The De Giorgi variational principle for gradient flows: a direct approach

Giuseppe Savaré

Department of Decision Sciences, Bocconi University, Milan, Italy

Based on joint work with Alessandro Pinzi & Filippo Riva

Multiscale Stochastics, Patterns, and Analysis of Combinatorial Environments

Bocconi University, Milan, March 16–19, 2026

Funded by the European Union

European Research Council
Established by the European Commission

Plan of the talk

1. **Variational principles for evolution equations**

2. **De Giorgi's principle and main result**

3. **Three key arguments**
   - Relaxation and continuity equation
   - Von Neumann duality
   - Backward estimates and maximum principle for first order HJ equations

4. **The infinite-dimensional setting**

5. **Extensions and open problems**

Outline

Gradient flows and doubly nonlinear equations: a simple smooth model

Let $X = \mathbb{R}^d$, $\varphi\colon \mathbb{R}^d \to \mathbb{R}$ be a smooth energy function, $x_0 \in \mathbb{R}^d$ be an initial datum.

Gradient flow

$$\dot{x}(t) = -\nabla\varphi(x(t)), \qquad x(0) = x_0,$$

Gradient flows and doubly nonlinear equations: a simple smooth model

Let $X = \mathbb{R}^d$, $\varphi\colon \mathbb{R}^d \to \mathbb{R}$ be a smooth energy function, $x_0 \in \mathbb{R}^d$ be an initial datum.

Gradient flow

$$\dot{x}(t) = -\nabla\varphi(x(t)), \qquad x(0) = x_0,$$

More generally:

Doubly nonlinear equation

$$D\Psi(\dot{x}(t)) = -D\varphi(x(t)), \qquad x(0) = x_0,$$

where Ψ is a smooth convex *dissipation potential*, superlinear, $\Psi(0) = 0$.

Gradient flow: the choice $\Psi(v) = \frac{1}{2}|v|^2$ (Hilbert setting), $D\Psi = I$.

Gradient flows and doubly nonlinear equations: a simple smooth model

Let $X = \mathbb{R}^d$, $\varphi\colon \mathbb{R}^d \to \mathbb{R}$ be a smooth energy function, $x_0 \in \mathbb{R}^d$ be an initial datum.

Gradient flow

$$\dot{x}(t) = -\nabla\varphi(x(t)), \qquad x(0) = x_0,$$

More generally:

Doubly nonlinear equation

$$D\Psi(\dot{x}(t)) = -D\varphi(x(t)), \qquad x(0) = x_0,$$

where Ψ is a smooth convex *dissipation potential*, superlinear, $\Psi(0) = 0$.

Gradient flow: the choice $\Psi(v) = \frac{1}{2}|v|^2$ (Hilbert setting), $D\Psi = I$.

If $D\Psi$ is invertible ($\rightsquigarrow$ superlinearity of Ψ):

$$\dot{x}(t) = (D\Psi)^{-1}\big(-D\varphi(x(t))\big).$$

Simple finite dimensional, smooth setting: **existence by Cauchy-Lipschitz Theorem.**

Gradient flows and doubly nonlinear equations: a simple smooth model

Let $X = \mathbb{R}^d$, $\varphi\colon \mathbb{R}^d \to \mathbb{R}$ be a smooth energy function, $x_0 \in \mathbb{R}^d$ be an initial datum.

Gradient flow

$$\dot{x}(t) = -\nabla\varphi(x(t)), \qquad x(0) = x_0,$$

More generally:

Doubly nonlinear equation

$$D\Psi(\dot{x}(t)) = -D\varphi(x(t)), \qquad x(0) = x_0,$$

where Ψ is a smooth convex *dissipation potential*, superlinear, $\Psi(0) = 0$.

Gradient flow: the choice $\Psi(v) = \frac{1}{2}|v|^2$ (Hilbert setting), $D\Psi = I$.

If $D\Psi$ is invertible ($\rightsquigarrow$ superlinearity of Ψ):

$$\dot{x}(t) = (D\Psi)^{-1}\big(-D\varphi(x(t))\big).$$

Simple finite dimensional, smooth setting: **existence by Cauchy-Lipschitz Theorem.**

Key question: how to prove existence of solutions via *global-in-time variational methods*?

The Brezis–Ekeland principle (1976)

For **convex, superlinear** φ in a **Euclidean-Hilbert** space.

The Brezis–Ekeland principle (1976)

For **convex, superlinear** φ in a **Euclidean-Hilbert** space.

Fenchel conjugate: $\Phi^*(v) := \sup_x \{\langle v, x\rangle - \Phi(x)\}$.

Fenchel inequality: $\Phi(x) + \Phi^*(v) \geqslant \langle v, x\rangle$,

$$\text{equality} \quad \Leftrightarrow \quad v = D\Phi(x) \quad \Leftrightarrow \quad x = D\Phi^*(v) = (D\Phi)^{-1}(v)$$

The Brezis–Ekeland principle (1976)

For **convex, superlinear** φ in a **Euclidean-Hilbert** space.

Fenchel conjugate: $\Phi^*(v) := \sup_x \left\{ \langle v, x \rangle - \Phi(x) \right\}$.

Fenchel inequality: $\Phi(x) + \Phi^*(v) \geqslant \langle v, x \rangle$,

$$\text{equality} \quad \Leftrightarrow \quad v = D\Phi(x) \quad \Leftrightarrow \quad x = D\Phi^*(v) = (D\Phi)^{-1}(v)$$

Examples:

- $\Phi(x) = \frac{1}{2}|x|^2 \Rightarrow \Phi^*(v) = \frac{1}{2}|v|^2$.
- $\Phi(x) = \frac{1}{p}\|x\|^p \Rightarrow \Phi^*(v) = \frac{1}{q}\|v\|_*^q, \quad \frac{1}{p} + \frac{1}{q} = 1$.
- $\Phi(x) = x \log x - x \Rightarrow \Phi^*(v) = e^v$.

The Brezis–Ekeland principle (1976)

For **convex, superlinear** φ in a **Euclidean-Hilbert** space.

Fenchel conjugate: $\Phi^*(v) := \sup_x \left\{\langle v, x\rangle - \Phi(x)\right\}$.

Fenchel inequality: $\Phi(x) + \Phi^*(v) \geqslant \langle v, x\rangle$,

$$\text{equality} \quad \Leftrightarrow \quad v = D\Phi(x) \quad \Leftrightarrow \quad x = D\Phi^*(v) = (D\Phi)^{-1}(v)$$

Examples:

- $\Phi(x) = \frac{1}{2}|x|^2 \Rightarrow \Phi^*(v) = \frac{1}{2}|v|^2$.
- $\Phi(x) = \frac{1}{p}\|x\|^p \Rightarrow \Phi^*(v) = \frac{1}{q}\|v\|_*^q, \qquad \frac{1}{p} + \frac{1}{q} = 1$.
- $\Phi(x) = x \log x - x \Rightarrow \Phi^*(v) = e^v$.

Rewriting the gradient flow: $\dot{x} = -D\varphi(x) \iff -\dot{x} = D\varphi(x) \iff x = D\varphi^*(-\dot{x})$, i.e.

$$-D\varphi^*(-\dot{x}(t)) = -x(t) \quad \iff \quad D\Psi(\dot{x}(t)) = -D\Phi(x(t))$$

with $\Psi(v) := \varphi^*(-v)$ and $\Phi(x) := \frac{1}{2}|x|^2$.

The Brezis–Ekeland functional

Brezis–Ekeland functional

$$\mathcal{BE}(T, x) := \frac{1}{2}|x(T)|^2 + \int_0^T \Big(\varphi^*(-\dot{x}(\tau)) + \varphi(x(\tau))\Big)\, d\tau - \frac{1}{2}|x_0|^2.$$

The Brezis–Ekeland functional

Brezis–Ekeland functional

$$\mathcal{BE}(T, x) := \frac{1}{2}|x(T)|^2 + \int_0^T \Big(\varphi^*(-\dot{x}(\tau)) + \varphi(x(\tau))\Big)\, d\tau - \frac{1}{2}|x_0|^2.$$

$$\mathcal{BE}(T, x) = \int_0^T \Big(\varphi^*(-\dot{x}(\tau)) + \varphi(x(\tau)) - \langle -\dot{x}(t), x(t)\rangle\Big)\, d\tau$$

The Brezis–Ekeland functional

Brezis–Ekeland functional

$$\mathcal{BE}(T,x) := \frac{1}{2}|x(T)|^2 + \int_0^T \Big(\varphi^*(-\dot{x}(\tau)) + \varphi(x(\tau))\Big)\, d\tau - \frac{1}{2}|x_0|^2.$$

$$\mathcal{BE}(T,x) = \int_0^T \Big(\varphi^*(-\dot{x}(\tau)) + \varphi(x(\tau)) - \langle -\dot{x}(t), x(t)\rangle\Big)\, d\tau$$

Observation: Fenchel's inequality gives

$$\varphi(x(t)) + \varphi^*(-\dot{x}(t)) \geqslant -\langle x(t), \dot{x}(t)\rangle = -\frac{d}{dt}\frac{1}{2}|x(t)|^2,$$

with **equality** $\iff$ $-\dot{x}(t) = D\varphi(x(t))$ (i.e. x is a gradient flow).

The Brezis–Ekeland functional

Brezis–Ekeland functional

$$\mathcal{BE}(T, x) := \frac{1}{2}|x(T)|^2 + \int_0^T \Big(\varphi^*(-\dot{x}(\tau)) + \varphi(x(\tau))\Big)\,d\tau - \frac{1}{2}|x_0|^2.$$

$$\mathcal{BE}(T, x) = \int_0^T \Big(\varphi^*(-\dot{x}(\tau)) + \varphi(x(\tau)) - \langle -\dot{x}(t), x(t)\rangle\Big)\,d\tau$$

Observation: Fenchel's inequality gives

$$\varphi(x(t)) + \varphi^*(-\dot{x}(t)) \geqslant -\langle x(t), \dot{x}(t)\rangle = -\frac{d}{dt}\frac{1}{2}|x(t)|^2,$$

with **equality** $\iff$ $-\dot{x}(t) = D\varphi(x(t))$ (i.e. x is a gradient flow).

Hence: $\mathcal{BE}(T, x) \geqslant 0$, and $\mathcal{BE}(T, x) = 0 \iff x$ solves the gradient flow.

Null-minimization: solutions = null-minimizers of $\mathcal{BE}$.

Proving $\min \mathcal{BE} = 0$: Ghoussoub–Tzou/McCann

Challenge: prove that $\min \mathcal{BE} = 0$ without knowing *a priori* that a solution exists.

Proving $\min \mathcal{BE} = 0$: Ghoussoub–Tzou/McCann

Challenge: prove that $\min \mathcal{BE} = 0$ without knowing *a priori* that a solution exists.

Ghoussoub–Tzou (2004), Ghoussoub–McCann (2006):

Via *self-dual Lagrangians* and duality, they prove $\min \mathcal{BE} = 0$ for:

- convex energies φ (Ghoussoub–Tzou),
- λ-convex energies φ (Ghoussoub–McCann).

Proving $\min \mathcal{BE} = 0$: Ghoussoub–Tzou/McCann

Challenge: prove that $\min \mathcal{BE} = 0$ without knowing *a priori* that a solution exists.

Ghoussoub–Tzou (2004), Ghoussoub–McCann (2006):

Via *self-dual Lagrangians* and duality, they prove $\min \mathcal{BE} = 0$ for:

- convex energies φ (Ghoussoub–Tzou),
- λ-convex energies φ (Ghoussoub–McCann).

Key features of Brezis–Ekeland:

- The space is **Hilbert**.
- The "Lyapunov functional" is always $\frac{1}{2}|x|^2$ (the *energy* φ enters through $\varphi + \varphi^*$).
- The self-dual structure requires **convexity** of φ.

Outline

1. Variational principles for evolution equations
2. **De Giorgi's principle and main result**
3. Three key arguments
   - Relaxation and continuity equation
   - Von Neumann duality
   - Backward estimates and maximum principle for first order HJ equations
4. The infinite-dimensional setting
5. Extensions and open problems

A different variational characterization, where the Lyapunov functional is the **energy** φ itself.

For gradient flows ($\dot{x} = -\mathrm{D}\varphi(x)$):

De Giorgi's energy-dissipation functional

$$\mathcal{J}(T, x) := \varphi(x(T)) + \int_0^T \left(\frac{1}{2}|\dot{x}|^2 + \frac{1}{2}\left| -\mathrm{D}\varphi(x) \right|^2 \right) \mathrm{d}\tau - \varphi(x_0).$$

The De Giorgi functional

A different variational characterization, where the Lyapunov functional is the **energy** φ itself.

For gradient flows ($\dot{\boldsymbol{x}} = -D\varphi(\boldsymbol{x})$):

De Giorgi's energy-dissipation functional

$$\mathcal{J}(T, \boldsymbol{x}) := \varphi(\boldsymbol{x}(T)) + \int_0^T \left(\tfrac{1}{2}|\dot{\boldsymbol{x}}|^2 + \tfrac{1}{2}\big| - D\varphi(\boldsymbol{x})\big|^2\right) d\tau - \varphi(x_0).$$

$$\mathcal{J}(T, \boldsymbol{x}) = \int_0^T \left(\tfrac{1}{2}|\dot{\boldsymbol{x}}|^2 + \tfrac{1}{2}\big| - D\varphi(\boldsymbol{x})\big|^2 - \langle \dot{\boldsymbol{x}}(t), -D\varphi(\boldsymbol{x}(t))\rangle\right) d\tau.$$

The De Giorgi functional

A different variational characterization, where the Lyapunov functional is the **energy** φ itself.

For gradient flows ($\dot{x} = -D\varphi(x)$):

De Giorgi's energy-dissipation functional

$$\mathcal{J}(T, x) := \varphi(x(T)) + \int_0^T \left(\tfrac{1}{2}|\dot{x}|^2 + \tfrac{1}{2}\big| - D\varphi(x)\big|^2 \right) d\tau - \varphi(x_0).$$

$$\mathcal{J}(T, x) = \int_0^T \left(\tfrac{1}{2}|\dot{x}|^2 + \tfrac{1}{2}\big| - D\varphi(x)\big|^2 - \langle \dot{x}(t), -D\varphi(x(t))\rangle \right) d\tau.$$

Why $\mathcal{J} \geqslant 0$? By Young's inequality $\frac{1}{2}|a|^2 + \frac{1}{2}|b|^2 \geqslant \langle a, b\rangle$:

$$\tfrac{1}{2}\big|\dot{x}\big|^2 + \tfrac{1}{2}\big| - D\varphi(x)\big|^2 \geqslant \langle - D\varphi(x), \dot{x}\rangle = -\frac{d}{dt}\varphi(x(t)),$$

with **equality** $\iff$ $\dot{x} = -D\varphi(x)$. Integrating: $\mathcal{J} \geqslant 0$, and $\mathcal{J} = 0$ $\iff$ x solves the gradient flow.

The De Giorgi functional

A different variational characterization, where the Lyapunov functional is the **energy** φ itself.

For gradient flows ($\dot{x} = -D\varphi(x)$):

De Giorgi's energy-dissipation functional

$$\mathcal{J}(T, x) := \varphi(x(T)) + \int_0^T \left(\tfrac{1}{2}|\dot{x}|^2 + \tfrac{1}{2}\big|-D\varphi(x)\big|^2\right) d\tau - \varphi(x_0).$$

$$\mathcal{J}(T, x) = \int_0^T \left(\tfrac{1}{2}|\dot{x}|^2 + \tfrac{1}{2}\big|-D\varphi(x)\big|^2 - \langle \dot{x}(t), -D\varphi(x(t))\rangle\right) d\tau.$$

Why $\mathcal{J} \geqslant 0$? By Young's inequality $\frac{1}{2}|a|^2 + \frac{1}{2}|b|^2 \geqslant \langle a, b\rangle$:

$$\tfrac{1}{2}\big|\dot{x}\big|^2 + \tfrac{1}{2}\big|-D\varphi(x)\big|^2 \geqslant \langle -D\varphi(x), \dot{x}\rangle = -\frac{d}{dt}\varphi(x(t)),$$

with **equality** $\iff$ $\dot{x} = -D\varphi(x)$. Integrating: $\mathcal{J} \geqslant 0$, and $\mathcal{J} = 0 \iff x$ solves the gradient flow.

Null-minimization: solutions = null-minimizers of $\mathcal{J}$.

The hard part: prove $\min \mathcal{J} \leqslant 0$ without knowing a solution exists!

The De Giorgi functional for doubly nonlinear equations

For $D\Psi(\dot{x}) = -D\varphi(x)$, the same construction with **Fenchel inequality** replacing Young:

$$\mathcal{J}(T, x) := \varphi(x(T)) + \int_0^T \Big(\Psi(\dot{x}(\tau)) + \Psi^*(-D\varphi(x(\tau))) \Big) \, d\tau - \varphi(x_0).$$

For $D\Psi(\dot{x}) = -D\varphi(x)$, the same construction with **Fenchel inequality** replacing Young:

$$\mathcal{J}(T, x) := \varphi(x(T)) + \int_0^T \Big(\Psi(\dot{x}(\tau)) + \Psi^*(-D\varphi(x(\tau)))\Big)\, d\tau - \varphi(x_0).$$

Indeed: $\Psi(v) + \Psi^*(-z) \geqslant \langle -z, v\rangle$ gives

$$\Psi(\dot{x}) + \Psi^*(-D\varphi(x)) \geqslant -\langle D\varphi(x), \dot{x}\rangle = -\frac{d}{dt}\varphi(x(t)),$$

with equality $\iff$ $D\Psi(\dot{x}) = -D\varphi(x)$.

For $D\Psi(\dot{x}) = -D\varphi(x)$, the same construction with **Fenchel inequality** replacing Young:

$$\mathcal{J}(T, x) := \varphi(x(T)) + \int_0^T \Big(\Psi(\dot{x}(\tau)) + \Psi^*(-D\varphi(x(\tau)))\Big)\, d\tau - \varphi(x_0).$$

Indeed: $\Psi(v) + \Psi^*(-z) \geqslant \langle -z, v\rangle$ gives

$$\Psi(\dot{x}) + \Psi^*(-D\varphi(x)) \geqslant -\langle D\varphi(x), \dot{x}\rangle = -\frac{d}{dt}\varphi(x(t)),$$

with equality $\iff$ $D\Psi(\dot{x}) = -D\varphi(x)$.

The **slope** $S(x) := \Psi^*(-D\varphi(x))$ (for GF: $\frac{1}{2}|a|^2 + \frac{1}{2}|b|^2 \rightsquigarrow S = \frac{1}{2}|D\varphi|^2$).

Proposition (De Giorgi's principle)

(a) $\mathcal{J}(T, \boldsymbol{x}) \geqslant 0$ *for every* $\boldsymbol{x}$ *with* $\boldsymbol{x}(0) = x_0$.

(b) $\mathcal{J}(T, \boldsymbol{x}) = 0 \iff \boldsymbol{x}$ *solves* $D\Psi(\dot{\boldsymbol{x}}) = -D\varphi(\boldsymbol{x})$.

De Giorgi's principle

Proposition (De Giorgi's principle)

(a) $\mathcal{J}(T, x) \geqslant 0$ *for every* x *with* $x(0) = x_0$.

(b) $\mathcal{J}(T, x) = 0 \iff x$ *solves* $D\Psi(\dot{x}) = -D\varphi(x)$.

Null-minimization: solutions = null-minimizers of $\mathcal{J}$.

The hard part: prove $\min \mathcal{J} \leqslant 0$ without knowing a solution exists!

Proposition (De Giorgi's principle)

(a) $\mathcal{J}(T, \boldsymbol{x}) \geqslant 0$ *for every* $\boldsymbol{x}$ *with* $\boldsymbol{x}(0) = x_0$.

(b) $\mathcal{J}(T, \boldsymbol{x}) = 0 \iff \boldsymbol{x}$ *solves* $D\Psi(\dot{\boldsymbol{x}}) = -D\varphi(\boldsymbol{x})$.

Null-minimization: solutions = null-minimizers of $\mathcal{J}$.

The hard part: prove $\min \mathcal{J} \leqslant 0$ without knowing a solution exists!

Fundamental differences with Brezis–Ekeland:

- **No convexity** of φ required.
- The Lyapunov functional is **the energy** φ, not $\frac{1}{2}|x|^2$.
- Works in **Banach** and **metric** spaces.
- However $\mathcal{J}$ is **not convex** in general, even for convex φ, due to the presence of $\Psi^*(-D\varphi)$.

The Minimizing Movement / JKO scheme

The classical approach to prove $\min \mathcal{J} = 0$: **time discretization**. a variational form of the **Implicit Euler Scheme**.

The Minimizing Movement / JKO scheme

The classical approach to prove $\min \mathcal{J} = 0$: **time discretization**. a variational form of the **Implicit Euler Scheme**.

Fix a time step $\tau > 0$. Given x_0, solve iteratively:

$$x_k^\tau \in \arg\min_x \Big\{ \tau \Psi\Big(\frac{x - x_{k-1}^\tau}{\tau}\Big) + \varphi(x) \Big\}, \quad k = 1, 2, \ldots \qquad \rightsquigarrow \quad D\Psi\Big(\frac{x - x_{k-1}^\tau}{\tau}\Big) + D\varphi(x) = 0.$$

The Minimizing Movement / JKO scheme

The classical approach to prove $\min \mathcal{J} = 0$: **time discretization**. a variational form of the **Implicit Euler Scheme**.

Fix a time step $\tau > 0$. Given x_0, solve iteratively:

$$x_k^\tau \in \arg\min_x \left\{ \tau \Psi\Big(\frac{x - x_{k-1}^\tau}{\tau}\Big) + \varphi(x) \right\}, \quad k = 1, 2, \ldots \qquad \rightsquigarrow \qquad D\Psi\Big(\frac{x - x_{k-1}^\tau}{\tau}\Big) + D\varphi(x) = 0.$$

For gradient flows ($\Psi = \frac{1}{2}|\cdot|^2$): $\quad x_k^\tau \in \arg\min_x \left\{ \frac{|x - x_{k-1}^\tau|^2}{2\tau} + \varphi(x) \right\}.$

This is the **JKO/Minimizing Movement scheme** (Jordan–Kinderlehrer–Otto, 1998; De Giorgi, 1993).

The Minimizing Movement / JKO scheme

The classical approach to prove $\min \mathcal{J} = 0$: **time discretization**. a variational form of the **Implicit Euler Scheme**.

Fix a time step $\tau > 0$. Given x_0, solve iteratively:

$$x_k^\tau \in \arg\min_x \left\{ \tau \Psi\Big(\frac{x - x_{k-1}^\tau}{\tau}\Big) + \varphi(x) \right\}, \quad k = 1, 2, \ldots \qquad \rightsquigarrow \quad D\Psi\Big(\frac{x - x_{k-1}^\tau}{\tau}\Big) + D\varphi(x) = 0.$$

For gradient flows ($\Psi = \frac{1}{2}|\cdot|^2$): $\quad x_k^\tau \in \arg\min_x \left\{ \frac{|x - x_{k-1}^\tau|^2}{2\tau} + \varphi(x) \right\}.$

This is the **JKO/Minimizing Movement scheme** (Jordan–Kinderlehrer–Otto, 1998; De Giorgi, 1993).

As $\tau \to 0$: piecewise constant interpolants $\to$ solution.

The discrete energy identity gives $\mathcal{J} \approx 0$ along the approximation.

The Minimizing Movement / JKO scheme

The classical approach to prove $\min \mathcal{J} = 0$: **time discretization**. a variational form of the **Implicit Euler Scheme**.

Fix a time step $\tau > 0$. Given x_0, solve iteratively:

$$x_k^\tau \in \arg\min_x \left\{ \tau \Psi\Big(\frac{x - x_{k-1}^\tau}{\tau}\Big) + \varphi(x) \right\}, \quad k = 1, 2, \dots \qquad \rightsquigarrow \qquad D\Psi\Big(\frac{x - x_{k-1}^\tau}{\tau}\Big) + D\varphi(x) = 0.$$

For gradient flows ($\Psi = \frac{1}{2}|\cdot|^2$): $\quad x_k^\tau \in \arg\min_x \left\{ \frac{|x - x_{k-1}^\tau|^2}{2\tau} + \varphi(x) \right\}.$

This is the **JKO/Minimizing Movement scheme** (Jordan–Kinderlehrer–Otto, 1998; De Giorgi, 1993).

As $\tau \to 0$: piecewise constant interpolants $\to$ solution.

The discrete energy identity gives $\mathcal{J} \approx 0$ along the approximation.

Our approach: prove $\min \mathcal{J} \leqslant 0$ *directly*, without time discretization, via convexification + duality.

The main theorem (smooth Euclidean setting)

Additional assumptions: $\varphi\colon \mathbb{R}^d \to [0,+\infty)$ smooth, coercive: $\varphi(x) \to +\infty$, $|D\varphi(x)| \to +\infty$ as $|x| \to +\infty$. **No convexity!** $\Rightarrow S(x) = \Psi^*(-D\varphi(x))$ coercive.

Variational argument (Pinzi–Riva–S.)

For all $x_0 \in \mathbb{R}^d$:

$$\boxed{\min\left\{\mathcal{J}(T,\boldsymbol{x}) : \boldsymbol{x} \in AC([0,T];\mathbb{R}^d),\ \boldsymbol{x}(0) = x_0\right\} = 0.}$$

In particular, any minimizer solves $D\Psi(\dot{\boldsymbol{x}}) = -D\varphi(\boldsymbol{x})$.

The main theorem (smooth Euclidean setting)

Additional assumptions: $\varphi\colon \mathbb{R}^d \to [0,+\infty)$ smooth, coercive: $\varphi(x) \to +\infty$, $|D\varphi(x)| \to +\infty$ as $|x| \to +\infty$. **No convexity!** $\Rightarrow S(x) = \Psi^*(-D\varphi(x))$ coercive.

Variational argument (Pinzi–Riva–S.)

For all $x_0 \in \mathbb{R}^d$:

$$\boxed{\min\left\{\mathcal{J}(T,\boldsymbol{x}) : \boldsymbol{x} \in AC([0,T];\mathbb{R}^d),\ \boldsymbol{x}(0) = x_0\right\} = 0.}$$

In particular, any minimizer solves $D\Psi(\dot{\boldsymbol{x}}) = -D\varphi(\boldsymbol{x})$.

Existence of minimizers: direct method + coercivity. $\mathcal{J} \geqslant 0$. The core is to prove $\min \mathcal{J} \leqslant 0$.

The main theorem (smooth Euclidean setting)

Additional assumptions: $\varphi\colon \mathbb{R}^d \to [0,+\infty)$ smooth, coercive: $\varphi(x) \to +\infty$, $|D\varphi(x)| \to +\infty$ as $|x| \to +\infty$. **No convexity!** $\Rightarrow S(x) = \Psi^*(-D\varphi(x))$ coercive.

Variational argument (Pinzi–Riva–S.)

For all $x_0 \in \mathbb{R}^d$:

$$\boxed{\min\left\{\mathcal{J}(T,\boldsymbol{x}) : \boldsymbol{x} \in AC([0,T];\mathbb{R}^d),\ \boldsymbol{x}(0) = x_0\right\} = 0.}$$

In particular, any minimizer solves $D\Psi(\dot{\boldsymbol{x}}) = -D\varphi(\boldsymbol{x})$.

Existence of minimizers: direct method + coercivity. $\mathcal{J} \geqslant 0$. The core is to prove $\min \mathcal{J} \leqslant 0$.

Three ingredients:

1. Relaxation via continuity equation and superposition principle.
2. Von Neumann minimax and the dual problem.
3. Backward (upper) estimates for solutions to first order Hamilton-Jacobi equations.

Outline

1. Variational principles for evolution equations
2. De Giorgi's principle and main result
3. **Three key arguments**
   - Relaxation and continuity equation
   - Von Neumann duality
   - Backward estimates and maximum principle for first order HJ equations
4. The infinite-dimensional setting
5. Extensions and open problems

Convexification

$\mathcal{J}$ is **not convex** (it depends on φ, which is not convex).

To apply a **duality argument**, we need convexity.

Convexification

$\mathcal{J}$ is **not convex** (it depends on φ, which is not convex).

To apply a **duality argument**, we need convexity.

Natural convexification: consider probability measures $\lambda \in \mathscr{P}(AC([0,T];\mathbb{R}^d))$ on curves with $\boldsymbol{x}(0) = x_0$, and define

$$\bar{\mathcal{J}}(\lambda) := \int_{AC} \mathcal{J}(T, \boldsymbol{x})\, d\lambda(\boldsymbol{x}).$$

$\bar{\mathcal{J}}$ is **linear** in λ, hence convex.

$\mathcal{J}$ is **not convex** (it depends on φ, which is not convex).

To apply a **duality argument**, we need convexity.

Natural convexification: consider probability measures $\lambda \in \mathscr{P}(AC([0,T];\mathbb{R}^d))$ on curves with $\mathbf{x}(0) = x_0$, and define

$$\overline{\mathcal{J}}(\lambda) := \int_{AC} \mathcal{J}(T, \mathbf{x})\, d\lambda(\mathbf{x}).$$

$\overline{\mathcal{J}}$ is **linear** in λ, hence convex.

Since $\mathcal{J}(\mathbf{x}) = \overline{\mathcal{J}}(\delta_{\mathbf{x}})$: $\quad \min_\lambda \overline{\mathcal{J}} \leqslant \min_{\mathbf{x}} \mathcal{J}$.

Conversely, for any λ: $\overline{\mathcal{J}}(\lambda) = \int \mathcal{J}\, d\lambda \geqslant \min \mathcal{J}$.

Therefore: $\min_\lambda \overline{\mathcal{J}} = \min_{\mathbf{x}} \mathcal{J}$. **Convexification does not change the minimum.**

From measures on curves to curves of measures

A curve x with $x(0) = x_0$ induces a triplet of measures on $X_T = [0,T] \times \mathbb{R}^d$:

$$m_x = \delta_{x(T)}, \quad \mu_x = \mathcal{L}^1 \otimes \delta_{x(t)}, \quad \nu_x = \dot{x}(t)\, \mathcal{L}^1 \otimes \delta_{x(t)} = \boldsymbol{v} \cdot \mu_x$$

From measures on curves to curves of measures

A curve x with $x(0) = x_0$ induces a triplet of measures on $X_T = [0, T] \times \mathbb{R}^d$:

$$m_x = \delta_{x(T)}, \quad \mu_x = \mathcal{L}^1 \otimes \delta_{x(t)}, \quad \nu_x = \dot{x}(t)\, \mathcal{L}^1 \otimes \delta_{x(t)} = \boldsymbol{v} \cdot \mu_x$$

They satisfy the **continuity equation** in the sense of distributions:

$$\partial_t \mu_x + \operatorname{div} \nu_x = 0 \quad \text{in } (0, T) \times \mathbb{R}^d, \quad \mu_x(0, \cdot) = \delta_{x_0}, \ \mu_x(T, \cdot) = m_x;$$

$$\xi(0, x_0) + \iint \Big(\partial_t \xi(t, x) + \langle D\xi(t, x), \boldsymbol{v}(t, x) \rangle \Big)\, d\mu(t, x) - \int \xi(T, x)\, dm(x) = 0$$

From measures on curves to curves of measures

A curve x with $x(0) = x_0$ induces a triplet of measures on $X_T = [0, T] \times \mathbb{R}^d$:

$$m_x = \delta_{x(T)}, \quad \mu_x = \mathcal{L}^1 \otimes \delta_{x(t)}, \quad \nu_x = \dot{x}(t)\,\mathcal{L}^1 \otimes \delta_{x(t)} = \boldsymbol{v} \cdot \mu_x$$

They satisfy the **continuity equation** in the sense of distributions:

$$\partial_t \mu_x + \operatorname{div} \nu_x = 0 \quad \text{in } (0, T) \times \mathbb{R}^d, \quad \mu_x(0, \cdot) = \delta_{x_0}, \ \mu_x(T, \cdot) = m_x;$$

$$\xi(0, x_0) + \iint \Big(\partial_t \xi(t, x) + \langle D\xi(t, x), \boldsymbol{v}(t, x)\rangle\Big)\, d\mu(t, x) - \int \xi(T, x)\, dm(x) = 0$$

Given $\lambda \in \mathscr{P}(AC)$, integrate: $\quad m_\lambda = \int m_x \, d\lambda, \ \ \mu_\lambda = \int \mu_x \, d\lambda, \ \ \nu_\lambda = \int \nu_x \, d\lambda.$

$(m_\lambda, \mu_\lambda, \nu_\lambda)$ still satisfies the CE (linearity); $\nu = \boldsymbol{v}\mu$.

The link between position and velocity is **transferred to the continuity equation**.

A curve x with $x(0) = x_0$ induces a triplet of measures on $X_T = [0, T] \times \mathbb{R}^d$:

$$m_x = \delta_{x(T)}, \quad \mu_x = \mathcal{L}^1 \otimes \delta_{x(t)}, \quad \nu_x = \dot{x}(t)\,\mathcal{L}^1 \otimes \delta_{x(t)} = \boldsymbol{v} \cdot \mu_x$$

They satisfy the **continuity equation** in the sense of distributions:

$$\partial_t \mu_x + \operatorname{div} \nu_x = 0 \quad \text{in } (0, T) \times \mathbb{R}^d, \quad \mu_x(0, \cdot) = \delta_{x_0}, \;\; \mu_x(T, \cdot) = m_x;$$

$$\xi(0, x_0) + \iint \Big(\partial_t \xi(t, x) + \langle D\xi(t, x), \boldsymbol{v}(t, x) \rangle \Big)\, d\mu(t, x) - \int \xi(T, x)\, dm(x) = 0$$

Given $\lambda \in \mathscr{P}(AC)$, integrate: $\quad m_\lambda = \int m_x \, d\lambda, \;\; \mu_\lambda = \int \mu_x \, d\lambda, \;\; \nu_\lambda = \int \nu_x \, d\lambda.$

$(m_\lambda, \mu_\lambda, \nu_\lambda)$ still satisfies the CE (linearity); $\nu = \boldsymbol{v}\mu$.

The link between position and velocity is **transferred to the continuity equation**.

Relaxed functional: over triplets (m, μ, ν), $\nu = \boldsymbol{v}\mu$, satisfying the CE,

$$\mathcal{E}(m, \mu, \nu) := \int_{\mathbb{R}^d} \varphi \, dm + \int_{X_T} \big(\Psi(\boldsymbol{v}(t, x)) + \Psi^*(-D\varphi(x))\big)\, d\mu - \varphi(x_0); \qquad \boldsymbol{v} = \frac{d\nu}{d\mu}$$

By Jensen: $\mathcal{E}(m_\lambda, \mu_\lambda, \nu_\lambda) \leqslant \overline{\mathcal{J}}(\lambda)$. **But could we lose something?**

Superposition principle (Smirnov; Ambrosio; Ambrosio–Gigli–S.):

Any solution (m, μ, ν), with $\nu = \boldsymbol{v}\mu$, of the continuity equation can be *decomposed* into absolutely continuous curves: there exists $\lambda \in \mathscr{P}(AC([0,T];\mathbb{R}^d))$ such that

$$m = m_\lambda, \quad \mu = \mu_\lambda, \quad \nu = \nu_\lambda,$$

and the curves satisfy $\dot{x}(t) = \boldsymbol{v}(t, \boldsymbol{x}(t)) \quad \lambda\text{-a.e.}$

Superposition principle (Smirnov; Ambrosio; Ambrosio–Gigli–S.):

Any solution (m, μ, ν), with $\nu = \boldsymbol{v}\mu$, of the continuity equation can be *decomposed* into absolutely continuous curves: there exists $\lambda \in \mathscr{P}(AC([0,T];\mathbb{R}^d))$ such that

$$m = m_\lambda, \quad \mu = \mu_\lambda, \quad \nu = \nu_\lambda,$$

and the curves satisfy $\dot{\boldsymbol{x}}(t) = \boldsymbol{v}(t, \boldsymbol{x}(t)) \quad \lambda\text{-a.e.}$

Therefore: $\mathcal{E}(m, \mu, \nu) = \displaystyle\int_{AC} \mathcal{J}(T, \boldsymbol{x})\, d\lambda(\boldsymbol{x}) = \overline{\mathcal{J}}(\lambda) \geqslant \min \overline{\mathcal{J}}$.

Superposition principle (Smirnov; Ambrosio; Ambrosio–Gigli–S.):

Any solution (m, μ, ν), with $\nu = \boldsymbol{\nu}\mu$, of the continuity equation can be *decomposed* into absolutely continuous curves: there exists $\lambda \in \mathscr{P}(AC([0, T]; \mathbb{R}^d))$ such that

$$m = m_\lambda, \quad \mu = \mu_\lambda, \quad \nu = \nu_\lambda,$$

and the curves satisfy $\dot{\boldsymbol{x}}(t) = \boldsymbol{\nu}(t, \boldsymbol{x}(t)) \quad \lambda$-a.e.

Therefore: $\mathcal{E}(m, \mu, \nu) = \displaystyle\int_{AC} \mathcal{J}(T, \boldsymbol{x})\, d\lambda(\boldsymbol{x}) = \overline{\mathcal{J}}(\lambda) \geqslant \min \overline{\mathcal{J}}$.

Putting everything together:

$$\boxed{\min_{\boldsymbol{x}} \mathcal{J} = \min_{\lambda} \overline{\mathcal{J}} = \min_{(m,\mu,\nu)\in CE} \mathcal{E}.}$$

Superposition principle (Smirnov; Ambrosio; Ambrosio–Gigli–S.):

Any solution (m, μ, ν), with $\nu = \boldsymbol{\nu}\mu$, of the continuity equation can be *decomposed* into absolutely continuous curves: there exists $\lambda \in \mathscr{P}(AC([0,T];\mathbb{R}^d))$ such that

$$m = m_\lambda, \quad \mu = \mu_\lambda, \quad \nu = \nu_\lambda,$$

and the curves satisfy $\dot{\boldsymbol{x}}(t) = \boldsymbol{\nu}(t, \boldsymbol{x}(t)) \quad \lambda$-a.e.

Therefore: $\mathcal{E}(m, \mu, \nu) = \int_{AC} \mathcal{J}(T, \boldsymbol{x})\, d\lambda(\boldsymbol{x}) = \overline{\mathcal{J}}(\lambda) \geqslant \min \overline{\mathcal{J}}$.

Putting everything together:

$$\boxed{\min_{\boldsymbol{x}} \mathcal{J} = \min_{\lambda} \overline{\mathcal{J}} = \min_{(m,\mu,\nu)\in \mathrm{CE}} \mathcal{E}.}$$

No loss! It suffices to show $\min \mathcal{E} \leqslant 0$ for the **convex** functional $\mathcal{E}$.

The saddle-point formulation

Encode the continuity equation constraint via a **Lagrange multiplier** $\xi \in C^1_b(X_T)$:

$$\mathcal{C}(\xi, m, \mu, \nu) := \int_{X_T} \partial_t \xi \, d\mu + \int_{X_T} D_x \xi \cdot d\nu - \int_{\mathbb{R}^d} \xi(T) \, dm + \xi(0, x_0).$$

The saddle-point formulation

Encode the continuity equation constraint via a **Lagrange multiplier** $\xi \in C_b^1(X_T)$:

$$\mathcal{C}(\xi, m, \mu, \nu) := \int_{X_T} \partial_t \xi \, d\mu + \int_{X_T} D_x \xi \cdot d\nu - \int_{\mathbb{R}^d} \xi(T) \, dm + \xi(0, x_0).$$

Lagrangian: $\mathcal{L}(\xi, m, \mu, \nu) = \mathcal{E}(m, \mu, \nu) + \mathcal{C}(\xi, m, \mu, \nu)$.

Then:

$$\min \mathcal{E} \text{ (subject to CE)} = \min_{(m,\mu,\nu)} \sup_{\xi} \mathcal{L}(\xi, m, \mu, \nu).$$

The saddle-point formulation

Encode the continuity equation constraint via a **Lagrange multiplier** $\xi \in C^1_b(X_T)$:

$$\mathcal{C}(\xi, m, \mu, \nu) := \int_{X_T} \partial_t \xi \, d\mu + \int_{X_T} D_x \xi \cdot d\nu - \int_{\mathbb{R}^d} \xi(T) \, dm + \xi(0, x_0).$$

Lagrangian: $\mathcal{L}(\xi, m, \mu, \nu) = \mathcal{E}(m, \mu, \nu) + \mathcal{C}(\xi, m, \mu, \nu)$.

Then:

$$\min \mathcal{E} \text{ (subject to CE)} = \min_{(m,\mu,\nu)} \sup_{\xi} \mathcal{L}(\xi, m, \mu, \nu).$$

In general: $\quad \min_B \sup_A \mathcal{L} \geqslant \sup_A \inf_B \mathcal{L} \qquad$ (weak duality).

Encode the continuity equation constraint via a **Lagrange multiplier** $\xi \in C_b^1(X_T)$:

$$\mathcal{C}(\xi, m, \mu, \nu) := \int_{X_T} \partial_t \xi \, d\mu + \int_{X_T} D_x \xi \cdot d\nu - \int_{\mathbb{R}^d} \xi(T) \, dm + \xi(0, x_0).$$

Lagrangian: $\mathcal{L}(\xi, m, \mu, \nu) = \mathcal{E}(m, \mu, \nu) + \mathcal{C}(\xi, m, \mu, \nu)$.

Then:

$$\min \mathcal{E} \text{ (subject to CE)} = \min_{(m,\mu,\nu)} \sup_{\xi} \mathcal{L}(\xi, m, \mu, \nu).$$

In general: $\quad \min_B \sup_A \mathcal{L} \geqslant \sup_A \inf_B \mathcal{L} \qquad$ (weak duality).

Von Neumann minimax theorem: under structural conditions on $\mathcal{L}$ (convexity/concavity + compactness), equality holds:

$$\min_{(m,\mu,\nu)} \sup_{\xi} \mathcal{L} = \sup_{\xi} \inf_{(m,\mu,\nu)} \mathcal{L}.$$

The saddle-point formulation

Encode the continuity equation constraint via a **Lagrange multiplier** $\xi \in C_b^1(X_T)$:

$$\mathcal{C}(\xi, m, \mu, \nu) := \int_{X_T} \partial_t \xi \, d\mu + \int_{X_T} D_x \xi \cdot d\nu - \int_{\mathbb{R}^d} \xi(T) \, dm + \xi(0, x_0).$$

Lagrangian: $\mathcal{L}(\xi, m, \mu, \nu) = \mathcal{E}(m, \mu, \nu) + \mathcal{C}(\xi, m, \mu, \nu)$.

Then:

$$\min \mathcal{E} \text{ (subject to CE)} = \min_{(m,\mu,\nu)} \sup_{\xi} \mathcal{L}(\xi, m, \mu, \nu).$$

In general: $\min_B \sup_A \mathcal{L} \geqslant \sup_A \inf_B \mathcal{L}$ (weak duality).

Von Neumann minimax theorem: under structural conditions on $\mathcal{L}$ (convexity/concavity + compactness), equality holds:

$$\min_{(m,\mu,\nu)} \sup_{\xi} \mathcal{L} = \sup_{\xi} \inf_{(m,\mu,\nu)} \mathcal{L}.$$

Here: $\mathcal{L}$ is **convex** in (m, μ, ν) and **linear** (hence concave) in ξ. ✓

Verification of the minimax hypotheses

Theorem (Von Neumann minimax)

Let $\mathcal{L}(\xi, \theta)$ *with* $\xi \in \mathrm{A}$, $\theta \in \mathrm{B}$ *convex sets,* B *endowed with a Hausdorff topology. Assume:*

- $\xi \mapsto \mathcal{L}(\xi, \theta)$ *is* **concave** *for every* θ*;*
- $\theta \mapsto \mathcal{L}(\xi, \theta)$ *is* **convex** *and* **l.s.c.** *for every* ξ*;*
- **coercivity:** *there exist* $\overline{\mathrm{C}} > \sup_{\mathrm{A}} \inf_{\mathrm{B}} \mathcal{L}$ *and* $\bar{\xi} \in \mathrm{A}$ *such that* $\{\mathcal{L}(\bar{\xi}, \cdot) \leqslant \overline{\mathrm{C}}\}$ *is nonempty and compact.*

Then: $\displaystyle \min_{\theta \in \mathrm{B}} \sup_{\xi \in \mathrm{A}} \mathcal{L}(\xi, \theta) = \sup_{\xi \in \mathrm{A}} \inf_{\theta \in \mathrm{B}} \mathcal{L}(\xi, \theta).$

No topology on A needed!

Verification of the minimax hypotheses

Theorem (Von Neumann minimax)

Let $\mathcal{L}(\xi, \theta)$ with $\xi \in A$, $\theta \in B$ convex sets, B endowed with a Hausdorff topology. Assume:

- $\xi \mapsto \mathcal{L}(\xi, \theta)$ *is* **concave** *for every* θ;
- $\theta \mapsto \mathcal{L}(\xi, \theta)$ *is* **convex** *and* **l.s.c.** *for every* ξ;
- **coercivity:** *there exist* $\overline{C} > \sup_A \inf_B \mathcal{L}$ *and* $\bar{\xi} \in A$ *such that* $\{\mathcal{L}(\bar{\xi}, \cdot) \leqslant \overline{C}\}$ *is nonempty and compact.*

Then: $$\min_{\theta \in B} \sup_{\xi \in A} \mathcal{L}(\xi, \theta) = \sup_{\xi \in A} \inf_{\theta \in B} \mathcal{L}(\xi, \theta).$$

No topology on A needed!

In our setting: $\xi \in A = C^1_b(X_T)$, $\quad \theta = (m, \mu, \nu) \in B$ with narrow topology.

- Concavity in ξ: $\mathcal{L}$ is linear in ξ. ✓
- Convexity + l.s.c. in θ: convexity of Ψ. ✓
- **Compactness of sublevels** (Prokhorov): coercivity of φ and $S \Rightarrow$ tightness of m, μ; superlinearity of $\Psi \Rightarrow$ tightness of ν. ✓

The dual problem

After switching inf and sup and computing:

$$\inf_{(m,\mu,\nu)} \mathcal{L}(\xi, m, \mu, \nu) = \begin{cases} \xi(0, x_0) - \varphi(x_0), & \text{if } \xi \in \mathsf{HJ}, \\ -\infty, & \text{otherwise,} \end{cases}$$

The dual problem

After switching inf and sup and computing:

$$\inf_{(m,\mu,\nu)} \mathcal{L}(\xi, m, \mu, \nu) = \begin{cases} \xi(0, x_0) - \varphi(x_0), & \text{if } \xi \in \mathrm{HJ}, \\ -\infty, & \text{otherwise}, \end{cases}$$

where $\xi \in \mathrm{HJ}$ means that ξ satisfies the **Hamilton–Jacobi backward differential inequality**:

$$\boxed{\begin{cases} -\partial_t \xi(t,x) + \Psi^*(-D_x \xi(t,x)) \leqslant \Psi^*(-D\varphi(x)), \\ \xi(T,x) \leqslant \varphi(x). \end{cases}}$$

The dual problem

After switching inf and sup and computing:

$$\inf_{(m,\mu,\nu)} \mathcal{L}(\xi, m, \mu, \nu) = \begin{cases} \xi(0, x_0) - \varphi(x_0), & \text{if } \xi \in \mathrm{HJ}, \\ -\infty, & \text{otherwise,} \end{cases}$$

where $\xi \in \mathrm{HJ}$ means that ξ satisfies the **Hamilton–Jacobi backward differential inequality**:

$$\boxed{\begin{cases} -\partial_t \xi(t,x) + \Psi^*(-D_x\xi(t,x)) \leqslant \Psi^*(-D\varphi(x)), \\ \xi(T,x) \leqslant \varphi(x). \end{cases}}$$

Hence the problem reduces to:

$$\min \mathcal{J} = \sup \Big\{ \xi(0,x_0) - \varphi(x_0) : \xi \in \mathrm{HJ} \Big\}.$$

It suffices to show the **r.h.s.** $\leqslant 0$.

[Links with optimal transport, first order mean field games, mean field planning, (mean field) optimal control: Otto-Villani, Cardaliaguet, Santambrogio, Porretta, ...]

The backward estimate (comparison/maximum principle for HJ)

Backward (upper) estimate

Any $\xi \in \mathrm{HJ}$ satisfies $\xi(t,x) \leqslant \varphi(x)$ for *all* $(t,x) \in X_T$.

The backward estimate (comparison/maximum principle for HJ)

Backward (upper) estimate

Any $\xi \in HJ$ satisfies $\xi(t,x) \leqslant \varphi(x)$ for *all* $(t,x) \in X_T$.

Interpretation (maximum principle): φ is a trivial, constant in time, solution of the associated Hamilton–Jacobi *equation*.

$$-\partial_t \varphi(x) + \Psi^*(-D_x\varphi(x)) = \Psi^*(-D\varphi(x)),$$

If the subsolution ξ satisfies $\xi(T,\cdot) \leqslant \varphi$ at the final time, the HJ dynamics propagates this bound *backward to all earlier times*.

Here ξ is **smooth**, so the argument is elementary.

The backward estimate (comparison/maximum principle for HJ)

Backward (upper) estimate

Any $\xi \in \mathrm{HJ}$ satisfies $\xi(t,x) \leqslant \varphi(x)$ for *all* $(t,x) \in X_T$.

Interpretation (maximum principle): φ is a trivial, constant in time, solution of the associated Hamilton–Jacobi *equation*.

$$-\partial_t \varphi(x) + \Psi^*(-D_x \varphi(x)) = \Psi^*(-D\varphi(x)),$$

If the subsolution ξ satisfies $\xi(T,\cdot) \leqslant \varphi$ at the final time, the HJ dynamics propagates this bound *backward to all earlier times*.

Here ξ is **smooth**, so the argument is elementary.

Consequence: $\xi(0,x_0) \leqslant \varphi(x_0)$, hence

$$\sup\left\{\xi(0,x_0) - \varphi(x_0) : \xi \in \mathrm{HJ}\right\} \leqslant 0 \quad \Longrightarrow \quad \boxed{\min \mathcal{J} \leqslant 0.}$$

Proof of the backward estimate

By contradiction. Perturbation: for $\varepsilon > 0$, set $\xi_\varepsilon(t, x) := \xi(t, x) + \varepsilon(t - T - 1)$ (strict HJ inequality).

Proof of the backward estimate

By contradiction. Perturbation: for $\varepsilon > 0$, set $\xi_\varepsilon(t,x) := \xi(t,x) + \varepsilon(t - T - 1)$ (strict HJ inequality).

Suppose $\xi_\varepsilon(\tilde{t},\tilde{x}) \geqslant \varphi(\tilde{x})$ for some $(\tilde{t},\tilde{x})$. Define

$$\bar{t} := \sup\left\{t \in [0,T] : \exists x \in \mathbb{R}^d \text{ with } \xi_\varepsilon(t,x) \geqslant \varphi(x)\right\}.$$

Proof of the backward estimate

By contradiction. Perturbation: for $\varepsilon > 0$, set $\xi_\varepsilon(t,x) := \xi(t,x) + \varepsilon(t - T - 1)$ (strict HJ inequality).

Suppose $\xi_\varepsilon(\tilde{t},\tilde{x}) \geqslant \varphi(\tilde{x})$ for some $(\tilde{t},\tilde{x})$. Define

$$\bar{t} := \sup\left\{t \in [0,T] : \exists x \in \mathbb{R}^d \text{ with } \xi_\varepsilon(t,x) \geqslant \varphi(x)\right\}.$$

By coercivity of φ, $\bar{t}$ is attained at some $\bar{x}$. Then:

- $\bar{t} < T$ (from the terminal condition $\xi_\varepsilon(T,x) \leqslant \varphi(x) - \varepsilon$);
- $\xi_\varepsilon(\bar{t},\bar{x}) = \varphi(\bar{x})$;
- $\partial_t \xi_\varepsilon(\bar{t},\bar{x}) \leqslant 0$ and $D_x \xi_\varepsilon(\bar{t},\bar{x}) = D\varphi(\bar{x})$.

Proof of the backward estimate

By contradiction. Perturbation: for $\varepsilon > 0$, set $\xi_\varepsilon(t,x) := \xi(t,x) + \varepsilon(t - T - 1)$ (strict HJ inequality).

Suppose $\xi_\varepsilon(\tilde{t},\tilde{x}) \geqslant \varphi(\tilde{x})$ for some $(\tilde{t},\tilde{x})$. Define

$$\bar{t} := \sup\left\{t \in [0,T] : \exists x \in \mathbb{R}^d \text{ with } \xi_\varepsilon(t,x) \geqslant \varphi(x)\right\}.$$

By coercivity of φ, $\bar{t}$ is attained at some $\bar{x}$. Then:

- $\bar{t} < T$ (from the terminal condition $\xi_\varepsilon(T,x) \leqslant \varphi(x) - \varepsilon$);
- $\xi_\varepsilon(\bar{t},\bar{x}) = \varphi(\bar{x})$;
- $\partial_t \xi_\varepsilon(\bar{t},\bar{x}) \leqslant 0$ and $D_x \xi_\varepsilon(\bar{t},\bar{x}) = D\varphi(\bar{x})$.

Plugging into the HJ inequality:

$$0 \leqslant -\partial_t \xi_\varepsilon(\bar{t},\bar{x}) = -\partial_t \xi_\varepsilon(\bar{t},\bar{x}) + \underbrace{\Psi^*(-D_x\xi_\varepsilon(\bar{t},\bar{x})) - \Psi^*(-D\varphi(\bar{x}))}_{=0}$$

$< 0.$ Contradiction!

Dynamic programming viewpoint

The connection with HJ is natural from **optimal control**.

The **value function**

$$V(t,x) := \inf \left\{ \int_t^T \Big(\Psi(v) + \Psi^*(-D\varphi(\mathbf{x})) \Big) d\tau + \varphi(\mathbf{x}(T)) : \dot{\mathbf{x}} = v,\ \mathbf{x}(t) = x \right\}$$

solves the Hamilton–Jacobi equation

$$\begin{cases} -\partial_t V + \Psi^*(-D_x V) = \Psi^*(-D\varphi(x)), \\ V(T,x) = \varphi(x). \end{cases}$$

Dynamic programming viewpoint

The connection with HJ is natural from **optimal control**.

The **value function**

$$V(t,x) := \inf\left\{\int_t^T \Big(\Psi(v) + \Psi^*(-D\varphi(\boldsymbol{x}))\Big)d\tau + \varphi(\boldsymbol{x}(T)) : \dot{\boldsymbol{x}} = v,\ \boldsymbol{x}(t) = x\right\}$$

solves the Hamilton–Jacobi equation

$$\begin{cases} -\partial_t V + \Psi^*(-D_x V) = \Psi^*(-D\varphi(x)), \\ V(T,x) = \varphi(x). \end{cases}$$

Since φ solves the HJ equation, uniqueness gives $V \equiv \varphi$, hence $\min \mathcal{J} = V(0,x_0) - \varphi(x_0) = 0$.

Dynamic programming viewpoint

The connection with HJ is natural from **optimal control**.

The **value function**

$$V(t,x) := \inf \left\{ \int_t^T \Big(\Psi(v) + \Psi^*(-D\varphi(\boldsymbol{x})) \Big) d\tau + \varphi(\boldsymbol{x}(T)) : \dot{\boldsymbol{x}} = v,\ \boldsymbol{x}(t) = x \right\}$$

solves the Hamilton–Jacobi equation

$$\begin{cases} -\partial_t V + \Psi^*(-D_x V) = \Psi^*(-D\varphi(x)), \\ V(T,x) = \varphi(x). \end{cases}$$

Since φ solves the HJ equation, uniqueness gives $V \equiv \varphi$, hence $\min \mathcal{J} = V(0,x_0) - \varphi(x_0) = 0$.

But: this requires smoothness and uniqueness of HJ.

Our approach replaces *uniqueness* with a *comparison principle* for **subsolutions** — works in nonsmooth, infinite-dimensional settings.

Outline

1. Variational principles for evolution equations
2. De Giorgi's principle and main result
3. Three key arguments
   - Relaxation and continuity equation
   - Von Neumann duality
   - Backward estimates and maximum principle for first order HJ equations
4. **The infinite-dimensional setting**
5. Extensions and open problems

The general setting

So far: $\mathbb{R}^d$, smooth φ and Ψ.

The general setting

So far: $\mathbb{R}^d$, smooth φ and Ψ.

Now: X reflexive separable Banach space, $\varphi\colon X \to (-\infty, +\infty]$ l.s.c., proper, $\Psi\colon X \to [0, +\infty]$ convex, l.s.c., superlinear, 0 is an interior point of the domain of Ψ.

Doubly nonlinear equation

$$\partial\Psi(\dot{x}(t)) + \partial\varphi(x(t)) \ni 0, \qquad x(0) = x_0.$$

So far: $\mathbb{R}^d$, smooth φ and Ψ.

Now: X reflexive separable Banach space, $\varphi\colon X \to (-\infty, +\infty]$ l.s.c., proper, $\Psi\colon X \to [0, +\infty]$ convex, l.s.c., superlinear, 0 is an interior point of the domain of Ψ.

Doubly nonlinear equation

$$\partial\Psi(\dot{\boldsymbol{x}}(t)) + \partial\varphi(\boldsymbol{x}(t)) \ni 0, \qquad \boldsymbol{x}(0) = x_0.$$

New difficulties:

- φ is not smooth $\Rightarrow$ need a notion of **subdifferential** $\partial\varphi$.
- The slope $S(x) = \min_{z\in\partial\varphi(x)} \Psi^*(-z)$ replaces $\Psi^*(-D\varphi(x))$.
- The chain rule $\frac{d}{dt}\varphi(\boldsymbol{x}(t)) = \langle z(t), \dot{\boldsymbol{x}}(t)\rangle$, $z(t) \in \partial\varphi(\boldsymbol{x}(t))$, is no longer automatic.

The general setting

So far: $\mathbb{R}^d$, smooth φ and Ψ.

Now: X reflexive separable Banach space, $\varphi\colon X \to (-\infty, +\infty]$ l.s.c., proper, $\Psi\colon X \to [0, +\infty]$ convex, l.s.c., superlinear, 0 is an interior point of the domain of Ψ.

Doubly nonlinear equation

$$\partial\Psi(\dot{x}(t)) + \partial\varphi(x(t)) \ni 0, \qquad x(0) = x_0.$$

New difficulties:

- φ is not smooth $\Rightarrow$ need a notion of **subdifferential** $\partial\varphi$.
- The slope $S(x) = \min_{z\in\partial\varphi(x)} \Psi^*(-z)$ replaces $\Psi^*(-D\varphi(x))$.
- The chain rule $\frac{d}{dt}\varphi(x(t)) = \langle z(t), \dot{x}(t)\rangle$, $z(t) \in \partial\varphi(x(t))$, is no longer automatic.

Even the Hilbert gradient flow $\dot{x} + \partial\varphi(x) \ni 0$

is nontrivial: φ is **not assumed convex** nor a quadratic perturbation of a convex function.

(Colli–Visintin; Colli; DiBenedetto–Showalter; Rossi–S.; Mielke–Rossi–S.)

The subdifferential in Banach spaces

X reflexive separable Banach space, $\varphi\colon X \to (-\infty, +\infty]$ l.s.c.

Fréchet (viscosity) subdifferential

$X^* \ni z \in \partial\varphi(x)$ if there exists a C^1 function g with $Dg(x) = z$ and $\varphi - g$ has a local minimum at x.

X reflexive separable Banach space, $\varphi\colon X \to (-\infty, +\infty]$ l.s.c.

Fréchet (viscosity) subdifferential

$X^* \ni z \in \partial\varphi(x)$ if there exists a C^1 function g with $Dg(x) = z$ and $\varphi - g$ has a local minimum at x.

Equivalently: $z \in \partial\varphi(x) \iff \liminf_{y \to x} \dfrac{\varphi(y) - \varphi(x) - \langle z, y - x \rangle}{\|y - x\|} \geqslant 0.$

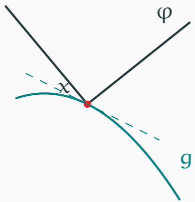
The subdifferential in Banach spaces

X reflexive separable Banach space, $\varphi\colon X \to (-\infty, +\infty]$ l.s.c.

Fréchet (viscosity) subdifferential

$X^* \ni z \in \partial\varphi(x)$ if there exists a C^1 function g with $Dg(x) = z$ and $\varphi - g$ has a local minimum at x.

Equivalently: $z \in \partial\varphi(x) \iff \liminf_{y\to x} \dfrac{\varphi(y) - \varphi(x) - \langle z, y-x\rangle}{\|y-x\|} \geqslant 0.$



Viscosity viewpoint: $z \in \partial\varphi(x)$ means that a smooth function g *touches* φ *from below* at x with $Dg(x) = z$.

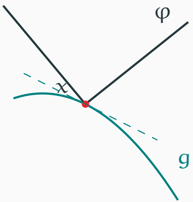
In the proof, the test function ξ touches φ from below at the critical point, giving $D_x\xi \in \partial\varphi(\bar{x})$.

X reflexive separable Banach space, $\varphi\colon X \to (-\infty, +\infty]$ l.s.c.

Fréchet (viscosity) subdifferential

$X^* \ni z \in \partial\varphi(x)$ if there exists a C^1 function g with $Dg(x) = z$ and $\varphi - g$ has a local minimum at x.

Equivalently: $z \in \partial\varphi(x) \iff \displaystyle\liminf_{y\to x} \frac{\varphi(y) - \varphi(x) - \langle z, y - x\rangle}{\|y - x\|} \geqslant 0.$



Viscosity viewpoint: $z \in \partial\varphi(x)$ means that a smooth function g *touches* φ *from below* at x with $Dg(x) = z$.

In the proof, the test function ξ touches φ from below at the critical point, giving $D_x\xi \in \partial\varphi(\bar{x})$.

Assumption: $\partial\varphi$ is *strong-weak** *sequentially closed*.

Then $S(x) = \min_{z\in\partial\varphi(x)} \Psi^*(-z)$ is l.s.c.

General assumptions and the slope

$\varphi\colon X \to (-\infty, +\infty]$ l.s.c., with boundedly compact sublevels, bounded below.

$\Psi\colon X \to (-\infty, +\infty]$ convex, l.s.c., superlinear, *finite in a neighborhood of* 0: Ψ^* is **weakly $*$ coercive** (bounded sublevels).

With closed $\partial\varphi$: $S(x) = \min_{z \in \partial\varphi(x)} \Psi^*(-z)$ is l.s.c.

General assumptions and the slope

$\varphi \colon X \to (-\infty, +\infty]$ l.s.c., with boundedly compact sublevels, bounded below.

$\Psi \colon X \to (-\infty, +\infty]$ convex, l.s.c., superlinear, *finite in a neighborhood of* 0: Ψ^* is **weakly * coercive** (bounded sublevels).

With closed $\partial\varphi$: $S(x) = \min_{z\in\partial\varphi(x)} \Psi^*(-z)$ is l.s.c.

Theorem (Pinzi–Riva–S., first version)

If S *has* **compact sublevels**, *then*

$$\min\left\{\mathcal{J}(T, \boldsymbol{x}) : \boldsymbol{x} \in AC([0,T]; X),\ \boldsymbol{x}(0) = x_0\right\} \leqslant 0.$$

General assumptions and the slope

$\varphi: X \to (-\infty, +\infty]$ l.s.c., with boundedly compact sublevels, bounded below.

$\Psi: X \to (-\infty, +\infty]$ convex, l.s.c., superlinear, *finite in a neighborhood of* 0: Ψ^* is **weakly * coercive** (bounded sublevels).

With closed $\partial\varphi$: $S(x) = \min_{z \in \partial\varphi(x)} \Psi^*(-z)$ is l.s.c.

Theorem (Pinzi–Riva–S., first version)

If S *has* **compact sublevels**, *then*

$$\min \Big\{ \mathcal{J}(T, \boldsymbol{x}) : \boldsymbol{x} \in AC([0,T]; X),\ \boldsymbol{x}(0) = x_0 \Big\} \leqslant 0.$$

The proof follows the same three steps.

Compactness of sublevels of S ensures tightness of μ in the Von Neumann argument (replacing the coercivity of $|D\varphi|$ used in $\mathbb{R}^d$).

The exponentially weighted functional

When S *does not* have compact sublevels, we gain coercivity via an **exponential weight** ($a > 0$).

Trick: differentiate $e^{-at}\varphi(\mathbf{x}(t))$ instead of $\varphi(\mathbf{x}(t))$: $\frac{d}{dt}\big(e^{-at}\varphi(\mathbf{x})\big) = e^{-at}\big(\langle z, \dot{\mathbf{x}}\rangle \boldsymbol{-a\,\varphi(\mathbf{x})}\big)$.

Weighted functional:

$$\mathcal{J}^a(t, \mathbf{x}) := e^{-at}\varphi(\mathbf{x}(t)) + \int_0^t e^{-a\tau}\Big(\Psi(\dot{\mathbf{x}}(\tau)) + S(\mathbf{x}(\tau)) \boldsymbol{+a\,\varphi(\mathbf{x}(\tau))}\Big)\,d\tau - \varphi(x_0).$$

The exponentially weighted functional

When S *does not* have compact sublevels, we gain coercivity via an **exponential weight** $(a > 0)$.

Trick: differentiate $e^{-at}\varphi(\mathbf{x}(t))$ instead of $\varphi(\mathbf{x}(t))$: $\frac{d}{dt}\big(e^{-at}\varphi(\mathbf{x})\big) = e^{-at}\big(\langle z, \dot{\mathbf{x}}\rangle \boldsymbol{-a\,\varphi(\mathbf{x})}\big)$.

Weighted functional:

$$\mathcal{J}^a(t,\mathbf{x}) := e^{-at}\varphi(\mathbf{x}(t)) + \int_0^t e^{-a\tau}\Big(\Psi(\dot{\mathbf{x}}(\tau)) + S(\mathbf{x}(\tau)) \boldsymbol{+a\,\varphi(\mathbf{x}(\tau))}\Big)\,d\tau - \varphi(x_0).$$

The integrand now contains $S + a\varphi$: for $a > 0$ large enough, $S + a\varphi$ **has compact sublevels** even if S alone does not.

The exponentially weighted functional

When S *does not* have compact sublevels, we gain coercivity via an **exponential weight** ($a > 0$).

Trick: differentiate $e^{-at}\varphi(x(t))$ instead of $\varphi(x(t))$: $\frac{d}{dt}\big(e^{-at}\varphi(x)\big) = e^{-at}\big(\langle z, \dot{x}\rangle - \mathbf{a}\,\boldsymbol{\varphi}(\mathbf{x})\big)$.

Weighted functional:

$$\mathcal{J}^a(t, x) := e^{-at}\varphi(x(t)) + \int_0^t e^{-a\tau}\Big(\Psi(\dot{x}(\tau)) + S(x(\tau)) + \mathbf{a}\,\boldsymbol{\varphi}(\mathbf{x}(\boldsymbol{\tau}))\Big)\,d\tau - \varphi(x_0).$$

The integrand now contains $S + a\varphi$: for $a > 0$ large enough, $S + a\varphi$ **has compact sublevels** even if S alone does not.

Theorem (Pinzi–Riva–S., general version)

For all $a \geqslant 0$ such that $S + a\varphi$ has compact sublevels:

$$\min\Big\{\mathcal{J}^a(T, x) : x \in AC([0, T]; X),\ x(0) = x_0\Big\} \leqslant 0.$$

The exponentially weighted functional

When S *does not* have compact sublevels, we gain coercivity via an **exponential weight** ($a > 0$).

Trick: differentiate $e^{-at}\varphi(x(t))$ instead of $\varphi(x(t))$: $\frac{d}{dt}\big(e^{-at}\varphi(x)\big) = e^{-at}\big(\langle z, \dot{x}\rangle \boldsymbol{- a\,\varphi(x)}\big)$.

Weighted functional:

$$\mathcal{J}^a(t, x) := e^{-at}\varphi(x(t)) + \int_0^t e^{-a\tau}\Big(\Psi(\dot{x}(\tau)) + S(x(\tau)) \boldsymbol{+ a\,\varphi(x(\tau))}\Big)\,d\tau - \varphi(x_0).$$

The integrand now contains $S + a\varphi$: for $a > 0$ large enough, $S + a\varphi$ **has compact sublevels** even if S alone does not.

Theorem (Pinzi–Riva–S., general version)

For all $a \geqslant 0$ *such that* $S + a\varphi$ *has compact sublevels:*

$$\min\Big\{\mathcal{J}^a(T, x) : x \in AC([0,T]; X),\ x(0) = x_0\Big\} \leqslant 0.$$

For $a = 0$: recovers the first version.

For $\boldsymbol{a > 0}$: coercivity of φ is sufficient.

The chain rule and solutions

Chain Rule (CR)

For every $x \in AC([0,T];X)$ with $\int_0^T (\Psi(\dot{x}) + S(x))\,d\tau < +\infty$,

$t \mapsto \varphi(x(t))$ is absolutely continuous and $\frac{d}{dt}\varphi(x(t)) = \langle z(t), \dot{x}(t)\rangle$ for any selection $z(t) \in \partial\varphi(x(t))$.

The chain rule and solutions

Chain Rule (CR)

For every $x \in AC([0,T];X)$ with $\int_0^T (\Psi(\dot{x}) + S(x))\, d\tau < +\infty$,

$t \mapsto \varphi(x(t))$ is absolutely continuous and $\dfrac{d}{dt}\varphi(x(t)) = \langle z(t), \dot{x}(t)\rangle$ for any selection $z(t) \in \partial\varphi(x(t))$.

The upper bound $\min \mathcal{J}^{a} \leqslant 0$ **holds without (CR).**

Chain Rule (CR)

For every $x \in AC([0,T];X)$ with $\int_0^T (\Psi(\dot{x}) + S(x))\,d\tau < +\infty$,

$t \mapsto \varphi(x(t))$ is absolutely continuous and $\frac{d}{dt}\varphi(x(t)) = \langle z(t), \dot{x}(t)\rangle$ for any selection $z(t) \in \partial\varphi(x(t))$.

The upper bound $\min \mathcal{J}^a \leqslant 0$ holds without (CR).

The chain rule is an *independent* property:

- If (CR) holds: $\mathcal{J}^a \geqslant 0$ always $\Rightarrow \min \mathcal{J}^a = 0 \Rightarrow$ minimizers are solutions.
- If (CR) fails: $\min \mathcal{J}^a$ could be < 0 (no guarantee that minimizers solve the equation).

Chain Rule (CR)

For every $x \in AC([0,T];X)$ with $\int_0^T (\Psi(\dot{x}) + S(x))\, d\tau < +\infty$,

$t \mapsto \varphi(x(t))$ is absolutely continuous and $\dfrac{d}{dt}\varphi(x(t)) = \langle z(t), \dot{x}(t)\rangle$ for any selection $z(t) \in \partial\varphi(x(t))$.

The upper bound $\min \mathcal{J}^a \leqslant 0$ holds without (CR).

The chain rule is an *independent* property:

- If (CR) holds: $\mathcal{J}^a \geqslant 0$ always $\Rightarrow \min \mathcal{J}^a = 0 \Rightarrow$ minimizers are solutions.
- If (CR) fails: $\min \mathcal{J}^a$ could be < 0 (no guarantee that minimizers solve the equation).

In the smooth $\mathbb{R}^d$ case, (CR) holds automatically.

In infinite dimensions, (CR) requires additional structure (e.g. λ-convexity or chain rule conditions of Rossi–Mielke–S.).

Key issues in infinite dimensions

The three-step strategy extends, with additional difficulties:

The three-step strategy extends, with additional difficulties:

1. **Continuity equation in Banach spaces.**

- Test functions: **cylinder functions** $\xi(t, x) = \zeta(t, \langle z_1, x\rangle, \ldots, \langle z_k, x\rangle)$, $\zeta \in C^1_b$.
- Superposition principle still holds (Ambrosio–Trevisan).

Key issues in infinite dimensions

The three-step strategy extends, with additional difficulties:

1. Continuity equation in Banach spaces.

- Test functions: **cylinder functions** $\xi(t,x) = \zeta(t, \langle z_1, x\rangle, \dots, \langle z_k, x\rangle)$, $\zeta \in C_b^1$.
- Superposition principle still holds (Ambrosio–Trevisan).

2. Narrow topology for Banach-valued measures.

- $\nu \in \mathcal{M}(X_T; X)$ — need a topology weak enough for compactness, strong enough for l.s.c.
- Duality with $\zeta(t,x)\, z$, $\zeta \in C_b$, $z \in X^*$ (inspired by Schwartz).
- Local metrizability and compactness criterion.

Key issues in infinite dimensions

The three-step strategy extends, with additional difficulties:

1. Continuity equation in Banach spaces.

- Test functions: **cylinder functions** $\xi(t,x) = \zeta(t, \langle z_1, x\rangle, \dots, \langle z_k, x\rangle)$, $\zeta \in C_b^1$.
- Superposition principle still holds (Ambrosio–Trevisan).

2. Narrow topology for Banach-valued measures.

- $\nu \in \mathcal{M}(X_T; X)$ — need a topology weak enough for compactness, strong enough for l.s.c.
- Duality with $\zeta(t,x)\,z$, $\zeta \in C_b$, $z \in X^*$ (inspired by Schwartz).
- Local metrizability and compactness criterion.

3. Hamilton–Jacobi inequality and backward estimates.

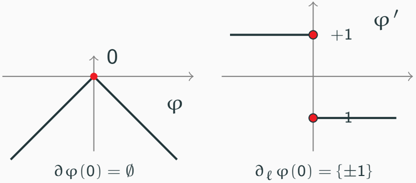
- Cylinder test functions $\Rightarrow$ the HJ constraint involves *finitely many* dual variables.
- The viscosity interpretation of $\partial\varphi$ is crucial: ξ touches φ from below $\Rightarrow D_x\xi \in \partial\varphi(\bar{x})$.

When the subdifferential is not closed

If $\partial\varphi$ is not strong-weak* closed, one uses:

- The **limiting subdifferential** $\partial_\ell\varphi(x) = \{z : \exists\, x_n \to x,\ z_n \in \partial\varphi(x_n),\ z_n \rightharpoonup z\}$.
- The **relaxed slope** $S^-(x) = \inf\left\{ \liminf_n \Psi^*(-z_n) : z_n \in \partial\varphi(x_n),\ x_n \to x \right\}$.

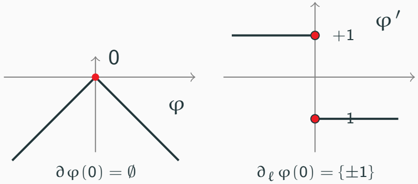
S^- is l.s.c. by construction and $S^- \leqslant S$.



When the subdifferential is not closed

If $\partial\varphi$ is not strong-weak* closed, one uses:

- The **limiting subdifferential**
  $\partial_\ell\varphi(x) = \{z : \exists\, x_n \to x,\ z_n \in \partial\varphi(x_n),\ z_n \rightharpoonup z\}$.
- The **relaxed slope** $S^-(x) = \inf\left\{ \liminf_n \Psi^*(-z_n) :\right.$
  $\left. z_n \in \partial\varphi(x_n),\ x_n \to x \right\}$.



S^- is l.s.c. by construction and $S^- \leqslant S$.

The full theory applies with S^- replacing S and $\partial_\ell\varphi$ replacing $\partial\varphi$ in the doubly nonlinear equation.

Outline

1. Variational principles for evolution equations
2. De Giorgi's principle and main result
3. Three key arguments
   - Relaxation and continuity equation
   - Von Neumann duality
   - Backward estimates and maximum principle for first order HJ equations
4. The infinite-dimensional setting
5. **Extensions and open problems**

The method naturally extends to **non-autonomous** problems:

$$\partial_v \Psi(t, \boldsymbol{x}(t), \dot{\boldsymbol{x}}(t)) + \partial \varphi(t, \boldsymbol{x}(t)) \ni 0.$$

Time- and state-dependent dissipation

The method naturally extends to **non-autonomous** problems:

$$\partial_v \Psi(t, \boldsymbol{x}(t), \dot{\boldsymbol{x}}(t)) + \partial \varphi(t, \boldsymbol{x}(t)) \ni 0.$$

Example: quadratic dissipation with varying friction,

$$\Psi(t, x, v) = \frac{1}{2} a(t, x)|v|^2, \qquad \Psi^*(t, x, z) = \frac{1}{2a(t, x)}|z|^2.$$

Time- and state-dependent dissipation

The method naturally extends to **non-autonomous** problems:

$$\partial_v \Psi(t, x(t), \dot{x}(t)) + \partial \varphi(t, x(t)) \ni 0.$$

Example: quadratic dissipation with varying friction,

$$\Psi(t, x, v) = \tfrac{1}{2} a(t, x) |v|^2, \qquad \Psi^*(t, x, z) = \tfrac{1}{2a(t,x)} |z|^2.$$

The proof carries over: all three ingredients apply without essential modifications.

The method naturally extends to **non-autonomous** problems:

$$\partial_v \Psi(t, \mathbf{x}(t), \dot{\mathbf{x}}(t)) + \partial \varphi(t, \mathbf{x}(t)) \ni 0.$$

Example: quadratic dissipation with varying friction,

$$\Psi(t, x, v) = \tfrac{1}{2} a(t, x) |v|^2, \qquad \Psi^*(t, x, z) = \tfrac{1}{2a(t,x)} |z|^2.$$

The proof carries over: all three ingredients apply without essential modifications.

When φ depends on time, a further term $\partial_t \varphi(t, \mathbf{x}(t))$ appears in the De Giorgi functional.

Open problems and perspectives

1. **Gradient flows in metric spaces.**

   The De Giorgi functional makes sense in (X, d) metric (via metric derivative and slope).
   Can the duality approach be adapted? The continuity equation and superposition principle are available (Lisini)

Open problems and perspectives

1. **Gradient flows in metric spaces.**

   The De Giorgi functional makes sense in (X, d) metric (via metric derivative and slope).
   Can the duality approach be adapted? The continuity equation and superposition principle are available (Lisini)

2. **Monotone evolutions.**

   Can one formulate analogous variational principles for evolutions governed by maximal monotone operators $A\dot{x}(t) + \partial\varphi(x(t)) \ni 0$ without gradient structure?

Open problems and perspectives

1. **Gradient flows in metric spaces.**

   The De Giorgi functional makes sense in (X, d) metric (via metric derivative and slope). Can the duality approach be adapted? The continuity equation and superposition principle are available (Lisini)

2. **Monotone evolutions.**

   Can one formulate analogous variational principles for evolutions governed by maximal monotone operators $A\dot{x}(t) + \partial\varphi(x(t)) \ni 0$ without gradient structure?

3. **Rate-independent systems.**

   The case of 1-homogeneous dissipation Ψ (linear growth) leads to BV solutions and discontinuities — a fundamentally different regime.

Thank you!

Reference: A. Pinzi, F. Riva, G. Savaré, *A direct method for doubly nonlinear equations via convexification in spaces of measures and duality*, preprint 2026.