Evolution of probability measures and Optimal Transport

Giuseppe Savaré
Department of Decision Sciences, Bocconi University, Milan, Italy

In collaboration with L. Ambrosio, N. Gigli; G. Cavagnari, G. Sodini

Analytical Methods in Quantum and Continuum Mechanics
Torino, November 29, 2021.

Table of contents

Outline

Let us consider the evolution of a system $\mathbf{X}(t) = (X_\omega(t))_{\omega\in\Omega} \in H^\Omega$ of N indistinguishable particles in the Euclidean space $H = \mathbb{R}^d$, parametrized by a finite set $\Omega := \{\omega_1, \omega_2, \cdots, \omega_N\}$, governed by a first order system of differential equations

$$\boxed{\dot{X}_\omega(t) = V_\omega(t) = F(\omega, \mathbf{X}(t)), \qquad \omega \in \Omega}$$

Particle models

Let us consider the evolution of a system $\mathbf{X}(t) = (X_\omega(t))_{\omega\in\Omega} \in H^\Omega$ of N indistinguishable particles in the Euclidean space $H = \mathbb{R}^d$, parametrized by a finite set $\Omega := \{\omega_1, \omega_2, \cdots, \omega_N\}$, governed by a first order system of differential equations

$$\boxed{\dot{X}_\omega(t) = V_\omega(t) = F(\omega, \mathbf{X}(t)), \qquad \omega \in \Omega}$$

where $F : \Omega \times H^\Omega \to H^\Omega$ is invariant with respect to permutations $\sigma \in \mathrm{Sym}(\Omega)$ of the coordinates:

$$F(\sigma(\omega), \mathbf{X} \circ \sigma) = F(\omega, \mathbf{X}) \quad \text{for every permutation } \sigma \in \mathrm{Sym}(\Omega),\ \sigma : \Omega \to \Omega.$$

Particle models

Let us consider the evolution of a system $\mathbf{X}(t) = (X_\omega(t))_{\omega\in\Omega} \in H^\Omega$ of N indistinguishable particles in the Euclidean space $H = \mathbb{R}^d$, parametrized by a finite set $\Omega := \{\omega_1, \omega_2, \cdots, \omega_N\}$, governed by a first order system of differential equations

$$\boxed{\dot{X}_\omega(t) = V_\omega(t) = F(\omega, \mathbf{X}(t)), \qquad \omega \in \Omega}$$

where $F : \Omega \times H^\Omega \to H^\Omega$ is invariant with respect to permutations $\sigma \in \mathrm{Sym}(\Omega)$ of the coordinates:

$$F(\sigma(\omega), \mathbf{X} \circ \sigma) = F(\omega, \mathbf{X}) \quad \text{for every permutation } \sigma \in \mathrm{Sym}(\Omega),\ \sigma : \Omega \to \Omega.$$

We may look for a robust description, suitable to pass to the limit as $N \uparrow \infty$.

Examples

- Finite dimensional Cauchy-Lipschitz theory [1]

$$F(\omega, \mathbf{X}) = A(X_\omega) + \frac{1}{N}\sum_{\eta\in\Omega} B(X_\omega - X_\eta), \quad A, B : H \to H \text{ (one-sided) Lipschitz.}$$

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

[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

Examples

- Finite dimensional Cauchy-Lipschitz theory [1]

$$F(\omega, \mathbf{X}) = A(X_\omega) + \frac{1}{N}\sum_{\eta\in\Omega} B(X_\omega - X_\eta), \quad A, B : H \to H \text{ (one-sided) Lipschitz.}$$

- Gradient flows: [2] given a permutation invariant function

$$F(\omega, \mathbf{X}) := -\partial_{X_\omega}\Phi(\mathbf{X}), \quad \Phi : H^\Omega \to \mathbb{R} \quad \text{invariant by permutations}$$

Φ typically satisfies a suitable convexity condition but can be nonsmooth (subdifferential calculus): e.g. [3]

$$\Phi(\mathbf{X}) = R\Big(\frac{1}{N}\sum_{\omega} T(X_\omega)\Big) + \frac{1}{N^2}\sum_{\eta,\omega} W(X_\omega - X_\eta) + \frac{1}{N}\sum_{\omega} V(X_\omega)$$

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

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

[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

Examples

- Finite dimensional Cauchy-Lipschitz theory [1]

$$F(\omega, \mathbf{X}) = A(X_\omega) + \frac{1}{N}\sum_{\eta\in\Omega} B(X_\omega - X_\eta), \quad A, B : H \to H \text{ (one-sided) Lipschitz.}$$

- Gradient flows: [2] given a permutation invariant function

$$F(\omega, \mathbf{X}) := -\partial_{X_\omega}\Phi(\mathbf{X}), \quad \Phi : H^\Omega \to \mathbb{R} \quad \text{invariant by permutations}$$

  Φ typically satisfies a suitable convexity condition but can be nonsmooth (subdifferential calculus): e.g. [3]

$$\Phi(\mathbf{X}) = R\Big(\frac{1}{N}\sum_{\omega} T(X_\omega)\Big) + \frac{1}{N^2}\sum_{\eta,\omega} W(X_\omega - X_\eta) + \frac{1}{N}\sum_{\omega} V(X_\omega)$$

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

- Dissipative evolution, contraction semigroups: H Hilbert or Banach space, F multivalued. E.g. the Lipschitz perturbation of a multivalued subgradient.

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

[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

Measure description

Ω can be endowed with the normalized counting measure $\mathbb{P} = \frac{1}{N}\sum_{\omega\in\Omega}\delta_\omega$ and the evolution of the system can be described by the discrete probability measures $\mu_t \in \mathcal{P}^N(H)$

$$\mu_t := \frac{1}{N}\sum_{\omega\in\Omega}\delta_{X_\omega(t)} = \mathbf{X}(t)_\sharp\mathbb{P}, \qquad \int \varphi \, d\mu_t = \frac{1}{N}\sum_{\omega\in\Omega}\varphi(X_\omega(t)) = \mathbb{E}[\varphi\circ\mathbf{X}(t)]$$

Recall that a map $Z : \Omega \to H$ acts on discrete measures as

$$Z_\sharp\delta_\omega = \delta_{Z(\omega)}, \quad Z_\sharp\Big(\frac{1}{N}\sum_\omega \delta_\omega\Big) = \frac{1}{N}\sum_\omega \delta_{Z(\omega)}.$$

Measure description

Ω can be endowed with the normalized counting measure $\mathbb{P} = \frac{1}{N}\sum_{\omega\in\Omega}\delta_\omega$ and the evolution of the system can be described by the discrete probability measures $\mu_t \in \mathcal{P}^N(H)$

$$\mu_t := \frac{1}{N}\sum_{\omega\in\Omega}\delta_{X_\omega(t)} = \mathbf{X}(t)_\sharp\mathbb{P}, \qquad \int \varphi\, d\mu_t = \frac{1}{N}\sum_{\omega\in\Omega}\varphi(X_\omega(t)) = \mathbb{E}[\varphi\circ\mathbf{X}(t)]$$

Recall that a map $Z : \Omega \to H$ acts on discrete measures as

$$Z_\sharp\delta_\omega = \delta_{Z(\omega)}, \quad Z_\sharp\Big(\frac{1}{N}\sum_\omega \delta_\omega\Big) = \frac{1}{N}\sum_\omega \delta_{Z(\omega)}.$$

The evolution law involves the velocity vector measure

$$\boldsymbol{\nu}_t = \frac{1}{N}\sum_{\omega\in\Omega}\dot{X}_\omega(t)\delta_{X_\omega(t)} = \underline{v}_t\mu_t,$$

satisfying the continuity equation

$$\boxed{\partial_t\mu_t + \nabla\cdot\boldsymbol{\nu}_t = 0 \quad \text{in the distributional sense of } (0,\infty)\times H}$$

Measure description

Ω can be endowed with the normalized counting measure $\mathbb{P} = \frac{1}{N}\sum_{\omega\in\Omega}\delta_\omega$ and the evolution of the system can be described by the discrete probability measures $\mu_t \in \mathcal{P}^N(H)$

$$\mu_t := \frac{1}{N}\sum_{\omega\in\Omega}\delta_{X_\omega(t)} = \mathbf{X}(t)_\sharp\mathbb{P}, \qquad \int \varphi\,d\mu_t = \frac{1}{N}\sum_{\omega\in\Omega}\varphi(X_\omega(t)) = \mathbb{E}[\varphi\circ\mathbf{X}(t)]$$

Recall that a map $Z:\Omega\to H$ acts on discrete measures as

$$Z_\sharp\delta_\omega = \delta_{Z(\omega)}, \quad Z_\sharp\Big(\frac{1}{N}\sum_\omega\delta_\omega\Big) = \frac{1}{N}\sum_\omega\delta_{Z(\omega)}.$$

The evolution law involves the velocity vector measure

$$\boldsymbol{\nu}_t = \frac{1}{N}\sum_{\omega\in\Omega}\dot{X}_\omega(t)\delta_{X_\omega(t)} = \underline{v}_t\mu_t,$$

satisfying the continuity equation

$$\boxed{\partial_t\mu_t + \nabla\cdot\boldsymbol{\nu}_t = 0 \quad \text{in the distributional sense of } (0,\infty)\times H}$$

$$\frac{d}{dt}\int\varphi\,d\mu_t = \int D\varphi\cdot\underline{v}_t\,d\mu_t \quad \forall\,\varphi\in C_c^\infty(H).$$

Measure description

Ω can be endowed with the normalized counting measure $\mathbb{P} = \frac{1}{N}\sum_{\omega\in\Omega}\delta_\omega$ and the evolution of the system can be described by the discrete probability measures $\mu_t \in \mathcal{P}^N(H)$

$$\mu_t := \frac{1}{N}\sum_{\omega\in\Omega}\delta_{X_\omega(t)} = \mathbf{X}(t)_\sharp\mathbb{P}, \qquad \int \varphi\,d\mu_t = \frac{1}{N}\sum_{\omega\in\Omega}\varphi(X_\omega(t)) = \mathbb{E}[\varphi\circ\mathbf{X}(t)]$$

Recall that a map $Z:\Omega\to H$ acts on discrete measures as

$$Z_\sharp\delta_\omega = \delta_{Z(\omega)}, \quad Z_\sharp\Big(\frac{1}{N}\sum_\omega\delta_\omega\Big) = \frac{1}{N}\sum_\omega\delta_{Z(\omega)}.$$

The evolution law involves the velocity vector measure

$$\boldsymbol{\nu}_t = \frac{1}{N}\sum_{\omega\in\Omega}\dot{X}_\omega(t)\delta_{X_\omega(t)} = \underline{v}_t\mu_t,$$

satisfying the continuity equation

$$\boxed{\partial_t\mu_t + \nabla\cdot\boldsymbol{\nu}_t = 0 \quad \text{in the distributional sense of } (0,\infty)\times H}$$

$$\frac{d}{dt}\int\varphi\,d\mu_t = \int D\varphi\cdot\underline{v}_t\,d\mu_t \quad \forall\,\varphi\in C_c^\infty(H).$$

$$\frac{d}{dt}\int\varphi\,d\mu_t = \frac{1}{N}\sum_\omega D\varphi(X_\omega(t))\dot{X}_\omega(t) = \int D\varphi\cdot\underline{v}_t\,d\mu_t.$$

Velocity vector measure

$$\frac{d}{dt}\int \varphi \, d\mu_t = \frac{1}{N}\sum_{\omega} D\varphi(X_\omega(t))\dot{X}_\omega(t) = \int D\varphi \cdot \underline{v}_t \, d\mu_t.$$

Notice that $\dot{X}_\omega(t) = \dot{X}_\eta(t)$ for a.e. t such that $X_\omega(t) = X_\eta(t)$, so that $\underline{v}_t$ is well defined a.e. on the support of μ_t.

Velocity vector measure

$$\frac{d}{dt}\int \varphi\, d\mu_t = \frac{1}{N}\sum_{\omega} D\varphi(X_\omega(t))\dot{X}_\omega(t) = \int D\varphi \cdot \underline{v}_t\, d\mu_t.$$

Notice that $\dot{X}_\omega(t) = \dot{X}_\eta(t)$ for a.e. t such that $X_\omega(t) = X_\eta(t)$, so that $\underline{v}_t$ is well defined a.e. on the support of μ_t.

A PDE description: suppose that there exists a smooth map $\underline{F} : X \times \mathcal{P}^N(H) \to H$ such that

$$\underline{F}\Big(y, \frac{1}{N}\sum_{\omega} \delta_{X_\omega}\Big) = F(\omega, \mathbf{X}) \quad \text{whenever } X_\omega = y.$$

$$\begin{cases} \partial_t \mu_t + \nabla \cdot (\underline{v}_t \mu_t) = 0 & \text{in the distributional sense of } (0,\infty) \times X, \\ \underline{v}_t(x) = \underline{F}(x, \mu_t). \end{cases}$$

Weak convergence

Definition (Weak convergence)

A sequence of probability measures $(\mu^N)_{N\in\mathbb{N}} \subset \mathcal{P}(H)$ converges to $\mu \in \mathcal{P}(H)$ if

$$\boxed{\lim_{N\to\infty} \int \varphi \, d\mu^N = \int \varphi \, d\mu \quad \text{for every } \varphi \in C_b(H) \text{ (or, equivalently, } C_c^\infty(H))}$$

Definition (Weak convergence)

A sequence of probability measures $(\mu^N)_{N\in\mathbb{N}} \subset \mathcal{P}(H)$ converges to $\mu \in \mathcal{P}(H)$ if

$$\lim_{N\to\infty} \int \varphi \, d\mu^N = \int \varphi \, d\mu \quad \text{for every } \varphi \in C_b(H) \text{ (or, equivalently, } C_c^\infty(H))$$

If the map $\underline{F}$ can be extended to $X \times \mathcal{P}(H)$, so that

$$(\mu^N)_{N\in\mathbb{N}} \subset \mathcal{P}^N(H), \quad \mu^N \stackrel{N\uparrow\infty}{\longrightarrow} \mu \quad \Rightarrow \quad \underline{F}(\cdot, \mu^N) \to \underline{F}(\cdot, \mu) \quad \text{uniformly}$$

then one can hope to pass to the limit in the discrete PDE to obtain

$$\begin{cases} \partial_t \mu_t + \nabla \cdot (\underline{v}_t \mu_t) = 0 \quad \text{in the distributional sense of } (0,\infty) \times X, \\ \underline{v}_t(x) = \underline{F}(x, \mu_t). \end{cases}$$

Definition (Weak convergence)

A sequence of probability measures $(\mu^N)_{N\in\mathbb{N}} \subset \mathcal{P}(H)$ converges to $\mu \in \mathcal{P}(H)$ if

$$\boxed{\lim_{N\to\infty} \int \varphi \, d\mu^N = \int \varphi \, d\mu \quad \text{for every } \varphi \in C_b(H) \text{ (or, equivalently, } C_c^\infty(H))}$$

If the map $\underline{F}$ can be extended to $X \times \mathcal{P}(H)$, so that

$$(\mu^N)_{N\in\mathbb{N}} \subset \mathcal{P}^N(H), \quad \mu^N \xrightarrow{N\uparrow\infty} \mu \quad \Rightarrow \quad \underline{F}(\cdot, \mu^N) \to \underline{F}(\cdot, \mu) \quad \text{uniformly}$$

then one can hope to pass to the limit in the discrete PDE to obtain

$$\begin{cases} \partial_t \mu_t + \nabla \cdot (\underline{v}_t \mu_t) = 0 & \text{in the distributional sense of } (0,\infty) \times X, \\ \underline{v}_t(x) = \underline{F}(x, \mu_t). \end{cases}$$

Questions:

1. How to provide a quantitative description of weak convergence.
2. How to provide a quantitative description of continuity of velocity vector fields and study the stability of the system with respect to the initial datum.
3. How to deal with “discontinuous” velocity vector fields.

Evolution of discrete measures

$\mu = \frac{1}{N}\sum_\omega \delta_{X_\omega}$, $\nu = \frac{1}{N}\sum_\omega \delta_{y_\omega}$.

$$\boxed{W_2^2(\mu,\nu) := \min_{\sigma\in\mathrm{Sym}(\Omega)} \frac{1}{N}\sum_\omega |X_\omega - y_{\sigma(\omega)}|^2}$$

W_2 **is a distance between discrete measures.**

$\mu = \frac{1}{N}\sum_{\omega} \delta_{X_\omega}$, $\nu = \frac{1}{N}\sum_{\omega} \delta_{y_\omega}$.

$$\boxed{W_2^2(\mu,\nu) := \min_{\sigma \in \mathrm{Sym}(\Omega)} \frac{1}{N}\sum_{\omega} |X_\omega - y_{\sigma(\omega)}|^2}$$

W_2 is a distance between discrete measures.

Permutations $\leftrightarrow$ couplings: $\gamma_\sigma = \frac{1}{N}\sum_{\omega} \delta_{(X_\omega, y_{\sigma(\omega)})} \in \Gamma^N(\mu,\nu) \subset \mathcal{P}(H \times H)$,

$$\frac{1}{N}\sum_{\omega} |X_\omega - y_{\sigma(\omega)}|^2 = \int |x - y|^2 \, d\gamma$$

Evolution of discrete measures

$\mu = \frac{1}{N}\sum_\omega \delta_{X_\omega}$, $\nu = \frac{1}{N}\sum_\omega \delta_{y_\omega}$.

$$\boxed{W_2^2(\mu,\nu) := \min_{\sigma\in \mathrm{Sym}(\Omega)} \frac{1}{N}\sum_\omega |X_\omega - y_{\sigma(\omega)}|^2}$$

W_2 is a distance between discrete measures.

Permutations $\leftrightarrow$ couplings: $\gamma_\sigma = \frac{1}{N}\sum_\omega \delta_{(X_\omega, y_{\sigma(\omega)})} \in \Gamma^N(\mu,\nu) \subset \mathcal{P}(H\times H)$,

$$\frac{1}{N}\sum_\omega |X_\omega - y_{\sigma(\omega)}|^2 = \int |x-y|^2\, d\gamma$$

Optimal couplings:

$$\Gamma_o(\mu,\nu) := \left\{\gamma \in \Gamma^N(\mu,\nu) : W_2^2(\mu,\nu) = \int |x-y|^2\, d\gamma\right\}$$

Outline

Intrinsic characterization of curves

Let $t \mapsto \mu_t$, $t \in [0, T]$, be dscrete measures (with N particles).

Theorem

The following properties are equivalent:

1. $(\mu_t)_{t\in[0,T]}$ *satisfy the Lipschitz condition:*

$$\boxed{W_2(\mu_s, \mu_t) \leqslant L|t-s| \quad \textit{for every } s, t \in [0, T].}$$

Intrinsic characterization of curves

Let $t \mapsto \mu_t$, $t \in [0,T]$, be dscrete measures (with N particles).

Theorem

The following properties are equivalent:

1. $(\mu_t)_{t\in[0,T]}$ *satisfy the Lipschitz condition:*

$$\boxed{W_2(\mu_s,\mu_t) \leqslant L|t-s| \quad \textit{for every } s,t \in [0,T].}$$

2. *There exists a family of vector measures* $\boldsymbol{\nu}_t = \underline{v}_t \mu_t$, $t \in (0,T)$, *such that* $\int |\underline{v}_t|^2 \, d\mu_t \leqslant L^2$ *a.e. in* $(0,T)$ *and*

$$\boxed{\partial_t \mu_t + \nabla \cdot \boldsymbol{\nu}_t = 0 \quad \textit{in the distributional sense of } (0,T) \times X.}$$

Intrinsic characterization of curves

Let $t \mapsto \mu_t$, $t \in [0, T]$, be dscrete measures (with N particles).

Theorem

The following properties are equivalent:

1. $(\mu_t)_{t\in[0,T]}$ *satisfy the Lipschitz condition:*

$$\boxed{W_2(\mu_s, \mu_t) \leqslant L|t-s| \quad \textit{for every } s, t \in [0, T].}$$

2. *There exists a family of vector measures* $\boldsymbol{\nu}_t = \underline{v}_t \mu_t$, $t \in (0, T)$, *such that* $\int |\underline{v}_t|^2 \, d\mu_t \leqslant L^2$ *a.e. in* $(0, T)$ *and*

$$\boxed{\partial_t \mu_t + \nabla \cdot \boldsymbol{\nu}_t = 0 \quad \textit{in the distributional sense of } (0, T) \times X.}$$

3. *There exist Lipschitz maps* $X_\omega : [0, T] \to X$ *such that* $\mu_t = \frac{1}{N} \sum_{\omega\in\Omega} \delta_{X_\omega(t)}$ *and*

$$\boxed{\frac{1}{N} \sum_{\omega} |\dot{X}_\omega(t)|^2 = |\dot{\mu}_t|_W^2 := \lim_{h\to 0} \left(\frac{W_2(\mu_{t+h}, \mu_t)}{|h|}\right)^2 \leqslant L^2 \quad \textit{a.e. in (0,T).}}$$

The derivative of the Wasserstein distance

Let $\mu_t = \frac{1}{N}\sum_\omega \delta_{X_\omega(t)}$ be a Lipschitz map of discrete measures and let $\nu = \frac{1}{N}\sum_\omega \delta_{Y_\omega}$ be another discrete measure.

For every $t \in [0,T]$ let us select an **optimal coupling** $\gamma_t \in \Gamma_o(\mu_t,\nu)$ associated to an **optimal permutation** σ_t. For a.e. $t \in (0,T)$ we have

$$\frac{d}{dt}\frac{1}{2}W_2^2(\mu_t,\nu) = \frac{1}{N}\sum_\omega \dot{X}_\omega(t)(X_\omega(t) - Y_{\sigma_t(\omega)}) = \boxed{\int \underline{v}_t(x)\cdot(x-y)\,d\gamma_t(x,y).}$$

Let $\mu_t = \frac{1}{N}\sum_\omega \delta_{X_\omega(t)}$ be a Lipschitz map of discrete measures and let $\nu = \frac{1}{N}\sum_\omega \delta_{Y_\omega}$ be another discrete measure.

For every $t \in [0, T]$ let us select an **optimal coupling** $\gamma_t \in \Gamma_o(\mu_t, \nu)$ associated to an **optimal permutation** σ_t. For a.e. $t \in (0, T)$ we have

$$\frac{d}{dt}\frac{1}{2}W_2^2(\mu_t, \nu) = \frac{1}{N}\sum_\omega \dot{X}_\omega(t)(X_\omega(t) - Y_{\sigma_t(\omega)}) = \boxed{\int \underline{v}_t(x)\cdot(x-y)\,d\gamma_t(x,y).}$$

Simlarly, if $\nu_t = \frac{1}{N}\sum_\omega \delta_{Y_\omega(t)}$ is another Lipschitz curve with velocity vector $\underline{w}_t$ we have

$$\begin{aligned}\frac{d}{dt}\frac{1}{2}W_2^2(\mu_t, \nu_t) &= \frac{1}{N}\sum_\omega (\dot{X}_\omega(t) - \dot{Y}_{\sigma_t(\omega)}(t))(X_\omega(t) - Y_{\sigma_t(\omega)}(t)) \\ &= \boxed{\int (\underline{v}_t(x) - \underline{w}_t(y))\cdot(x-y)\,d\gamma_t(x,y).}\end{aligned}$$

Dissipative fields

Suppose that $(\mu_t)_{t\in[0,T]}, (\nu_t)_{t\in[0,T]}$ are solutions of the evolution system

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

$$\partial_t \nu_t + \nabla \cdot (\underline{w}_t \nu_t) = 0, \quad \underline{w}_t(x) = \underline{F}(x, \nu_t)$$

The system is dissipative if the distance between two solutions is not increasing w.r.t. time.

Dissipative fields

Suppose that $(\mu_t)_{t\in[0,T]}, (\nu_t)_{t\in[0,T]}$ are solutions of the evolution system

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

The system is dissipative if the distance between two solutions is not increasing w.r.t. time.

This is equivalent to say that for every μ,ν every optimal permutation $\sigma \in \mathrm{Sym}(\Omega)$ and corresponding $\gamma \in \Gamma_o(\mu,\nu)$

$$\frac{1}{N}\sum_{\omega}(\underline{F}(X_\omega,\mu) - \underline{F}(Y_{\sigma(\omega)},\nu))(X_\omega - Y_{\sigma(\omega)})$$
$$= \boxed{\int(\underline{F}(x,\mu) - \underline{F}(y,\nu))\cdot(x-y)\,d\gamma(x,y) \leqslant 0}$$

Dissipative fields

Suppose that $(\mu_t)_{t\in[0,T]}, (\nu_t)_{t\in[0,T]}$ are solutions of the evolution system

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

The system is dissipative if the distance between two solutions is not increasing w.r.t. time.

This is equivalent to say that for every μ, ν every optimal permutation $\sigma \in \mathrm{Sym}(\Omega)$ and corresponding $\gamma \in \Gamma_o(\mu, \nu)$

$$\frac{1}{N} \sum_{\omega} (\underline{F}(X_\omega, \mu) - \underline{F}(Y_{\sigma(\omega)}, \nu))(X_\omega - Y_{\sigma(\omega)})$$
$$= \boxed{\int (\underline{F}(x, \mu) - \underline{F}(y, \nu)) \cdot (x - y)\, d\gamma(x, y) \leqslant 0}$$

More generally, λ-dissipativity

$$\boxed{\int (\underline{F}(x, \mu) - \underline{F}(y, \nu)) \cdot (x - y)\, d\gamma(x, y) \leqslant \lambda \int |x - y|^2\, d\gamma = \lambda W_2^2(\mu, \nu)}$$

corresponds to the exponential behaviour

$$\boxed{W_2^2(\mu_t, \nu_t) \leqslant e^{2\lambda t} W_2^2(\mu_0, \nu_0) \quad \text{for every } t \geqslant 0.}$$

Dissipative fields and geodesically convex functionals

Displacement interpolation: given $\vartheta \in [0, 1]$, $\sigma \in \Gamma_o(\mu, \nu)$

$$\mu_\vartheta := \frac{1}{N}\sum_\omega \delta_{X_\omega(\vartheta)}, \quad X_\omega(\vartheta) := (1-\vartheta)X_\omega + \vartheta y_{\sigma(\omega)}.$$

$\vartheta \mapsto \mu_\vartheta$ is a constant speed, minimal geodesic connecting μ to ν:

$$\boxed{W_2(\mu_{\vartheta_1}, \mu_{\vartheta_2}) = |\vartheta_1 - \vartheta_2| W_2(\mu, \nu) \quad \text{for every } \vartheta_1, \vartheta_2 \in [0, 1].}$$

Dissipative fields and geodesically convex functionals

Displacement interpolation: given $\vartheta \in [0,1]$, $\sigma \in \Gamma_o(\mu,\nu)$

$$\mu_\vartheta := \frac{1}{N}\sum_\omega \delta_{X_\omega(\vartheta)}, \quad X_\omega(\vartheta) := (1-\vartheta)X_\omega + \vartheta y_{\sigma(\omega)}.$$

$\vartheta \mapsto \mu_\vartheta$ is a constant speed, minimal geodesic connecting μ to ν:

$$\boxed{W_2(\mu_{\vartheta_1}, \mu_{\vartheta_2}) = |\vartheta_1 - \vartheta_2| W_2(\mu,\nu) \quad \text{for every } \vartheta_1, \vartheta_2 \in [0,1].}$$

A functional $\mathscr{F}$ defined on discrete measures,

$$\mathscr{F}(\mu) = \Phi(\mathbf{X}) \quad \text{if } \mu = \frac{1}{N}\sum_\omega \delta_{X_\omega}$$

is displacement convex (or, equivalently, geodesically convex) if

$$\boxed{\mathscr{F}(\mu_\vartheta) \leqslant (1-\vartheta)\mathscr{F}(\mu) + \vartheta\mathscr{F}(\nu).}$$

In this case the field $\underline{F}$ satisfying $\underline{F}(X_\omega, \frac{1}{N}\sum_\omega \delta_{X_\omega}) = -\partial_{X_\omega}\Phi(\frac{1}{N}\sum_\omega \delta_{X_\omega})$ is dissipative.

Explicit Euler method

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 $\bar{\mu} = \frac{1}{N}\sum_\omega \delta_{\bar{X}_\omega}$, and we consider themeasures $\mu^{\tau,n} = \frac{1}{N}\sum_\omega \delta_{X_\omega^{\tau,n}}$ where the sequence of maps $(\mathbf{X}^{\tau,n})_{n\in\mathbb{N}}$ is defined by induction:

$$\boxed{X_\omega^{\tau,0} = \bar{X}_\omega, \quad X_\omega^{\tau,n+1} = X_\omega^{\tau,n} + \tau F(X_\omega^{\tau,n}, \mathbf{X}^{\tau,n})}$$

Explicit Euler method

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 $\bar{\mu} = \frac{1}{N}\sum_\omega \delta_{\bar{X}_\omega}$, and we consider themeasures $\mu^{\tau,n} = \frac{1}{N}\sum_\omega \delta_{X_\omega^{\tau,n}}$ where the sequence of maps $(\mathbf{X}^{\tau,n})_{n\in\mathbb{N}}$ is defined by induction:

$$\boxed{X_\omega^{\tau,0} = \bar{X}_\omega, \quad X_\omega^{\tau,n+1} = X_\omega^{\tau,n} + \tau F(X_\omega^{\tau,n}, \mathbf{X}^{\tau,n})}$$

In term of measures, this is equivalent to

$$\boxed{\mu^{\tau,n+1} = (E^{\tau,n})_\sharp \mu^{\tau,n}, \quad E^{\tau,n}(x) := x + \tau \underline{F}(x, \mu^{n,\tau})}$$

Explicit Euler method

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 $\bar{\mu} = \frac{1}{N}\sum_\omega \delta_{\bar{X}_\omega}$, and we consider themeasures $\mu^{\tau,n} = \frac{1}{N}\sum_\omega \delta_{X^{\tau,n}_\omega}$ where the sequence of maps $(\mathbf{X}^{\tau,n})_{n\in\mathbb{N}}$ is defined by induction:

$$\boxed{X^{\tau,0}_\omega = \bar{X}_\omega, \quad X^{\tau,n+1}_\omega = X^{\tau,n}_\omega + \tau F(X^{\tau,n}_\omega, \mathbf{X}^{\tau,n})}$$

In term of measures, this is equivalent to

$$\boxed{\mu^{\tau,n+1} = (E^{\tau,n})_\sharp \mu^{\tau,n}, \quad E^{\tau,n}(x) := x + \tau \underline{F}(x, \mu^{n,\tau})}$$

If $\mu^\tau_t := \mu^{\tau,\lfloor t/\tau \rfloor}$ is the **piecewise constant interpolation,** $\mu^\tau_t = \mu^{\tau,n}$ if $n\tau \leqslant t < (n+1)\tau$, we get $\lim_{\tau\downarrow 0} \mu^\tau_t = \mu_t$ solving

$$\partial_t \mu_t + \nabla \cdot (\underline{v}_t \mu_t) = 0, \quad \underline{v}_t(x) = \underline{F}(x, \mu_t), \quad \mu_0 = \bar{\mu}.$$

Implicit Euler method

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.

In this case at each step one has to solve the equation

$$\boxed{(E^{-\tau,n+1})_\sharp \mu^{\tau,n+1} = \mu^{\tau,n}, \quad E^{-\tau,n+1}(x) := x - \tau \underline{F}(x, \mu^{n+1,\tau})}$$

[4] 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

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.

In this case at each step one has to solve the equation

$$\boxed{(E^{-\tau,n+1})_\sharp \mu^{\tau,n+1} = \mu^{\tau,n}, \quad E^{-\tau,n+1}(x) := x - \tau \underline{F}(x, \mu^{n+1,\tau})}$$

According to the JKO -Minimizing Movement approach

[4], it is sufficient to select $\mu^{\tau,n+1}$ among the minimizers of

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

[4] 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),

Diffusion PDE's

The implicit Euler method is well adapted to deal with important example arising in diffusion equations [5] as the Fokker-Planck equation

$$\boxed{\partial_t \mu_t + \nabla \cdot (\underline{v}_t \mu_t) = 0, \quad \underline{v}_t = -\nabla(\log u_t + V), \quad \mu_t = u_t \mathscr{L}^d \ll \mathscr{L}^d}$$

[5] R. Jordan, D. Kinderlehrer, F. Otto, The variational formulation of the Fokker-Planck equation. SIAM J. Math. Anal. (1998)

The implicit Euler method is well adapted to deal with important example arising in diffusion equations [5] as the Fokker-Planck equation

$$\boxed{\partial_t \mu_t + \nabla \cdot (\underline{v}_t \mu_t) = 0, \quad \underline{v}_t = -\nabla(\log u_t + V), \quad \mu_t = u_t \mathscr{L}^d \ll \mathscr{L}^d}$$

which can be interpreted as the gradient flow of the Relative Entropy functional

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

[5] R. Jordan, D. Kinderlehrer, F. Otto, The variational formulation of the Fokker-Planck equation. SIAM J. Math. Anal. (1998)

Outline

Arbitrary probability measures

We can extend the notion of (Kantorovich-Rubinstein-)Wasserstein distance to arbitrary probability measures with finite quadratic moment:

$$\mu \in \mathcal{P}_2(H), \quad \int |x|^2 \, d\mu(x) < +\infty$$

The role of couplings associated to permutations is now replaced by **arbitrary probability measures** $\gamma \in \mathcal{P}(H \times H)$ **with assigned marginals** μ **and** ν**:**

$$\Gamma(\mu, \nu) := \Big\{ \gamma \in \mathcal{P}(H \times H) : \pi^1_\sharp(\gamma) = \mu, \ \pi^2_\sharp(\gamma) = \nu \Big\}$$

where $\pi^i(x_1, x_2) := x_i$, $i = 1, 2$.

$$\boxed{W_2^2(\mu, \nu) := \min \Big\{ \int |x_1 - x_2|^2 \, d\gamma(x_1, x_2) : \gamma \in \Gamma(\mu, \nu) \Big\}}$$

Arbitrary probability measures

We can extend the notion of (Kantorovich-Rubinstein-)Wasserstein distance to arbitrary probability measures with finite quadratic moment:

$$\mu \in \mathcal{P}_2(H), \quad \int |x|^2 \, d\mu(x) < +\infty$$

The role of couplings associated to permutations is now replaced by **arbitrary probability measures** $\gamma \in \mathcal{P}(H \times H)$ **with assigned marginals** μ **and** ν**:**

$$\Gamma(\mu,\nu) := \Big\{\gamma \in \mathcal{P}(H \times H) : \pi^1_\sharp(\gamma) = \mu, \ \pi^2_\sharp(\gamma) = \nu\Big\}$$

where $\pi^i(x_1, x_2) := x_i$, $i = 1, 2$.

$$\boxed{W_2^2(\mu,\nu) := \min\Big\{\int |x_1 - x_2|^2 \, d\gamma(x_1, x_2) : \gamma \in \Gamma(\mu,\nu)\Big\}}$$

Optimal couplings:

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

Arbitrary probability measures

We can extend the notion of (Kantorovich-Rubinstein-)Wasserstein distance to arbitrary probability measures with finite quadratic moment:

$$\mu \in \mathcal{P}_2(H), \quad \int |x|^2 \, d\mu(x) < +\infty$$

The role of couplings associated to permutations is now replaced by **arbitrary probability measures** $\gamma \in \mathcal{P}(H \times H)$ **with assigned marginals** μ **and** ν**:**

$$\Gamma(\mu, \nu) := \Big\{ \gamma \in \mathcal{P}(H \times H) : \pi^1_\sharp(\gamma) = \mu, \ \pi^2_\sharp(\gamma) = \nu \Big\}$$

where $\pi^i(x_1, x_2) := x_i$, $i = 1, 2$.

$$\boxed{W_2^2(\mu, \nu) := \min \Big\{ \int |x_1 - x_2|^2 \, d\gamma(x_1, x_2) : \gamma \in \Gamma(\mu, \nu) \Big\}}$$

Optimal couplings:

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

$(\mathcal{P}_2(H), W_2)$ is a **complete and separable metric space** and the collection of discrete measures is a **dense** subset of $\mathcal{P}_2(H)$ (it can be considered as the natural metric completion of set of discrete measures).

Coupling and permutations

Couplings allow for splitting of mass: if $\mu = \frac{1}{N}\sum_\omega \delta_{X_\omega}$, $\nu = \frac{1}{N}\sum_\omega \delta_{Y_\omega}$ then an element of $\Gamma(\mu,\nu)$ can be expressed as

$$\gamma = \sum_{\omega,\eta\in\Omega} m_{\omega,\eta}\delta_{(X_\omega,Y_\eta)}, \quad \sum_\eta m_{\omega,\eta} = \sum_\omega m_{\omega,\eta} = 1/N.$$

and it is a remarkable consequence of Birkhoff theorem that for discrete measures

$$W_2^2(\mu,\nu) = \min_{\sigma\in \mathrm{Sym}(\Omega)} \frac{1}{N}\sum_\omega |X_\omega - Y_{\sigma(\omega)}|^2$$

Let now $\Omega := AC^2([0,T];H)$, the space of **absolutely continuous curves** $\omega : [0,T] \to H$ such that $\int_0^T |\dot\omega(t)|^2\,dt < \infty$ endowed with the topology of uniform convergence. For every $t \in [0,T]$ we consider the evaluation maps $X(t) : \Omega \to H$, $X(t,\omega) := \omega(t)$.

Theorem

The following properties are equivalent:

1. $(\mu_t)_{t\in[0,T]}$ *is* **Lipschitz:** $\quad W_2(\mu_s,\mu_t) \leqslant L|t-s| \quad$ *for every* $s,t \in [0,T]$.

Lipschitz curves in $\mathcal{P}_2(H)$

Let now $\Omega := AC^2([0,T];H)$, the space of **absolutely continuous curves** $\omega : [0,T] \to H$ such that $\int_0^T |\dot\omega(t)|^2\,dt < \infty$ endowed with the topology of uniform convergence. For every $t \in [0,T]$ we consider the evaluation maps $X(t) : \Omega \to H$, $X(t,\omega) := \omega(t)$.

Theorem

The following properties are equivalent:

1. $(\mu_t)_{t\in[0,T]}$ *is* **Lipschitz:** $\quad W_2(\mu_s,\mu_t) \leqslant L|t-s| \quad$ *for every* $s,t \in [0,T]$.
2. *There exists a* **Borel vector field** $\underline{v} : (0,T)\times H \to H$ *such that*

$$\int |\underline{v}_t|^2\,d\mu_t \leqslant L^2 \text{ a.e. in } (0,T) \text{ and } \quad \partial_t\mu_t + \nabla\cdot(\underline{v}_t\mu_t) = 0 \quad \text{in } \mathscr{D}'((0,T)\times H).$$

Let now $\Omega := AC^2([0,T];H)$, the space of **absolutely continuous curves** $\omega : [0,T] \to H$ such that $\int_0^T |\dot\omega(t)|^2\,dt < \infty$ endowed with the topology of uniform convergence. For every $t \in [0,T]$ we consider the evaluation maps $X(t) : \Omega \to H$, $X(t,\omega) := \omega(t)$.

Theorem

The following properties are equivalent:

1. $(\mu_t)_{t\in[0,T]}$ *is* **Lipschitz:** $W_2(\mu_s,\mu_t) \leqslant L|t-s|$ *for every* $s,t \in [0,T]$.
2. *There exists a* **Borel vector field** $\underline{v} : (0,T) \times H \to H$ *such that*

$$\int |\underline{v}_t|^2\,d\mu_t \leqslant L^2 \text{ a.e. in } (0,T) \text{ and } \quad \partial_t\mu_t + \nabla\cdot(\underline{v}_t\mu_t) = 0 \quad \text{in } \mathscr{D}'((0,T)\times H).$$

3. *There exist a* **Radon probability measure** $\mathbb{P}$ *on* Ω *such that for every* $t \in [0,T]$ $\mu_t = X(t)_\sharp\mathbb{P}$ *and*

$$\int |\dot X(t,\omega)|^2\,d\mathbb{P}(\omega) = |\dot\mu_t|^2_W := \lim_{h\to 0}\left(\frac{W_2(\mu_{t+h},\mu_t)}{|h|}\right)^2 \leqslant L^2 \quad \text{a.e. in } (0,T).$$

Moreover $\dot X(t,\omega) = \underline{v}_t(X(t,\omega))$ *for* $\mathbb{P}$*-a.e.* ω.

Variational characterization of the Wasserstein velocity field

If $(\mu_t)_{t\in[0,T]}$ is Lipschitz: $\quad W_2(\mu_s,\mu_t) \leqslant L|t-s| \quad$ for every $s,t\in[0,T]$ and $\underline{v}:(0,T)\times H\to H$ is a Borel vector field such that

$$\int_0^T\int |\underline{v}_t|^2\,d\mu_t\,dt<\infty \text{ and } \quad \partial_t\mu_t+\nabla\cdot(\underline{v}_t\mu_t)=0 \quad \text{in } \mathscr{D}'((0,T)\times H),$$

then $\underline{v}$ satisfies the **optimal condition**

$$\boxed{\int |\underline{v}_t|^2\,d\mu_t=|\dot\mu_t|_W^2=\lim_{h\to 0}\left(\frac{W_2(\mu_{t+h},\mu_t)}{|h|}\right)^2 \quad \text{a.e. in } (0,T)}$$

if and only if

$$\boxed{\underline{v}_t\in \mathrm{Tan}_{\mu_t}(\mathcal{P}_2(H))=\overline{\left\{\nabla\varphi:\varphi\in C_c^\infty(H)\right\}}^{L^2(\mu_t;H)} \qquad \text{for a.e. } t\in(0,T).}$$

Variational characterization of the Wasserstein velocity field

If $(\mu_t)_{t\in[0,T]}$ is Lipschitz: $W_2(\mu_s,\mu_t) \leqslant L|t-s|$ for every $s,t\in[0,T]$ and $\underline{v}:(0,T)\times H\to H$ is a Borel vector field such that

$$\int_0^T\int |\underline{v}_t|^2\,d\mu_t\,dt<\infty \text{ and } \quad \partial_t\mu_t+\nabla\cdot(\underline{v}_t\mu_t)=0 \quad \text{in } \mathscr{D}'((0,T)\times H),$$

then $\underline{v}$ satisfies the **optimal condition**

$$\boxed{\int |\underline{v}_t|^2\,d\mu_t = |\dot\mu_t|^2_W = \lim_{h\to 0}\Big(\frac{W_2(\mu_{t+h},\mu_t)}{|h|}\Big)^2 \quad \text{a.e. in } (0,T)}$$

if and only if

$$\boxed{\underline{v}_t\in \mathrm{Tan}_{\mu_t}(\mathcal{P}_2(H)) = \overline{\Big\{\nabla\varphi:\varphi\in C^\infty_c(H)\Big\}}^{L^2(\mu_t;H)} \qquad \text{for a.e. } t\in(0,T).}$$

Moreover for almost every $t\in(0,T)$ we have

$$W_2(\mu_{t+h},(I+h\underline{v}_t)_\sharp\mu_t)=o(h) \quad \text{as } h\to 0.$$

"Splitting times" are negligible in $(0,T)$.

The derivative of the Wasserstein distance

Let $(\mu_t)_{t\in[0,T]}$ be a Lipschitz curve in $\mathcal{P}_2(H)$ let $\nu \in \mathcal{P}_2(H)$. For every $t \in [0,T]$ let us select an optimal coupling $\gamma_t \in \Gamma_o(\mu_t, \nu)$. For a.e. $t \in (0,T)$ we have

$$\boxed{\frac{d}{dt}\frac{1}{2}W_2^2(\mu_t,\nu) = \int \underline{v}_t(x)\cdot(x-y)\,d\gamma_t(x,y).}$$

The derivative of the Wasserstein distance

Let $(\mu_t)_{t\in[0,T]}$ be a Lipschitz curve in $\mathcal{P}_2(H)$ let $\nu \in \mathcal{P}_2(H)$. For every $t \in [0, T]$ let us select an optimal coupling $\gamma_t \in \Gamma_o(\mu_t, \nu)$. For a.e. $t \in (0, T)$ we have

$$\boxed{\frac{d}{dt}\frac{1}{2}W_2^2(\mu_t, \nu) = \int \underline{v}_t(x) \cdot (x - y)\, d\gamma_t(x, y).}$$

Simlarly, if $(\nu_t)_{t\in[0,T]}$ is another Lipschitz curve with velocity vector field $\underline{w}_t$ we have

$$\boxed{\frac{d}{dt}\frac{1}{2}W_2^2(\mu_t, \nu_t) = \int (\underline{v}_t(x) - \underline{w}_t(y)) \cdot (x - y)\, d\gamma_t(x, y).}$$

A metric definition of a gradient flow

If $u(t)$ is a solution of the gradient flow in $\mathbb{R}^d$

$$\dot{u}(t) = -\nabla\phi(u(t)) \tag{GF}$$

where $\phi : \mathbb{R}^d \to \mathbb{R}$ is a smooth convex function, then for every test vector $w \in \mathbb{R}^d$ we have

$$\frac{d}{dt}\frac{1}{2}|u(t) - w|^2 = \langle \dot{u}(t), u(t) - w\rangle = \langle \nabla\phi(u(t)), w - u(t)\rangle \leqslant \phi(w) - \phi(u(t))$$

where we used the subgradient inequality

$$\langle \nabla(u), w - u\rangle \leqslant \phi(w) - \phi(u).$$

A metric definition of a gradient flow

If $u(t)$ is a solution of the gradient flow in $\mathbb{R}^d$

$$\dot{u}(t) = -\nabla\phi(u(t)) \tag{GF}$$

where $\phi : \mathbb{R}^d \to \mathbb{R}$ is a smooth convex function, then for every test vector $w \in \mathbb{R}^d$ we have

$$\frac{d}{dt}\frac{1}{2}|u(t) - w|^2 = \langle \dot{u}(t), u(t) - w\rangle = \langle \nabla\phi(u(t)), w - u(t)\rangle \leqslant \phi(w) - \phi(u(t))$$

where we used the subgradient inequality

$$\langle \nabla(u), w - u\rangle \leqslant \phi(w) - \phi(u).$$

In fact, u solves (GF) if and only if

$$\frac{d}{dt}\frac{1}{2}|u(t) - w|^2 \leqslant \phi(w) - \phi(u(t)) \quad \text{for every } w \in \mathbb{R}^d \tag{EVI}$$

The Wasserstein case

Let $\mathscr{F} : \mathcal{P}_2(H) \to [0, +\infty]$ 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 $\nu \in D(\mathscr{F}) \subset \mathcal{P}_2(H)$

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

Let $\mathscr{F} : \mathcal{P}_2(H) \to [0, +\infty]$ 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 $\nu \in D(\mathscr{F}) \subset \mathcal{P}_2(H)$

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

Theorem

For every initial datum $\bar{\mu} \in D(\mathscr{F})$ *there exists a unique EVI solution to* (EVI) *satisfying* $\lim_{t \downarrow 0} \mu_t = \bar{\mu}$*. Moreover,* μ_t *is the uniform limit of the JKO-Miinimizing Movement approximations, obtained by solving*

$$\mu^{\tau,n+1} \in \operatorname*{argmin}_{\mu} \left\{ \frac{1}{2\tau} W_2^2(\mu, \mu^{\tau,n}) + \mathscr{F}(\mu) \right\}, \quad \mu^{\tau,0} := \bar{\mu}.$$

Let $\mathscr{F} : \mathcal{P}_2(H) \to [0, +\infty]$ 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 $\nu \in D(\mathscr{F}) \subset \mathcal{P}_2(H)$

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

Theorem

For every initial datum $\bar{\mu} \in D(\mathscr{F})$ there exists a unique EVI solution to (EVI) *satisfying* $\lim_{t\downarrow 0} \mu_t = \bar{\mu}$*. Moreover, μ_t is the uniform limit of the JKO-Miinimizing Movement approximations, obtained by solving*

$$\mu^{\tau,n+1} \in \operatorname*{argmin}_{\mu} \left\{ \frac{1}{2\tau} W_2^2(\mu, \mu^{\tau,n}) + \mathscr{F}(\mu) \right\}, \quad \mu^{\tau,0} := \bar{\mu}.$$

The EVI formulation is equivalent to the continuity equation

$$\partial_t \mu_t + \nabla \cdot (\underline{v}_t \mu_t) = 0$$

where the Wasserstein velocity vector field $\underline{v}$ satisfies for a.e. $t > 0$

$$\int \langle \underline{v}_t(x), x - y \rangle \, d\gamma_t(x, y) \leqslant \mathscr{F}(\nu) - \mathscr{F}(\mu_t) \quad \text{for every } \gamma_t \in \Gamma_o(\mu_t, \nu).$$

Outline

Dissipative vector fields in $\mathcal{P}_2(H)$

In $\mathcal{P}_2(H)$ a general dissipative vector field F can be represented by a map (possibly multivalued) from $\mathcal{P}_2(H)$ to $\mathcal{P}_2(H \times H)$:

$$F(\mu) \in \mathcal{P}_2(H;\mu) := \left\{ \underline{F} \in \mathcal{P}_2(H \times H) : \pi^1_\sharp \underline{F} = \mu \right\}.$$

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

In the "regular case" $\underline{F}_x$ 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 $H \times \mathcal{P}_2(H)$ into H.

Dissipative vector fields in $\mathcal{P}_2(H)$

In $\mathcal{P}_2(H)$ a general dissipative vector field F can be represented by a map (possibly multivalued) from $\mathcal{P}_2(H)$ to $\mathcal{P}_2(H \times H)$:

$$F(\mu) \in \mathcal{P}_2(H;\mu) := \Big\{ \underline{F} \in \mathcal{P}_2(H \times H) : \pi^1_\sharp \underline{F} = \mu \Big\}.$$

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

In the "regular case" $\underline{F}_x$ 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 $H \times \mathcal{P}_2(H)$ into H.

In the general case, we can allow for a **general probability measure** $\underline{F}_x$ **depending on** x. Denoting by (x, v) the variables in $H \times H$ and by $\exp^\tau(x, v) := x + \tau v$, F acts on μ by the family of maps

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

In $\mathcal{P}_2(\mathrm{H})$ a general dissipative vector field F can be represented by a map (possibly multivalued) from $\mathcal{P}_2(\mathrm{H})$ to $\mathcal{P}_2(\mathrm{H}\times\mathrm{H})$:

$$\mathrm{F}(\mu) \in \mathcal{P}_2(\mathrm{H};\mu) := \Big\{\underline{\mathrm{F}} \in \mathcal{P}_2(\mathrm{H}\times\mathrm{H}) : \pi^1_\sharp \underline{\mathrm{F}} = \mu\Big\}.$$

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

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

In the general case, we can allow for a **general probability measure** $\underline{\mathrm{F}}_x$ **depending on** x. Denoting by (x,v) the variables in $\mathrm{H}\times\mathrm{H}$ and by $\exp^\tau(x,v) := x+\tau v$, F acts on μ by the family of maps

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

Notice that in the regular case we have one step of the explicit Euler method:

$$\mathrm{F}^\tau(\mu) = \mathrm{E}^\tau_\sharp \mu, \quad \mathrm{E}^\tau(x) = x + \tau\underline{\mathrm{F}}(x,\mu).$$

Metric dissipativity

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

F is dissipative if

$$\boxed{\frac{d}{d\tau} W_2(F^\tau(\mu), F^\tau(\nu))\Big|_{\tau=0+} \leqslant 0 \quad \text{for every } \mu, \nu \in \mathcal{P}_2(H).}$$

In the regular case this condition reads as

$$\int \langle \underline{F}(x, \mu) - \underline{F}(y, \nu), x - y \rangle \, d\gamma(x, y) \leqslant 0 \quad \text{for some } \gamma \in \Gamma_o(\mu, \nu).$$

The explicit Euler method

If F is a dissipative vector field and $\tau > 0$ is a time step, we can construct the map $F^\tau : \mathcal{P}_2(H) \to \mathcal{P}_2(H)$

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

and therefore the sequence of explicit Euler approximations:

$$\mu^{\tau,0} := \bar{\mu} \text{ given}, \quad \mu^{\tau,n+1} := F^\tau(\mu^{\tau,n}), \quad \mu^\tau_t := \mu^{\tau,\lfloor t/\tau \rfloor}$$

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

A metric definition of dissipative evolutions

Let us now consider a solution $u(t)$ of the differential equation in $\mathbb{R}^d$

$$\dot{u}(t) = F(u(t)) \tag{DE}$$

where $F : \mathbb{R}^d \to \mathbb{R}^d$ is a dissipative vector field. For every vector $w \in \mathbb{R}^d$ we have

$$\begin{aligned}\frac{d}{dt}\frac{1}{2}|u(t) - w|^2 = \langle \dot{u}(t), u(t) - w\rangle &= \langle \dot{u}(t) - Fw, u(t) - w\rangle + \langle Fw, u(t) - w\rangle \\ &\leqslant \langle Fw, u(t) - w\rangle.\end{aligned}$$

where we used the **dissipativity condition**

$$\langle Fu - Fw, u - w\rangle \leqslant 0 \quad \text{for every } u, w \in \mathbb{R}^d.$$

[6] P. Bénilan, Solutions intégrales d'équations d'évolution dans un espace de Banach. C.R.A.S. (1972)

A metric definition of dissipative evolutions

Let us now consider a solution $u(t)$ of the differential equation in $\mathbb{R}^d$

$$\dot{u}(t) = F(u(t)) \tag{DE}$$

where $F : \mathbb{R}^d \to \mathbb{R}^d$ is a dissipative vector field. For every vector $w \in \mathbb{R}^d$ we have

$$\begin{aligned}\frac{d}{dt}\frac{1}{2}|u(t) - w|^2 &= \langle \dot{u}(t), u(t) - w\rangle = \langle \dot{u}(t) - Fw, u(t) - w\rangle + \langle Fw, u(t) - w\rangle \\ &\leqslant \langle Fw, u(t) - w\rangle.\end{aligned}$$

where we used the **dissipativity condition**

$$\langle Fu - Fw, u - w\rangle \leqslant 0 \quad \text{for every } u, w \in \mathbb{R}^d.$$

Thus u solves (DE) if and only if [6]

$$\frac{d}{dt}\frac{1}{2}|u(t) - w|^2 \leqslant \langle Fw, u(t) - w\rangle \quad \text{for every } w \in \mathbb{R}^d \tag{EVI}$$

<u>How can we interpret $\langle Fw, u(t) - w\rangle$</u> for probability vector fields?

[6] P. Bénilan, Solutions intégrales d'équations d'évolution dans un espace de Banach. C.R.A.S. (1972)

A pseudo-scalar product

Notice that $\langle Fw, u(t) - w \rangle$ can be expressed as

$$\langle Fw, w - u \rangle = \lim_{\tau \downarrow 0} \frac{1}{2\tau}\Big(|w + \tau Fw - u|^2 - |w - u|^2 \Big)$$

We can analogously define a **pseudo-scalar product**

$$\boxed{\langle F(\nu), \mu \rangle_- := \lim_{\tau \downarrow 0} \frac{1}{2\tau}\Big(W_2^2(F^\tau(\nu), \mu) - W_2^2(\nu, \mu) \Big)}$$

It is possibile to prove that the limit always exists and can be expressed by a suitable coupling between $F(\nu)$ and $\gamma \in \Gamma_o(\nu, \mu)$.

A pseudo-scalar product

Notice that $\langle Fw, u(t) - w\rangle$ can be expressed as

$$\langle Fw, w - u\rangle = \lim_{\tau\downarrow 0} \frac{1}{2\tau}\Big(|w + \tau Fw - u|^2 - |w - u|^2\Big)$$

We can analogously define a **pseudo-scalar product**

$$\boxed{\langle F(\nu), \mu\rangle_- := \lim_{\tau\downarrow 0} \frac{1}{2\tau}\Big(W_2^2(F^\tau(\nu), \mu) - W_2^2(\nu, \mu)\Big)}$$

It is possibile to prove that the limit always exists and can be expressed by a suitable coupling between $F(\nu)$ and $\gamma \in \Gamma_o(\nu, \mu)$.

Moreover, F **is dissipative if and only if**

$$\boxed{\langle F(\mu), \nu\rangle_- + \langle F(\nu), \mu\rangle_- \leqslant 0.}$$

Convergence

Let $F : \mathcal{P}_2(H) \to \mathcal{P}_2(H \times H)$ be a dissipative probability vector field satisfying the bound

$$\sup\left\{ \int |v|^2 \, dF(\mu) : \int |x|^2 \, d\mu \leqslant C \right\} < \infty \quad \text{for every } C > 0.$$

Theorem (Cavagnari-Sodini-S.)

- *For every $\bar{\mu} \in \mathcal{P}_2(H)$ the discrete solutions μ^τ of the explicit Euler method uniformly converge to a unique limit $\mu : [0, \infty) \to \mathcal{P}_2(H)$.*

Convergence

Let $F : \mathcal{P}_2(H) \to \mathcal{P}_2(H \times H)$ be a dissipative probability vector field satisfying the bound

$$\sup \left\{ \int |v|^2 \, dF(\mu) : \int |x|^2 \, d\mu \leqslant C \right\} < \infty \quad \text{for every } C > 0.$$

Theorem (Cavagnari-Sodini-S.)

- *For every $\bar{\mu} \in \mathcal{P}_2(H)$ the discrete solutions μ^τ of the explicit Euler method uniformly converge to a unique limit $\mu : [0, \infty) \to \mathcal{P}_2(H)$.*
- *μ is a Lipschitz curve and it is the unique solution of the disspative EVI*

$$\boxed{\frac{d}{dt}\frac{1}{2}W_2^2(\mu_t, \nu) \leqslant -\langle F(\nu), \mu \rangle_- \quad \textit{for every } \nu \in \mathcal{P}_2(H), \textit{ a.e. in } (0, \infty).}$$

Convergence

Let $F : \mathcal{P}_2(H) \to \mathcal{P}_2(H \times H)$ be a dissipative probability vector field satisfying the bound

$$\sup \left\{ \int |v|^2 \, dF(\mu) : \int |x|^2 \, d\mu \leqslant C \right\} < \infty \quad \text{for every } C > 0.$$

Theorem (Cavagnari-Sodini-S.)

- *For every $\bar{\mu} \in \mathcal{P}_2(H)$ the discrete solutions μ^τ of the explicit Euler method uniformly converge to a unique limit $\mu : [0, \infty) \to \mathcal{P}_2(H)$.*
- *μ is a Lipschitz curve and it is the unique solution of the disspative EVI*

$$\boxed{\frac{d}{dt}\frac{1}{2}W_2^2(\mu_t, \nu) \leqslant -\langle F(\nu), \mu \rangle_- \quad \text{for every } \nu \in \mathcal{P}_2(H), \text{ a.e. in } (0, \infty).}$$

- *Moreover, for every $T > 0$ we have the optimal error estimate*

$$W_2(\mu_t, \mu_t^\tau) \leqslant C\sqrt{\tau} \quad \text{for every } t \in [0, T].$$