Lecture 20

Recall:

<u>Theorem</u> (characterization of AC^2 curves and solns of (CE))

(i) Suppose $\mu \in AC^2(0,T;\mathcal{P}_2(\mathbb{R}^d))$.

Then $\exists\, v$ s.t. (μ, v) solve (CE) and

$$\left(\int_{\mathbb{R}^d} |v(x,t)|^2 d\mu_t(x)\right)^{1/2} \leq |\mu'|(t), \text{ a.e. } t$$

(ii) Suppose (μ, v) solve (CE) and

$$\int_0^T \int_{\mathbb{R}^d} |v(x,t)|^2 d\mu_t(x)\, dt < +\infty.$$

Then $\mu \in AC^2(0,T;\mathcal{P}_2(\mathbb{R}^d))$ and

$$|\mu'|(t) \leq \left(\int_{\mathbb{R}^d} |v(x,t)|^2 d\mu_t(x)\right)^{1/2}, \text{ a.e. } t.$$

Rmk: If the result holds for $T=1$, then, by reparametrizing in time, it holds for all $T>0$.

Pf part (ii):
Suppose (μ, v) solves (CE) and

$$\int_0^1 \int_{\mathbb{R}^d} |v(x,t)|^2 d\mu_t(x)\, dt < +\infty.$$

First, we will show $\mu \in AC^2(0,1; \mathcal{P}_2(\mathbb{R}^d))$.

Let $\varphi_\varepsilon : \mathbb{R}^d \to [0,+\infty)$ be a Gaussian with standard deviation $\varepsilon > 0$.

Let $\eta : \mathbb{R}^d \to [0,1]$ be a smooth, radially decreasing cutoff fn w/ $\eta \equiv 1$ on B_1, $\eta \equiv 0$ on B_2^c, $\|\nabla \eta\|_\infty \leq 2$.
Let $\eta_R(x) = \eta\left(\frac{x}{R}\right)$.

Then, for any $f \in C_c^\infty(\mathbb{R}^d)$,
$(\varphi_\varepsilon * f)\eta_R \in C_c^\infty(\mathbb{R}^d)$, so

$$\int_{\mathbb{R}^d} (\varphi_\varepsilon * f)\eta_R \, d\mu_{t_1} - \int_{\mathbb{R}^d} (\varphi_\varepsilon * f)\eta_R \, d\mu_{t_0} = \int_{t_0}^{t_1} \int_{\mathbb{R}^d} \nabla\big((\varphi_\varepsilon * f)\eta_R\big) \cdot v_t \, d\mu_t \, dt$$

$$\Big\downarrow \begin{array}{l} R \to \infty \\ \text{DCT} \end{array}$$

$$\int_{\mathbb{R}^d} (\varphi_\varepsilon * f) \, d\mu_{t_1} - \int_{\mathbb{R}^d} (\varphi_\varepsilon * f) \, d\mu_{t_0} = \int_{t_0}^{t_1} \int_{\mathbb{R}^d} \overbrace{\nabla(\varphi_\varepsilon * f)}^{= \varphi_\varepsilon * \nabla f} \cdot v_t \, d\mu_t \, dt$$

$$\Big\downarrow \text{Fubini}$$

$$\int_{\mathbb{R}^d} f \, (\varphi_\varepsilon * \mu_{t_1}) \, dx - \int_{\mathbb{R}^d} f \, (\varphi_\varepsilon * \mu_{\cancel{t_1} t_0}) = \int_{t_0}^{t_1} \int_{\mathbb{R}^d} \nabla f \cdot \overbrace{\varphi_\varepsilon * (v_t \mu_t)}^{\overbrace{\frac{\varphi_\varepsilon * (v_t \mu_t)}{\varphi_\varepsilon * \mu_t}}^{v_t^\varepsilon} \varphi_\varepsilon * \mu_t \, dx} \, dx \, dt$$

$$\varphi_\varepsilon * \mu_t(x) = \int \varphi_\varepsilon(x-y) \, d\mu_t(y)$$

Thus, $(\varphi_\varepsilon * \mu_t \, dx, v_t^\varepsilon)$ is a weak soln of (CE) and

$$f_B(x,r) \underset{r>0}{=} \frac{|x|^2}{r}$$

$$\int |v_t^\varepsilon|^2 \varphi_\varepsilon * \mu_t \, dx$$

$$= \int \frac{|\varphi_\varepsilon * (v_t \mu_t)|^2}{\varphi_\varepsilon * \mu_t} dx$$

$$= \int f_B(\varphi_\varepsilon * (v_t \mu_t), \varphi_\varepsilon * \mu_t) \, dx$$

$$= \int f_B\Big(\frac{c_\varepsilon}{c_\varepsilon} \int v_t(y) \overbrace{\varphi_\varepsilon(x-y) \, d\mu_t(y)}^{d\nu(y), \text{ let } c_\varepsilon = \int d\nu(y)}, \frac{c_\varepsilon}{c_\varepsilon} \int \varphi_\varepsilon(x-y) \, d\mu_t(y) \Big) dx$$

1-homogeneity of f_B

$$\leq \int \int \frac{1}{\cancel{c_\varepsilon}} f_B(\cancel{c_\varepsilon} v_t(y), \cancel{c_\varepsilon}) \varphi_\varepsilon(x-y) \, d\mu_t(y) \, dx$$

$\downarrow$ Fubini

$$= \int f_B(v_t(y), 1) \, d\mu_t(y)$$

$$= \int |v_t|^2 d\mu_t < +\infty$$

Important fact:
Since $\forall \varepsilon > 0$, v_t^ε is locally bounded and Lipschitz, unif in $t \in [0,1]$, the solution of (CE) with initial data $(\varphi_\varepsilon * \mu_0)$ is unique and is of the form

$$\varphi_\varepsilon * \mu_t = X_t^\varepsilon \# (\varphi_\varepsilon * \mu_0)$$

for

$$\begin{cases} \dot{X}_t^\varepsilon = v_t^\varepsilon (X_{\varepsilon, t}) \\ X_0^\varepsilon = id \end{cases}$$

(See AGS Prop 8.18.)

Thus $(X_t^\varepsilon, X_s^\varepsilon) \# (\varphi_\varepsilon * \mu_0) \in \Gamma(\varphi_\varepsilon * \mu_t, \varphi_\varepsilon * \mu_s)$,
so

$$W_2^2(\varphi_\varepsilon * \mu_t, \varphi_\varepsilon * \mu_s) \leq \int_{\mathbb{R}^d} |X_t^\varepsilon - X_s^\varepsilon|^2 d(\varphi_\varepsilon * \mu_0)$$

$$\ldots \int_{\mathbb{R}^d} \left| \int_s^t \dot{x}_r^\varepsilon \, dr \right|^2 d(\varphi_\varepsilon * \mu_0)$$

$$= \int_{\mathbb{R}^d} \left| \int_s^t v_r^\varepsilon(x_r^\varepsilon) \, dr \right|^2 d\varphi_\varepsilon * \mu_0$$

Jensen
Fubini

$$\leq |t-s| \int_s^t \int_{\mathbb{R}^d} |v_r^\varepsilon(x_r^\varepsilon)|^2 d\varphi_\varepsilon * \mu_0 \, dr$$

$$= |t-s| \int_s^t \int_{\mathbb{R}^d} |v_r^\varepsilon(x)|^2 d\varphi_\varepsilon * \mu_r(x) \, dr$$

$$\leq |t-s| \int_s^t \int_{\mathbb{R}^d} |v_r(x)|^2 d\mu_r(x) \, dr$$

To send $\varepsilon \to 0$, note that $\varphi_\varepsilon * \mu_t \xrightarrow{\varepsilon \to 0} \mu_t$ narrowly, by lsc of W_2 wrt narrow convergence,

$$\frac{W_2^2(\mu_t, \mu_s)}{|t-s|^2} \leq \fint_s^t \int_{\mathbb{R}^d} |v_r(x)|^2 d\mu_r(x) \, dr$$

By Lebesgue differentiation, for $\mathcal{L}^1$-a.e. $t \in [0,1]$, the RHS is locally bdd as $s \to t$. Thus, μ_t is locally Lipschitz, so $|\mu'|(t)$ exists for a.e. $t \in [0,1]$. Thus, sending $s \to t$ in ☆ gives

$$|\mu'|^2(t) \le \underbrace{\int_{\mathbb{R}^d} |v_t(x)|^2 d\mu_t(x)}_{\in L^1([0,1])}$$

so $\mu \in AC^2(0,1;\mathcal{P}_2(\mathbb{R}^d))$.

☆ Dynamic characterization of W_2 ☆

<u>Cor</u> (Benamou-Brenier): For all $\mu_0, \mu_1 \in \mathcal{P}_2(\mathbb{R}^d)$,

$$W_2^2(\mu_0,\mu_1) = \min_{\substack{(\mu,v) \text{ solve (CE)} \\ \mu_t|_{t=0}=\mu_0,\ \mu_t|_{t=1}=\mu_1}} \left\{ \int_0^1 \int_{\mathbb{R}^d} |v(x,t)|^2 d\mu_t \, dt \right\}.$$

Pf: We already showed that, $\forall\, \mu_0, \mu_1 \in \mathcal{P}_2(\mathbb{R}^d)$, there exists a geodesic μ between them and a velocity v s.t. (μ, v) satisfy (CE) and $\forall\, t \in [0,1]$,

$$\int_0^1 \int_{\mathbb{R}^d} |v_t|^2 d\mu_t\, dt = \int |v_t|^2 d\mu_t = W_2^2(\mu_0, \mu_1)$$

Thus, it suffices to show that, for any (μ, v) as in the constraint set, we have $\int_0^1 \int_{\mathbb{R}^d} |v_t|^2 d\mu_t\, dt \geq W_2^2(\mu_0, \mu_1)$

By the previous thm, $\mu \in AC^2(0,1; \mathcal{P}_2(\mathbb{R}^d))$ and

$$W_2^2(\mu_0, \mu_1) \leq \left(\int_0^1 |\mu'|(t)\, dt\right)^2 \leq \int_0^1 |\mu'|^2(t)\, dt$$

$$\leq \int_0^1 \int_{\mathbb{R}^d} |v_t|^2 d\mu_t\, dt. \qquad \square$$

Heuristic Riemannian structure of W_2

Recall...

M smooth manifold, $p \in M$

Riemannian metric $g_p : \mathrm{Tan}_p M \times \mathrm{Tan}_p M \to \mathbb{R}$
smooth, positive definite inner product

For $\rho_0, \rho_1 \in M$,

$$d_M^2(\rho_0, \rho_1) = \inf_{\substack{\rho : [0,1] \to M \text{ piecewise smooth} \\ \rho(0) = \rho_0,\ \rho(1) = \rho_1}} \int_0^1 \overbrace{\|\dot{\rho}(t)\|^2_{\mathrm{Tan}_{\rho(t)} M}}^{g_\rho(\dot{\rho}(t), \dot{\rho}(t))} \, dt.$$

$$= \inf_{\substack{\rho \in AC^2(0,1;M) \\ \rho(0) = \rho_0,\ \rho(1) = \rho_1}} \int_0^1 \overbrace{\|\dot{\rho}(t)\|^2_{\mathrm{Tan}_{\rho(t)} M}}^{|\rho'|(t) \text{ for a.e. } t} \, dt$$

Thus, Benamou and Brenier's dynamic characterization of W_2 suggests the following analogy...

Smooth Riemannian Mfld	2-Wasserstein Space
(M, d_M)	$(\mathcal{P}_2(\mathbb{R}^d), W_2)$
$AC^2(0,1;M)$	Solutions (μ,v) of (CE) with finite kinetic energy
	in principle, this may only exists a.e. t
$\mathrm{Tan}_p M$ $= \{ \dot{p}(t)\vert_{t=0} : p$ p.w. smooth, $p(0)=p \}$	$\mathrm{Tan}_\mu \mathcal{P}_2(\mathbb{R}^d)$ $= \{ \partial_t \mu_t \vert_{t=0} : \mu_t$ solves (CE), $\mu_0 = \mu \}$ $= \{ v_t\vert_{t=0} : (\mu,v)$ solve (E), $\mu_0=\mu \}$
$\Vert \dot{p}(t)\vert_{t=0} \Vert^2_{\mathrm{Tan}_p M}$ $= g_p(\dot{p}(0), \dot{p}(0))$	$\Vert v_0 \Vert^2_{\mathrm{Tan}_{\mu_0} \mathcal{P}_2(\mathbb{R}^d)}$ $= \int \vert v_0 \vert^2 d\mu_0$

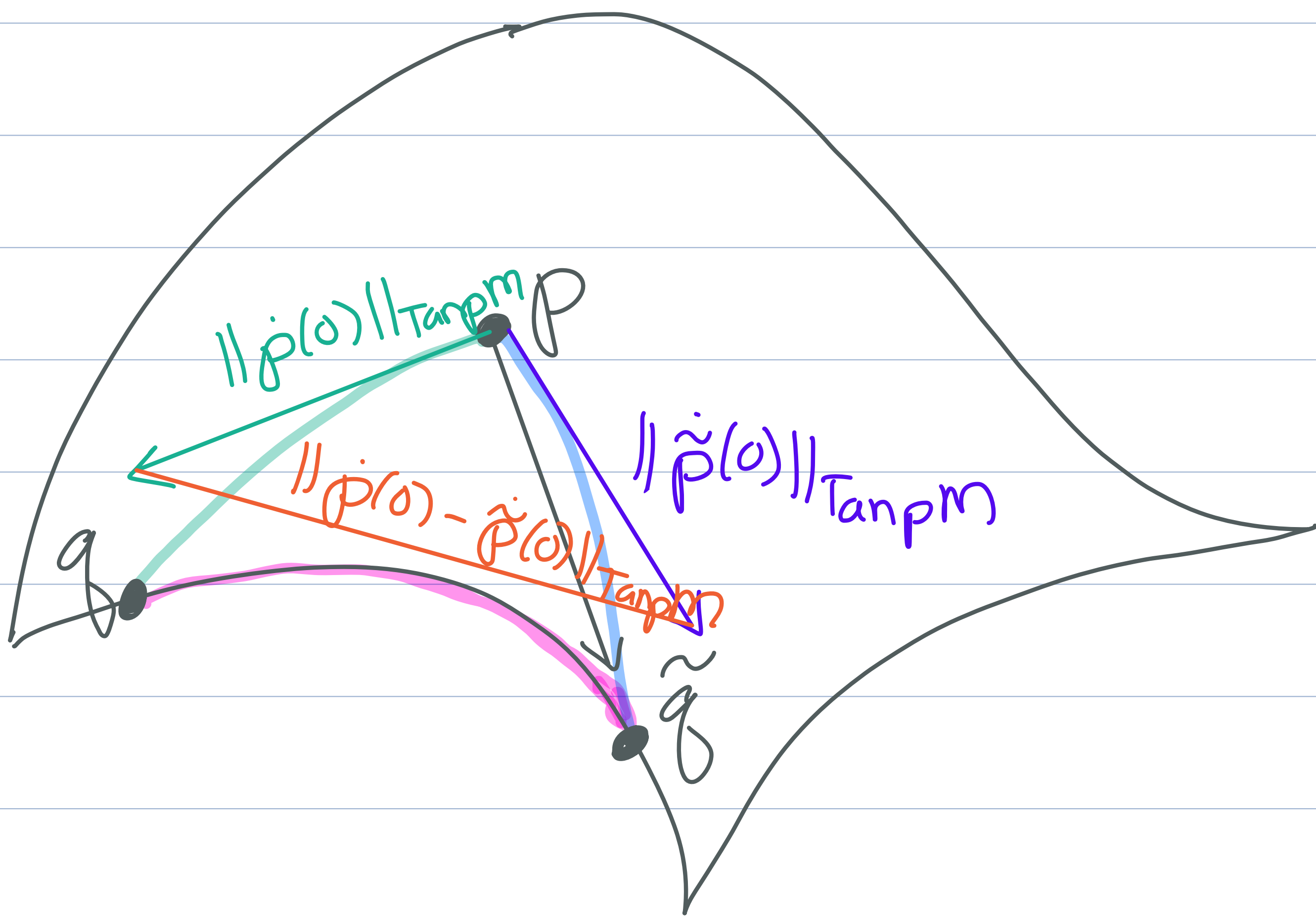
This suggests

$$\mathrm{Tan}_\mu \mathcal{P}_2(\mathbb{R}^d) = L^2_\mu(\mathbb{R}^d; \mathbb{R}^d)$$

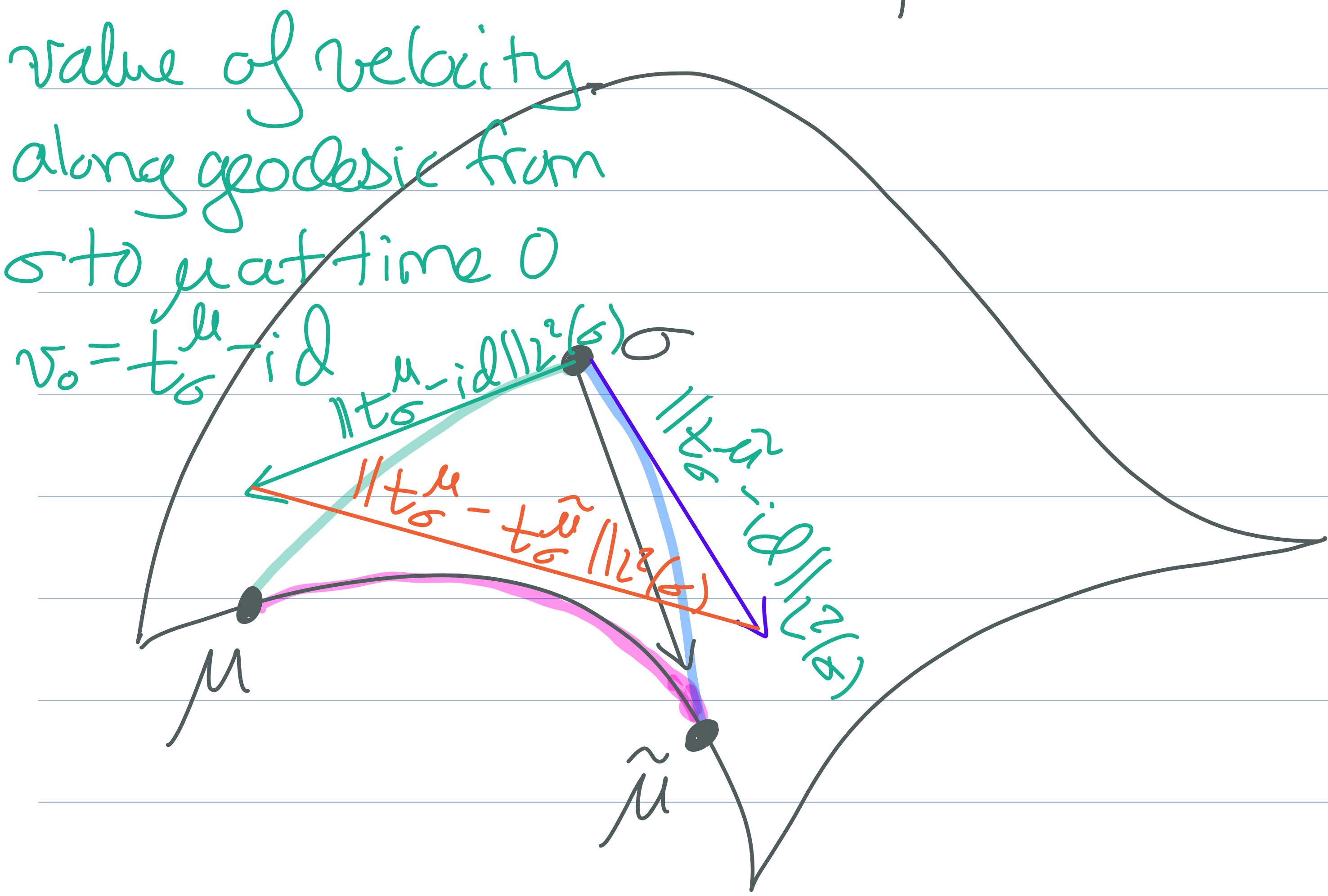
$$g_\mu(f, g) = \int f \cdot g \, d\mu$$

Application: Linearized OT

Wang, et al. "A linear optimal transportation framework," 2012.



Assume $\sigma \ll \mathcal{L}^d$ and let t_σ^μ denote the OT map from σ to μ.



Thus, to approximate pairwise W_2 distances between $\{\mu_i\}_{i=1}^N$, it suffices to fix $\sigma \ll \mathcal{L}^d$ and compute

$$\{t_\sigma^{\mu_i}\}_{i=1}^N$$

← N expensive OT computations

Then,

$$W_2(\mu_i, \mu_j) \approx \| t_\sigma^{\mu_i} - t_\sigma^{\mu_j} \|_{L^2(\sigma)}.$$

$O(N^2)$ cheap $L^2(\sigma)$ computations

Furthermore

$$\mu \mapsto t_\sigma^{\mu}$$

provides an "almost" isometric embedding of $(\mathbb{P}_2(\mathbb{R}^d), W_2)$ into $L^2(\sigma)$.

See many works by Delalande and Mérigot

Heuristic computation of W_2 gradient:

Given $E : \mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R} \cup \{+\infty\}$,
fix $\mu_0 \in \mathcal{P}_2(\mathbb{R}^d)$

We will begin by taking directional derivative of E in the direction $v \in \mathrm{Tan}_{\mu_0} \mathcal{P}_2(\mathbb{R}^d)$.

Let (μ, v) be soln of (CE) w/ i.c. μ_0

$$\lim_{t \to 0} \frac{E(\mu_t) - E(\mu_0)}{t}$$

Assuming $\frac{\delta E}{\delta \mu_0}$ exists

$$= \int \frac{\delta E}{\delta \mu_0} \partial_t \mu_t \Big|_{t=0}$$

$$= \int \frac{\delta E}{\delta \mu_0} \left(- \nabla \cdot (v_t \mu_t) \Big|_{t=0} \right)$$

$$= \int \nabla \frac{\delta E}{\delta \mu_0} \cdot v_0 \, d\mu_0 = g_{\mu_0}\left(\nabla \frac{\delta E}{\delta \mu_0}, v_0 \right)$$

This suggests $\nabla_{W_2} E(\mu_0) = \nabla \frac{\delta E}{\delta \mu_0}$.

AGS

OTOH,

$$\mathrm{Tan}_\mu \mathcal{P}_2(\mathbb{R}^d) = \left\{ \partial_t \mu_t \big|_{t=0} : \begin{array}{l} \mu_t \text{ solves (CE)} \\ \mu_0 = \mu \end{array} \right\}$$

Then $\nabla \frac{\delta E}{\delta \mu_0}$ is the velocity field and

$$\partial_t \mu_t \Big|_{t=0} = -\nabla \cdot \left(\left(\nabla \frac{\delta E}{\delta \mu_0} \right) \mu_t \right) \Big|_{t=0}$$

$$= \underbrace{-\nabla \cdot \left(\left(\nabla \frac{\delta E}{\delta \mu_0} \right) \mu_0 \right)}_{\nabla_{W_2} E(\mu_0) \text{ in Otto.}}$$

Thanks for a great quarter!