |v_t|^2 \rho_t \, dx : \operatorname{div}(v_t \rho_t) = -\partial_t \rho_t ,\ v_t \cdot n|_{\partial\Omega} = 0 \right\} .$$

$$\Rightarrow W_2^2(\bar\rho_0, \bar\rho_1) = \inf_{\rho_t} \left\{ \int_0^1 \|\partial_t \rho_t\|^2_{\rho_t} \, dt : \rho_0 = \bar\rho_0 ,\ \rho_1 = \bar\rho_1 \right\} .$$

We want the v_t that realizes the infimum.

Hence, given ρ_t, $\partial_t\rho_t$, let v_t be a minimizer, ω be a vector field s.t. $\operatorname{div}(\omega) = 0$ and $\omega \cdot n|_{\partial\Omega} = 0$.

$\forall \varepsilon > 0 \quad \operatorname{div}\left(\left(v_t + \varepsilon \frac{\omega}{\rho_t} \right) \rho_t \right) = -\partial_t \rho_t$

Thus $v_t + \varepsilon \frac{\omega}{\rho_t}$ is an admissible vector field.

$$\int_\Omega |v_t|^2 \rho_t \, dx \le \int_\Omega \left| v_t + \varepsilon \frac{\omega}{\rho_t} \right|^2 \rho_t \, dx = \int_\Omega |v_t|^2 \rho_t \, dx + 2\varepsilon \int_\Omega \langle v_t, \omega \rangle \, dx + \varepsilon^2 \int_\Omega \frac{|\omega|^2}{\rho_t} \, dx$$

Dividing by ε and letting it $\to 0$ yields $\int_\Omega \langle v_t, \omega \rangle = 0$ $\forall \omega$ s.t. $\operatorname{div}(\omega) = 0$ and $\omega \cdot n|_{\partial\Omega} = 0$

By the Helmholtz decomp $v_t \in \{ \omega : \operatorname{div}(\omega) = 0 ,\ \omega \cdot n|_{\partial\Omega} = 0 \}^\perp = \{ \nabla g : g : \Omega \to \mathbb{R} \}$

$\Rightarrow \exists$ function ψ_t s.t. $v_t = \nabla \psi_t$. Since $\operatorname{div}(v_t \rho_t) = -\partial_t \rho_t$ and $v_t \cdot n|_{\partial\Omega} = 0$, then ψ_t is a sol. of

$$\begin{cases} \operatorname{div}(\rho_t \nabla \psi_t) = -\partial_t \rho_t & \text{in } \Omega . \\ \dfrac{\partial \psi_t}{\partial n} = 0 & \text{on } \partial\Omega \end{cases}$$

If ρ_t is 'nice', this is the uniformly elliptic equation with Neumann boundary conditions for ψ_t, the solution ψ_t is unique up to a constant. So one can define

$$\|\partial \rho_t\|^2_{\rho_t} = \int_\Omega |\nabla \psi_t|^2 \rho_t \, dx . \qquad \psi_t \text{ solves the DE} .$$

More generally, given $\rho \in \mathcal{P}_2(\Omega)$ $h : \Omega \to \mathbb{R}$ s.t. $\int_\Omega h = 0$

this is needed for the solvability of the DE.

Rmk. $\int_\Omega h \, dx = \int_\Omega \operatorname{div}(\rho \nabla \psi) \, dx = \int_{\partial\Omega} \frac{\partial \psi}{\partial n} \rho = 0$.

Whenever ρ_t is a curve of prob.

$\int_\Omega \partial_t \rho_t \, dx = \frac{d}{dt} \int_\Omega \rho_t \, dx = \frac{d}{dt} 1 = 0$.

We can define $\|h\|^2_\rho := \int_\Omega |\nabla \psi|^2 \rho \, dx$ where $\begin{cases} \operatorname{div}(\rho \nabla \psi) = -h & \text{in } \Omega \\ \dfrac{\partial \psi}{\partial n} = 0 & \text{on } \partial\Omega \end{cases}$

Construction of norm is done.

Def^n Given two functions $h_1, h_2 : \Omega \to \mathbb{R}$ with $\int_\Omega h_1 = 0$, $\int_\Omega h_2 = 0$.

$$\langle h_1, h_2 \rangle_\rho := \int_\Omega \nabla\psi_1 \cdot \nabla\psi_2 \, \rho \, dx \quad \text{where} \quad \begin{cases} \operatorname{div}(\rho \nabla\psi_i) = -h_i & \text{in } \Omega \\ \rho \frac{\partial\psi_i}{\partial n} = 0 & \text{on } \partial\Omega \end{cases}$$

Def^n Given a functional $\mathcal{F} \in \mathcal{P}_2(\Omega)^*$ (dual), its gradient w.r.t. the Wasserstein scalar product at $\bar\rho \in \mathcal{P}_2(\Omega)$ is the unique function $\operatorname{grad}_{W_2} \mathcal{F}[\bar\rho]$ s.t.

$$\frac{d}{d\varepsilon} \mathcal{F}[\rho_\varepsilon]\Big|_{\varepsilon=0} = \left\langle \operatorname{grad}_{W_2} \mathcal{F}[\bar\rho], \frac{\partial\rho_\varepsilon}{\partial\varepsilon}\Big|_{\varepsilon=0} \right\rangle_{\bar\rho}$$

$\forall$ smooth curve $\rho_\varepsilon : (-\varepsilon_0, \varepsilon_0) \to \mathcal{P}(\Omega)$ with $\rho_0 = \bar\rho$.

What are explicit formula? Given $\mathcal{F} \in \mathcal{P}_2(\Omega)^*$ and $\bar\rho \in \mathcal{P}_2(\Omega)$.

Denote by $\frac{\delta\mathcal{F}[\bar\rho]}{\delta\rho}$ its first L^2 variation

$$\frac{d}{d\varepsilon}\Big|_{\varepsilon=0} \mathcal{F}[\rho_\varepsilon] = \int_\Omega \frac{\delta\mathcal{F}[\bar\rho]}{\delta\rho}(x) \frac{\partial\rho_\varepsilon(x)}{\partial\varepsilon}\Big|_{\varepsilon=0} dx. \quad \forall \text{ smooth } \rho_\varepsilon : (-\varepsilon_0, \varepsilon_0) \to \mathcal{P}_2(\Omega) \text{ st } \rho_0 = \bar\rho$$

Then by the def^n of Wasserstein gradient,

$$\left\langle \operatorname{grad}_{W_2} \mathcal{F}[\bar\rho], \frac{\partial\rho_\varepsilon}{\partial\varepsilon}\Big|_{\varepsilon=0} \right\rangle_{\bar\rho} = \int_\Omega \frac{\delta\mathcal{F}[\bar\rho]}{\delta\rho} \frac{\partial\rho_\varepsilon}{\partial\varepsilon}\Big|_{\varepsilon=0} dx$$

Thus denoting by ψ the solution of $\operatorname{div}(\nabla\psi\bar\rho) = -\frac{\partial\rho_\varepsilon}{\partial\varepsilon}\Big|_{\varepsilon=0}$ with zero Neumann boundary conditions,

$$\left\langle \operatorname{grad}_{W_2} \mathcal{F}[\bar\rho], \frac{\partial\rho_\varepsilon}{\partial\varepsilon}\Big|_{\varepsilon=0} \right\rangle_{\bar\rho} = -\int_\Omega \frac{\delta\mathcal{F}[\bar\rho]}{\delta\rho} \operatorname{div}(\nabla\psi\bar\rho) dx$$

$$= \int_\Omega \nabla \frac{\delta\mathcal{F}[\bar\rho]}{\delta\rho} \cdot \nabla\psi \bar\rho \, dx$$

If $\bar\rho \, \partial_n \left(\frac{\delta\mathcal{F}[\bar\rho]}{\delta\rho}\right) = 0$ on $\partial\Omega$, $\operatorname{grad}_{W_2} \mathcal{F}[\bar\rho] = -\operatorname{div}\left(\nabla\left(\frac{\delta\mathcal{F}[\bar\rho]}{\delta\rho}\right)\bar\rho\right)$.

Now we give an e.g. that the Wasserstein gradient doesn't exist if $\bar\rho \, \partial_n\left(\frac{\delta\mathcal{F}[\bar\rho]}{\delta\rho}\right) \neq 0$ on $\partial\Omega$.

If $\mathcal{F}[\rho] = \int_\Omega U(\rho(x)) dx$, $U : \mathbb{R} \to \mathbb{R}$. $\forall$ smooth variation $\varepsilon \mapsto \rho_\varepsilon$

$$\frac{d}{d\varepsilon} \int_\Omega U(\rho_\varepsilon(x)) dx = \int_\Omega U'(\bar\rho(x)) \frac{\partial\rho_\varepsilon(x)}{\partial\varepsilon}\Big|_{\varepsilon=0} dx.$$

The L^2-variation of $\mathcal{F}[\rho]$ at $\bar\rho \in \mathcal{P}_2(\Omega)$ is given by $\frac{\delta\mathcal{F}[\bar\rho]}{\delta\rho}(x) = U'(\bar\rho(x))$

$\Rightarrow \operatorname{grad}_{W_2} \mathcal{F}[\bar\rho] = -\operatorname{div}(\bar\rho \nabla[U'(\bar\rho)]) = -\operatorname{div}(\bar\rho U''(\bar\rho) \nabla\bar\rho)$ if $\bar\rho U''(\bar\rho) \partial_\nu \bar\rho = 0$ on $\partial\Omega$

In the special case $U(s) = s\log(s)$, $U''(s) = \frac{1}{s}$. $\Rightarrow \operatorname{grad}_{W_2} \mathcal{F}[\bar\rho] = -\Delta\bar\rho$ if $\partial_\nu \bar\rho = 0$ on $\partial\Omega$

If $U(s) = \frac{s^m}{m-1}$, $\exists m \neq 1$, $\operatorname{grad}_{W_2} \mathcal{F}[\bar\rho] = -\operatorname{div}(\bar\rho m \bar\rho^{m-2} \nabla\bar\rho) = -\Delta(\bar\rho^m)$

if $\partial_\nu(\bar\rho^m) = 0$ on $\partial\Omega$.

(Cont)

If $\mathcal{F}[\rho] = \int_\Omega \rho(x) V(x)\, dx$ with $V: \Omega \to \mathbb{R}$ it's first L^2-variation at $\bar\rho \in \mathcal{P}(\Omega)$ is

$$\frac{\delta \mathcal{F}[\bar\rho]}{\delta \rho}(x) = V(x)$$

Therefore the Wasserstein gradient of $\mathcal{F}$ is $\quad \mathrm{grad}_{W_2} \mathcal{F}[\bar\rho] = -\mathrm{div}(\bar\rho \nabla V)$ given $\bar\rho\, \partial_n V = 0$ on $\partial\Omega$.

If $\mathcal{F}[\rho] = \frac{1}{2}\iint \rho(x)\rho(y) W(x-y)\, dx\, dy \quad W: \mathbb{R}^d \to \mathbb{R}$ s.t. $W(z) = W(-z)$.

$$\frac{\delta \mathcal{F}[\bar\rho]}{\delta \rho}(x) = W * \bar\rho(x) = \int_\Omega W(x-y)\, \bar\rho(y)\, dy.$$

$$\Rightarrow \mathrm{grad}_{W_2} \mathcal{F}[\bar\rho] = -\mathrm{div}((\nabla W * \bar\rho)\bar\rho) \quad \text{given } \bar\rho\, \partial_n \frac{\delta \mathcal{F}[\bar\rho]}{\delta \rho} = 0 \text{ on } \partial\Omega$$

Def^n: Given a functional $\mathcal{F}: \mathcal{P}(\Omega) \to \mathbb{R}$, $\rho: [0,T) \to \mathcal{P}_2(\Omega)$ is a gradient flow of $\mathcal{F}$ wrt. W_2 and with starting point $\bar\rho_0$ if

$$\begin{cases} \partial_t \rho_t = -\mathrm{grad}_{W_2} \mathcal{F}[\rho_t] \\ \rho_0 = \bar\rho_0 \end{cases}$$

The Wasserstein gradient flow of the entropy functional $\mathcal{F}[\rho] = \int_\Omega \rho \log \rho\, dx$ is the heat equation with Neumann boundary conditions

$$\begin{cases} \partial_t \rho = -\mathrm{grad}_{W_2} \mathcal{F}[\rho] = \Delta \rho & \text{in } \Omega \\ \partial_n \rho = 0 & \text{on } \partial\Omega \end{cases}$$

If $\mathcal{F}[\rho] = \frac{1}{m-1}\int_\Omega \rho^m$ for $m \neq 1$ with $m > 0$, then the gradient flow is

$$\begin{cases} \partial_t \rho = -\mathrm{grad}_{W_2} \mathcal{F}[\rho] = \Delta(\rho^m) & \text{in } \Omega \\ \partial_n(\rho^m) = 0 & \text{on } \partial\Omega. \end{cases}$$

that is the porous medium eqt $m > 1$ or the fast diffusion equation (if $m \in (0,1)$)