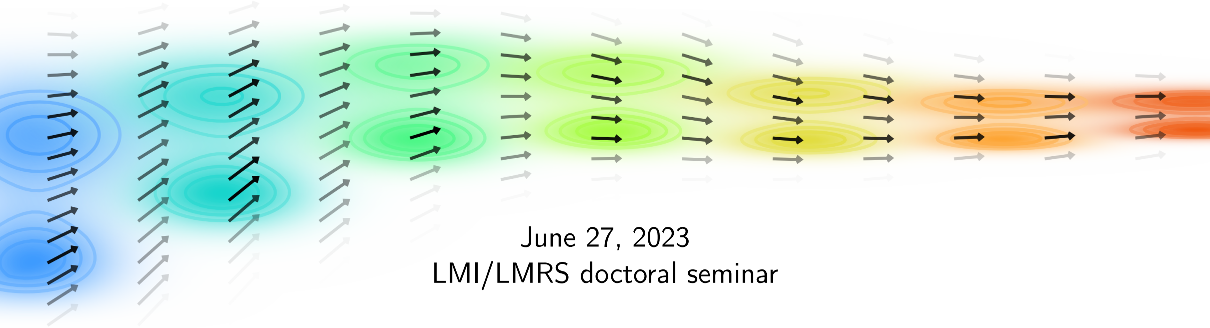
The D in PDE

Strategies for first-order differentiation in the space of measures

Averil Prost



June 27, 2023
LMI/LMRS doctoral seminar

Let $\mathscr{P}_2(\mathbb{R}^d)$ be the space of (nonnegative Borel) probability measures on the space $\mathbb{R}^d$ with

$$\int_{x\in\mathbb{R}^d} |x|^2 \, d\mu(x) < \infty.$$

Let $\mathscr{P}_2(\mathbb{R}^d)$ be the space of (nonnegative Borel) probability measures on the space $\mathbb{R}^d$ with

$$\int_{x\in\mathbb{R}^d} |x|^2 \, d\mu(x) < \infty.$$

Aim For certain $u : \mathscr{P}_2(\mathbb{R}^d) \to \mathbb{R}$, give a meaning to $\partial_\mu u(\mu)$.

Let $\mathscr{P}_2(\mathbb{R}^d)$ be the space of (nonnegative Borel) probability measures on the space $\mathbb{R}^d$ with

$$\int_{x\in\mathbb{R}^d} |x|^2 \, d\mu(x) < \infty.$$

Aim For certain $u : \mathscr{P}_2(\mathbb{R}^d) \to \mathbb{R}$, give a meaning to $\partial_\mu u(\mu)$.

For instance, we would like that the equation

$$\partial_t u(t,\mu) + \langle \partial_\mu u(t,\mu), b\rangle = 0, \qquad u(0,\mu) = u_0(\mu)$$

admits as solution $u(t,\mu) = u_0\left((id - tb)\#\mu\right)$.

Let $\mathscr{P}_2(\mathbb{R}^d)$ be the space of (nonnegative Borel) probability measures on the space $\mathbb{R}^d$ with

$$\int_{x\in\mathbb{R}^d} |x|^2 \, d\mu(x) < \infty.$$

Aim For certain $u : \mathscr{P}_2(\mathbb{R}^d) \to \mathbb{R}$, give a meaning to $\partial_\mu u(\mu)$.

For instance, we would like that the equation

$$\partial_t u(t,\mu) + \langle \partial_\mu u(t,\mu), b\rangle = 0, \qquad u(0,\mu) = u_0(\mu)$$

admits as solution $u(t,\mu) = u_0\left((id - tb)\#\mu\right)$.

This talk will review the definitions of the literature, going from smoothest to most general.

Table of Contents

Some definitions

Let $\mathcal{D} := \mathcal{C}_c^1\left(\mathbb{R}^d, \mathbb{R}\right)$. For each initial measure μ, denote $(\mu_s^{\mu,p})_{s\geqslant 0}$ the unique solution of the continuity equation

$$\partial_s \mu_s + \text{div}\left(\nabla p\, \mu_s\right) = 0, \qquad \mu_0 = \mu.$$

Some definitions

Let $\mathcal{D} := \mathcal{C}_c^1\left(\mathbb{R}^d, \mathbb{R}\right)$. For each initial measure μ, denote $(\mu_s^{\mu,p})_{s\geqslant 0}$ the unique solution of the continuity equation

$$\partial_s \mu_s + \text{div}\left(\nabla p\, \mu_s\right) = 0, \qquad \mu_0 = \mu.$$

Def 1 – Distributional derivative A map $u : \mathscr{P}(\mathbb{R}^d) \to \mathbb{R}$ admit a distributional derivative if there exist a distribution $\text{grad}_\mu u \in \mathcal{D}'$ such that for all $p \in \mathcal{D}$,

$$\lim_{s\searrow 0} \frac{u(\mu_s^{\mu,p}) - u(\mu)}{s} = \left\langle \text{grad}_\mu u(\mu), p \right\rangle_{\mathcal{D}',\mathcal{D}}.$$

Some definitions

Let $\mathcal{D} := \mathcal{C}_c^1\left(\mathbb{R}^d, \mathbb{R}\right)$. For each initial measure μ, denote $(\mu_s^{\mu,p})_{s\geqslant 0}$ the unique solution of the continuity equation

$$\partial_s \mu_s + \operatorname{div}\left(\nabla p\, \mu_s\right) = 0, \qquad \mu_0 = \mu.$$

Def 1 – Distributional derivative A map $u : \mathscr{P}(\mathbb{R}^d) \to \mathbb{R}$ admit a distributional derivative if there exist a distribution $\operatorname{grad}_\mu u \in \mathcal{D}'$ such that for all $p \in \mathcal{D}$,

$$\lim_{s \searrow 0} \frac{u(\mu_s^{\mu,p}) - u(\mu)}{s} = \left\langle \operatorname{grad}_\mu u(\mu), p \right\rangle_{\mathcal{D}',\mathcal{D}}.$$

This definition is used in [FK09, FN12] to adress Hamilton-Jacobi equations.

Example: the linear map

Consider the map $u : \mu \mapsto \int_{x\in\mathbb{R}^d} \ell(x) d\mu(x)$.

Example: the linear map

Consider the map $u : \mu \mapsto \int_{x\in\mathbb{R}^d} \ell(x)d\mu(x)$. If ℓ is Lipschitz and $\mathcal{C}^1$, then u is Lipschitz from $\mathscr{P}_2(\mathbb{R}^d)$ to $\mathbb{R}$, and

$$\lim_{s\searrow 0} \frac{u(\mu_s^{\mu,p}) - u(\mu)}{s} = \lim_{s\searrow 0} \frac{u((id + s\nabla p)\#\mu) - u(\mu)}{s}$$

Example: the linear map

Consider the map $u : \mu \mapsto \int_{x\in\mathbb{R}^d} \ell(x)d\mu(x)$. If ℓ is Lipschitz and $\mathcal{C}^1$, then u is Lipschitz from $\mathscr{P}_2(\mathbb{R}^d)$ to $\mathbb{R}$, and

$$\lim_{s\searrow 0} \frac{u(\mu_s^{\mu,p}) - u(\mu)}{s} = \lim_{s\searrow 0} \frac{u((id + s\nabla p)\#\mu) - u(\mu)}{s} = \lim_{s\searrow 0} \int_{x\in\mathbb{R}^d} \frac{\ell(x + s\nabla p(x)) - \ell(x)}{s} d\mu(x)$$

Example: the linear map

Consider the map $u : \mu \mapsto \int_{x\in\mathbb{R}^d} \ell(x)d\mu(x)$. If ℓ is Lipschitz and $\mathcal{C}^1$, then u is Lipschitz from $\mathscr{P}_2(\mathbb{R}^d)$ to $\mathbb{R}$, and

$$\lim_{s\searrow 0} \frac{u(\mu_s^{\mu,p}) - u(\mu)}{s} = \lim_{s\searrow 0} \frac{u((id + s\nabla p)\#\mu) - u(\mu)}{s} = \lim_{s\searrow 0} \int_{x\in\mathbb{R}^d} \frac{\ell(x + s\nabla p(x)) - \ell(x)}{s} d\mu(x)$$
$$= \int_{x\in\mathbb{R}^d} \langle \nabla\ell(x), \nabla p(x)\rangle \, d\mu(x)$$

Example: the linear map

Consider the map $u : \mu \mapsto \int_{x\in\mathbb{R}^d} \ell(x)d\mu(x)$. If ℓ is Lipschitz and $\mathcal{C}^1$, then u is Lipschitz from $\mathscr{P}_2(\mathbb{R}^d)$ to $\mathbb{R}$, and

$$\lim_{s\searrow 0} \frac{u(\mu_s^{\mu,p}) - u(\mu)}{s} = \lim_{s\searrow 0} \frac{u((id + s\nabla p)\#\mu) - u(\mu)}{s} = \lim_{s\searrow 0} \int_{x\in\mathbb{R}^d} \frac{\ell(x + s\nabla p(x)) - \ell(x)}{s} d\mu(x)$$
$$= \int_{x\in\mathbb{R}^d} \langle \nabla \ell(x), \nabla p(x) \rangle \, d\mu(x) = - \langle \operatorname{div} (\mu \nabla \ell) \, , p \rangle \, ,$$

so that the distributional derivative of u is the distribution $\operatorname{grad}_\mu u(\mu) := - \operatorname{div} (\mu \nabla \ell)$.

Example: the linear map

Consider the map $u : \mu \mapsto \int_{x\in\mathbb{R}^d} \ell(x) d\mu(x)$. If ℓ is Lipschitz and $\mathcal{C}^1$, then u is Lipschitz from $\mathscr{P}_2(\mathbb{R}^d)$ to $\mathbb{R}$, and

$$\lim_{s\searrow 0} \frac{u(\mu_s^{\mu,p}) - u(\mu)}{s} = \lim_{s\searrow 0} \frac{u((id + s\nabla p)\#\mu) - u(\mu)}{s} = \lim_{s\searrow 0} \int_{x\in\mathbb{R}^d} \frac{\ell(x + s\nabla p(x)) - \ell(x)}{s} d\mu(x)$$
$$= \int_{x\in\mathbb{R}^d} \langle \nabla \ell(x), \nabla p(x)\rangle \, d\mu(x) = - \langle \text{div } (\mu \nabla \ell) \, , p\rangle \, ,$$

so that the distributional derivative of u is the distribution $\text{grad}_\mu u(\mu) := -\text{div } (\mu\nabla\ell)$.

Remark 1 – Meaning of the divergence Here $\text{div } (\mu\cdot)$ is a *notation* for the adjoint operator of the gradient, i.e. $\langle \text{div}\, (\mu F), p\rangle := \langle F, \nabla p\rangle_\mu = \int_{x\in\mathbb{R}^d} \langle F(x), \nabla p(x)\rangle \, d\mu(x)$. In particular, if μ is the Lebesgue measure, it contains the boundary terms.

The Otto calculus

The work of Otto [JKO98, Ott01] contributed to raise the interest in this family of derivatives. The "formal Otto calculus" allows to recast canonical equations as gradient flows in the Wasserstein space, as summarized in [Vil09].

The Otto calculus

The work of Otto [JKO98, Ott01] contributed to raise the interest in this family of derivatives. The "formal Otto calculus" allows to recast canonical equations as gradient flows in the Wasserstein space, as summarized in [Vil09].

Consider the map $u(\mu) := \int_{x \in \mathbb{R}^d} U(\rho(x))\, dx$, where $\mu = \rho dx$. Denote $\mu_t = \rho(t,x)\nu$ the solution of $\partial_t \mu + \operatorname{div}(\mu \nabla p) = 0$.

The Otto calculus

The work of Otto [JKO98, Ott01] contributed to raise the interest in this family of derivatives. The "formal Otto calculus" allows to recast canonical equations as gradient flows in the Wasserstein space, as summarized in [Vil09].

Consider the map $u(\mu) := \int_{x\in\mathbb{R}^d} U\left(\rho(x)\right) dx$, where $\mu = \rho dx$. Denote $\mu_t = \rho(t,x)\nu$ the solution of $\partial_t\mu + \operatorname{div}\left(\mu\nabla p\right) = 0$. Then, at least formally,

$$\frac{d}{dt}_{|t=0} u(\mu_t) = \int_{x\in\mathbb{R}^d} U'(\rho_0)\partial_t\rho_0 dx$$

The Otto calculus

The work of Otto [JKO98, Ott01] contributed to raise the interest in this family of derivatives. The "formal Otto calculus" allows to recast canonical equations as gradient flows in the Wasserstein space, as summarized in [Vil09].

Consider the map $u(\mu) := \int_{x\in\mathbb{R}^d} U(\rho(x))\,dx$, where $\mu = \rho dx$. Denote $\mu_t = \rho(t,x)\nu$ the solution of $\partial_t \mu + \operatorname{div}(\mu\nabla p) = 0$. Then, at least formally,

$$\frac{d}{dt}_{|t=0} u(\mu_t) = \int_{x\in\mathbb{R}^d} U'(\rho_0)\partial_t\rho_0 dx = \int_{x\in\mathbb{R}^d} U'(\rho_0)\left(-\operatorname{div}(\rho_0\nabla p)\right)dx$$

The Otto calculus

The work of Otto [JKO98, Ott01] contributed to raise the interest in this family of derivatives. The "formal Otto calculus" allows to recast canonical equations as gradient flows in the Wasserstein space, as summarized in [Vil09].

Consider the map $u(\mu) := \int_{x\in\mathbb{R}^d} U\left(\rho(x)\right) dx$, where $\mu = \rho dx$. Denote $\mu_t = \rho(t,x)\nu$ the solution of $\partial_t \mu + \operatorname{div}\left(\mu \nabla p\right) = 0$. Then, at least formally,

$$\begin{aligned} \frac{d}{dt}_{|t=0} u(\mu_t) &= \int_{x\in\mathbb{R}^d} U'(\rho_0)\partial_t \rho_0 dx = \int_{x\in\mathbb{R}^d} U'(\rho_0)\left(-\operatorname{div}\left(\rho_0 \nabla p\right)\right) dx \\ &= \int_{x\in\mathbb{R}^d} \left\langle \rho_0 \nabla [U' \circ \rho_0], \nabla p \right\rangle dx \end{aligned}$$

The Otto calculus

The work of Otto [JKO98, Ott01] contributed to raise the interest in this family of derivatives. The "formal Otto calculus" allows to recast canonical equations as gradient flows in the Wasserstein space, as summarized in [Vil09].

Consider the map $u(\mu) := \int_{x\in\mathbb{R}^d} U(\rho(x))\,dx$, where $\mu = \rho dx$. Denote $\mu_t = \rho(t,x)\nu$ the solution of $\partial_t \mu + \text{div}\,(\mu\nabla p) = 0$. Then, at least formally,

$$\begin{aligned}\frac{d}{dt}_{|t=0} u(\mu_t) &= \int_{x\in\mathbb{R}^d} U'(\rho_0)\partial_t\rho_0 dx = \int_{x\in\mathbb{R}^d} U'(\rho_0)\left(-\,\text{div}\,(\rho_0\nabla p)\right)dx \\ &= \int_{x\in\mathbb{R}^d} \left\langle \rho_0\nabla[U'\circ\rho_0], \nabla p\right\rangle dx = \int_{x\in\mathbb{R}^d} -\,\text{div}\,\left(\rho_0\nabla[U'\circ\rho_0]\right)p(x)dx.\end{aligned}$$

The Otto calculus

The work of Otto [JKO98, Ott01] contributed to raise the interest in this family of derivatives. The "formal Otto calculus" allows to recast canonical equations as gradient flows in the Wasserstein space, as summarized in [Vil09].

Consider the map $u(\mu) := \int_{x\in\mathbb{R}^d} U(\rho(x))\,dx$, where $\mu = \rho dx$. Denote $\mu_t = \rho(t,x)\nu$ the solution of $\partial_t \mu + \operatorname{div}(\mu\nabla p) = 0$. Then, at least formally,

$$\begin{aligned}\frac{d}{dt}_{|t=0} u(\mu_t) &= \int_{x\in\mathbb{R}^d} U'(\rho_0)\partial_t\rho_0 dx = \int_{x\in\mathbb{R}^d} U'(\rho_0)\left(-\operatorname{div}(\rho_0\nabla p)\right)dx \\ &= \int_{x\in\mathbb{R}^d} \langle \rho_0\nabla[U'\circ\rho_0], \nabla p\rangle\,dx = \int_{x\in\mathbb{R}^d} -\operatorname{div}\left(\rho_0\nabla[U'\circ\rho_0]\right)p(x)dx.\end{aligned}$$

Hence $\operatorname{grad}_\mu u = -\operatorname{div}\left(\rho\nabla[U'\circ\rho]\right)$. For instance, $U(r) = r\ln(r)$ gives $\operatorname{grad}_\mu u = -\Delta\rho(x)$.

Table of Contents

The lift

History in two parts: the original extrinsic *lifted* formulation, and the (quite new) *intrinsic* one.

The lift

History in two parts: the original extrinsic *lifted* formulation, and the (quite new) *intrinsic* one.

Fundamental theorem of simulation (name from [BL94], [CD18a, Lemma 5.29])
Let $(\Omega, \mathcal{A}, \mathbb{P})$ be an atomless probability space, and $\mu \in \mathscr{P}_2\left(\mathbb{R}^d\right)$. Then there exist $X \in L^2\left(\Omega, \mathcal{A}, \mathbb{P}; \mathbb{R}^d\right)$ such that

the law of X is μ, i.e. $\mu = X\#\mathbb{P}$, i.e. $\mu(A) = \mathbb{P}(X^{-1}(A)) \quad \forall A \in \mathcal{A}$.

The lift

History in two parts: the original extrinsic *lifted* formulation, and the (quite new) *intrinsic* one.

Fundamental theorem of simulation (name from [BL94], [CD18a, Lemma 5.29])
Let $(\Omega, \mathcal{A}, \mathbb{P})$ be an atomless probability space, and $\mu \in \mathscr{P}_2\left(\mathbb{R}^d\right)$. Then there exist $X \in L^2\left(\Omega, \mathcal{A}, \mathbb{P}; \mathbb{R}^d\right)$ such that

the law of X is μ, i.e. $\mu = X\#\mathbb{P}$, i.e. $\mu(A) = \mathbb{P}(X^{-1}(A)) \quad \forall A \in \mathcal{A}$.

Def 2 – Lift Let $u : \mathscr{P}_2(\mathbb{R}^d) \to \mathbb{R}$. Its *lift* is a map $U : L^2\left(\Omega, \mathcal{A}, \mathbb{P}; \mathbb{R}^d\right) \to \mathbb{R}$ given by

$$U(X) = u(\mathcal{L}(\mathcal{X})) = u(X\#\mathbb{P}).$$

Gradient using the Hilbert structure

Def 3 – L-derivative Assume that U is F-differentiable in $L^2\left(\Omega, \mathcal{A}, \mathbb{P}; \mathbb{R}^d\right)$. Then for all $\mu \in \mathscr{P}_2(\mathbb{R}^d)$, there exist an element $\xi_\mu \in L^2_\mu\left(\mathbb{R}^d; \mathbb{R}^d\right)$ such that

$$\forall X \in L^2\left(\Omega, \mathcal{A}, \mathbb{P}; \mathbb{R}^d\right) \text{ s.t. } \mathcal{L}(X) = \mu, \qquad \nabla U(X)(\omega) = \xi_\mu(X(\omega)) \quad \forall \omega \in \Omega.$$

We then denote $\partial_\mu u(\mu) := \xi_\mu$. (Here, $\partial_\mu u(\mu)$ is a function in $L^2_\mu(\mathbb{R}^d; \mathbb{R}^d)$.)

Gradient using the Hilbert structure

Def 3 – L-derivative Assume that U is F-differentiable in $L^2\left(\Omega, \mathcal{A}, \mathbb{P}; \mathbb{R}^d\right)$. Then for all $\mu \in \mathscr{P}_2(\mathbb{R}^d)$, there exist an element $\xi_\mu \in L^2_\mu\left(\mathbb{R}^d; \mathbb{R}^d\right)$ such that

$$\forall X \in L^2\left(\Omega, \mathcal{A}, \mathbb{P}; \mathbb{R}^d\right) \text{ s.t. } \mathcal{L}(X) = \mu, \qquad \nabla U(X)(\omega) = \xi_\mu(X(\omega)) \quad \forall \omega \in \Omega.$$

We then denote $\partial_\mu u(\mu) := \xi_\mu$. (Here, $\partial_\mu u(\mu)$ is a function in $L^2_\mu(\mathbb{R}^d; \mathbb{R}^d)$.)

Idea launched by P.L. Lions in [Lio06], transcripted in [Car13]. Very popular notion, used (in particular) in [CCD15, PW17, PW18, BY19, CGK$^+$22, CGK$^+$22, MZ22] to make the link between SDEs and PDEs, with focus on the master equation. Higher order derivatives are also defined (see [Sal23] for arbitrary order).

Example

Consider again $u(\mu) := \int_{x\in\mathbb{R}^d} \ell(x) d\mu(x)$. Then its lift U is defined as

$$U(X) = \int_{\omega\in\Omega} \ell(X(\omega)) d\mathbb{P}(\omega).$$

Example

Consider again $u(\mu) := \int_{x\in\mathbb{R}^d} \ell(x)d\mu(x)$. Then its lift U is defined as

$$U(X) = \int_{\omega\in\Omega} \ell(X(\omega))d\mathbb{P}(\omega).$$

If $\ell \in \mathcal{C}^1$ and Lipschitz, then

$$\lim_{h\searrow 0} \frac{U(X+hY)-U(X)}{h} = \int_{\omega\in\Omega} \langle \nabla\ell(X(\omega)), Y(\omega)\rangle \, d\mathbb{P}(\omega) = \langle \nabla\ell \circ X, Y\rangle_{L^2_{\mathbb{P}}}.$$

Example

Consider again $u(\mu) := \int_{x\in\mathbb{R}^d} \ell(x)d\mu(x)$. Then its lift U is defined as

$$U(X) = \int_{\omega\in\Omega} \ell(X(\omega))d\mathbb{P}(\omega).$$

If $\ell \in \mathcal{C}^1$ and Lipschitz, then

$$\lim_{h\searrow 0} \frac{U(X+hY)-U(X)}{h} = \int_{\omega\in\Omega} \langle \nabla\ell(X(\omega)), Y(\omega)\rangle \, d\mathbb{P}(\omega) = \langle \nabla\ell\circ X, Y\rangle_{L^2_{\mathbb{P}}}.$$

Hence $DU(X) = \nabla\ell\circ X$, and $\partial_\mu u(\mu) = \nabla\ell : \mathbb{R}^d \to \mathbb{R}^d$ (here independant of μ).

Example

Consider again $u(\mu) := \int_{x\in\mathbb{R}^d} \ell(x)d\mu(x)$. Then its lift U is defined as

$$U(X) = \int_{\omega\in\Omega} \ell(X(\omega))d\mathbb{P}(\omega).$$

If $\ell \in \mathcal{C}^1$ and Lipschitz, then

$$\lim_{h\searrow 0} \frac{U(X+hY)-U(X)}{h} = \int_{\omega\in\Omega} \langle \nabla\ell(X(\omega)), Y(\omega)\rangle \, d\mathbb{P}(\omega) = \langle \nabla\ell \circ X, Y\rangle_{L^2_{\mathbb{P}}}.$$

Hence $DU(X) = \nabla\ell \circ X$, and $\partial_\mu u(\mu) = \nabla\ell : \mathbb{R}^d \to \mathbb{R}^d$ (here independant of μ).

Remark 2 – Insatisfaction This "delocalization" procedure does not seem really natural.

The linear derivative

Def 4 – Linear derivative A map $u : \mathscr{P}_2(\mathbb{R}^d) \to \mathbb{R}$ is said to admit a linear (functional) derivative if there exist a function $(\mu, x) \mapsto \frac{\delta u}{\delta \mu}(\mu, x)$ satisfying

- for all $\nu \in \mathscr{P}_2(\mathbb{R}^d)$, $\lim_{s \searrow 0} \frac{u(\mu+s(\nu-\mu))-u(\mu)}{s} = \int_{x\in\mathbb{R}^d} \frac{\delta u}{\delta \mu}(\mu, x) d\left[\nu - \mu\right](x)$,

The linear derivative

Def 4 – Linear derivative A map $u : \mathscr{P}_2(\mathbb{R}^d) \to \mathbb{R}$ is said to admit a linear (functional) derivative if there exist a function $(\mu, x) \mapsto \frac{\delta u}{\delta \mu}(\mu, x)$ satisfying

- for all $\nu \in \mathscr{P}_2(\mathbb{R}^d)$, $\lim_{s \searrow 0} \frac{u(\mu+s(\nu-\mu))-u(\mu)}{s} = \int_{x\in\mathbb{R}^d} \frac{\delta u}{\delta \mu}(\mu, x) d\left[\nu - \mu\right](x)$,
- the normalizing convention $\int_{x\in\mathbb{R}^d} \frac{\delta u}{\delta \mu}(\mu, x) d\mu(x) = 0$.

The linear derivative

Def 4 – Linear derivative A map $u : \mathscr{P}_2(\mathbb{R}^d) \to \mathbb{R}$ is said to admit a linear (functional) derivative if there exist a function $(\mu, x) \mapsto \frac{\delta u}{\delta \mu}(\mu, x)$ satisfying

- for all $\nu \in \mathscr{P}_2(\mathbb{R}^d)$, $\lim_{s \searrow 0} \frac{u(\mu+s(\nu-\mu))-u(\mu)}{s} = \int_{x\in\mathbb{R}^d} \frac{\delta u}{\delta \mu}(\mu, x) d\left[\nu - \mu\right](x)$,
- the normalizing convention $\int_{x\in\mathbb{R}^d} \frac{\delta u}{\delta \mu}(\mu, x) d\mu(x) = 0$.

This formulation goes back to Fleming-Viot processes [FV79], and is used outside of the Wasserstein context (see [CLS18] for references). It corresponds to the Fréchet derivative in the Banach space $(\mathcal{M}, |\cdot|_{TV})$, restricted to $\mathscr{P}_2(\mathbb{R}^d)$. Used in viscosity [BIRS19] and for the master equation [CD18a, CDLL19].

Example and chain rule

Let $u(\mu) := \int_{x\in\mathbb{R}^d} \ell(x)d\mu(x)$. Then one simply has

$$\frac{u(\mu + s(\nu - \mu)) - u(\mu)}{s} = \int_{x\in\mathbb{R}^d} \ell(x)d[\nu - \mu](x),$$

Example and chain rule

Let $u(\mu) := \int_{x\in\mathbb{R}^d} \ell(x) d\mu(x)$. Then one simply has

$$\frac{u(\mu + s(\nu - \mu)) - u(\mu)}{s} = \int_{x\in\mathbb{R}^d} \ell(x) d[\nu - \mu](x),$$

so that $\frac{\delta u}{\delta \mu}(\mu, x) = \ell(x) - \langle \ell, \mu \rangle$ (for the normalization).

Example and chain rule

Let $u(\mu) := \int_{x\in\mathbb{R}^d} \ell(x) d\mu(x)$. Then one simply has

$$\frac{u(\mu + s(\nu-\mu)) - u(\mu)}{s} = \int_{x\in\mathbb{R}^d} \ell(x) d[\nu-\mu](x),$$

so that $\frac{\delta u}{\delta \mu}(\mu, x) = \ell(x) - \langle \ell, \mu \rangle$ (for the normalization).

Chain rule 1 If $(x,\mu) \mapsto \frac{\delta u}{\delta \mu}(\mu, x)$ is Lipschitz in $\mathscr{P}_2(\mathbb{R}^d) \times \mathbb{R}^d$ and the curve $(\mu_t)_{t\in[0,T]} \subset \mathscr{P}_2(\mathbb{R}^d)$ is Lipschitz in time, then

$$u(\mu_T) - u(\mu_0) = \int_{t=0}^{T} \langle \frac{\delta u}{\delta \mu}(\mu_t, \cdot), \partial_t \mu_t \rangle dt. \qquad \text{(CR-}\delta/\delta\mu\text{)}$$

The (natural) derivative

Def 5 – Natural derivative Assume that $u : \mathscr{P}_2(\mathbb{R}^d) \to \mathbb{R}$ admits a linear derivative that is jointly continuous, and such that for all fixed $\mu \in \mathscr{P}_2(\mathbb{R}^d)$, the map $x \mapsto \frac{\delta u}{\delta \mu}(\mu, x)$ is differentiable in $\mathbb{R}^d$. Then one defines the natural derivative of u as

$$D_\mu u : \mathscr{P}_2(\mathbb{R}^d) \times \mathbb{R}^d \to \mathbb{R}^d, \qquad D_\mu u(\mu, x) = \nabla_x \frac{\delta u}{\delta \mu}(\mu, x).$$

The (natural) derivative

Def 5 – Natural derivative Assume that $u : \mathscr{P}_2(\mathbb{R}^d) \to \mathbb{R}$ admits a linear derivative that is jointly continuous, and such that for all fixed $\mu \in \mathscr{P}_2(\mathbb{R}^d)$, the map $x \mapsto \frac{\delta u}{\delta \mu}(\mu, x)$ is differentiable in $\mathbb{R}^d$. Then one defines the natural derivative of u as

$$D_\mu u : \mathscr{P}_2(\mathbb{R}^d) \times \mathbb{R}^d \to \mathbb{R}^d, \qquad D_\mu u(\mu, x) = \nabla_x \frac{\delta u}{\delta \mu}(\mu, x).$$

The notations are taken from [CD18a, CD18b], and this definition is used in [CDLL19]. The terminology is not clear, and we called $D_\mu u$ "natural derivative" in waiting of a better name.

Example and chain rule

According to Frame 12, the map $u(\mu) := \int_{x\in\mathbb{R}^d} \ell(x)d\mu(x)$ has a linear derivative $\frac{\delta u}{\delta\mu}(\mu,x) = \ell(x) - \langle \ell, \mu\rangle$. Hence we directly have

$$D_\mu u(\mu,x) = \nabla_x \frac{\delta u}{\delta\mu}(\mu,x) = \nabla\ell(x).$$

Example and chain rule

According to Frame 12, the map $u(\mu) := \int_{x\in\mathbb{R}^d} \ell(x)d\mu(x)$ has a linear derivative $\frac{\delta u}{\delta\mu}(\mu,x) = \ell(x) - \langle \ell, \mu\rangle$. Hence we directly have

$$D_\mu u(\mu,x) = \nabla_x \frac{\delta u}{\delta \mu}(\mu,x) = \nabla \ell(x).$$

Chain rule 2 Assume that $D_\mu u(\mu)$ is jointly continuous, and let the measure curve $(\mu_t)_{t\in[0,T]}$ solve $\partial_t \mu_t = -\operatorname{div}(g(t,\cdot,\mu_t)\#\mu_t)$. Then, from (CR-$\delta/\delta\mu$), we obtain

$$\begin{aligned} u(\mu_T) - u(\mu_0) &= \int_{t=0}^{T} \left\langle \nabla_x \frac{\delta u}{\delta\mu}(\mu,x), g(t,\cdot,\mu_t) \right\rangle_{\mu_t} dt \\ &= \int_{t=0}^{T} \int_{x\in\mathbb{R}^d} \langle D_\mu u(\mu_t,x), g(t,x,\mu_t)\rangle \, d\mu_t(x) dt. \end{aligned}$$

Links

Link ([CD18a, Prop. 5.48] and [CDLL19, Appx A]) Let $u : \mathscr{P}_2(\mathbb{R}^d) \to \mathbb{R}$, and assume that either

- u admits a jointly continuous Lions-differential $\partial_\mu u$ in the sense of Def 3 that has linear growth in x uniformly in μ,
- u admits a jointly continuous natural derivative $D_\mu u$ in the sense of Def 5 that has linear growth in x uniformly in μ.

Then the other point stands and

$$\partial_\mu u(\mu, x) = D_\mu u(\mu, x) \qquad \forall (\mu, x) \in \mathscr{P}_2(\mathbb{R}^d) \times \mathbb{R}^d.$$

Hence the two definitions are gathered under the vocabulary of "Lions differentiability".

Table of Contents

The regular case

Define a family of "tangent vectors" to μ as

$$T_\mu \mathscr{P}_2(\mathbb{R}^d) := \overline{\{\nabla\varphi \mid \varphi \in \mathcal{C}_c^1(\mathbb{R}^d;\mathbb{R})\}}^{L^2_\mu}.$$

The regular case

Define a family of "tangent vectors" to μ as

$$T_\mu \mathscr{P}_2(\mathbb{R}^d) := \overline{\{\nabla\varphi \mid \varphi \in \mathcal{C}_c^1(\mathbb{R}^d;\mathbb{R})\}}^{L^2_\mu}.$$

Def 6 – Regular semidifferentials Let $u : \mathscr{P}_2(\mathbb{R}^d) \to \mathbb{R}$ and $\mu \in \mathscr{P}_2(\mathbb{R}^d)$. An element $\xi \in T_\mu \mathscr{P}_2(\mathbb{R}^d)$ is said to belong to the subdifferential of u at μ if for all $\nu \in \mathscr{P}_2(\mathbb{R}^d)$,

$$u(\nu) - u(\mu) \geqslant \sup_{\eta \in \Gamma_o(\mu,\nu)} \int_{(x,y)\in(\mathbb{R}^d)^2} \langle \xi(x), y - x\rangle \, d\eta(x,y) + o\left(d_{\mathcal{W}}(\mu,\nu)\right).$$

The set of such ξ is denoted $\partial_{.} u(\mu)$. The superdifferential writes $\partial^{\cdot} u(\mu) := -\partial_{.}(-u)(\mu)$.

The regular case

Define a family of "tangent vectors" to μ as

$$T_\mu \mathscr{P}_2(\mathbb{R}^d) := \overline{\{\nabla\varphi \mid \varphi \in \mathcal{C}_c^1(\mathbb{R}^d;\mathbb{R})\}}^{L^2_\mu}.$$

Def 6 – Regular semidifferentials Let $u : \mathscr{P}_2(\mathbb{R}^d) \to \mathbb{R}$ and $\mu \in \mathscr{P}_2(\mathbb{R}^d)$. An element $\xi \in T_\mu\mathscr{P}_2(\mathbb{R}^d)$ is said to belong to the subdifferential of u at μ if for all $\nu \in \mathscr{P}_2(\mathbb{R}^d)$,

$$u(\nu) - u(\mu) \geqslant \sup_{\eta \in \Gamma_o(\mu,\nu)} \int_{(x,y)\in(\mathbb{R}^d)^2} \langle \xi(x), y - x\rangle \, d\eta(x,y) + o\left(d_{\mathcal{W}}(\mu,\nu)\right).$$

The set of such ξ is denoted $\partial_. u(\mu)$. The superdifferential writes $\partial^{\cdot} u(\mu) := -\partial_.(-u)(\mu)$.

This definition inspired the $\delta-$semidifferentials of [CQ08, MQ18, JMQ20, JMQ22].

Example

Let $u(\mu) = \int_{x \in \mathbb{R}^d} \ell(x) d\mu(x)$, and assume that $\ell \in \mathcal{C}^1$ is $\lambda-$semiconvex, i.e.

$$\ell(y) - \ell(x) \geqslant \langle \nabla \ell(x), y - x \rangle - \frac{\lambda}{2} |y - x|^2 .$$

Example

Let $u(\mu) = \int_{x\in\mathbb{R}^d} \ell(x) d\mu(x)$, and assume that $\ell \in \mathcal{C}^1$ is $\lambda-$semiconvex, i.e.

$$\ell(y) - \ell(x) \geqslant \langle \nabla \ell(x), y - x \rangle - \frac{\lambda}{2} |y - x|^2 .$$

Then, for any $(\mu, \nu) \in (\mathscr{P}_2(\mathbb{R}^d))^2$, integrating the above against $\eta \in \Gamma_o(\mu, \nu)$ yields

$$\underbrace{\int_{y\in\mathbb{R}^d} \ell(y) d\nu(y)}_{u(\nu)} - \underbrace{\int_{x\in\mathbb{R}^d} \ell(x) d\mu(x)}_{u(\mu)} \geqslant \int_{(x,y)\in(\mathbb{R}^d)^2} \langle \nabla \ell(x), y - x \rangle \, d\eta(x, y) - \frac{\lambda}{2} d^2_{\mathcal{W}}(\mu, \nu).$$

Example

Let $u(\mu) = \int_{x\in\mathbb{R}^d} \ell(x)d\mu(x)$, and assume that $\ell \in \mathcal{C}^1$ is $\lambda-$semiconvex, i.e.

$$\ell(y) - \ell(x) \geqslant \langle \nabla \ell(x), y - x\rangle - \frac{\lambda}{2}\left|y - x\right|^2.$$

Then, for any $(\mu,\nu) \in (\mathscr{P}_2(\mathbb{R}^d))^2$, integrating the above against $\eta \in \Gamma_o(\mu,\nu)$ yields

$$\underbrace{\int_{y\in\mathbb{R}^d} \ell(y)d\nu(y)}_{u(\nu)} - \underbrace{\int_{x\in\mathbb{R}^d} \ell(x)d\mu(x)}_{u(\mu)} \geqslant \int_{(x,y)\in(\mathbb{R}^d)^2} \langle \nabla \ell(x), y - x\rangle \, d\eta(x,y) - \frac{\lambda}{2} d_{\mathcal{W}}^2(\mu,\nu).$$

Since η is arbitrary, we conclude that $x \mapsto \nabla \ell(x)$ belongs to the subdifferential of u at μ.

Link with the Lions differentiability

Def 7 – W-differential Let $u : \mathscr{P}_2(\mathbb{R}^d) \to \mathbb{R}$ such that $\partial_{\cdot} u(\mu) \neq \emptyset$ and $\partial^{\cdot} u(\mu) \neq \emptyset$. Then $\partial_{\cdot} u(\mu) = \partial^{\cdot} u(\mu) = \{\xi\}$, and the Wasserstein gradient of u at μ is $\nabla_W u(\mu) := \xi$.

Link with the Lions differentiability

Def 7 – W-differential Let $u : \mathscr{P}_2(\mathbb{R}^d) \to \mathbb{R}$ such that $\partial_\cdot u(\mu) \neq \emptyset$ and $\partial^\cdot u(\mu) \neq \emptyset$. Then $\partial_\cdot u(\mu) = \partial^\cdot u(\mu) = \{\xi\}$, and the Wasserstein gradient of u at μ is $\nabla_W u(\mu) := \xi$.

The map u admits a W-gradient $\nabla_W u(\mu)$ at μ if and only if its lift $U(X) := u(\mathcal{L}(X))$ is differentiable at some X such that $\mathcal{L}(X) = \mu$. In this case, one has $\nabla_W u(\mu) = \partial_\mu u(\mu)$.

Link with the Lions differentiability

Def 7 – W-differential Let $u : \mathscr{P}_2(\mathbb{R}^d) \to \mathbb{R}$ such that $\partial_{\cdot} u(\mu) \neq \emptyset$ and $\partial^{\cdot} u(\mu) \neq \emptyset$. Then $\partial_{\cdot} u(\mu) = \partial^{\cdot} u(\mu) = \{\xi\}$, and the Wasserstein gradient of u at μ is $\nabla_W u(\mu) := \xi$.

The map u admits a W-gradient $\nabla_W u(\mu)$ at μ if and only if its lift $U(X) := u(\mathcal{L}(X))$ is differentiable at some X such that $\mathcal{L}(X) = \mu$. In this case, one has $\nabla_W u(\mu) = \partial_\mu u(\mu)$.

The geometric approach of Wasserstein gradients originated in [AGS05], followed by [GNT08]. [AG08, GŚ14] make a direct use of this definition in viscosity solutions. The above link was shown in [GT19, Corollary 3.22] (see also [CD18a, Theorem 5.64]).

The general tangent cone

Problem: the regular tangent cone does not split mass.

The general tangent cone

Problem: the regular tangent cone does not split mass. Define the general tangent cone as

$$\boldsymbol{T}_\mu \mathscr{P}_2(\mathbb{R}^d) := \overline{\left\{\xi \in \mathscr{P}(T\mathbb{R}^d)_\mu \;\middle|\; \exists \varepsilon > 0, t \mapsto \exp_\mu(t \cdot \xi) \text{ is a geodesic on } t \in [0,\varepsilon]\right\}}^{W_\mu},$$

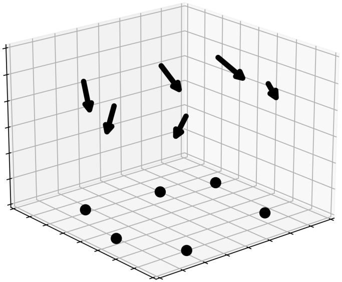
where $W_\mu(\xi,\eta) := \int_{x\in\mathbb{R}^d} d_{\mathcal{W}}(\xi_x,\eta_x) d\mu(x)$ is a generalization of the L^2_μ distance on plans.

The general tangent cone

Problem: the regular tangent cone does not split mass. Define the general tangent cone as

$$\boldsymbol{T}_\mu \mathscr{P}_2(\mathbb{R}^d) := \overline{\left\{\xi \in \mathscr{P}(T\mathbb{R}^d)_\mu \mid \exists \varepsilon > 0, t \mapsto \exp_\mu(t \cdot \xi) \text{ is a geodesic on } t \in [0, \varepsilon]\right\}}^{W_\mu},$$

where $W_\mu(\xi, \eta) := \int_{x \in \mathbb{R}^d} d_{\mathcal{W}}(\xi_x, \eta_x) d\mu(x)$ is a generalization of the L^2_μ distance on plans.

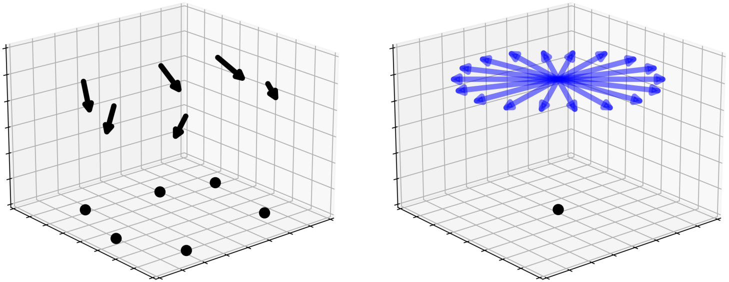


The general tangent cone

Problem: the regular tangent cone does not split mass. Define the general tangent cone as

$$\boldsymbol{T}_\mu \mathscr{P}_2(\mathbb{R}^d) := \overline{\{\xi \in \mathscr{P}(T\mathbb{R}^d)_\mu \mid \exists \varepsilon > 0, t \mapsto \exp_\mu(t \cdot \xi) \text{ is a geodesic on } t \in [0,\varepsilon]\}}^{W_\mu},$$

where $W_\mu(\xi, \eta) := \int_{x \in \mathbb{R}^d} d_{\mathcal{W}}(\xi_x, \eta_x) d\mu(x)$ is a generalization of the L^2_μ distance on plans.

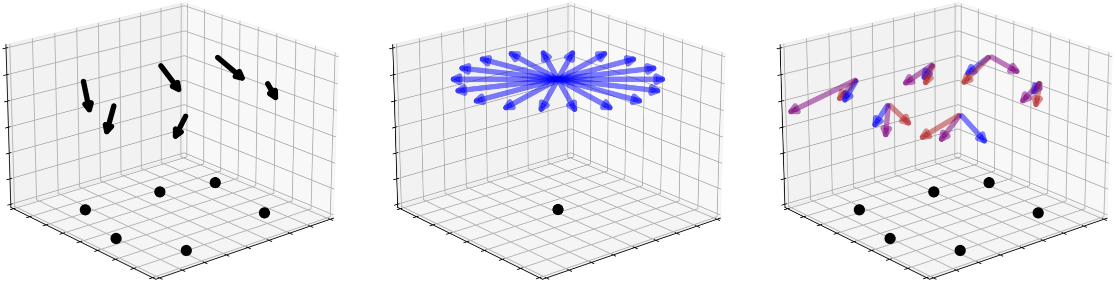


The general tangent cone

Problem: the regular tangent cone does not split mass. Define the general tangent cone as

$$\boldsymbol{T}_\mu \mathscr{P}_2(\mathbb{R}^d) := \overline{\{\xi \in \mathscr{P}(T\mathbb{R}^d)_\mu \mid \exists \varepsilon > 0, t \mapsto \exp_\mu(t \cdot \xi) \text{ is a geodesic on } t \in [0,\varepsilon]\}}^{W_\mu},$$

where $W_\mu(\xi, \eta) := \int_{x \in \mathbb{R}^d} d_{\mathcal{W}}(\xi_x, \eta_x) d\mu(x)$ is a generalization of the L^2_μ distance on plans.



Generalized semidifferentials

For any $\xi \in \boldsymbol{T}_\mu \mathscr{P}_2(\mathbb{R}^d)$ and $\nu \in \mathscr{P}_2(\mathbb{R}^d)$, denote $\Gamma_o(\xi, \nu)$ the set of plans

$$\eta \in \mathscr{P}\left(\left\{(x, v_1, v_2) \;\middle|\; x \in \mathbb{R}^d, v_i \in T_x\mathbb{R}^d\right\}\right) \quad \text{s.t.} \quad \begin{cases} \pi_{x,v_1}\#\eta = \xi, \\ (\pi_x, \pi_x + \pi_{v_2})\#\eta \in \Gamma_o(\mu, \nu). \end{cases}$$

Generalized semidifferentials

For any $\xi \in \boldsymbol{T}_\mu \mathscr{P}_2(\mathbb{R}^d)$ and $\nu \in \mathscr{P}_2(\mathbb{R}^d)$, denote $\Gamma_o(\xi,\nu)$ the set of plans

$$\eta \in \mathscr{P}\left(\left\{(x,v_1,v_2) \;\middle|\; x \in \mathbb{R}^d, v_i \in T_x\mathbb{R}^d\right\}\right) \quad \text{s.t.} \quad \begin{cases} \pi_{x,v_1}\#\eta = \xi, \\ (\pi_x, \pi_x + \pi_{v_2})\#\eta \in \Gamma_o(\mu,\nu). \end{cases}$$

Def 8 Let $u : \mathscr{P}_2(\mathbb{R}^d) \to \mathbb{R}$ and $\mu \in \mathscr{P}_2(\mathbb{R}^d)$. A tangent vector $\xi \in \boldsymbol{T}_\mu \mathscr{P}_2(\mathbb{R}^d)$ belongs to the generalized subdifferential of u at μ, denoted $\boldsymbol{\partial}_{.}u(\mu)$, if for all $\nu \in \mathscr{P}_2(\mathbb{R}^d)$,

$$u(\nu) - u(\mu) \geqslant \sup_{\eta \in \Gamma_o(\xi,\nu)} \int_{x \in \mathbb{R}^d, (v_1,v_2) \in (T_x\mathbb{R}^d)^2} \langle v_1, v_2 \rangle \, d\eta(x,v_1,v_2) + o\left(d_{\mathcal{W}}(\mu,\nu)\right),$$

The generalized superdifferential is defined as $\boldsymbol{\partial}^{\cdot}u(\mu) := -\boldsymbol{\partial}_{.}(-u)(\mu)$.

Example

Let $u(\mu) := \int_{x\in\mathbb{R}^d} \ell(x)d\mu(x)$, and assume that ℓ is $\lambda-$semiconvex (but not $\mathcal{C}^1$ anymore). Denote $\partial_x\ell$ the subdifferential of ℓ at x (a set of vectors).

Example

Let $u(\mu) := \int_{x\in\mathbb{R}^d} \ell(x)d\mu(x)$, and assume that ℓ is $\lambda-$semiconvex (but not $\mathcal{C}^1$ anymore). Denote $\partial_x\ell$ the subdifferential of ℓ at x (a set of vectors). Let $\xi \in \mathscr{P}\left(\bigcup_{x\in\mathbb{R}^d}\{x\}\times\partial_x\ell\right)$ be such that $\pi_x\#\xi = \mu$ (ξ only gives mass to the subdifferential of ℓ). Then, for any $x\in\mathbb{R}^d$, any $v_1\in\partial_x\ell$ and any $v_2\in T_x\mathbb{R}^d$,

$$\ell(x+v_2) - \ell(x) \geqslant \langle v_1, v_2\rangle - \frac{\lambda}{2}\left|v_2\right|^2.$$

Example

Let $u(\mu) := \int_{x\in\mathbb{R}^d} \ell(x) d\mu(x)$, and assume that ℓ is $\lambda-$semiconvex (but not $\mathcal{C}^1$ anymore). Denote $\partial_x \ell$ the subdifferential of ℓ at x (a set of vectors). Let $\xi \in \mathscr{P}\left(\bigcup_{x\in\mathbb{R}^d}\{x\} \times \partial_x \ell\right)$ be such that $\pi_x \# \xi = \mu$ (ξ only gives mass to the subdifferential of ℓ). Then, for any $x \in \mathbb{R}^d$, any $v_1 \in \partial_x \ell$ and any $v_2 \in T_x\mathbb{R}^d$,

$$\ell(x+v_2) - \ell(x) \geqslant \langle v_1, v_2 \rangle - \frac{\lambda}{2} |v_2|^2 .$$

Let $(\mu,\nu) \in (\mathscr{P}_2(\mathbb{R}^d))^2$, and $\eta \in \Gamma_o(\xi,\nu)$. Integrating the above against η,

$$u(\nu) - u(\mu) \geqslant \int_{x\in\mathbb{R}^d,(v_1,v_2)\in(T_x\mathbb{R}^d)^2} \langle v_1, v_2 \rangle \, d\eta(x, v_1, v_2) - \frac{\lambda}{2} d^2_{\mathcal{W}}(\mu,\nu).$$

Example

Let $u(\mu) := \int_{x\in\mathbb{R}^d} \ell(x)d\mu(x)$, and assume that ℓ is $\lambda-$semiconvex (but not $\mathcal{C}^1$ anymore). Denote $\partial_x \ell$ the subdifferential of ℓ at x (a set of vectors). Let $\xi \in \mathscr{P}\left(\bigcup_{x\in\mathbb{R}^d}\{x\} \times \partial_x \ell\right)$ be such that $\pi_x \# \xi = \mu$ (ξ only gives mass to the subdifferential of ℓ). Then, for any $x \in \mathbb{R}^d$, any $v_1 \in \partial_x \ell$ and any $v_2 \in T_x\mathbb{R}^d$,

$$\ell(x+v_2) - \ell(x) \geqslant \langle v_1, v_2 \rangle - \frac{\lambda}{2}\left|v_2\right|^2.$$

Let $(\mu,\nu) \in (\mathscr{P}_2(\mathbb{R}^d))^2$, and $\eta \in \Gamma_o(\xi,\nu)$. Integrating the above against η,

$$u(\nu) - u(\mu) \geqslant \int_{x\in\mathbb{R}^d,(v_1,v_2)\in(T_x\mathbb{R}^d)^2} \langle v_1, v_2 \rangle \, d\eta(x,v_1,v_2) - \frac{\lambda}{2} d^2_{\mathcal{W}}(\mu,\nu).$$

Since η is arbitrary, we obtain that $\xi \in \boldsymbol{\partial}.u(\mu)$.

Table of Contents

Differentiate in length spaces

Def 9 – Metric slope Let (X, d) be a metric space. The metric slope of a map $u : X \to \mathbb{R}$ at the point x is given by

$$|\nabla u(x)| := \overline{\lim_{y \to x}} \frac{|u(y) - u(x)|}{d(x, y)}.$$

Differentiate in length spaces

Def 9 – Metric slope Let (X,d) be a metric space. The metric slope of a map $u : X \to \mathbb{R}$ at the point x is given by

$$|\nabla u(x)| := \varlimsup_{y \to x} \frac{|u(y)-u(x)|}{d(x,y)}.$$

Metric slopes are used to formulate equations in (length) metric spaces, for instance in [AGS05, Vil09, Oht09] on gradient flows, of [GNT08, HK15, GŚ15a, GŚ15b, GHN15] on eikonal-type equations.

Differentiate in length spaces

Def 9 – Metric slope Let (X, d) be a metric space. The metric slope of a map $u : X \to \mathbb{R}$ at the point x is given by

$$|\nabla u(x)| := \overline{\lim_{y \to x}} \frac{|u(y) - u(x)|}{d(x, y)}.$$

Metric slopes are used to formulate equations in (length) metric spaces, for instance in [AGS05, Vil09, Oht09] on gradient flows, of [GNT08, HK15, GŚ15a, GŚ15b, GHN15] on eikonal-type equations.

(Last) example: let $u(\mu) = \int_{x \in \mathbb{R}^d} \ell(x) d\mu$ with $\ell \in \mathcal{C}_b^2$. Then $|\nabla^+ u(\mu)| = \int_{x \in \mathbb{R}^d} |\nabla \ell(x)| \, d\mu(x)$.

Gradient flows

In [AGS05], general gradient flows are studied in metric spaces. They want to give a meaning to curves satisfying

$$\frac{d}{dt}y(t) = -\nabla\Phi(y(t)), \qquad y(0) = y_0.$$

[1]Under the assumptions of [AGS05, Theorem 11.3.2].

Gradient flows

In [AGS05], general gradient flows are studied in metric spaces. They want to give a meaning to curves satisfying

$$\frac{d}{dt}y(t) = -\nabla\Phi(y(t)), \qquad y(0) = y_0.$$

To this aim, a numerical scheme is designed, and an approximating sequence $(y^N)_N$ is computed. In the case of the Wasserstein space, Ambrosio, Gigli and Savaré showed that[1]

- the limit $\overline{y}$ exists and satisfies an axiomatic definition of gradient curve,

[1] Under the assumptions of [AGS05, Theorem 11.3.2].

Gradient flows

In [AGS05], general gradient flows are studied in metric spaces. They want to give a meaning to curves satisfying

$$\frac{d}{dt}y(t) = -\nabla\Phi(y(t)), \qquad y(0) = y_0.$$

To this aim, a numerical scheme is designed, and an approximating sequence $(y^N)_N$ is computed. In the case of the Wasserstein space, Ambrosio, Gigli and Savaré showed that[1]

- the limit $\overline{y}$ exists and satisfies an axiomatic definition of gradient curve,
- an appropriate generalization of $\frac{d}{dt}\overline{y}(t)$ converges to an element of $\boldsymbol{\partial}.\Phi(\overline{y}(t))$.

[1] Under the assumptions of [AGS05, Theorem 11.3.2].

Gradient flows

In [AGS05], general gradient flows are studied in metric spaces. They want to give a meaning to curves satisfying

$$\frac{d}{dt}y(t) = -\nabla\Phi(y(t)), \qquad y(0) = y_0.$$

To this aim, a numerical scheme is designed, and an approximating sequence $(y^N)_N$ is computed. In the case of the Wasserstein space, Ambrosio, Gigli and Savaré showed that[1]

- the limit $\overline{y}$ exists and satisfies an axiomatic definition of gradient curve,
- an appropriate generalization of $\frac{d}{dt}\overline{y}(t)$ converges to an element of $\boldsymbol{\partial}.\Phi(\overline{y}(t))$.

The regular tangent space $\partial.\Phi$ may be two small (case of $\Phi = d^2_{\mathcal{W}}(\cdot,\sigma)$ for instance).

[1]Under the assumptions of [AGS05, Theorem 11.3.2].

Eikonal-type equations (HJ depending only on the norm of ∇u)

Canonical example: a minimal time problem

$$-\partial_t u(t,\mu) + \frac{1}{2}\left|\nabla^+ u(t,\mu)\right|^2 = 1, \qquad u(T,\mu) = 0.$$

Eikonal-type equations (HJ depending only on the norm of ∇u)

Canonical example: a minimal time problem

$$-\partial_t u(t,\mu) + \frac{1}{2}\left|\nabla^+ u(t,\mu)\right|^2 = 1, \qquad u(T,\mu) = 0.$$

In [AF14], such equations is studied in geodesic/length spaces by first using metric slopes. They show that

- a definition of viscosity using the generalized semidifferentials is compatible with their metric definition (a solution for the former is a solution for the latter).

Eikonal-type equations (HJ depending only on the norm of ∇u)

Canonical example: a minimal time problem

$$-\partial_t u(t,\mu) + \frac{1}{2}\left|\nabla^+ u(t,\mu)\right|^2 = 1, \qquad u(T,\mu) = 0.$$

In [AF14], such equations is studied in geodesic/length spaces by first using metric slopes. They show that

- a definition of viscosity using the generalized semidifferentials is compatible with their metric definition (a solution for the former is a solution for the latter).
- it is no longer the case when restricted to the regular semidifferentials.

Eikonal-type equations (HJ depending only on the norm of ∇u)

Canonical example: a minimal time problem

$$-\partial_t u(t,\mu) + \frac{1}{2}\left|\nabla^+ u(t,\mu)\right|^2 = 1, \qquad u(T,\mu) = 0.$$

In [AF14], such equations is studied in geodesic/length spaces by first using metric slopes. They show that

- a definition of viscosity using the generalized semidifferentials is compatible with their metric definition (a solution for the former is a solution for the latter).
- it is no longer the case when restricted to the regular semidifferentials.

The construction of generalized subdifferentials in [AF14] is linked to the tangent cone for curved spaces, explored for the Wasserstein case in [Gig08] (see [AKP22] for material on curved spaces).

The derivatives of the linear map in one glance

Recall that $u : \mathscr{P}_2(\mathbb{R}^d)$ is defined as

$$u(\mu) := \int_{x\in\mathbb{R}^d} \ell(x) d\mu(x).$$

Distributional derivative	Lions derivative	Linear derivative	Natural derivative	Regular subdifferential	General subdifferential
$\text{grad}_\mu u(\mu)$	$\partial_\mu u(\mu)$	$\frac{\delta u}{\delta \mu}(\mu, \cdot)$	$D_\mu u(\mu, \cdot)$	$\partial . u(\mu)$, $\nabla_W u$	$\boldsymbol{\partial} . u(\mu)$
$-\operatorname{div}(\mu \nabla \ell)$	$\nabla \ell$	ℓ	$\nabla \ell$	$\nabla \ell$, select° of $\partial \ell$	$\nabla \ell \# \mu$, $\mathscr{P}(\text{Gr}(\partial \ell))$
distribution, duality with $\mathcal{C}_c^1(\mathbb{R}^d, \mathbb{R})$	element of $L^2_\mu(\mathbb{R}^d, \mathbb{R}^d)$	element of $L^2_\mu(\mathbb{R}^d, \mathbb{R})$	element of $L^2_\mu(\mathbb{R}^d, \mathbb{R}^d)$	element of $T_\mu \mathscr{P}_2(\mathbb{R}^d)$	element of $\boldsymbol{T}_\mu \mathscr{P}_2(\mathbb{R}^d)$

Conclusion

- The L-differentiability (lift, intrinsic/extrinsic) is well-adapted to some problems (for instance the master equation with smooth coefficients), and emerges naturally from the modelization.

Conclusion

- The L-differentiability (lift, intrinsic/extrinsic) is well-adapted to some problems (for instance the master equation with smooth coefficients), and emerges naturally from the modelization.
- The W-differentiability, constructed in a more geometric fashion, has been reconciled with the L-differentiability since they coincide on sufficiently smooth functions.

Conclusion

- The L-differentiability (lift, intrinsic/extrinsic) is well-adapted to some problems (for instance the master equation with smooth coefficients), and emerges naturally from the modelization.
- The W-differentiability, constructed in a more geometric fashion, has been reconciled with the L-differentiability since they coincide on sufficiently smooth functions.
- Whenever the map u is not differentiable, generalized subdifferentials (although less maniable) are more suited than regular ones.

Conclusion

- The L-differentiability (lift, intrinsic/extrinsic) is well-adapted to some problems (for instance the master equation with smooth coefficients), and emerges naturally from the modelization.
- The W-differentiability, constructed in a more geometric fashion, has been reconciled with the L-differentiability since they coincide on sufficiently smooth functions.
- Whenever the map u is not differentiable, generalized subdifferentials (although less maniable) are more suited than regular ones.

Open questions:

- How to get out of vector spaces?

Conclusion

- The L-differentiability (lift, intrinsic/extrinsic) is well-adapted to some problems (for instance the master equation with smooth coefficients), and emerges naturally from the modelization.
- The W-differentiability, constructed in a more geometric fashion, has been reconciled with the L-differentiability since they coincide on sufficiently smooth functions.
- Whenever the map u is not differentiable, generalized subdifferentials (although less maniable) are more suited than regular ones.

Open questions:

- How to get out of vector spaces?
- Is there an existence theorem to dig for continuity equations written as $\partial_t \mu_t = -\operatorname{div}(\mu_t F(\mu_t))$, where $F[\mu_t]$ is a plan in $\boldsymbol{T}_\mu \mathscr{P}_2(\mathbb{R}^d)$? Can this be posed pointwise in time, and under which condition does existence hold?

Thank you!

[AF14] Luigi Ambrosio and Jin Feng.
On a class of first order Hamilton–Jacobi equations in metric spaces.
Journal of Differential Equations, 256(7):2194–2245, April 2014.

[AG08] Luigi Ambrosio and Wilfrid Gangbo.
Hamiltonian ODEs in the Wasserstein space of probability measures.
Communications on Pure and Applied Mathematics, 61(1):18–53, 2008.

[AGS05] Luigi Ambrosio, Nicola Gigli, and Guiseppe Savaré.
Gradient Flows.
Lectures in Mathematics ETH Zürich. Birkhäuser-Verlag, Basel, 2005.

[AKP22] Stephanie Alexander, Vitali Kapovitch, and Anton Petrunin.
Alexandrov geometry: Foundations, October 2022.

[BIRS19] Matteo Burzoni, Vicenzo Ignazio, A. Max Reppen, and H. Mete Soner.
Viscosity Solutions for Controlled McKean-Vlasov Jump-Diffusions.
2019.

[BL94] Nicolas Bouleau and Dominique Lépingle.
Numerical Methods for Stochastic Processes.
1994.

[BY19] Alain Bensoussan and Sheung Chi Phillip Yam.
Control problem on space of random variables and master equation.
ESAIM: Control, Optimisation and Calculus of Variations, 25:10, 2019.

[Car13] Pierre Cardaliaguet.
Notes on Mean Field Games.
page 59, 2013.

[CCD15] Jean-François Chassagneux, Dan Crisan, and François Delarue.
A Probabilistic approach to classical solutions of the master equation for large population equilibria, April 2015.

[CD18a] René Carmona and François Delarue.
Probabilistic Theory of Mean Field Games with Applications I, volume 83 of *Probability Theory and Stochastic Modelling*.
Springer International Publishing, Cham, 2018.

[CD18b] René Carmona and François Delarue.
Probabilistic Theory of Mean Field Games with Applications II, volume 84 of *Probability Theory and Stochastic Modelling*.
Springer International Publishing, Cham, 2018.

[CDLL19] Pierre Cardaliaguet, François Delarue, Jean-Michel Lasry, and Pierre-Louis Lions.
The Master Equation and the Convergence Problem in Mean Field Games.
Number 201 in Annals of Mathematics Studies. Princeton University Press, Princeton, NJ, 2019.

[CGK+22] Andrea Cosso, Fausto Gozzi, Idris Kharroubi, Huyên Pham, and Mauro Rosestolato.
Master Bellman equation in the Wasserstein space: Uniqueness of viscosity solutions, February 2022.

[CLS18] Christa Cuchiero, Martin Larsson, and Sara Svaluto-Ferro.
Probability measure-valued polynomial diffusions, July 2018.

[CQ08] P. Cardaliaguet and M. Quincampoix.
Deterministic differential games under probability knowledge of initial condition.
International Game Theory Review, 10(01):1–16, March 2008.

[FK09] Jin Feng and Markos Katsoulakis.
A Comparison Principle for Hamilton–Jacobi Equations Related to Controlled Gradient Flows in Infinite Dimensions.
Archive for Rational Mechanics and Analysis, 192(2):275–310, May 2009.

[FN12] Jin Feng and Truyen Nguyen.
Hamilton–Jacobi equations in space of measures associated with a system of conservation laws.
Journal de Mathématiques Pures et Appliquées, 97(4):318–390, April 2012.

[FV79] Wendell H. Fleming and Michel Viot.
Some Measure-Valued Markov Processes in Population Genetics Theory.
Indiana University Mathematics Journal, 28(5):817–843, 1979.

[GHN15] Yoshikazu Giga, Nao Hamamuki, and Atsushi Nakayasu.
Eikonal equations in metric spaces.
Transactions of the American Mathematical Society, 367(1):49–66, January 2015.

[Gig08] Nicola Gigli.
On the Geometry of the Space of Probability Measures Endowed with the Quadratic Optimal Transport Distance.
PhD thesis, Scuola Normale Superiore di Pisa, Pisa, 2008.

[GNT08] Wilfrid Gangbo, Truyen Nguyen, and Adrian Tudorascu.
Hamilton-Jacobi Equations in the Wasserstein Space.
Methods and Applications of Analysis, 15(2):155–184, 2008.

[GŚ14] Wilfrid Gangbo and Andrzej Świech.
Optimal transport and large number of particles.
Discrete and Continuous Dynamical Systems, 34(4):1397–1441, 2014.

[GŚ15a] Wilfrid Gangbo and Andrzej Święch.
Existence of a solution to an equation arising from the theory of Mean Field Games.
Journal of Differential Equations, 259(11):6573–6643, December 2015.

[GŚ15b] Wilfrid Gangbo and Andrzej Święch.
Metric viscosity solutions of Hamilton–Jacobi equations depending on local slopes.
Calculus of Variations and Partial Differential Equations, 54(1):1183–1218, September 2015.

[GT19] Wilfrid Gangbo and Adrian Tudorascu.
On differentiability in the Wasserstein space and well-posedness for Hamilton–Jacobi equations.
Journal de Mathématiques Pures et Appliquées, 125:119–174, May 2019.

[HK15] Ryan Hynd and Hwa Kil Kim.
Value functions in the Wasserstein spaces: Finite time horizons.
Journal of Functional Analysis, 269(4):968–997, August 2015.

[JKO98] Richard Jordan, David Kinderlehrer, and Felix Otto.
The Variational Formulation of the Fokker-Planck Equation.
SIAM Journal on Mathematical Analysis, 29(1):1–17, January 1998.

[JMQ20] Chloé Jimenez, Antonio Marigonda, and Marc Quincampoix.
Optimal control of multiagent systems in the Wasserstein space.
Calculus of Variations and Partial Differential Equations, 59, March 2020.

[JMQ22] Chloé Jimenez, Antonio Marigonda, and Marc Quincampoix.
Dynamical systems and Hamilton-Jacobi-Bellman equations on the Wasserstein space and their L2 representations.
2022.

[Lio06] Pierre-Louis Lions.
Jeux à champ moyen, 2006.

[MQ18] Antonio Marigonda and Marc Quincampoix.
Mayer control problem with probabilistic uncertainty on initial positions.
Journal of Differential Equations, 264(5):3212–3252, March 2018.

[MZ22] Chenchen Mou and Jianfeng Zhang.
Wellposedness of Second Order Master Equations for Mean Field Games with Nonsmooth Data, January 2022.

[Oht09] Shin-Ichi Ohta.
Gradient Flows on Wasserstein Spaces over Compact Alexandrov Spaces.
American Journal of Mathematics, 131(2):475–516, 2009.

[Ott01] Felix Otto.
The Geometry of Dissipative Evolution Equations: The Porous Medium Equation.
Communications in Partial Differential Equations, 26(1-2):101–174, January 2001.

[PW17] Huyên Pham and Xiaoli Wei.
Dynamic programming for optimal control of stochastic McKean-Vlasov dynamics, January 2017.

[PW18] Huyên Pham and Xiaoli Wei.
Bellman equation and viscosity solutions for mean-field stochastic control problem.
ESAIM: Control, Optimisation and Calculus of Variations, 24(1):437–461, January 2018.

[Sal23] William Salkeld.
Higher order Lions-Taylor expansions, March 2023.

[Vil09] Cédric Villani.
Optimal Transport, volume 338 of *Grundlehren Der Mathematischen Wissenschaften*.
Springer Berlin Heidelberg, Berlin, Heidelberg, 2009.