Dissipative evolution of Probability Measures

Giuseppe Savaré

Dynamics and Discretization: PDEs, Sampling, and Optimization,
October 26, 2021

Department of Decision Sciences, Bocconi University, Milan, Italy

In collaboration with *G. Cavagnari, G. Sodini*

Outline

Probability vector fields and evolution

Dissipative operators and contraction semigroups in Hilbert spaces

Convergence of the Explicit Euler method and contraction semigroups

Borel probability measures

Problem: generation of contraction semigroups in
the space of Borel probability measures $\mathcal{P}_2(E)$
$\big(E = \mathbb{R}^d$ Euclidean space or $E = \mathbf{H}$ Hilbert$\big)$
driven by dissipative probability vector fields.

$\mathcal{P}_2(E)$: probability measures $\mu \in \mathcal{P}(E)$ with finite quadratic moment $\int |x|^2 \, d\mu < \infty$.

Borel probability measures

Problem: generation of contraction semigroups in
the space of Borel probability measures $\mathcal{P}_2(E)$
$\big(E = \mathbb{R}^d$ Euclidean space or $E = \mathbf{H}$ Hilbert$\big)$
driven by dissipative probability vector fields.

$\mathcal{P}_2(E)$: probability measures $\mu \in \mathcal{P}(E)$ with finite quadratic moment $\int |x|^2 \, d\mu < \infty$.

Push forward: if $X : \Omega \to E$ is a Borel map, $\mathbb{P} \in \mathcal{P}(\Omega)$,

$X_\sharp \mathbb{P}$ is the law of the Random variable X: $\quad X_\sharp \mathbb{P}(E) := \mathbb{P}[X^{-1}(E)]$

Borel probability measures

Problem: generation of contraction semigroups in
the space of Borel probability measures $\mathcal{P}_2(E)$
$\big(E = \mathbb{R}^d$ Euclidean space or $E = \mathbf{H}$ Hilbert$\big)$
driven by dissipative probability vector fields.

$\mathcal{P}_2(E)$: probability measures $\mu \in \mathcal{P}(E)$ with finite quadratic moment $\displaystyle\int |x|^2 \, d\mu < \infty$.

Push forward: if $X : \Omega \to E$ is a Borel map, $\mathbb{P} \in \mathcal{P}(\Omega)$,

$X_\sharp\mathbb{P}$ is the law of the Random variable X: $\quad X_\sharp\mathbb{P}(E) := \mathbb{P}[X^{-1}(E)]$

$\Gamma(\mu, \nu) :=$ couplings between $\mu \in \mathcal{P}(E)$, $\nu \in \mathcal{P}(\mathbf{F})$, measures $\boldsymbol{\gamma} \in \mathcal{P}(E \times \mathbf{F})$ whose marginals are μ and ν, e.g. $\boldsymbol{\gamma} = (X, Y)_\sharp\mathbb{P}$, $X_\sharp\mathbb{P} = \mu$, $Y_\sharp\mathbb{P} = \nu$.

Borel probability measures

Problem: generation of contraction semigroups in
the space of Borel probability measures $\mathcal{P}_2(E)$
$\big(E = \mathbb{R}^d$ Euclidean space or $E = \mathbf{H}$ Hilbert$\big)$
driven by dissipative probability vector fields.

$\mathcal{P}_2(E)$: probability measures $\mu \in \mathcal{P}(E)$ with finite quadratic moment $\int |x|^2 \, d\mu < \infty$.

Push forward: if $X : \Omega \to E$ is a Borel map, $\mathbb{P} \in \mathcal{P}(\Omega)$,

$X_\sharp \mathbb{P}$ is the law of the Random variable X: $\quad X_\sharp \mathbb{P}(E) := \mathbb{P}[X^{-1}(E)]$

$\Gamma(\mu, \nu) :=$ couplings between $\mu \in \mathcal{P}(E)$, $\nu \in \mathcal{P}(\mathbf{F})$, measures $\boldsymbol{\gamma} \in \mathcal{P}(E \times \mathbf{F})$ whose marginals are μ and ν, e.g. $\boldsymbol{\gamma} = (X, Y)_\sharp \mathbb{P}$, $X_\sharp \mathbb{P} = \mu$, $Y_\sharp \mathbb{P} = \nu$.

$\Gamma_o(\mu, \nu)$: optimal couplings for the L^2-Wasserstein distance. $\gamma_o \in \Gamma_o(\mu, \nu)$ iff

$$W_2^2(\mu, \nu) = \int |x - y|^2 \, d\gamma_o = \min \Big\{ \int |x - y|^2 \, d\gamma : \gamma \in \Gamma(\mu, \nu) \Big\}.$$

Probability vector fields in $\mathcal{P}_2(\mathsf{E})$

Tangent space: $\mathbf{T}\mathsf{E} = \{(x, v) : x, v \in \mathsf{E}\} \approx \mathsf{E} \times \mathsf{E}$, $\boldsymbol{x}(x, v) = x$, $\boldsymbol{v}(x, v) = v$.

In $\mathcal{P}_2(\mathsf{E})$ a probability vector field $\mathfrak{F}$ can be represented by a map (possibly multivalued) from $\mathrm{D}(\mathfrak{F}) \subset \mathcal{P}_2(\mathsf{E})$ to $\mathcal{P}_2(\mathbf{T}\mathsf{E})$ such that

$$\text{for every } \underline{\mathrm{F}} \in \mathfrak{F}(\mu) : \quad \boldsymbol{x}_\sharp \underline{\mathrm{F}} = \mu.$$

Probability vector fields in $\mathcal{P}_2(E)$

Tangent space: $\mathsf{TE} = \{(x, v) : x, v \in E\} \approx E \times E$, $\boldsymbol{x}(x, v) = x$, $\boldsymbol{v}(x, v) = v$.

In $\mathcal{P}_2(E)$ a probability vector field $\mathfrak{F}$ can be represented by a map (possibly multivalued) from $D(\mathfrak{F}) \subset \mathcal{P}_2(E)$ to $\mathcal{P}_2(\mathsf{TE})$ such that

$$\text{for every } \underline{F} \in \mathfrak{F}(\mu) : \quad \boldsymbol{x}_\sharp \underline{F} = \mu.$$

By **disintegrating** $\underline{F} \in \mathfrak{F}(\mu)$ w.r.t. μ we obtain a family of measures $F_x \in \mathcal{P}_2(E)$ which represent **probability laws on directions starting from** x**.**

In the "regular case" F_x is **concentrated on a single vector** $\delta_{\underline{F}(x,\mu)}$ and therefore can be represented by a vector field $\underline{F}(x, \mu)$ mapping $E \times \mathcal{P}_2(E)$ into E.

Probability vector fields in $\mathcal{P}_2(E)$

Tangent space: $\mathsf{TE} = \{(x, v) : x, v \in E\} \approx E \times E$, $\boldsymbol{x}(x, v) = x$, $\boldsymbol{v}(x, v) = v$.

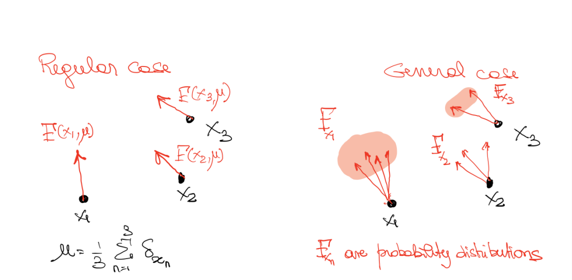
In $\mathcal{P}_2(E)$ a probability vector field $\mathfrak{F}$ can be represented by a map (possibly multivalued) from $D(\mathfrak{F}) \subset \mathcal{P}_2(E)$ to $\mathcal{P}_2(\mathsf{TE})$ such that

$$\text{for every } \underline{F} \in \mathfrak{F}(\mu) : \quad \boldsymbol{x}_\sharp \underline{F} = \mu.$$

By **disintegrating** $\underline{F} \in \mathfrak{F}(\mu)$ w.r.t. μ we obtain a family of measures $F_x \in \mathcal{P}_2(E)$ which represent **probability laws on directions starting from** x.

In the "regular case" F_x is **concentrated on a single vector** $\delta_{\underline{F}(x,\mu)}$ and therefore can be represented by a vector field $\underline{F}(x, \mu)$ mapping $E \times \mathcal{P}_2(E)$ into E.

In the general case, we can allow for a **general probability measure** F_x **depending on** x.



Evolution driven by $\mathfrak{F}$

We want to study the evolution of probability measures driven by a PVF $\mathfrak{F}$, formally

$$\dot{\mu}_t = \mathfrak{F}(\mu_t) \quad t > 0.$$

[1] B. Bonnet, H. Frankowska Differential inclusions in Wasserstein spaces: The Cauchy-Lipschitz framework, JDE,2021

Evolution driven by $\mathfrak{F}$

We want to study the evolution of probability measures driven by a PVF $\mathfrak{F}$, formally

$$\dot{\mu}_t = \mathfrak{F}(\mu_t) \quad t > 0.$$

Example: finite dimensional Cauchy-Lipschitz theory [1]

$\mathfrak{F}$ does not split particles and it is concentrated on the vector field $\underline{F}(x, \mu)$,

$$\mathfrak{F}(\mu) = (\mathrm{Id} \times \underline{F}(\cdot, \mu))_\sharp \mu.$$

Examples are

$$\underline{F}(x, \mu) = A(x) + \int B(x - y)\, d\mu(y), \quad A, B : E \to E \text{ dissipative.}$$

[1] B. Bonnet, H. Frankowska Differential inclusions in Wasserstein spaces: The Cauchy-Lipschitz framework, JDE,2021

Evolution driven by $\mathfrak{F}$

We want to study the evolution of probability measures driven by a PVF $\mathfrak{F}$, formally

$$\dot{\mu}_t = \mathfrak{F}(\mu_t) \quad t > 0.$$

Example: finite dimensional Cauchy-Lipschitz theory [1]

$\mathfrak{F}$ does not split particles and it is concentrated on the vector field $\underline{F}(x, \mu)$,

$$\mathfrak{F}(\mu) = (\mathrm{Id} \times \underline{F}(\cdot, \mu))_\sharp \mu.$$

Examples are

$$\underline{F}(x, \mu) = A(x) + \int B(x - y)\, d\mu(y), \quad A, B : E \to E \text{ dissipative.}$$

The curve $(\mu_t)_{t>0}$ solves the continuity equation

$$\partial_t \mu_t + \nabla \cdot (\mu_t \boldsymbol{v}_t) = 0, \quad \boldsymbol{v}_t(x) = \underline{F}(x, \mu_t).$$

[1] B. Bonnet, H. Frankowska Differential inclusions in Wasserstein spaces: The Cauchy-Lipschitz framework, JDE,2021

Examples

- Gradient flows [2] generated by a λ-geodesically convex functional $\mathscr{F} : \mathcal{P}_2(E) \to (-\infty, +\infty]$. $\mathscr{F}$ can be nonsmooth (subdifferential calculus): e.g. [3]

$$\mathscr{F}(\mu) = R\left(\int T(x)\,d\mu(x)\right) + \iint W(x-y)\,d\mu(x)\,d\mu(y) + \int V\,d\mu$$

$T : E \to \tilde{E}$ is a vector valued map, $R : \tilde{E} \to \mathbb{R}$, $W, V : E \to \mathbb{R}$.

$-\mathfrak{F}$ is the multivalued Wasserstein subdifferential of $\mathscr{F}$.

[2] L. Ambrosio, N. Gigli, G. S., Gradient flows in metric spaces and in the space of probability measures, Birkäuser, 2008

[3] L. Chizat, F. Bach, On the global convergence of gradient descent for over-parameterized models using optimal transport, 2018

[4] B. Piccoli, Measure differential equations. Arch. Ration. Mech. Anal. (2019)

Examples

- Gradient flows [2] generated by a λ-geodesically convex functional $\mathscr{F} : \mathcal{P}_2(E) \to (-\infty, +\infty]$. $\mathscr{F}$ can be nonsmooth (subdifferential calculus): e.g. [3]

$$\mathscr{F}(\mu) = R\Big(\int T(x)\,d\mu(x)\Big) + \iint W(x-y)\,d\mu(x)\,d\mu(y) + \int V\,d\mu$$

  $T : E \to \tilde{E}$ is a vector valued map, $R : \tilde{E} \to \mathbb{R}$, $W, V : E \to \mathbb{R}$.
  $-\mathfrak{F}$ is the multivalued Wasserstein subdifferential of $\mathscr{F}$.

- Dissipative evolution, contraction semigroups: E Hilbert space, F multivalued. E.g. the Lipschitz perturbation of a multivalued subgradient. This case has been studied by Piccoli [4] in finite dimension with a different approach.

[2] L. Ambrosio, N. Gigli, G. S., Gradient flows in metric spaces and in the space of probability measures, Birkäuser, 2008

[3] L. Chizat, F. Bach, On the global convergence of gradient descent for over-parameterized models using optimal transport, 2018

[4] B. Piccoli, Measure differential equations. Arch. Ration. Mech. Anal. (2019)

Outline

Dissipative oprators in Hilbert space

In a Hilbert space H a (multivalued) map $\mathrm{B} : \mathrm{D}(\mathrm{B}) \subset \mathrm{H} \rightrightarrows \mathrm{H}$ is dissipative if

$$\boxed{\langle v - w, x - y \rangle \leqslant 0 \quad \text{for every } v \in \mathrm{B}x,\ \ w \in \mathrm{B}y}$$

Dissipative oprators in Hilbert space

In a Hilbert space $\mathbb{H}$ a (multivalued) map $B : D(B) \subset \mathbb{H} \rightrightarrows \mathbb{H}$ is dissipative if

$$\boxed{\langle v - w, x - y\rangle \leqslant 0 \quad \text{for every } v \in Bx,\ w \in By}$$

This property has a natural metric interpretation: if we consider the curves

$$x(\tau) := x + \tau v, \quad y(\tau) := y + \tau w, \quad v \in Bx,\ w \in By$$

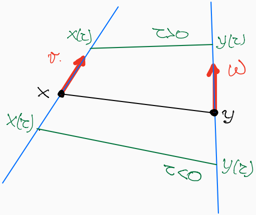
and their squared distance $\qquad D(\tau) := \frac{1}{2}|x(\tau) - y(\tau)|^2$

then

$$\langle v - w, x - y\rangle = D'(0) = \frac{1}{2}\frac{d}{d\tau}|x(\tau) - y(\tau)|^2\Big|_{\tau=0} \leqslant 0$$

so that

$$|x(\tau) - y(\tau)|^2 \leqslant |x - y|^2 + \tau^2|v - w|^2 = |x - y|^2 + o(\tau) \quad \text{as } \tau \downarrow 0$$



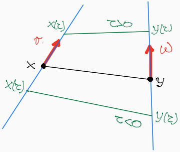
The resolvent

$D(\tau) = \frac{1}{2}|x(\tau) - y(\tau)|^2$ is convex, $D'(0) \leqslant 0$ yields

$|x - y|^2 \leqslant |x(s) - y(s)|^2$ for every $s < 0$

If $x' - \tau B x' = x$ and $y' - \tau B y' = y$

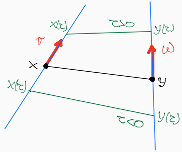
then $\boxed{|x' - y'|^2 \leqslant |x - y|^2}$



The resolvent

$D(\tau) = \frac{1}{2}|x(\tau) - y(\tau)|^2$ is convex, $D'(0) \leqslant 0$ yields

$|x - y|^2 \leqslant |x(s) - y(s)|^2$ for every $s < 0$



If $x' - \tau B x' = x$ and $y' - \tau B y' = y$

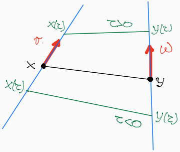
then $\boxed{|x' - y'|^2 \leqslant |x - y|^2}$

It follows that for $\tau > 0$ the resolvent

$$\boxed{J_\tau := (\mathrm{Id} - \tau B)^{-1}, \quad x' = J_\tau(x) \Leftrightarrow x' - \tau B x' = x \quad \text{is a contraction}}$$

This property can be used to define dissipative operators in Banach spaces.

$D(\tau) = \frac{1}{2}|x(\tau) - y(\tau)|^2$ is convex, $D'(0) \leqslant 0$ yields



$|x - y|^2 \leqslant |x(s) - y(s)|^2$ for every $s < 0$

If $x' - \tau Bx' = x$ and $y' - \tau By' = y$

then $\boxed{|x' - y'|^2 \leqslant |x - y|^2}$

It follows that for $\tau > 0$ the resolvent

$$\boxed{J_\tau := (\mathrm{Id} - \tau B)^{-1}, \quad x' = J_\tau(x) \;\Leftrightarrow\; x' - \tau Bx' = x \quad \text{is a contraction}}$$

This property can be used to define dissipative operators in Banach spaces.

B is m-dissipative (or maximal dissipative) if J_τ is defined in all the space $\mathbf{H}$: for every $x \in \mathbf{H}$ the equation

$$\boxed{y - \tau By \ni x \quad \text{has a (unique) solution } y = J_\tau x.}$$

A particular case is the subdifferential $B = -\partial\Phi$ of a convex l.s.c. function $\Phi : \mathbf{H} \to (-\infty, +\infty]$: $x_\tau = J_\tau(x)$ if and only if

$$x_\tau \text{ minimizes} \qquad y \mapsto \frac{1}{2\tau}|y - x|^2 + \Phi(y).$$

The Explicit and Implicit Euler methods

If B is everywhere defined, for every $\boldsymbol{x}_\tau^0 \in \mathbf{H}$ one can solve the Explicit Euler method

$$\frac{\boldsymbol{x}_\tau^n - \boldsymbol{x}_\tau^{n-1}}{\tau} \in \mathrm{B}\boldsymbol{x}_\tau^{n-1} \quad \Leftrightarrow \quad \boldsymbol{x}_\tau^n = \boldsymbol{x}_\tau^{n-1} + \tau\mathrm{B}\boldsymbol{x}_\tau^{n-1} = (\mathrm{Id} + \tau\mathrm{B})^n\boldsymbol{x}_\tau^0, \quad n = 1, \cdots$$

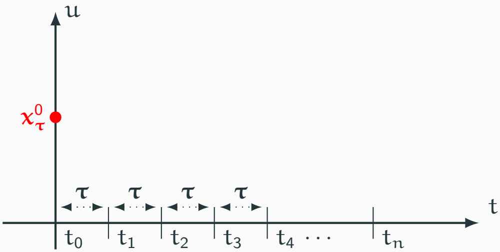
The Explicit and Implicit Euler methods

If B is everywhere defined, for every $\mathbf{x}_\tau^0 \in \mathbf{H}$ one can solve the Explicit Euler method

$$\frac{\mathbf{x}_\tau^n - \mathbf{x}_\tau^{n-1}}{\tau} \in B\mathbf{x}_\tau^{n-1} \quad \Leftrightarrow \quad \mathbf{x}_\tau^n = \mathbf{x}_\tau^{n-1} + \tau B\mathbf{x}_\tau^{n-1} = (\mathrm{Id} + \tau B)^n \mathbf{x}_\tau^0, \quad n = 1, \cdots$$

If B is m-dissipative, for every $\mathbf{x}_\tau^0 \in \mathbf{H}$ one can solve the Implicit Euler method

$$\frac{\mathbf{x}_\tau^n - \mathbf{x}_\tau^{n-1}}{\tau} \in B\mathbf{x}_\tau^{n} \quad \Leftrightarrow \quad \mathbf{x}_\tau^n = J_\tau \mathbf{x}_\tau^{n-1} = J_\tau^n \mathbf{x}_\tau^0, \quad n = 1, \cdots$$



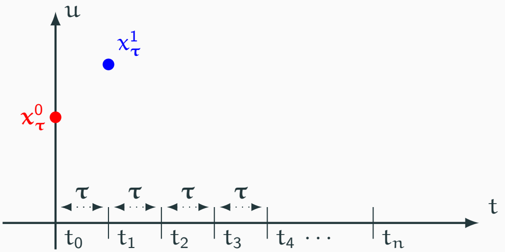
The Explicit and Implicit Euler methods

If B is everywhere defined, for every $x_\tau^0 \in H$ one can solve the Explicit Euler method

$$\frac{x_\tau^n - x_\tau^{n-1}}{\tau} \in Bx_\tau^{n-1} \quad \Leftrightarrow \quad x_\tau^n = x_\tau^{n-1} + \tau Bx_\tau^{n-1} = (\mathrm{Id} + \tau B)^n x_\tau^0, \quad n = 1, \cdots$$

If B is m-dissipative, for every $x_\tau^0 \in H$ one can solve the Implicit Euler method

$$\frac{x_\tau^n - x_\tau^{n-1}}{\tau} \in Bx_\tau^n \quad \Leftrightarrow \quad x_\tau^n = J_\tau x_\tau^{n-1} = J_\tau^n x_\tau^0, \quad n = 1, \cdots$$



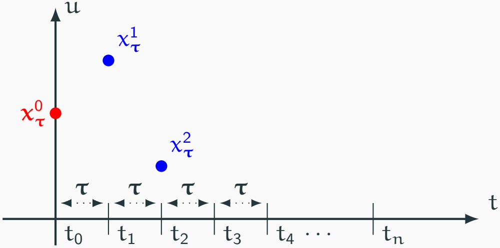
The Explicit and Implicit Euler methods

If B is everywhere defined, for every $\mathbf{x}_\tau^0 \in \mathbf{H}$ one can solve the Explicit Euler method

$$\frac{\mathbf{x}_\tau^n - \mathbf{x}_\tau^{n-1}}{\tau} \in B\mathbf{x}_\tau^{n-1} \quad \Leftrightarrow \quad \mathbf{x}_\tau^n = \mathbf{x}_\tau^{n-1} + \tau B\mathbf{x}_\tau^{n-1} = (\mathrm{Id} + \tau B)^n \mathbf{x}_\tau^0, \quad n = 1, \cdots$$

If B is m-dissipative, for every $\mathbf{x}_\tau^0 \in \mathbf{H}$ one can solve the Implicit Euler method

$$\frac{\mathbf{x}_\tau^n - \mathbf{x}_\tau^{n-1}}{\tau} \in B\mathbf{x}_\tau^n \quad \Leftrightarrow \quad \mathbf{x}_\tau^n = J_\tau \mathbf{x}_\tau^{n-1} = J_\tau^n \mathbf{x}_\tau^0, \quad n = 1, \cdots$$



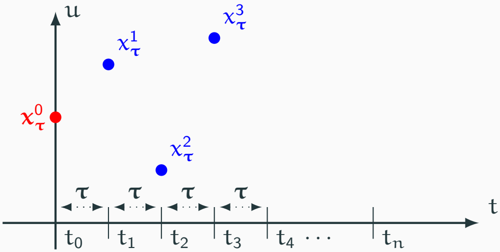
The Explicit and Implicit Euler methods

If B is everywhere defined, for every $\boldsymbol{x}_\tau^0 \in \mathbf{H}$ one can solve the Explicit Euler method

$$\frac{\boldsymbol{x}_\tau^n - \boldsymbol{x}_\tau^{n-1}}{\tau} \in B\boldsymbol{x}_\tau^{n-1} \quad \Leftrightarrow \quad \boldsymbol{x}_\tau^n = \boldsymbol{x}_\tau^{n-1} + \tau B\boldsymbol{x}_\tau^{n-1} = (\mathrm{Id} + \tau B)^n \boldsymbol{x}_\tau^0, \quad n = 1, \cdots$$

If B is m-dissipative, for every $\boldsymbol{x}_\tau^0 \in \mathbf{H}$ one can solve the Implicit Euler method

$$\frac{\boldsymbol{x}_\tau^n - \boldsymbol{x}_\tau^{n-1}}{\tau} \in B\boldsymbol{x}_\tau^n \quad \Leftrightarrow \quad \boldsymbol{x}_\tau^n = J_\tau \boldsymbol{x}_\tau^{n-1} = J_\tau^n \boldsymbol{x}_\tau^0, \quad n = 1, \cdots$$



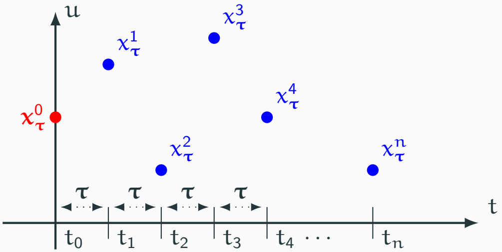
The Explicit and Implicit Euler methods

If B is everywhere defined, for every $\boldsymbol{x}_{\boldsymbol{\tau}}^{0} \in \mathbf{H}$ one can solve the Explicit Euler method

$$\frac{\boldsymbol{x}_{\boldsymbol{\tau}}^{\boldsymbol{n}} - \boldsymbol{x}_{\boldsymbol{\tau}}^{\boldsymbol{n-1}}}{\tau} \in B\boldsymbol{x}_{\boldsymbol{\tau}}^{\boldsymbol{n-1}} \quad \Leftrightarrow \quad \boldsymbol{x}_{\boldsymbol{\tau}}^{\boldsymbol{n}} = \boldsymbol{x}_{\boldsymbol{\tau}}^{\boldsymbol{n-1}} + \tau B\boldsymbol{x}_{\boldsymbol{\tau}}^{\boldsymbol{n-1}} = (\mathrm{Id} + \tau B)^n \boldsymbol{x}_{\boldsymbol{\tau}}^{0}, \quad n = 1, \cdots$$

If B is m-dissipative, for every $\boldsymbol{x}_{\boldsymbol{\tau}}^{0} \in \mathbf{H}$ one can solve the Implicit Euler method

$$\frac{\boldsymbol{x}_{\boldsymbol{\tau}}^{\boldsymbol{n}} - \boldsymbol{x}_{\boldsymbol{\tau}}^{\boldsymbol{n-1}}}{\tau} \in B\boldsymbol{x}_{\boldsymbol{\tau}}^{\boldsymbol{n}} \quad \Leftrightarrow \quad \boldsymbol{x}_{\boldsymbol{\tau}}^{\boldsymbol{n}} = J_\tau \boldsymbol{x}_{\boldsymbol{\tau}}^{\boldsymbol{n-1}} = J_\tau^n \boldsymbol{x}_{\boldsymbol{\tau}}^{0}, \quad n = 1, \cdots$$



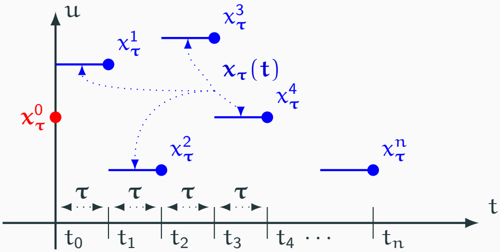
The Explicit and Implicit Euler methods

If B is everywhere defined, for every $\boldsymbol{x}_\tau^0 \in \mathbf{H}$ one can solve the Explicit Euler method

$$\frac{\boldsymbol{x}_\tau^n - \boldsymbol{x}_\tau^{n-1}}{\tau} \in \mathrm{B}\boldsymbol{x}_\tau^{n-1} \quad \Leftrightarrow \quad \boldsymbol{x}_\tau^n = \boldsymbol{x}_\tau^{n-1} + \tau \mathrm{B}\boldsymbol{x}_\tau^{n-1} = (\mathrm{Id} + \tau\mathrm{B})^n \boldsymbol{x}_\tau^0, \quad n = 1, \cdots$$

If B is m-dissipative, for every $\boldsymbol{x}_\tau^0 \in \mathbf{H}$ one can solve the Implicit Euler method

$$\frac{\boldsymbol{x}_\tau^n - \boldsymbol{x}_\tau^{n-1}}{\tau} \in \mathrm{B}\boldsymbol{x}_\tau^{n} \quad \Leftrightarrow \quad \boldsymbol{x}_\tau^n = \mathrm{J}_\tau \boldsymbol{x}_\tau^{n-1} = \mathrm{J}_\tau^n \boldsymbol{x}_\tau^0, \quad n = 1, \cdots$$



$\bar{\boldsymbol{x}}_\tau$ is the piecewise constant interpolant of the values $(\boldsymbol{x}_\tau^n)_{n\in\mathbb{N}}$.

Convergence and characterization of the limit solution

Theorem (Crandall-Liggett '71)

If B *is* m*-accretive, for every* $x_0 \in \overline{\mathrm{D}(\mathrm{B})}$ *the discrete solutions* $\bar{\boldsymbol{x}}_{\boldsymbol{\tau}}$ *of the implicit Eluer scheme converge uniformly to a limit curve* $\boldsymbol{x} \in \mathrm{C}([0,\infty);\mathbf{H})$.

$$\boxed{|\boldsymbol{x}(T) - \boldsymbol{x}_{\boldsymbol{\tau}}(\mathbf{T})| \leqslant |x_0 - \boldsymbol{x}_{\boldsymbol{\tau}}^0| + 2\sqrt{T\tau}\,|\mathrm{B}\boldsymbol{x}_{\boldsymbol{\tau}}^0|}$$

Convergence and characterization of the limit solution

Theorem (Crandall-Liggett '71)

If B *is* $\mathfrak{m}$*-accretive, for every* $x_0 \in \overline{\mathrm{D}(\mathrm{B})}$ *the discrete solutions* $\bar{\boldsymbol{x}}_\tau$ *of the implicit Eluer scheme converge uniformly to a limit curve* $\boldsymbol{x} \in \mathrm{C}([0,\infty);\mathbf{H})$.

$$\boxed{|\boldsymbol{x}(\mathrm{T}) - \boldsymbol{x}_\tau(\mathrm{T})| \leqslant |x_0 - \boldsymbol{x}_\tau^0| + 2\sqrt{\mathrm{T}\tau}\,|\mathrm{B}\boldsymbol{x}_\tau^0|}$$

Theorem (Bénilan, '72)

If $x_0 \in \mathrm{D}(\mathrm{B})$ *then* $\boldsymbol{x}$ *is Lipschitz and solves*

$$\dot{\boldsymbol{x}}(t) \in \mathrm{B}\boldsymbol{x}(t) \quad \textit{for a.e. } t > 0.$$

Convergence and characterization of the limit solution

Theorem (Crandall-Liggett '71)

If B *is* $\mathfrak{m}$*-accretive, for every* $x_0 \in \overline{\mathrm{D}(\mathrm{B})}$ *the discrete solutions* $\bar{\boldsymbol{x}}_\tau$ *of the implicit Eluer scheme converge uniformly to a limit curve* $\boldsymbol{x} \in C([0,\infty); \mathbf{H})$.

$$\boxed{|\boldsymbol{x}(T) - \boldsymbol{x}_\tau(T)| \leqslant |x_0 - \boldsymbol{x}_\tau^0| + 2\sqrt{T\tau}\,|\mathrm{B}\boldsymbol{x}_\tau^0|}$$

Theorem (Bénilan, '72)

If $x_0 \in \mathrm{D}(\mathrm{B})$ *then* $\boldsymbol{x}$ *is Lipschitz and solves*

$$\dot{\boldsymbol{x}}(t) \in \mathrm{B}\boldsymbol{x}(t) \quad \text{for a.e. } t > 0.$$

If $x_0 \in \overline{\mathrm{D}(\mathrm{B})}$ *then* $\boldsymbol{x}$ *is the unique integral solution:*

$$\frac{1}{2}\frac{d}{dt}|\boldsymbol{x}(t) - y|^2 \leqslant -\langle \mathrm{B}(y), y - \boldsymbol{x}(t)\rangle \quad \text{in } \mathscr{D}'(0,\infty), \ \text{for every } y \in \mathrm{D}(\mathrm{B}).$$

Convergence and characterization of the limit solution

Theorem (Crandall-Liggett '71)

If B *is* $\mathfrak{m}$*-accretive, for every* $x_0 \in \overline{\mathrm{D}(\mathrm{B})}$ *the discrete solutions* $\bar{\boldsymbol{x}}_\tau$ *of the implicit Eluer scheme converge uniformly to a limit curve* $\boldsymbol{x} \in C([0,\infty);\mathbf{H})$.

$$\boxed{|\boldsymbol{x}(T) - \boldsymbol{x}_\tau(T)| \leqslant |x_0 - \boldsymbol{x}_\tau^0| + 2\sqrt{T\tau}\,|B\boldsymbol{x}_\tau^0|}$$

Theorem (Bénilan, '72)

If $x_0 \in \mathrm{D}(\mathrm{B})$ *then* $\boldsymbol{x}$ *is Lipschitz and solves*

$$\dot{\boldsymbol{x}}(t) \in B\boldsymbol{x}(t) \quad \textit{for a.e. } t > 0.$$

If $x_0 \in \overline{\mathrm{D}(\mathrm{B})}$ *then* $\boldsymbol{x}$ *is the unique integral solution:*

$$\frac{1}{2}\frac{d}{dt}|\boldsymbol{x}(t) - y|^2 \leqslant -\langle B(y), y - \boldsymbol{x}(t)\rangle \quad \textit{in } \mathscr{D}'(0,\infty), \ \textit{for every } y \in \mathrm{D}(\mathrm{B}).$$

Formally

$$\frac{1}{2}\frac{d}{dt}|\boldsymbol{x}(t) - y|^2 = \langle B\boldsymbol{x}(t), \boldsymbol{x}(t) - y\rangle = \langle B\boldsymbol{x}(t) - By, \boldsymbol{x}(t) - y\rangle + \langle By, \boldsymbol{x}(t) - y\rangle$$
$$\leqslant \langle By, \boldsymbol{x}(t) - y\rangle.$$

PVFs and displacement extrapolation

In $\mathcal{P}_2(E)$ the role of the curve $x(\tau) := x + \tau B(x)$ is played by

$$\underline{F}(\tau) := \exp_\sharp^\tau \underline{F} = (\boldsymbol{x} + \tau \boldsymbol{v})_\sharp \underline{F}, \quad \underline{F} \in \mathfrak{F}(\mu).$$

If $(X, V)_\sharp \mathbb{P} = \underline{F}$ we have

$$\underline{F}(\tau) = (X + \tau V)_\sharp \mathbb{P}$$

Semiconcavity of the Wasserstein distance

If $\underline{\mathbf{F}} \in \mathfrak{F}(\mu)$ and $\underline{\mathbf{G}} \in \mathfrak{F}(\nu)$, the map

$$\mathrm{D}(\tau; \mu, \nu) := \frac{1}{2} W_2^2(\underline{\mathbf{F}}(\tau), \underline{\mathbf{G}}(\tau))$$

is not convex nor λ-convex for any $\lambda < 0$. In fact it is semiconcave, i.e.

$$\tau \mapsto \mathrm{D}(\tau; \mu, \nu) - \mathrm{C}\tau^2 \quad \text{is concave for a suitable C depending on } \underline{\mathbf{F}}, \underline{\mathbf{G}}.$$

Semiconcavity of the Wasserstein distance

If $\underline{\mathrm{F}} \in \mathfrak{F}(\mu)$ and $\underline{\mathbf{G}} \in \mathfrak{F}(\nu)$, the map

$$\mathrm{D}(\tau; \mu, \nu) := \frac{1}{2} W_2^2(\underline{\mathrm{F}}(\tau), \underline{\mathbf{G}}(\tau))$$

is not convex nor λ-convex for any $\lambda < 0$. In fact it is semiconcave, i.e.

$$\tau \mapsto \mathrm{D}(\tau; \mu, \nu) - \mathrm{C}\tau^2 \quad \text{is concave for a suitable C depending on } \underline{\mathrm{F}}, \underline{\mathbf{G}}.$$

We can still compute the derivative at $\tau = 0$ but

$$\frac{\mathrm{d}}{\mathrm{d}\tau} W_2^2(\underline{\mathrm{F}}(\tau), \underline{\mathbf{G}}(\tau))\Big|_{\tau=0+} \leqslant \frac{\mathrm{d}}{\mathrm{d}\tau} W_2^2(\underline{\mathrm{F}}(\tau), \underline{\mathbf{G}}(\tau))\Big|_{\tau=0-}.$$

Semiconcavity of the Wasserstein distance

If $\underline{\mathbf{F}} \in \mathfrak{F}(\mu)$ and $\underline{\mathbf{G}} \in \mathfrak{F}(\nu)$, the map

$$\mathrm{D}(\tau; \mu, \nu) := \frac{1}{2} W_2^2(\underline{\mathbf{F}}(\tau), \underline{\mathbf{G}}(\tau))$$

is not convex nor λ-convex for any $\lambda < 0$. In fact it is semiconcave, i.e.

$$\tau \mapsto \mathrm{D}(\tau; \mu, \nu) - C\tau^2 \quad \text{is concave for a suitable } C \text{ depending on } \underline{\mathbf{F}}, \underline{\mathbf{G}}.$$

We can still compute the derivative at $\tau = 0$ but

$$\frac{\mathrm{d}}{\mathrm{d}\tau} W_2^2(\underline{\mathbf{F}}(\tau), \underline{\mathbf{G}}(\tau))\Big|_{\tau=0+} \leqslant \frac{\mathrm{d}}{\mathrm{d}\tau} W_2^2(\underline{\mathbf{F}}(\tau), \underline{\mathbf{G}}(\tau))\Big|_{\tau=0-}.$$

In particular

The Wasserstein space is Positively Curved (PC)

In $\mathbb{E} = \mathcal{P}_2(\mathbb{R}^2)$ consider two point masses μ_0 and μ_1 ...

The Wasserstein space is Positively Curved (PC)

In $\mathbb{E} = \mathcal{P}_2(\mathbb{R}^2)$ consider two point masses μ_0 and μ_1 ...

The Wasserstein space is Positively Curved (PC)

In $E = \mathcal{P}_2(\mathbb{R}^2)$ consider two point masses μ_0 and μ_1 ...

The Wasserstein space is Positively Curved (PC)

In $\mathbb{E} = \mathcal{P}_2(\mathbb{R}^2)$ consider two point masses μ_0 and μ_1 ...

The Wasserstein space is Positively Curved (PC)

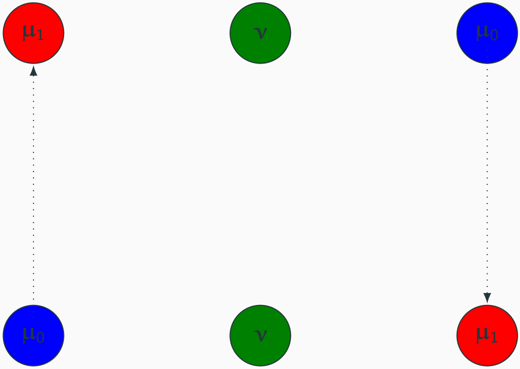
In $E = \mathcal{P}_2(\mathbb{R}^2)$ consider two point masses $\boldsymbol{\mu}_0$ and $\boldsymbol{\mu}_1$ ...

The Wasserstein space is Positively Curved (PC)

In $E = \mathcal{P}_2(\mathbb{R}^2)$ consider two point masses μ_0 and μ_1 ...

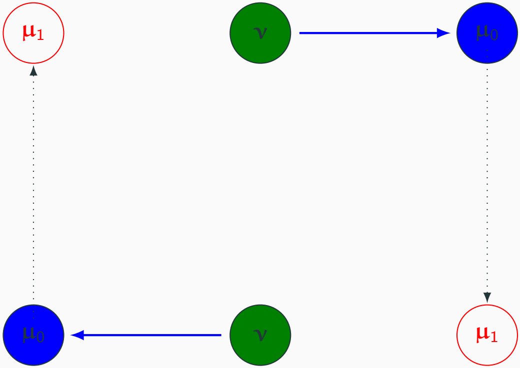
The Wasserstein space is Positively Curved (PC)

In $\mathbb{E} = \mathcal{P}_2(\mathbb{R}^2)$ consider two point masses $\boldsymbol{\mu}_0$ and $\boldsymbol{\mu}_1$ …and a third reference measure $\boldsymbol{\nu}$.



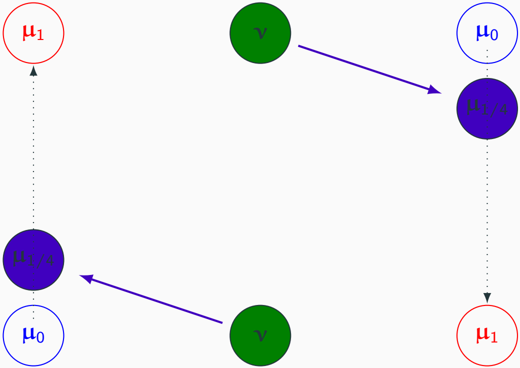
The Wasserstein space is Positively Curved (PC)

In $\mathbb{E} = \mathcal{P}_2(\mathbb{R}^2)$ consider two point masses μ_0 and μ_1 . . . and a third reference measure ν.



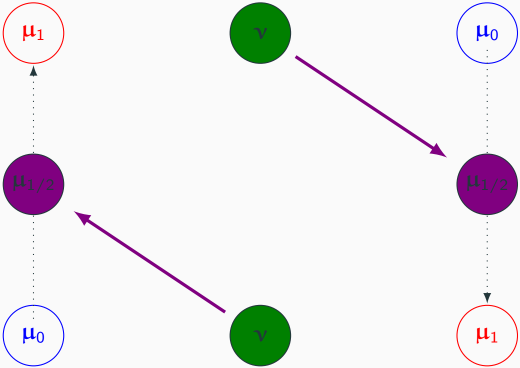
The Wasserstein space is Positively Curved (PC)

In $\mathbb{E} = \mathcal{P}_2(\mathbb{R}^2)$ consider two point masses μ_0 and μ_1 ...and a third reference measure ν.



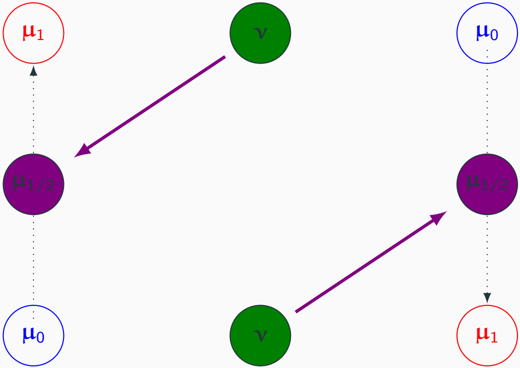
The Wasserstein space is Positively Curved (PC)

In $\mathbb{E} = \mathcal{P}_2(\mathbb{R}^2)$ consider two point masses μ_0 and μ_1 . . . and a third reference measure ν.



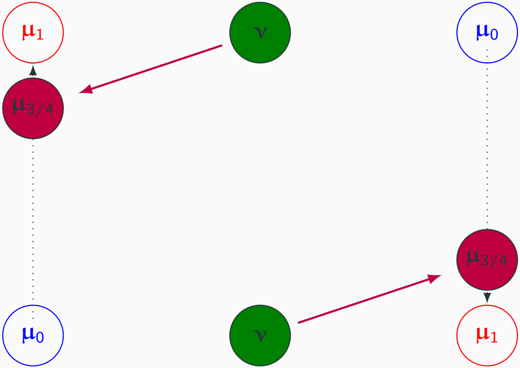
The Wasserstein space is Positively Curved (PC)

In $\mathbb{E} = \mathcal{P}_2(\mathbb{R}^2)$ consider two point masses $\boldsymbol{\mu}_0$ and $\boldsymbol{\mu}_1$ . . . and a third reference measure $\boldsymbol{\nu}$.



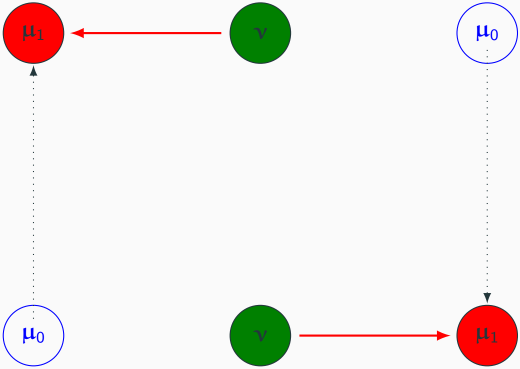
The Wasserstein space is Positively Curved (PC)

In $E = \mathcal{P}_2(\mathbb{R}^2)$ consider two point masses μ_0 and μ_1 ...and a third reference measure ν.



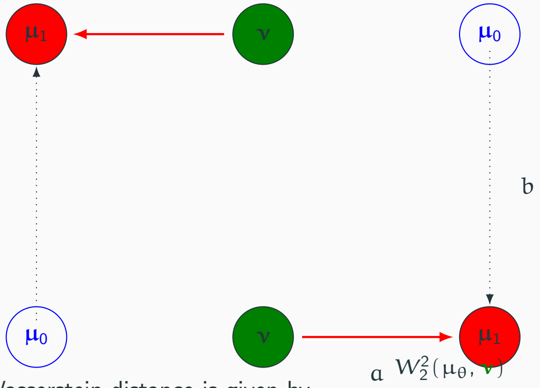
The Wasserstein space is Positively Curved (PC)

In $\mathbb{E} = \mathcal{P}_2(\mathbb{R}^2)$ consider two point masses μ_0 and μ_1 . . . and a third reference measure ν.



The Wasserstein space is Positively Curved (PC)

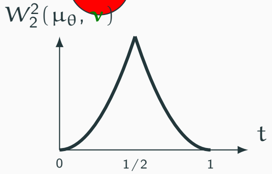
In $E = \mathcal{P}_2(\mathbb{R}^2)$ consider two point masses μ_0 and μ_1 ...and a third reference measure ν.



The Wasserstein distance is given by

$$W_2^2(\nu, \mu_\theta) = \min\left(a^2 + b^2\theta^2, a^2 + b^2(1-\theta)^2\right)$$

It is not λ-convex, for any λ.



Metric dissipativity

We first compute the right derivative

$$\frac{\mathrm{d}}{\mathrm{d}\tau}\frac{1}{2}W_2^2(\underline{\mathrm{F}}(\tau),\nu)\Big|_{\tau=0+}, \qquad \underline{\mathrm{F}}(\tau)=(\boldsymbol{x}+\tau\boldsymbol{v})_\sharp\underline{\mathrm{F}}, \quad \underline{\mathrm{F}}\in\mathfrak{F}(\mu)$$

keeping fixed ν.

Metric dissipativity

We first compute the right derivative

$$\frac{\mathrm{d}}{\mathrm{d}\tau}\frac{1}{2}W_2^2(\underline{\mathbf{F}}(\tau),\nu)\Big|_{\tau=0+}, \qquad \underline{\mathbf{F}}(\tau)=(\boldsymbol{x}+\tau\boldsymbol{v})_\sharp\underline{\mathbf{F}}, \quad \underline{\mathbf{F}}\in\mathfrak{F}(\mu)$$

keeping fixed ν.

It is usefult to keep in mind that in Hilbert spaces,

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{1}{2}\Big|(x+\tau v)-y\Big|^2=\langle v,x-y\rangle$$

Metric dissipativity

We first compute the right derivative

$$\frac{d}{d\tau}\frac{1}{2}W_2^2(\underline{F}(\tau),\nu)\Big|_{\tau=0+}, \qquad \underline{F}(\tau)=(\boldsymbol{x}+\tau\boldsymbol{v})_\sharp\underline{F}, \quad \underline{F}\in\mathfrak{F}(\mu)$$

keeping fixed ν.

It is usefult to keep in mind that in Hilbert spaces,

$$\frac{d}{dt}\frac{1}{2}\Big|(x+\tau v)-y\Big|^2=\langle v, x-y\rangle$$

We introduce the set $\Gamma_o(\underline{F},\nu)$ of couplings given by triple of random variables X, V, Y such that

$$(X,V)_\sharp\mathbb{P}=\underline{F}, \quad Y_\sharp\mathbb{P}=\nu \text{ and}$$
$$(X,Y)_\sharp\mathbb{P}\in\Gamma_o(\mu,\nu) \text{ is an optimal coupling between } \mu \text{ and } \nu.$$

$$[\underline{F},\nu]_r=\frac{d}{d\tau}\frac{1}{2}W_2^2(\underline{F}(\tau),\nu)\Big|_{\tau=0+}=\min\Big\{\mathbb{E}\Big[\langle V, X-Y\rangle\Big] : (X,V,Y)\in\Gamma_o(\underline{F},\nu)\Big\}$$

Metric dissipativity

We first compute the right derivative

$$\frac{d}{d\tau}\frac{1}{2}W_2^2(\underline{F}(\tau),\nu)\Big|_{\tau=0+}, \qquad \underline{F}(\tau)=(\boldsymbol{x}+\tau\boldsymbol{v})_\sharp\underline{F}, \quad \underline{F}\in\mathfrak{F}(\mu)$$

keeping fixed ν.

It is usefult to keep in mind that in Hilbert spaces,

$$\frac{d}{dt}\frac{1}{2}\Big|(x+\tau v)-y\Big|^2=\langle v, x-y\rangle$$

We introduce the set $\Gamma_o(\underline{F},\nu)$ of couplings given by triple of random variables X, V, Y such that

$$(X,V)_\sharp\mathbb{P}=\underline{F}, \quad Y_\sharp\mathbb{P}=\nu \text{ and}$$
$$(X,Y)_\sharp\mathbb{P}\in\Gamma_o(\mu,\nu) \text{ is an optimal coupling between } \mu \text{ and } \nu.$$

$$\boxed{[\underline{F},\nu]_r=\frac{d}{d\tau}\frac{1}{2}W_2^2(\underline{F}(\tau),\nu)\Big|_{\tau=0+}=\min\Big\{\mathbb{E}\Big[\langle V,X-Y\rangle\Big]:(X,V,Y)\in\Gamma_o(\underline{F},\nu)\Big\}}$$

On the other hand

$$[\underline{F},\nu]_l=\frac{d}{d\tau}\frac{1}{2}W_2^2(\mu(\tau),\nu(\tau))\Big|_{\tau=0-}=\max\Big\{\mathbb{E}\Big[\langle V,X-Y\rangle\Big]:(X,V,Y)\in\Gamma_o(\underline{F},\nu)\Big\}.$$

Dissipativity

If $v = \dot{x}(0)$, $w = \dot{y}(0)$

$$\begin{aligned} \langle v - w, x - y \rangle &= \langle v, x - y \rangle + \langle w, y - x \rangle \\ &= \frac{d}{d\tau} \frac{1}{2} |x(\tau) - y|^2 \Big|_{\tau=0} + \frac{d}{d\tau} \frac{1}{2} |y(\tau) - x|^2 \Big|_{\tau=0} \end{aligned}$$

Dissipativity

If $v = \dot{x}(0)$, $w = \dot{y}(0)$

$$\begin{aligned}\langle v - w, x - y\rangle &= \langle v, x - y\rangle + \langle w, y - x\rangle \\ &= \frac{\mathrm{d}}{\mathrm{d}\tau}\frac{1}{2}|x(\tau) - y|^2\Big|_{\tau=0} + \frac{\mathrm{d}}{\mathrm{d}\tau}\frac{1}{2}|y(\tau) - x|^2\Big|_{\tau=0}\end{aligned}$$

$$\boxed{[\underline{F}, \nu]_r = \frac{\mathrm{d}}{\mathrm{d}\tau}\frac{1}{2}W_2^2(\underline{F}(\tau), \nu)\Big|_{\tau=0+} = \min\left\{\mathbb{E}\Big[\langle V, X - Y\rangle\Big] : (X, V, Y) \in \Gamma_o(\underline{F}, \nu)\right\}}$$

Dissipativity

If $v = \dot{x}(0)$, $w = \dot{y}(0)$

$$\begin{aligned}\langle v - w, x - y\rangle &= \langle v, x - y\rangle + \langle w, y - x\rangle \\ &= \frac{\mathrm{d}}{\mathrm{d}\tau}\frac{1}{2}|x(\tau) - y|^2\Big|_{\tau=0} + \frac{\mathrm{d}}{\mathrm{d}\tau}\frac{1}{2}|y(\tau) - x|^2\Big|_{\tau=0}\end{aligned}$$

$$\boxed{[\underline{F}, \nu]_r = \frac{\mathrm{d}}{\mathrm{d}\tau}\frac{1}{2}W_2^2(\underline{F}(\tau), \nu)\Big|_{\tau=0+} = \min\Big\{\mathbb{E}\Big[\langle V, X - Y\rangle\Big] : (X, V, Y) \in \Gamma_o(\underline{F}, \nu)\Big\}}$$

In general

$$[\underline{F}, \nu]_r + [\underline{\mathbf{G}}, \mu]_r \leqslant \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}\tau}W_2^2(\underline{F}(\tau), \underline{\mathbf{G}}(\tau))\Big|_{\tau=0+}, \qquad \underline{F} \in \mathfrak{F}(\mu),\ \underline{\mathbf{G}} \in \mathfrak{F}(\nu).$$

Dissipativity

If $v = \dot{x}(0)$, $w = \dot{y}(0)$

$$\begin{aligned}\langle v - w, x - y\rangle &= \langle v, x - y\rangle + \langle w, y - x\rangle \\ &= \frac{d}{d\tau}\frac{1}{2}|x(\tau) - y|^2\Big|_{\tau=0} + \frac{d}{d\tau}\frac{1}{2}|y(\tau) - x|^2\Big|_{\tau=0}\end{aligned}$$

$$\boxed{[\underline{F}, \nu]_r = \frac{d}{d\tau}\frac{1}{2}W_2^2(\underline{F}(\tau), \nu)\Big|_{\tau=0+} = \min\left\{\mathbb{E}\Big[\langle V, X - Y\rangle\Big] : (X, V, Y) \in \Gamma_o(\underline{F}, \nu)\right\}}$$

In general

$$[\underline{F}, \nu]_r + [\underline{G}, \mu]_r \leqslant \frac{1}{2}\frac{d}{d\tau}W_2^2(\underline{F}(\tau), \underline{G}(\tau))\Big|_{\tau=0+}, \qquad \underline{F} \in \mathfrak{F}(\mu),\ \underline{G} \in \mathfrak{F}(\nu).$$

$\mathfrak{F}$ is (metrically) dissipative if [Cavagnari-Sodini-S.]

$$[\mathfrak{F}(\mu), \nu]_r + [\mathfrak{F}(\nu), \mu]_r \leqslant 0 \quad \text{for every } \mu, \nu \in \mathcal{P}_2(E).$$

Wasserstein subdifferential and the role of optimal plans

If $\mathscr{F} : \mathcal{P}_2(\mathsf{E}) \to (-\infty, +\infty]$ is a geodesically convex functional than its (opposite) Wasserstein subdifferential $\mathfrak{F} = -\partial_W \mathscr{F}$ is defined by

$$\boxed{\underline{\mathrm{F}} \in \mathfrak{F}(\mu) \quad \Leftrightarrow \quad [\underline{\mathrm{F}}, \nu]_r \leqslant \mathscr{F}(\nu) - \mathscr{F}(\mu) \quad \text{for every } \nu \in \mathrm{D}(\mathscr{F})}$$

If $\mathscr{F}$ is geodesically convex then $\mathfrak{F}$ is dissipative.

Wasserstein subdifferential and the role of optimal plans

If $\mathscr{F} : \mathcal{P}_2(E) \to (-\infty, +\infty]$ is a geodesically convex functional than its (opposite) Wasserstein subdifferential $\mathfrak{F} = -\partial_W \mathscr{F}$ is defined by

$$\boxed{\underline{F} \in \mathfrak{F}(\mu) \quad \Leftrightarrow \quad [\underline{F}, \nu]_r \leqslant \mathscr{F}(\nu) - \mathscr{F}(\mu) \quad \text{for every } \nu \in D(\mathscr{F})}$$

If $\mathscr{F}$ is geodesically convex then $\mathfrak{F}$ is dissipative.

The Relative Entropy functional

$$\boxed{\mathscr{F}(\mu) := \int u(\log u + V)\, dx = \mathrm{Ent}(\mu|\mathfrak{m}) \quad \text{if } \mu = u\mathscr{L}^d \ll \mathscr{L}^d, \quad \mathfrak{m} = e^{-V}\mathscr{L}^d.}$$

$\mathscr{F} \equiv +\infty$ on the discrete measures. $\mathscr{F}$ is geodesically convex (i.e. convex along displacement interpolations, [McCann '97]) but not convex along arbitrary interpolation of measures: optimal interpolations avoid collisions!

Implicit Euler method: the subgradient case

If $\underline{\mathrm{F}}$ arises as the gradient of a displacement convex functional $\mathscr{F}$, we can also use a variational formulation of the implicit Euler method.

[5] R. Jordan, D. Kinderlehrer, F. Otto, The variational formulation of the Fokker-Planck equation. SIAM J. Math. Anal. (1998)
L. Ambrosio, N. Gigli, G. S. (2008),

Implicit Euler method: the subgradient case

If $\underline{F}$ arises as the gradient of a displacement convex functional $\mathscr{F}$, we can also use a variational formulation of the implicit Euler method.

According to the JKO -Minimizing Movement approach [5], at each step it is sufficient to select $\boldsymbol{\mu}_{\boldsymbol{\tau}}^{\boldsymbol{n}}$ among the minimizers of

$$\boxed{\mu \mapsto \frac{1}{2\tau} W_2^2(\mu, \boldsymbol{\mu}_{\boldsymbol{\tau}}^{\boldsymbol{n}-1}) + \mathscr{F}(\mu)}$$

[5] R. Jordan, D. Kinderlehrer, F. Otto, The variational formulation of the Fokker-Planck equation. SIAM J. Math. Anal. (1998)
L. Ambrosio, N. Gigli, G. S. (2008),

Wasserstein gradient flows

Let $\mathscr{F} : \mathcal{P}_2(E) \to (-\infty, +\infty]$ be a lower semicontinuous and displacement convex functional. We say that a locally Lipschitz curve $(\mu_t)_{t>0}$ is an EVI-solution of the gradient flow of $\mathscr{F}$ if for every $\nu \in D(\mathscr{F}) \subset \mathcal{P}_2(E)$

$$\boxed{\frac{d}{dt}\frac{1}{2}W_2^2(\mu_t, \nu) \leqslant \mathscr{F}(\nu) - \mathscr{F}(\mu_t) \quad \text{a.e. in } (0, \infty).} \tag{EVI}$$

Wasserstein gradient flows

Let $\mathscr{F} : \mathcal{P}_2(E) \to (-\infty, +\infty]$ be a lower semicontinuous and displacement convex functional. We say that a locally Lipschitz curve $(\mu_t)_{t>0}$ is an EVI-solution of the gradient flow of $\mathscr{F}$ if for every $\nu \in D(\mathscr{F}) \subset \mathcal{P}_2(E)$

$$\boxed{\frac{d}{dt}\frac{1}{2}W_2^2(\mu_t, \nu) \leqslant \mathscr{F}(\nu) - \mathscr{F}(\mu_t) \quad \text{a.e. in } (0, \infty).} \tag{EVI}$$

Theorem (Ambrosio-Gigli-S.)

For every initial datum $\mu_0 \in \overline{D(\mathscr{F})}$ *there exists a unique EVI solution to* (EVI) *satisfying* $\lim_{t\downarrow 0} \mu_t = \mu_0$.

Moreover, μ *is the uniform limit of piecewise constant interpolant* $\boldsymbol{\mu}_\tau$ *of the JKO-Minimizing Movement approximations, obtained by solving*

$$\boldsymbol{\mu}_\tau^n \in \operatorname*{argmin}_\mu \Big\{ \frac{1}{2\tau} W_2^2(\mu, \boldsymbol{\mu}_\tau^{n-1}) + \mathscr{F}(\mu) \Big\}, \quad \boldsymbol{\mu}_\tau^0 := \mu_0$$

Uniform error estimate if $\mu_0 \in D(\mathscr{F})$*:*

$$W_2(\mu(t), \boldsymbol{\mu}_\tau(t)) \leqslant C\sqrt{\tau}.$$

Main problems in the general dissipative case

- Metric dissipativity does not imply contraction of the resolvent.

Main problems in the general dissipative case

- Metric dissipativity does not imply contraction of the resolvent.
- It is not clear how to solve the implicit Euler method even if $\mathfrak{F}$ is defined everywhere and it is single valued.

Main problems in the general dissipative case

- Metric dissipativity does not imply contraction of the resolvent.
- It is not clear how to solve the implicit Euler method even if $\mathfrak{F}$ is defined everywhere and it is single valued.
- Technical point: perturbations along $\mathfrak{F}$ can split particles.

Main problems in the general dissipative case

- Metric dissipativity does not imply contraction of the resolvent.
- It is not clear how to solve the implicit Euler method even if $\mathfrak{F}$ is defined everywhere and it is single valued.
- Technical point: perturbations along $\mathfrak{F}$ can split particles.
- $\mathfrak{F}$ behaves well only along (optimal) displacement interpolations.

Outline

The explicit Euler method for dissipative evoutions

We can construct solutions to the evolution equation by means of the Explicit Euler method:

we fix a step size $\tau > 0$, an initial measure μ_0. If $\mathfrak{F}$ is a dissipative vector field and $\tau > 0$ is a time step, we consider the curves $\underline{F}(\tau)$ in $\mathcal{P}_2(E)$, $\underline{F} \in \mathfrak{F}(\mu)$

$$\underline{F}(\tau) := \exp^{\tau}_{\sharp} \underline{F} = (x + \tau v)_{\sharp}\underline{F}, \quad \underline{F} \in \mathfrak{F}(\mu)$$

and therefore the sequence of explicit Euler approximations:

$$\boldsymbol{\mu}_{\boldsymbol{\tau}}^0 := \mu_0 \text{ given}, \quad \boldsymbol{\mu}_{\boldsymbol{\tau}}^{\mathbf{n}+1} := \underline{F}^n(\tau),\ \underline{F}^n \in \mathfrak{F}(\boldsymbol{\mu}_{\boldsymbol{\tau}}^{\mathbf{n}}), \quad \boldsymbol{\mu}_{\boldsymbol{\tau}}(\mathbf{t}) := \boldsymbol{\mu}_{\boldsymbol{\tau}}^{\lfloor \mathbf{t}/\boldsymbol{\tau} \rfloor}$$

$\boldsymbol{\mu}_{\boldsymbol{\tau}}$ is the **piecewise constant interpolation,** $\boldsymbol{\mu}_{\boldsymbol{\tau}}(\mathbf{t}) = \boldsymbol{\mu}_{\boldsymbol{\tau}}^{\mathbf{n}}$ if $n\tau \leqslant t < (n+1)\tau$.

Problems: convergence of the method and characterization of the limit.

The explicit Euler method for dissipative evoutions

We can construct solutions to the evolution equation by means of the Explicit Euler method:

we fix a step size $\tau > 0$, an initial measure μ_0. If $\mathfrak{F}$ is a dissipative vector field and $\tau > 0$ is a time step, we consider the curves $\underline{F}(\tau)$ in $\mathcal{P}_2(E)$, $\underline{F} \in \mathfrak{F}(\mu)$

$$\underline{F}(\tau) := \exp_\sharp^\tau \underline{F} = (x + \tau v)_\sharp \underline{F}, \quad \underline{F} \in \mathfrak{F}(\mu)$$

and therefore the sequence of explicit Euler approximations:

$$\boldsymbol{\mu}_{\boldsymbol{\tau}}^0 := \mu_0 \text{ given}, \quad \boldsymbol{\mu}_{\boldsymbol{\tau}}^{\mathbf{n}+1} := \underline{F}^n(\tau),\ \underline{F}^n \in \mathfrak{F}(\boldsymbol{\mu}_{\boldsymbol{\tau}}^{\mathbf{n}}), \quad \boldsymbol{\mu}_{\boldsymbol{\tau}}(\mathbf{t}) := \boldsymbol{\mu}_{\boldsymbol{\tau}}^{\lfloor \mathbf{t}/\boldsymbol{\tau} \rfloor}$$

$\boldsymbol{\mu}_{\boldsymbol{\tau}}$ is the **piecewise constant interpolation,** $\boldsymbol{\mu}_{\boldsymbol{\tau}}(\mathbf{t}) = \boldsymbol{\mu}_{\boldsymbol{\tau}}^{\mathbf{n}}$ if $n\tau \leqslant t < (n+1)\tau$.

Problems: convergence of the method and characterization of the limit.

Easy "Lipschitz" estimate:

$$\frac{W_2(\boldsymbol{\mu}_{\boldsymbol{\tau}}^{\mathbf{n}}, \boldsymbol{\mu}_{\boldsymbol{\tau}}^{\mathbf{n}-1})}{\tau} \leqslant \Big(\int |v|^2 \, d\underline{F}^n(x, v) \Big)^{1/2}$$

$\mathscr{M}(\mu_0, \tau, L, T) :=$ set of discrete solutions $\boldsymbol{\mu}_{\boldsymbol{\tau}}$ of the Explicit Euler method starting from μ_0, defined up to the final time T, such that

$$\int |v|^2 \, d\underline{F}^n(x, v) \leqslant L^2 \quad \text{for every } n \leqslant \lfloor T/\tau \rfloor.$$

Convergence

Theorem (Cavagnari-Sodini-S.)

Suppose that $\mathfrak{F}$ is a dissipative MPVF.

- *If* $k \mapsto \tau(k) \downarrow 0$ *is a vanishing sequence of step sizes and* $\boldsymbol{\mu}_k \in \mathscr{M}(\mu_0, \tau(k), L, T)$ *for some* $L \geqslant 0$*, then the sequence of discrete solutions* $\boldsymbol{\mu}_k$ *of the explicit Euler method uniformly converge to a unique limit* $\boldsymbol{\mu} : [0, T] \to \mathcal{P}_2(E)$.

Convergence

Theorem (Cavagnari-Sodini-S.)

Suppose that $\mathfrak{F}$ is a dissipative MPVF.

- *If* $k \mapsto \tau(k) \downarrow 0$ *is a vanishing sequence of step sizes and* $\boldsymbol{\mu}_k \in \mathscr{M}(\mu_0, \tau(k), L, T)$ *for some* $L \geqslant 0$*, then the sequence of discrete solutions* $\boldsymbol{\mu}_k$ *of the explicit Euler method uniformly converge to a unique limit* $\mu : [0, T] \to \mathcal{P}_2(E)$.
- μ *is a Lipschitz curve and it is the unique solution of the disspative EVI (in the distributional sense of* $\mathscr{D}'(0, T)$*)*

$$\boxed{\frac{d}{dt}\frac{1}{2}W_2^2(\boldsymbol{\mu}(t), \nu) \leqslant -[F(\nu), \mu]_r \quad \textit{for every } \nu \in D(\mathfrak{F}); \quad \boldsymbol{\mu}(0) = \mu_0,}$$

Convergence

Theorem (Cavagnari-Sodini-S.)

Suppose that $\mathfrak{F}$ is a dissipative MPVF.

- *If* $k \mapsto \tau(k) \downarrow 0$ *is a vanishing sequence of step sizes and* $\mu_k \in \mathscr{M}(\mu_0, \tau(k), L, T)$ *for some* $L \geqslant 0$*, then the sequence of discrete solutions* μ_k *of the explicit Euler method uniformly converge to a unique limit* $\mu : [0, T] \to \mathcal{P}_2(E)$.
- μ *is a Lipschitz curve and it is the unique solution of the disspative EVI (in the distributional sense of* $\mathscr{D}'(0, T)$*)*

$$\boxed{\frac{d}{dt}\frac{1}{2}W_2^2(\mu(t), \nu) \leqslant -[F(\nu), \mu]_r \quad \textit{for every } \nu \in D(\mathfrak{F}); \quad \mu(0) = \mu_0,}$$

- *We have the optimal error estimate*

$$W_2(\mu(t), \mu_\tau(t)) \leqslant CL\sqrt{T\tau} \quad \textit{for every } t \in [0, T].$$

Generation of a flow of contractions

Suppose that $\mathfrak{F}$ is a dissipative MPVF such that $\mathrm{D}(\mathfrak{F})$ contains all the measures with bounded support of $\mathcal{P}_b(E)$.

Suppose moreover that

- for every $\mu_0 \in \mathcal{P}_b(E)$ there exist $\rho, L > 0$ such that

  $$W_2(\mu, \mu_0) < \rho \quad \Rightarrow \quad \exists \underline{F} \in \mathfrak{F}(\mu) : \mathrm{supp}(v_\sharp \underline{F}) \subset B_L(0).$$

  (local solvability of the Explicit Euler method)
- every $\underline{F} \in \mathfrak{F}$ is concentrated on the set

  $$(x, v) \in E \times E : \quad \langle v, x \rangle \leqslant C(1 + |x|^2)$$

  for some constant C not depending on $\underline{F}$.

Generation of a flow of contractions

Suppose that $\mathfrak{F}$ is a dissipative MPVF such that $D(\mathfrak{F})$ contains all the measures with bounded support of $\mathcal{P}_b(E)$.

Suppose moreover that

- for every $\mu_0 \in \mathcal{P}_b(E)$ there exist $\rho, L > 0$ such that

$$W_2(\mu, \mu_0) < \rho \quad \Rightarrow \quad \exists \underline{F} \in \mathfrak{F}(\mu) : \mathrm{supp}(\nu_\sharp \underline{F}) \subset B_L(0).$$

  (local solvability of the Explicit Euler method)

- every $\underline{F} \in \mathfrak{F}$ is concentrated on the set

$$(x, v) \in E \times E : \quad \langle v, x \rangle \leqslant C(1 + |x|^2)$$

  for some constant C not depending on $\underline{F}$.

Theorem

Then $\mathfrak{F}$ generates a semigroup of contractions: for every $\mu_0 \in \mathcal{P}_2(E)$ there exists a unique continuous curve $\boldsymbol{\mu} = \mathbf{S}[\mu_0] \in C([0, \infty); \mathcal{P}_2(E))$ such that

$$\boxed{\frac{d}{dt}\frac{1}{2}W_2^2(\boldsymbol{\mu}(t), \nu) \leqslant -[F(\nu), \mu]_r \quad \textit{for every } \nu \in D(\mathfrak{F}) \quad \boldsymbol{\mu}(0) = \mu_0,}$$

$$W_2(\mathbf{S}_t[\mu_0], \mathbf{S}_t[\nu_0]) \leqslant W_2(\mu_0, \nu_0) \quad \textit{for every } \mu_0, \nu_0 \in \mathcal{P}_2(E), \ t > 0.$$

Barycentric property

Under the same conditions, let us also suppose that the sections $\mathfrak{F}(\mu)$ of $\mathfrak{F}$ are convex and the graph of $\mathfrak{F}$ is closed under strong-weak convergence: if a sequence $\underline{F}_n \in \mathfrak{F}(\mu_n)$ satisfies

$$\mu_n \to \mu \text{ in } \mathcal{P}_2(E), \quad \underline{F}_n \to \underline{F} \text{ in } \mathcal{P}(E \times E), \quad \sup_n \int |v|^2 \, d\underline{F}_n(x, v) < \infty$$

then $\underline{F} \in \mathfrak{F}(\mu)$.

Barycentric property

Under the same conditions, let us also suppose that the sections $\mathfrak{F}(\mu)$ of $\mathfrak{F}$ are convex and the graph of $\mathfrak{F}$ is closed under strong-weak convergence: if a sequence $\underline{F}_n \in \mathfrak{F}(\mu_n)$ satisfies

$$\mu_n \to \mu \text{ in } \mathcal{P}_2(E), \quad \underline{F}_n \to \underline{F} \text{ in } \mathcal{P}(E \times E), \quad \sup_n \int |v|^2 \, d\underline{F}_n(x,v) < \infty$$

then $\underline{F} \in \mathfrak{F}(\mu)$.

Theorem

Every EVI solution $\mu : (0,\infty) \to D(\mathfrak{F})$ *satisfies the barycentric property: for* $\mathscr{L}^1$*-a.e.* $t > 0$ *there exists* $\underline{F}_t \in \mathfrak{F}(\mu_t)$ *such that*

$$\frac{d}{dt} \int \zeta(x) \, d\mu_t(x) = \int \langle D\zeta(x), v \rangle \, d\underline{F}_t(x,v) \qquad (\star)$$

for every smooth bounded Lipschitz cylindrical function $\zeta : E \to \mathbb{R}$.

Conversely, if $\mu : (0,\infty) \to D(\mathfrak{F})$ *is absolutely continuous, it satisfies* $(\star)$*, and for a.e.* $t > 0$ $\mu_t \in \mathcal{P}_2^r(E)$ *or* $\mathfrak{F}(\mu_t)$ *contains a unique element concentrated on a map,* μ *is also an EVI solution.*

$(\star)$ is equivalent to $\quad \partial_t \mu_t + \nabla \cdot (\mu_t v_t) = 0, \quad v_t = \mathbf{proj}_{\mathrm{Tan}(\mu_t)}(\underline{F}_t)$

Extension and future developements

- Everything can be easily extended to λ-dissipative probability vector fields.
- Evolutions do not split particles in dimension $\geqslant 2$?
- Impose only local boundedness on $\mathfrak{F}$
- Implicit Euler scheme
- Stability and G-convergence
- ...