

# Lecture 2: Gradient Flows, Decay Rates and Contractions of Transport Distances

J. A. Carrillo

ICREA - Universitat Autònoma de Barcelona

OT Tutorials - IPAM

# Outline

- 1 Schematic Cartoon of Rate Arguments
  - Finite Dimensional Gradient Flows
  - First Example of Gradient Flow
- 2 Gradient Flow Structure
  - Gradient Flow
  - Convexity
- 3 Decay Rates
  - Conclusion
  - Theorems

# Outline

- 1 Schematic Cartoon of Rate Arguments
  - Finite Dimensional Gradient Flows
  - First Example of Gradient Flow
- 2 Gradient Flow Structure
  - Gradient Flow
  - Convexity
- 3 Decay Rates
  - Conclusion
  - Theorems

# Sliding down in a Convex Valley

Contraction / expansion bounds in a semi-convex valley:

Fix  $\alpha \in \mathbb{R}$ . If  $E \in C^2(\mathbb{R}^d)$  satisfies  $D^2E(x) \geq \alpha I_d$  throughout  $\mathbb{R}^d$ , and the curves  $t \in [0, \infty) \rightarrow x_t, y_t \in \mathbb{R}^d$  both solve the ODE:

$$\frac{dx_t}{dt} = -\nabla E(x_t),$$

corresponding to the *steepest descent or gradient flow* on the energy (entropy) landscape determined by  $E$ . Then  $|x_{t+t_0} - y_{t+t_0}| \leq e^{-\alpha t} |x_{t_0} - y_{t_0}|$ .

Proof.

Set  $\Psi(t) = |x_t - y_t|^2/2$ . Then

$$\begin{aligned} \Psi'(t) &= -\langle x_t - y_t, \nabla E(x_t) - \nabla E(y_t) \rangle \\ &= -\langle x_t - y_t, \int_0^1 D^2E[(1-s)x_t + sy_t] (y_t - x_t) ds \rangle \leq -2\alpha\Psi(t). \end{aligned}$$

Gronwall's inequality implies the desired result:  $\Psi(t + t_0) \leq e^{-2\alpha t} \Psi(t_0)$ . □

# Sliding down in a Convex Valley

Contraction / expansion bounds in a semi-convex valley:

Fix  $\alpha \in \mathbb{R}$ . If  $E \in C^2(\mathbb{R}^d)$  satisfies  $D^2E(x) \geq \alpha I_d$  throughout  $\mathbb{R}^d$ , and the curves  $t \in [0, \infty) \rightarrow x_t, y_t \in \mathbb{R}^d$  both solve the ODE:

$$\frac{dx_t}{dt} = -\nabla E(x_t),$$

corresponding to the *steepest descent or gradient flow* on the energy (entropy) landscape determined by  $E$ . Then  $|x_{t+t_0} - y_{t+t_0}| \leq e^{-\alpha t} |x_{t_0} - y_{t_0}|$ .

Proof.

Set  $\Psi(t) = |x_t - y_t|^2/2$ . Then

$$\begin{aligned} \Psi'(t) &= -\langle x_t - y_t, \nabla E(x_t) - \nabla E(y_t) \rangle \\ &= -\langle x_t - y_t, \int_0^1 D^2E[(1-s)x_t + sy_t] (y_t - x_t) ds \rangle \leq -2\alpha \Psi(t). \end{aligned}$$

Gronwall's inequality implies the desired result:  $\Psi(t + t_0) \leq e^{-2\alpha t} \Psi(t_0)$ . □

# Sliding down in a Convex Valley

Contraction / expansion bounds in a semi-convex valley:

Fix  $\alpha \in \mathbb{R}$ . If  $E \in C^2(\mathbb{R}^d)$  satisfies  $D^2E(x) \geq \alpha I_d$  throughout  $\mathbb{R}^d$ , and the curves  $t \in [0, \infty) \rightarrow x_t, y_t \in \mathbb{R}^d$  both solve the ODE:

$$\frac{dx_t}{dt} = -\nabla E(x_t),$$

corresponding to the *steepest descent or gradient flow* on the energy (entropy) landscape determined by  $E$ . Then  $|x_{t+t_0} - y_{t+t_0}| \leq e^{-\alpha t} |x_{t_0} - y_{t_0}|$ .

**Proof.**

Set  $\Psi(t) = |x_t - y_t|^2/2$ . Then

$$\begin{aligned} \Psi'(t) &= -\langle x_t - y_t, \nabla E(x_t) - \nabla E(y_t) \rangle \\ &= -\langle x_t - y_t, \int_0^1 D^2E[(1-s)x_t + sy_t] (y_t - x_t) ds \rangle \leq -2\alpha\Psi(t). \end{aligned}$$

Gronwall's inequality implies the desired result:  $\Psi(t + t_0) \leq e^{-2\alpha t} \Psi(t_0)$ . □

# Sliding down in a Convex Valley

## Convergence rates towards equilibrium: contraction

Assume  $\alpha > 0$ , the uniform convexity of  $E$  implies that the solution map

$$x_0 \in \mathbb{R}^d \longrightarrow X_t(x_0) = x_t$$

of the ODE defines a uniform contraction on  $\mathbb{R}^d$  for each  $t > 0$ . Since  $\mathbb{R}^d$  is complete, the contraction mapping principle dictates that this map has a unique fixed point  $X_t(x_\infty) = x_\infty \in \mathbb{R}^d$ , and each solution curve  $x_t = X_t(x_0)$  must converge to  $x_\infty$  in the long time limit  $t \rightarrow \infty$ .

# Outline

- 1 Schematic Cartoon of Rate Arguments
  - Finite Dimensional Gradient Flows
  - First Example of Gradient Flow
- 2 Gradient Flow Structure
  - Gradient Flow
  - Convexity
- 3 Decay Rates
  - Conclusion
  - Theorems

# Probability Measures sliding down in a Convex Valley

Let us consider the continuity equation with given velocity field  $-\nabla V$  with  $V$  uniformly convex potential on  $\mathbb{R}^d$ , i.e.,

$$\frac{\partial \rho}{\partial t} = \nabla \cdot (\rho \nabla V).$$

Assume without loss of generality the only minimum of  $V$  is 0.

We say that  $\rho \in C([0, T], \mathcal{P}(\mathbb{R}^d))$  is a solution with initial value  $\mu \in \mathcal{P}(\mathbb{R}^d)$  if for all test functions  $\psi \in C_0^\infty$ :

$$\begin{aligned} & \int_0^T \int_{\mathbb{R}^d} \frac{\partial \psi}{\partial t} d\rho(t) dt + \int_{\mathbb{R}^d} \psi(0) d\mu \\ &= \int_0^T \int_{\mathbb{R}^d} (\nabla \psi \cdot \nabla V) d\rho(t) dt + \int_{\mathbb{R}^d} \psi(T) d\rho(T). \end{aligned}$$

# Probability Measures sliding down in a Convex Valley

Let us consider the continuity equation with given velocity field  $-\nabla V$  with  $V$  uniformly convex potential on  $\mathbb{R}^d$ , i.e.,

$$\frac{\partial \rho}{\partial t} = \nabla \cdot (\rho \nabla V).$$

Assume without loss of generality the only minimum of  $V$  is 0.

We say that  $\rho \in C([0, T], \mathcal{P}(\mathbb{R}^d))$  is a solution with initial value  $\mu \in \mathcal{P}(\mathbb{R}^d)$  if for all test functions  $\psi \in C_0^\infty$ :

$$\begin{aligned} & \int_0^T \int_{\mathbb{R}^d} \frac{\partial \psi}{\partial t} d\rho(t) dt + \int_{\mathbb{R}^d} \psi(0) d\mu \\ &= \int_0^T \int_{\mathbb{R}^d} (\nabla \psi \cdot \nabla V) d\rho(t) dt + \int_{\mathbb{R}^d} \psi(T) d\rho(T). \end{aligned}$$

# Probability Measures sliding down in a Convex Valley

Let us consider the continuity equation with given velocity field  $-\nabla V$  with  $V$  uniformly convex potential on  $\mathbb{R}^d$ , i.e.,

$$\frac{\partial \rho}{\partial t} = \nabla \cdot (\rho \nabla V).$$

Assume without loss of generality the only minimum of  $V$  is 0.

We say that  $\rho \in C([0, T], \mathcal{P}(\mathbb{R}^d))$  is a solution with initial value  $\mu \in \mathcal{P}(\mathbb{R}^d)$  if for all test functions  $\psi \in C_o^\infty$ :

$$\begin{aligned} & \int_0^T \int_{\mathbb{R}^d} \frac{\partial \psi}{\partial t} d\rho(t) dt + \int_{\mathbb{R}^d} \psi(0) d\mu \\ &= \int_0^T \int_{\mathbb{R}^d} (\nabla \psi \cdot \nabla V) d\rho(t) dt + \int_{\mathbb{R}^d} \psi(T) d\rho(T). \end{aligned}$$

# Probability Measures sliding down in a Convex Valley

Let  $\Phi_{s,t}(x)$  be the **flow map** of the finite dimensional gradient flow:

$$\frac{\partial r}{\partial \tau} = u(r) = -\nabla V(r) \quad s < \tau < t,$$

with  $r(s) = x \in \mathbb{R}^d$ , giving a diffeomorphism family of  $\mathbb{R}^d$  onto itself. Let  $\Phi_t(x) = \Phi_{0,t}(x)$ .

Taking  $\psi(t, x) = \varphi(\Phi_{t,T}(x))$  in the definition of weak solution, we get

$$\int_{\mathbb{R}^d} \varphi(\Phi_T(x)) d\mu = \int_{\mathbb{R}^d} \varphi d\rho(T)$$

for all  $\varphi(x) \in C_0^\infty(\mathbb{R}^d)$ . Thus, the solution of the Cauchy problem with initial data  $\mu \in \mathcal{P}(\mathbb{R}^d)$  is given by

$$\rho(t) = \Phi_t \# \mu.$$

# Probability Measures sliding down in a Convex Valley

Let  $\Phi_{s,t}(x)$  be the **flow map** of the finite dimensional gradient flow:

$$\frac{\partial r}{\partial \tau} = u(r) = -\nabla V(r) \quad s < \tau < t,$$

with  $r(s) = x \in \mathbb{R}^d$ , giving a diffeomorphism family of  $\mathbb{R}^d$  onto itself. Let  $\Phi_t(x) = \Phi_{0,t}(x)$ .

Taking  $\psi(t, x) = \varphi(\Phi_{t,T}(x))$  in the definition of weak solution, we get

$$\int_{\mathbb{R}^d} \varphi(\Phi_T(x)) d\mu = \int_{\mathbb{R}^d} \varphi d\rho(T)$$

for all  $\varphi(x) \in C_0^\infty(\mathbb{R}^d)$ . Thus, the solution of the Cauchy problem with initial data  $\mu \in \mathcal{P}(\mathbb{R}^d)$  is given by

$$\rho(t) = \Phi_t \# \mu.$$

# Probability Measures sliding down in a Convex Valley

## Contraction Property

Let  $\alpha \in \mathbb{R}^+$  and  $V \in C^2(\mathbb{R}^d)$  such that  $D^2V(x) \geq \alpha I_d$  in  $\mathbb{R}^d$  and  $D^2V(x) \leq C(1 + |x|^{p-2})I_d$  with  $p \geq 2$ . Given any two probability measure solutions  $\rho_1(t)$  and  $\rho_2(t)$ , we have

$$W_2(\rho_1(t), \rho_2(t)) \leq e^{-\alpha t} W_2(\rho_1(0), \rho_2(0)).$$

# Probability Measures sliding down in a Convex Valley

Proof.

Fix any  $t_0 \geq 0$ . Let  $\gamma_o$  be the optimal transference plan between  $\rho_1(t_0)$  and  $\rho_2(t_0)$ :

$$W_2^2(\rho_1(t_0), \rho_2(t_0)) = \int_{\mathbb{R}^d \times \mathbb{R}^d} |x - y|^2 d\gamma_o(x, y).$$

The solutions are given by  $\rho_1(t) = \Phi_t \# \rho_1(0)$  and  $\rho_2(t) = \Phi_t \# \rho_2(0)$ . Take  $t \geq t_0$ . Define  $\gamma_t = (\Phi_{t_0, t} \times \Phi_{t_0, t}) \# \gamma_o$ , it is an admissible plan between  $\rho_1(t)$  and  $\rho_2(t)$ , and thus

$$W_2^2(\rho_1(t), \rho_2(t)) \leq \int_{\mathbb{R}^d \times \mathbb{R}^d} |x - y|^2 d\gamma_t(x, y) = \int_{\mathbb{R}^d \times \mathbb{R}^d} |\Phi_{t_0, t}(x) - \Phi_{t_0, t}(y)|^2 d\gamma_o(x, y).$$

As a consequence taking the derivative from the right, we have

$$\left. \frac{d}{dt} \right|_{t_0^+} W_2^2(\rho_1(t), \rho_2(t))/2 \leq - \int_{\mathbb{R}^d \times \mathbb{R}^d} (x - y) \cdot (\nabla V(x) - \nabla V(y)) d\gamma_o(x, y)$$

Uniform convexity of  $V$  implies

$$\left. \frac{d}{dt} \right|_{t_0^+} W_2^2(\rho_1(t), \rho_2(t)) \leq -2\alpha W_2^2(\rho_1(0), \rho_2(0)).$$



# Probability Measures sliding down in a Convex Valley

Proof.

Fix any  $t_0 \geq 0$ . Let  $\gamma_o$  be the optimal transference plan between  $\rho_1(t_0)$  and  $\rho_2(t_0)$ :

$$W_2^2(\rho_1(t_0), \rho_2(t_0)) = \int_{\mathbb{R}^d \times \mathbb{R}^d} |x - y|^2 d\gamma_o(x, y).$$

The solutions are given by  $\rho_1(t) = \Phi_t \# \rho_1(0)$  and  $\rho_2(t) = \Phi_t \# \rho_2(0)$ . Take  $t \geq t_0$ . Define  $\gamma_t = (\Phi_{t_0, t} \times \Phi_{t_0, t}) \# \gamma_o$ , it is an admissible plan between  $\rho_1(t)$  and  $\rho_2(t)$ , and thus

$$W_2^2(\rho_1(t), \rho_2(t)) \leq \int_{\mathbb{R}^d \times \mathbb{R}^d} |x - y|^2 d\gamma_t(x, y) = \int_{\mathbb{R}^d \times \mathbb{R}^d} |\Phi_{t_0, t}(x) - \Phi_{t_0, t}(y)|^2 d\gamma_o(x, y).$$

As a consequence taking the derivative from the right, we have

$$\left. \frac{d}{dt} \right|_{t_0^+} W_2^2(\rho_1(t), \rho_2(t))/2 \leq - \int_{\mathbb{R}^d \times \mathbb{R}^d} (x - y) \cdot (\nabla V(x) - \nabla V(y)) d\gamma_o(x, y)$$

Uniform convexity of  $V$  implies

$$\left. \frac{d}{dt} \right|_{t_0^+} W_2^2(\rho_1(t), \rho_2(t)) \leq -2\alpha W_2^2(\rho_1(0), \rho_2(0)).$$



# Probability Measures sliding down in a Convex Valley

Proof.

Fix any  $t_0 \geq 0$ . Let  $\gamma_o$  be the optimal transference plan between  $\rho_1(t_0)$  and  $\rho_2(t_0)$ :

$$W_2^2(\rho_1(t_0), \rho_2(t_0)) = \int_{\mathbb{R}^d \times \mathbb{R}^d} |x - y|^2 d\gamma_o(x, y).$$

The solutions are given by  $\rho_1(t) = \Phi_t \# \rho_1(0)$  and  $\rho_2(t) = \Phi_t \# \rho_2(0)$ . Take  $t \geq t_0$ . Define  $\gamma_t = (\Phi_{t_0, t} \times \Phi_{t_0, t}) \# \gamma_o$ , it is an admissible plan between  $\rho_1(t)$  and  $\rho_2(t)$ , and thus

$$W_2^2(\rho_1(t), \rho_2(t)) \leq \int_{\mathbb{R}^d \times \mathbb{R}^d} |x - y|^2 d\gamma_t(x, y) = \int_{\mathbb{R}^d \times \mathbb{R}^d} |\Phi_{t_0, t}(x) - \Phi_{t_0, t}(y)|^2 d\gamma_o(x, y).$$

As a consequence taking the derivative from the right, we have

$$\left. \frac{d}{dt} \right|_{t_0^+} W_2^2(\rho_1(t), \rho_2(t)) / 2 \leq - \int_{\mathbb{R}^d \times \mathbb{R}^d} (x - y) \cdot (\nabla V(x) - \nabla V(y)) d\gamma_o(x, y)$$

Uniform convexity of  $V$  implies

$$\left. \frac{d}{dt} \right|_{t_0^+} W_2^2(\rho_1(t), \rho_2(t)) \leq -2\alpha W_2^2(\rho_1(0), \rho_2(0)).$$



# Probability Measures sliding down in a Convex Valley

Proof.

Fix any  $t_0 \geq 0$ . Let  $\gamma_o$  be the optimal transference plan between  $\rho_1(t_0)$  and  $\rho_2(t_0)$ :

$$W_2^2(\rho_1(t_0), \rho_2(t_0)) = \int_{\mathbb{R}^d \times \mathbb{R}^d} |x - y|^2 d\gamma_o(x, y).$$

The solutions are given by  $\rho_1(t) = \Phi_t \# \rho_1(0)$  and  $\rho_2(t) = \Phi_t \# \rho_2(0)$ . Take  $t \geq t_0$ . Define  $\gamma_t = (\Phi_{t_0, t} \times \Phi_{t_0, t}) \# \gamma_o$ , it is an admissible plan between  $\rho_1(t)$  and  $\rho_2(t)$ , and thus

$$W_2^2(\rho_1(t), \rho_2(t)) \leq \int_{\mathbb{R}^d \times \mathbb{R}^d} |x - y|^2 d\gamma_t(x, y) = \int_{\mathbb{R}^d \times \mathbb{R}^d} |\Phi_{t_0, t}(x) - \Phi_{t_0, t}(y)|^2 d\gamma_o(x, y).$$

As a consequence taking the derivative from the right, we have

$$\left. \frac{d}{dt} \right|_{t_0^+} W_2^2(\rho_1(t), \rho_2(t)) / 2 \leq - \int_{\mathbb{R}^d \times \mathbb{R}^d} (x - y) \cdot (\nabla V(x) - \nabla V(y)) d\gamma_o(x, y)$$

Uniform convexity of  $V$  implies

$$\left. \frac{d}{dt} \right|_{t_0^+} W_2^2(\rho_1(t), \rho_2(t)) \leq -2\alpha W_2^2(\rho_1(0), \rho_2(0)).$$



# Probability Measures sliding down in a Convex Valley

Proof.

Fix any  $t_0 \geq 0$ . Let  $\gamma_o$  be the optimal transference plan between  $\rho_1(t_0)$  and  $\rho_2(t_0)$ :

$$W_2^2(\rho_1(t_0), \rho_2(t_0)) = \int_{\mathbb{R}^d \times \mathbb{R}^d} |x - y|^2 d\gamma_o(x, y).$$

The solutions are given by  $\rho_1(t) = \Phi_t \# \rho_1(0)$  and  $\rho_2(t) = \Phi_t \# \rho_2(0)$ . Take  $t \geq t_0$ . Define  $\gamma_t = (\Phi_{t_0, t} \times \Phi_{t_0, t}) \# \gamma_o$ , it is an admissible plan between  $\rho_1(t)$  and  $\rho_2(t)$ , and thus

$$W_2^2(\rho_1(t), \rho_2(t)) \leq \int_{\mathbb{R}^d \times \mathbb{R}^d} |x - y|^2 d\gamma_t(x, y) = \int_{\mathbb{R}^d \times \mathbb{R}^d} |\Phi_{t_0, t}(x) - \Phi_{t_0, t}(y)|^2 d\gamma_o(x, y).$$

As a consequence taking the derivative from the right, we have

$$\left. \frac{d}{dt} \right|_{t_0^+} W_2^2(\rho_1(t), \rho_2(t)) / 2 \leq - \int_{\mathbb{R}^d \times \mathbb{R}^d} (x - y) \cdot (\nabla V(x) - \nabla V(y)) d\gamma_o(x, y)$$

Uniform convexity of  $V$  implies

$$\left. \frac{d}{dt} \right|_{t_0^+} W_2^2(\rho_1(t), \rho_2(t)) \leq -2\alpha W_2^2(\rho_1(0), \rho_2(0)).$$



# Outline

- 1 Schematic Cartoon of Rate Arguments
  - Finite Dimensional Gradient Flows
  - First Example of Gradient Flow
- 2 Gradient Flow Structure
  - Gradient Flow
  - Convexity
- 3 Decay Rates
  - Conclusion
  - Theorems

# Heat Equation as a Gradient flow<sup>1</sup>

Consider

$$\begin{cases} \frac{\partial \rho}{\partial t} = \Delta \rho & \text{in } t > 0, x \in \mathbb{R}^d \\ \rho(0, x) = \mu \end{cases}$$

and the classical Boltzmann entropy functional

$$H(\rho) = \int_{\mathbb{R}^d} \rho \log \rho \, dx$$

defined on the space  $M = \mathcal{P}_2^{ac}(\mathbb{R}^d)$ . Variations in this space are formally

$T_\rho M = \{v \in L^1(\mathbb{R}^d) \text{ with zero mean}\}$ . Let us represent this tangent space by means of functions in  $\psi \in \mathcal{H} = W_\rho^{1,2} := W^{1,2}(\mathbb{R}^d, d\rho)$ , i.e., the closure of  $C_0^\infty(\mathbb{R}^d)$  with respect to

$$\langle \psi, \psi \rangle_\rho = \int_{\mathbb{R}^d} |\nabla \psi|^2 d\rho(x).$$

The representation of the tangent space is chosen by means of the elliptic equation  $-\operatorname{div}(\rho \nabla \psi) = v$ , and the metric is given by

$$\langle v_1, v_2 \rangle_\rho := \langle \psi_1, \psi_2 \rangle_\rho = \int_{\mathbb{R}^d} \nabla \psi_1 \cdot \nabla \psi_2 \, d\rho(x) = \int_{\mathbb{R}^d} \psi_1 v_2 \, dx$$

where  $\psi_i$  is the representation of  $v_i$ ,  $i = 1, 2$ .

<sup>1</sup>F. Otto, CPDE (2002).

# Heat Equation as a Gradient flow<sup>1</sup>

Consider

$$\begin{cases} \frac{\partial \rho}{\partial t} = \Delta \rho & \text{in } t > 0, x \in \mathbb{R}^d \\ \rho(0, x) = \mu \end{cases}$$

and the classical Boltzmann entropy functional

$$H(\rho) = \int_{\mathbb{R}^d} \rho \log \rho \, dx$$

defined on the space  $M = \mathcal{P}_2^{ac}(\mathbb{R}^d)$ . Variations in this space are formally

$T_\rho M = \{v \in L^1(\mathbb{R}^d) \text{ with zero mean}\}$ . Let us represent this tangent space by means of functions in  $\psi \in \mathcal{H} = W_\rho^{1,2} := W^{1,2}(\mathbb{R}^d, d\rho)$ , i.e., the closure of  $C_0^\infty(\mathbb{R}^d)$  with respect to

$$\langle \psi, \psi \rangle_\rho = \int_{\mathbb{R}^d} |\nabla \psi|^2 d\rho(x).$$

The representation of the tangent space is chosen by means of the elliptic equation  $-\operatorname{div}(\rho \nabla \psi) = v$ , and the metric is given by

$$\langle v_1, v_2 \rangle_\rho := \langle \psi_1, \psi_2 \rangle_\rho = \int_{\mathbb{R}^d} \nabla \psi_1 \cdot \nabla \psi_2 \, d\rho(x) = \int_{\mathbb{R}^d} \psi_1 v_2 \, dx$$

where  $\psi_i$  is the representation of  $v_i$ ,  $i = 1, 2$ .

<sup>1</sup>F. Otto, CPDE (2002).

# Heat Equation as a Gradient flow<sup>1</sup>

Consider

$$\begin{cases} \frac{\partial \rho}{\partial t} = \Delta \rho & \text{in } t > 0, x \in \mathbb{R}^d \\ \rho(0, x) = \mu \end{cases}$$

and the classical Boltzmann entropy functional

$$H(\rho) = \int_{\mathbb{R}^d} \rho \log \rho \, dx$$

defined on the space  $M = \mathcal{P}_2^{ac}(\mathbb{R}^d)$ . Variations in this space are formally

$T_\rho M = \{v \in L^1(\mathbb{R}^d) \text{ with zero mean}\}$ . Let us represent this tangent space by means of functions in  $\psi \in \mathcal{H} = W_\rho^{1,2} := W^{1,2}(\mathbb{R}^d, d\rho)$ , i.e., the closure of  $C_0^\infty(\mathbb{R}^d)$  with respect to

$$\langle \psi, \psi \rangle_\rho = \int_{\mathbb{R}^d} |\nabla \psi|^2 d\rho(x).$$

The representation of the tangent space is chosen by means of the elliptic equation  $-\operatorname{div}(\rho \nabla \psi) = v$ , and the metric is given by

$$\langle v_1, v_2 \rangle_\rho := \langle \psi_1, \psi_2 \rangle_\rho = \int_{\mathbb{R}^d} \nabla \psi_1 \cdot \nabla \psi_2 \, d\rho(x) = \int_{\mathbb{R}^d} \psi_1 v_2 \, dx$$

where  $\psi_i$  is the representation of  $v_i$ ,  $i = 1, 2$ .

<sup>1</sup>F. Otto, CPDE (2002).

# Heat Equation as a Gradient flow<sup>1</sup>

Consider

$$\begin{cases} \frac{\partial \rho}{\partial t} = \Delta \rho & \text{in } t > 0, x \in \mathbb{R}^d \\ \rho(0, x) = \mu \end{cases}$$

and the classical Boltzmann entropy functional

$$H(\rho) = \int_{\mathbb{R}^d} \rho \log \rho \, dx$$

defined on the space  $M = \mathcal{P}_2^{ac}(\mathbb{R}^d)$ . Variations in this space are formally  $T_\rho M = \{v \in L^1(\mathbb{R}^d) \text{ with zero mean}\}$ . Let us represent this tangent space by means of functions in  $\psi \in \mathcal{H} = W_\rho^{1,2} := W^{1,2}(\mathbb{R}^d, d\rho)$ , i.e., the closure of  $C_0^\infty(\mathbb{R}^d)$  with respect to

$$\langle \psi, \psi \rangle_\rho = \int_{\mathbb{R}^d} |\nabla \psi|^2 d\rho(x).$$

The representation of the tangent space is chosen by means of the elliptic equation  $-\operatorname{div}(\rho \nabla \psi) = v$ , and the metric is given by

$$\langle v_1, v_2 \rangle_\rho := \langle \psi_1, \psi_2 \rangle_\rho = \int_{\mathbb{R}^d} \nabla \psi_1 \cdot \nabla \psi_2 \, d\rho(x) = \int_{\mathbb{R}^d} \psi_1 v_2 \, dx$$

where  $\psi_i$  is the representation of  $v_i$ ,  $i = 1, 2$ .

<sup>1</sup>F. Otto, CPDE (2002).

# Heat Equation as a Gradient flow<sup>2</sup>

Formally, we can compute

$$DH_\rho(v) := \lim_{\epsilon \rightarrow 0} \frac{H(\rho + \epsilon v) - H(\rho)}{\epsilon} = \int_{\mathbb{R}^d} \frac{\delta H}{\delta \rho} v \, dx = \int_{\mathbb{R}^d} \nabla \frac{\delta H}{\delta \rho} \cdot \nabla \psi \, d\rho := \langle \nabla H_\rho, v \rangle_\rho$$

with  $\frac{\delta H}{\delta \rho} = \log \rho$ . If  $\frac{\delta H}{\delta \rho} \in \mathcal{H}$ , we can formally write that

$$\nabla H_\rho = -\operatorname{div} \left( \rho \nabla \frac{\delta H}{\delta \rho} \right).$$

In this way, the heat equation is seen as a gradient flow of the Boltzmann entropy  $H(\rho)$  in the space  $M$ ,

$$\frac{\partial \rho}{\partial t} = -\nabla H_\rho = \operatorname{div} \left( \rho \nabla \frac{\delta H}{\delta \rho} \right) = \Delta \rho$$

that dissipates the entropy

$$\frac{d}{dt} H(\rho) = -I(\rho) = - \int_{\mathbb{R}^d} \left| \nabla \frac{\delta H}{\delta \rho} \right|^2 \rho(x) \, dx = -\|\nabla H_\rho\|_\rho^2.$$

# Heat Equation as a Gradient flow<sup>2</sup>

Formally, we can compute

$$DH_\rho(v) := \lim_{\epsilon \rightarrow 0} \frac{H(\rho + \epsilon v) - H(\rho)}{\epsilon} = \int_{\mathbb{R}^d} \frac{\delta H}{\delta \rho} v \, dx = \int_{\mathbb{R}^d} \nabla \frac{\delta H}{\delta \rho} \cdot \nabla \psi \, d\rho := \langle \nabla H_\rho, v \rangle_\rho$$

with  $\frac{\delta H}{\delta \rho} = \log \rho$ . If  $\frac{\delta H}{\delta \rho} \in \mathcal{H}$ , we can formally write that

$$\nabla H_\rho = -\operatorname{div} \left( \rho \nabla \frac{\delta H}{\delta \rho} \right).$$

In this way, the heat equation is seen as a gradient flow of the Boltzmann entropy  $H(\rho)$  in the space  $M$ ,

$$\frac{\partial \rho}{\partial t} = -\nabla H_\rho = \operatorname{div} \left( \rho \nabla \frac{\delta H}{\delta \rho} \right) = \Delta \rho$$

that dissipates the entropy

$$\frac{d}{dt} H(\rho) = -I(\rho) = - \int_{\mathbb{R}^d} \left| \nabla \frac{\delta H}{\delta \rho} \right|^2 \rho(x) \, dx = -\|\nabla H_\rho\|_\rho^2.$$

# Heat Equation as a Gradient flow<sup>2</sup>

Formally, we can compute

$$DH_\rho(v) := \lim_{\epsilon \rightarrow 0} \frac{H(\rho + \epsilon v) - H(\rho)}{\epsilon} = \int_{\mathbb{R}^d} \frac{\delta H}{\delta \rho} v \, dx = \int_{\mathbb{R}^d} \nabla \frac{\delta H}{\delta \rho} \cdot \nabla \psi \, d\rho := \langle \nabla H_\rho, v \rangle_\rho$$

with  $\frac{\delta H}{\delta \rho} = \log \rho$ . If  $\frac{\delta H}{\delta \rho} \in \mathcal{H}$ , we can formally write that

$$\nabla H_\rho = -\operatorname{div} \left( \rho \nabla \frac{\delta H}{\delta \rho} \right).$$

In this way, the heat equation is seen as a gradient flow of the Boltzmann entropy  $H(\rho)$  in the space  $M$ ,

$$\frac{\partial \rho}{\partial t} = -\nabla H_\rho = \operatorname{div} \left( \rho \nabla \frac{\delta H}{\delta \rho} \right) = \Delta \rho$$

that dissipates the entropy

$$\frac{d}{dt} H(\rho) = -I(\rho) = - \int_{\mathbb{R}^d} \left| \nabla \frac{\delta H}{\delta \rho} \right|^2 \rho(x) \, dx = -\|\nabla H_\rho\|_\rho^2.$$

<sup>2</sup>F. Otto, CPDE (2002).

# Heat Equation as a Gradient flow<sup>2</sup>

Formally, we can compute

$$DH_\rho(v) := \lim_{\epsilon \rightarrow 0} \frac{H(\rho + \epsilon v) - H(\rho)}{\epsilon} = \int_{\mathbb{R}^d} \frac{\delta H}{\delta \rho} v \, dx = \int_{\mathbb{R}^d} \nabla \frac{\delta H}{\delta \rho} \cdot \nabla \psi \, d\rho := \langle \nabla H_\rho, v \rangle_\rho$$

with  $\frac{\delta H}{\delta \rho} = \log \rho$ . If  $\frac{\delta H}{\delta \rho} \in \mathcal{H}$ , we can formally write that

$$\nabla H_\rho = -\operatorname{div} \left( \rho \nabla \frac{\delta H}{\delta \rho} \right).$$

In this way, the heat equation is seen as a gradient flow of the Boltzmann entropy  $H(\rho)$  in the space  $M$ ,

$$\frac{\partial \rho}{\partial t} = -\nabla H_\rho = \operatorname{div} \left( \rho \nabla \frac{\delta H}{\delta \rho} \right) = \Delta \rho$$

that dissipates the entropy

$$\frac{d}{dt} H(\rho) = -I(\rho) = - \int_{\mathbb{R}^d} \left| \nabla \frac{\delta H}{\delta \rho} \right|^2 \rho(x) \, dx = -\|\nabla H_\rho\|_\rho^2.$$

<sup>2</sup>F. Otto, CPDE (2002).

# General Entropy Functional<sup>3</sup>

$$E(\rho) = \mathcal{U}(\rho) + \mathcal{V}(\rho) + \mathcal{W}(\rho)$$

with

$$\mathcal{U}(\rho) = \int_{\mathbb{R}^d} U(\rho(x)) dx \quad \text{internal energy}$$

$$\mathcal{V}(\rho) = \int_{\mathbb{R}^d} V(x)\rho(x) dx \quad \text{confinement energy}$$

$$\mathcal{W}(\rho) = \frac{1}{2} \int_{\mathbb{R}^{2d}} W(x-y)\rho(x)\rho(y) dx dy \quad \text{interaction energy}$$

Let us write the formal gradient flow equation as before:

$$\frac{\partial \rho}{\partial t} = \operatorname{div} \left( \rho \nabla \frac{\delta E}{\delta \rho} \right), \quad (x \in \mathbb{R}^d, t > 0).$$

and the dissipation of entropy is defined as

$$\frac{d}{dt} E(\rho) = -D(\rho) \quad \text{with} \quad D(\rho) = \int_{\mathbb{R}^d} |\xi|^2 \rho(x) dx,$$

with

$$\xi = \nabla [U'(\rho) + V + W * \rho] = \nabla \frac{\delta E}{\delta \rho}.$$

<sup>3</sup>J.A. Carrillo, R.J. McCann, C. Villani, Rev. Mat. Iberom. (2003).

# General Entropy Functional<sup>3</sup>

$$E(\rho) = \mathcal{U}(\rho) + \mathcal{V}(\rho) + \mathcal{W}(\rho)$$

with

$$\mathcal{U}(\rho) = \int_{\mathbb{R}^d} U(\rho(x)) dx \quad \text{internal energy}$$

$$\mathcal{V}(\rho) = \int_{\mathbb{R}^d} V(x)\rho(x) dx \quad \text{confinement energy}$$

$$\mathcal{W}(\rho) = \frac{1}{2} \int_{\mathbb{R}^{2d}} W(x-y)\rho(x)\rho(y) dx dy \quad \text{interaction energy}$$

Let us write the formal gradient flow equation as before:

$$\frac{\partial \rho}{\partial t} = \operatorname{div} \left( \rho \nabla \frac{\delta E}{\delta \rho} \right), \quad (x \in \mathbb{R}^d, t > 0).$$

and the dissipation of entropy is defined as

$$\frac{d}{dt} E(\rho) = -D(\rho) \quad \text{with} \quad D(\rho) = \int_{\mathbb{R}^d} |\xi|^2 \rho(x) dx,$$

with

$$\xi = \nabla [U'(\rho) + V + W * \rho] = \nabla \frac{\delta E}{\delta \rho}.$$

<sup>3</sup>J.A. Carrillo, R.J. McCann, C. Villani, Rev. Mat. Iberom. (2003).

# General Entropy Functional<sup>3</sup>

$$E(\rho) = \mathcal{U}(\rho) + \mathcal{V}(\rho) + \mathcal{W}(\rho)$$

with

$$\mathcal{U}(\rho) = \int_{\mathbb{R}^d} U(\rho(x)) dx \quad \text{internal energy}$$

$$\mathcal{V}(\rho) = \int_{\mathbb{R}^d} V(x)\rho(x) dx \quad \text{confinement energy}$$

$$\mathcal{W}(\rho) = \frac{1}{2} \int_{\mathbb{R}^{2d}} W(x-y)\rho(x)\rho(y) dx dy \quad \text{interaction energy}$$

Let us write the formal gradient flow equation as before:

$$\frac{\partial \rho}{\partial t} = \operatorname{div} \left( \rho \nabla \frac{\delta E}{\delta \rho} \right), \quad (x \in \mathbb{R}^d, t > 0).$$

and the dissipation of entropy is defined as

$$\frac{d}{dt} E(\rho) = -D(\rho) \quad \text{with} \quad D(\rho) = \int_{\mathbb{R}^d} |\xi|^2 \rho(x) dx,$$

with

$$\xi = \nabla [U'(\rho) + V + W * \rho] = \nabla \frac{\delta E}{\delta \rho}.$$

<sup>3</sup>J.A. Carrillo, R.J. McCann, C. Villani, Rev. Mat. Iberom. (2003).

# General Entropy Functional<sup>4</sup>

## Particular Examples:

- **Nonlinear Diffusions:**  $P(\rho) = \rho U'(\rho) - U(\rho)$ ,  $V(x) = 0$ ,  $W = 0$ .-

$$\frac{\partial u}{\partial t} = \Delta P(u), \quad (x \in \mathbb{R}^d, t > 0)$$

- **Nonlinear Fokker-Planck Equations:**  $P(\rho) = \rho U'(\rho) - U(\rho)$ ,  $V(x) = \frac{|x|^2}{2}$ ,  $W = 0$ .-

$$\frac{\partial \rho}{\partial t} = \operatorname{div}(x\rho + \nabla P(\rho)), \quad (x \in \mathbb{R}^d, t > 0)$$

- **Nonlinear friction equations:**  $P(\rho) = 0$ ,  $V(x) = 0$ ,  $W = \frac{|x|^{\gamma+2}}{\gamma+2}$ .- In 1D

$$\frac{\partial f}{\partial t} = \frac{\partial}{\partial v} \left[ \int_{\mathbb{R}} (v-w)|v-w|^\gamma f(w,t) dw f(v,t) \right]$$

<sup>4</sup>J.A. Carrillo, R.J. McCann, C. Villani, Rev. Mat. Iberom. (2003).

# General Entropy Functional<sup>4</sup>

## Particular Examples:

- **Nonlinear Diffusions:**  $P(\rho) = \rho U'(\rho) - U(\rho)$ ,  $V(x) = 0$ ,  $W = 0$ .-

$$\frac{\partial u}{\partial t} = \Delta P(u), \quad (x \in \mathbb{R}^d, t > 0)$$

- **Nonlinear Fokker-Planck Equations:**  $P(\rho) = \rho U'(\rho) - U(\rho)$ ,  $V(x) = \frac{|x|^2}{2}$ ,  $W = 0$ .-

$$\frac{\partial \rho}{\partial t} = \operatorname{div}(x\rho + \nabla P(\rho)), \quad (x \in \mathbb{R}^d, t > 0)$$

- **Nonlinear friction equations:**  $P(\rho) = 0$ ,  $V(x) = 0$ ,  $W = \frac{|x|^{\gamma+2}}{\gamma+2}$ .- In 1D

$$\frac{\partial f}{\partial t} = \frac{\partial}{\partial v} \left[ \int_{\mathbb{R}} (v-w)|v-w|^{\gamma} f(w,t) dw f(v,t) \right]$$

<sup>4</sup>J.A. Carrillo, R.J. McCann, C. Villani, Rev. Mat. Iberom. (2003).

# General Entropy Functional<sup>4</sup>

## Particular Examples:

- **Nonlinear Diffusions:**  $P(\rho) = \rho U'(\rho) - U(\rho)$ ,  $V(x) = 0$ ,  $W = 0$ .-

$$\frac{\partial u}{\partial t} = \Delta P(u), \quad (x \in \mathbb{R}^d, t > 0)$$

- **Nonlinear Fokker-Planck Equations:**  $P(\rho) = \rho U'(\rho) - U(\rho)$ ,  $V(x) = \frac{|x|^2}{2}$ ,  $W = 0$ .-

$$\frac{\partial \rho}{\partial t} = \operatorname{div}(x\rho + \nabla P(\rho)), \quad (x \in \mathbb{R}^d, t > 0)$$

- **Nonlinear friction equations:**  $P(\rho) = 0$ ,  $V(x) = 0$ ,  $W = \frac{|x|^{\gamma+2}}{\gamma+2}$ .- In 1D

$$\frac{\partial f}{\partial t} = \frac{\partial}{\partial v} \left[ \int_{\mathbb{R}} (v-w)|v-w|^\gamma f(w,t) dw f(v,t) \right]$$

<sup>4</sup>J.A. Carrillo, R.J. McCann, C. Villani, Rev. Mat. Iberom. (2003).

# General Entropy Functional

Let us consider two probability solutions  $\rho_1(t)$  and  $\rho_2(t)$  of the equation

$$\frac{\partial \rho}{\partial t} = \operatorname{div} \left( \rho \nabla \frac{\delta E}{\delta \rho} \right), \quad (x \in \mathbb{R}^d, t > 0).$$

They can be seen as solutions of the continuity equation

$$\frac{\partial \rho}{\partial t} + \operatorname{div}(\rho u_i) = 0, \quad (x \in \mathbb{R}^d, t > 0).$$

with velocity fields  $u_i = -\nabla \frac{\delta E}{\delta \rho}(\rho_i)$ ,  $i = 1, 2$ .

Assuming smoothness, taking the flow maps induced by both velocity fields and repeating analogous arguments as before

$$\frac{d}{dt} \Big|_{t_0^+} W_2^2(\rho_1(t), \rho_2(t))/2 \leq - \int_{\mathbb{R}^d \times \mathbb{R}^d} (x-y) \cdot \left( \nabla \frac{\delta E}{\delta \rho}(\rho_1(t_0, x)) - \nabla \frac{\delta E}{\delta \rho}(\rho_2(t_0, x)) \right) d\gamma_o(x, y)$$

where  $\gamma_o(x, y)$  is the optimal transference plan between  $\rho_1(t_0)$  and  $\rho_2(t_0)$ .

# General Entropy Functional

Let us consider two probability solutions  $\rho_1(t)$  and  $\rho_2(t)$  of the equation

$$\frac{\partial \rho}{\partial t} = \operatorname{div} \left( \rho \nabla \frac{\delta E}{\delta \rho} \right), \quad (x \in \mathbb{R}^d, t > 0).$$

They can be seen as solutions of the continuity equation

$$\frac{\partial \rho}{\partial t} + \operatorname{div}(\rho u_i) = 0, \quad (x \in \mathbb{R}^d, t > 0).$$

with velocity fields  $u_i = -\nabla \frac{\delta E}{\delta \rho}(\rho_i)$ ,  $i = 1, 2$ .

Assuming smoothness, taking the flow maps induced by both velocity fields and repeating analogous arguments as before

$$\frac{d}{dt} \Big|_{t_0^+} W_2^2(\rho_1(t), \rho_2(t))/2 \leq - \int_{\mathbb{R}^d \times \mathbb{R}^d} (x-y) \cdot \left( \nabla \frac{\delta E}{\delta \rho}(\rho_1(t_0, x)) - \nabla \frac{\delta E}{\delta \rho}(\rho_2(t_0, x)) \right) d\gamma_o(x, y)$$

where  $\gamma_o(x, y)$  is the optimal transference plan between  $\rho_1(t_0)$  and  $\rho_2(t_0)$ .

# General Entropy Functional

Let us consider two probability solutions  $\rho_1(t)$  and  $\rho_2(t)$  of the equation

$$\frac{\partial \rho}{\partial t} = \operatorname{div} \left( \rho \nabla \frac{\delta E}{\delta \rho} \right), \quad (x \in \mathbb{R}^d, t > 0).$$

They can be seen as solutions of the continuity equation

$$\frac{\partial \rho}{\partial t} + \operatorname{div}(\rho u_i) = 0, \quad (x \in \mathbb{R}^d, t > 0).$$

with velocity fields  $u_i = -\nabla \frac{\delta E}{\delta \rho}(\rho_i)$ ,  $i = 1, 2$ .

Assuming smoothness, taking the flow maps induced by both velocity fields and repeating analogous arguments as before

$$\frac{d}{dt} \Big|_{t_0^+} W_2^2(\rho_1(t), \rho_2(t))/2 \leq - \int_{\mathbb{R}^d \times \mathbb{R}^d} (x-y) \cdot \left( \nabla \frac{\delta E}{\delta \rho}(\rho_1(t_0, x)) - \nabla \frac{\delta E}{\delta \rho}(\rho_2(t_0, x)) \right) d\gamma_o(x, y)$$

where  $\gamma_o(x, y)$  is the optimal transference plan between  $\rho_1(t_0)$  and  $\rho_2(t_0)$ .

# Outline

- 1 Schematic Cartoon of Rate Arguments
  - Finite Dimensional Gradient Flows
  - First Example of Gradient Flow
- 2 Gradient Flow Structure
  - Gradient Flow
  - Convexity
- 3 Decay Rates
  - Conclusion
  - Theorems

# Geodesics

## Length Space and minimal geodesics:

A metric space  $(M, \text{dist})$  is called a **length space**<sup>a</sup> if each  $x, y \in M$  satisfy

$$\text{dist}(x, y) = \inf_{\substack{u_0=x \\ u_1=y}} \mathcal{L}(u),$$

where the infimum is over all continuous curves  $u_s \in M$  joining  $u_0 = x$  to  $u_1 = y$ . Suppose a continuous curve  $u_s \in M$  exists satisfying  $\text{dist}(u_s, u_{s+t}) = t \text{dist}(u_0, u_1)$  for  $0 \leq s \leq s+t \leq 1$  and linking any given pair of endpoints  $u_0, u_1 \in M$ . Then  $(M, \text{dist})$  is a length space and such curves are called the **minimal geodesics**.

---

<sup>a</sup>M. Gromov, Cedic/Fernand Nathan, 1981

# Geodesics

## Minimal geodesics in probability measures:

Given  $\mu, \nu \in \mathcal{P}_2^{ac}(\mathbb{R}^d)$ , let  $\gamma_0 \in \Gamma(\mu, \nu)$  denote the joint measure which achieves the infimum defining the Wasserstein distance:

$$\gamma_0 = (id \times (\nabla\varphi + id))\#\mu$$

where the function  $\varphi(x) + x^2/2$  is convex according to Brenier's theorem<sup>a</sup>.  
Therefore,

$$W_2^2(\mu, \nu) = \int_{\mathbb{R}^d} |\nabla\varphi(x)|^2 d\mu(x)$$

Minimal geodesics are given by

$$\mu_s := [id + s\nabla\varphi]\#\mu$$

<sup>a</sup>Y. Brenier, Comm. Pure Appl. Math. (1991)

# Geodesics

## Minimal geodesics in probability measures:

Given  $\mu, \nu \in \mathcal{P}_2^{ac}(\mathbb{R}^d)$ , let  $\gamma_0 \in \Gamma(\mu, \nu)$  denote the joint measure which achieves the infimum defining the Wasserstein distance:

$$\gamma_0 = (id \times (\nabla\varphi + id))\#\mu$$

where the function  $\varphi(x) + x^2/2$  is convex according to Brenier's theorem<sup>a</sup>.  
Therefore,

$$W_2^2(\mu, \nu) = \int_{\mathbb{R}^d} |\nabla\varphi(x)|^2 d\mu(x)$$

Minimal geodesics are given by

$$\mu_s := [id + s\nabla\varphi]\#\mu$$

---

<sup>a</sup>Y. Brenier, Comm. Pure Appl. Math. (1991)

# Geodesic Convexity

## Displacement Convexity:

A lower-semicontinuous energy  $E : M \rightarrow \mathbb{R} \cup \{+\infty\}$  on the length space  $M$  is said to be *displacement convex* if  $E(u_s)$  is a convex function on the interval  $[0, 1]$  along each minimal geodesic  $u_s \in M$  linking endpoints of finite energy.

## $\phi$ -uniform Convexity:

A lower-semicontinuous energy  $E : M \rightarrow \mathbb{R} \cup \{+\infty\}$  on the length space  $M$  is said to be  *$\phi$ -uniformly convex* ( $\phi$  is a modulus of convexity) if

$$E(u_0) - E(u_s) - E(u_{1-s}) + E(u_1) \geq \frac{1}{2} \int_{|1-2s|L}^L \phi(t) dt, \quad 0 \leq s \leq 1,$$

along each minimal geodesic  $u_s \in M$  of length  $L = \text{dist}(u_0, u_1)$  linking endpoints of finite energy.

# Geodesic Convexity

## Displacement Convexity:

A lower-semicontinuous energy  $E : M \rightarrow \mathbb{R} \cup \{+\infty\}$  on the length space  $M$  is said to be *displacement convex* if  $E(u_s)$  is a convex function on the interval  $[0, 1]$  along each minimal geodesic  $u_s \in M$  linking endpoints of finite energy.

## $\phi$ -uniform Convexity:

A lower-semicontinuous energy  $E : M \rightarrow \mathbb{R} \cup \{+\infty\}$  on the length space  $M$  is said to be  *$\phi$ -uniformly convex* ( $\phi$  is a modulus of convexity) if

$$E(u_0) - E(u_s) - E(u_{1-s}) + E(u_1) \geq \frac{1}{2} \int_{|1-2s|L}^L \phi(t) dt, \quad 0 \leq s \leq 1,$$

along each minimal geodesic  $u_s \in M$  of length  $L = \text{dist}(u_0, u_1)$  linking endpoints of finite energy.

# Geodesic Convexity

## Modulus of Convexity

Any function  $\phi$  taking a single sign on the positive reals and satisfying three conditions ( $\phi_0$ – $\phi_2$ ):

( $\phi_0$ )  $\phi : [0, \infty) \rightarrow \mathbb{R}$  is continuous and vanishes only at  $\phi(0) = 0$ ;

( $\phi_1$ )  $\phi(x) \geq -kx$  for some  $k < \infty$ ;

( $\phi_2$ )  $\phi(x) + \phi(y) \leq \phi(x + y)$  (superadditivity);

( $\phi_3$ )  $\chi_s(x) := \frac{1}{2} \int_{|1-2s|\sqrt{x}}^{\sqrt{x}} \phi(t) dt$  is convex on  $x \geq 0$  for each fixed  $s \in [0, 1]$ .

If  $\phi$  is convex then ( $\phi_0$ ) and ( $\phi_1$ ) together imply all four conditions ( $\phi_0$ – $\phi_3$ ) have been satisfied.

- ① *Displacement convexity*:  $\phi := 0$ .
- ② *2-uniform convexity*:  $\phi(s) = \alpha s \geq 0$ .
- ③ *Degenerate convexity*:  $\phi(s) = ks^{q-1} \geq 0$  with  $q > 2$ .
- ④ *Semiconvexity*:  $\phi(s) = -ks \leq 0$ .

# Geodesic Convexity

## Modulus of Convexity

Any function  $\phi$  taking a single sign on the positive reals and satisfying three conditions ( $\phi_0$ – $\phi_2$ ):

( $\phi_0$ )  $\phi : [0, \infty) \rightarrow \mathbb{R}$  is continuous and vanishes only at  $\phi(0) = 0$ ;

( $\phi_1$ )  $\phi(x) \geq -kx$  for some  $k < \infty$ ;

( $\phi_2$ )  $\phi(x) + \phi(y) \leq \phi(x + y)$  (superadditivity);

( $\phi_3$ )  $\chi_s(x) := \frac{1}{2} \int_{|1-2s|\sqrt{x}}^{\sqrt{x}} \phi(t) dt$  is convex on  $x \geq 0$  for each fixed  $s \in [0, 1]$ .

If  $\phi$  is convex then ( $\phi_0$ ) and ( $\phi_1$ ) together imply all four conditions ( $\phi_0$ – $\phi_3$ ) have been satisfied.

- ① *Displacement convexity*:  $\phi := 0$ .
- ② *2-uniform convexity*:  $\phi(s) = \alpha s \geq 0$ .
- ③ *Degenerate convexity*:  $\phi(s) = ks^{q-1} \geq 0$  with  $q > 2$ .
- ④ *Semiconvexity*:  $\phi(s) = -ks \leq 0$ .

# Geodesic Convexity

## Modulus of Convexity

Any function  $\phi$  taking a single sign on the positive reals and satisfying three conditions ( $\phi_0$ – $\phi_2$ ):

( $\phi_0$ )  $\phi : [0, \infty) \rightarrow \mathbb{R}$  is continuous and vanishes only at  $\phi(0) = 0$ ;

( $\phi_1$ )  $\phi(x) \geq -kx$  for some  $k < \infty$ ;

( $\phi_2$ )  $\phi(x) + \phi(y) \leq \phi(x + y)$  (superadditivity);

( $\phi_3$ )  $\chi_s(x) := \frac{1}{2} \int_{|1-2s|\sqrt{x}}^{\sqrt{x}} \phi(t) dt$  is convex on  $x \geq 0$  for each fixed  $s \in [0, 1]$ .

If  $\phi$  is convex then ( $\phi_0$ ) and ( $\phi_1$ ) together imply all four conditions ( $\phi_0$ – $\phi_3$ ) have been satisfied.

- ① *Displacement convexity*:  $\phi := 0$ .
- ② *2-uniform convexity*:  $\phi(s) = \alpha s \geq 0$ .
- ③ *Degenerate convexity*:  $\phi(s) = ks^{q-1} \geq 0$  with  $q > 2$ .
- ④ *Semiconvexity*:  $\phi(s) = -ks \leq 0$ .

# Geodesic Convexity

## Modulus of Convexity

Any function  $\phi$  taking a single sign on the positive reals and satisfying three conditions ( $\phi_0$ – $\phi_2$ ):

( $\phi_0$ )  $\phi : [0, \infty) \rightarrow \mathbb{R}$  is continuous and vanishes only at  $\phi(0) = 0$ ;

( $\phi_1$ )  $\phi(x) \geq -kx$  for some  $k < \infty$ ;

( $\phi_2$ )  $\phi(x) + \phi(y) \leq \phi(x + y)$  (superadditivity);

( $\phi_3$ )  $\chi_s(x) := \frac{1}{2} \int_{|1-2s|\sqrt{x}}^{\sqrt{x}} \phi(t) dt$  is convex on  $x \geq 0$  for each fixed  $s \in [0, 1]$ .

If  $\phi$  is convex then ( $\phi_0$ ) and ( $\phi_1$ ) together imply all four conditions ( $\phi_0$ – $\phi_3$ ) have been satisfied.

- 1 *Displacement convexity*:  $\phi := 0$ .
- 2 *2-uniform convexity*:  $\phi(s) = \alpha s \geq 0$ .
- 3 *Degenerate convexity*:  $\phi(s) = ks^{q-1} \geq 0$  with  $q > 2$ .
- 4 *Semiconvexity*:  $\phi(s) = -ks \leq 0$ .

# Geodesic Convexity

## Modulus of Convexity

Any function  $\phi$  taking a single sign on the positive reals and satisfying three conditions ( $\phi_0$ – $\phi_2$ ):

( $\phi_0$ )  $\phi : [0, \infty) \rightarrow \mathbb{R}$  is continuous and vanishes only at  $\phi(0) = 0$ ;

( $\phi_1$ )  $\phi(x) \geq -kx$  for some  $k < \infty$ ;

( $\phi_2$ )  $\phi(x) + \phi(y) \leq \phi(x + y)$  (superadditivity);

( $\phi_3$ )  $\chi_s(x) := \frac{1}{2} \int_{|1-2s|\sqrt{x}}^{\sqrt{x}} \phi(t) dt$  is convex on  $x \geq 0$  for each fixed  $s \in [0, 1]$ .

If  $\phi$  is convex then ( $\phi_0$ ) and ( $\phi_1$ ) together imply all four conditions ( $\phi_0$ – $\phi_3$ ) have been satisfied.

- ① *Displacement convexity*:  $\phi := 0$ .
- ② *2-uniform convexity*:  $\phi(s) = \alpha s \geq 0$ .
- ③ *Degenerate convexity*:  $\phi(s) = ks^{q-1} \geq 0$  with  $q > 2$ .
- ④ *Semiconvexity*:  $\phi(s) = -ks \leq 0$ .

# Geodesic Convexity

## Modulus of Convexity

Any function  $\phi$  taking a single sign on the positive reals and satisfying three conditions ( $\phi_0$ – $\phi_2$ ):

( $\phi_0$ )  $\phi : [0, \infty) \rightarrow \mathbb{R}$  is continuous and vanishes only at  $\phi(0) = 0$ ;

( $\phi_1$ )  $\phi(x) \geq -kx$  for some  $k < \infty$ ;

( $\phi_2$ )  $\phi(x) + \phi(y) \leq \phi(x + y)$  (superadditivity);

( $\phi_3$ )  $\chi_s(x) := \frac{1}{2} \int_{|1-2s|\sqrt{x}}^{\sqrt{x}} \phi(t) dt$  is convex on  $x \geq 0$  for each fixed  $s \in [0, 1]$ .

If  $\phi$  is convex then ( $\phi_0$ ) and ( $\phi_1$ ) together imply all four conditions ( $\phi_0$ – $\phi_3$ ) have been satisfied.

- ① *Displacement convexity*:  $\phi := 0$ .
- ② *2-uniform convexity*:  $\phi(s) = \alpha s \geq 0$ .
- ③ *Degenerate convexity*:  $\phi(s) = ks^{q-1} \geq 0$  with  $q > 2$ .
- ④ *Semiconvexity*:  $\phi(s) = -ks \leq 0$ .

# Convex Functionals on Probability Measures

Potential energy:

$\phi$ -uniform convexity of  $V$  on  $\mathbb{R}^d$  implies  $\phi$ -uniform convexity of

$$\mathcal{V}(\rho) = \int_{\mathbb{R}^d} V(x) d\rho(x)$$

on  $(\mathcal{P}_2(\mathbb{R}^d), W_2)$ .

Interaction energy:

$\sqrt{2}\phi(\cdot/\sqrt{2})$ -uniform convexity of  $W$  on  $\mathbb{R}^d$  implies  $\phi$ -uniform convexity of

$$\mathcal{W}(\rho) = \frac{1}{2} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} W(x-y) d\rho(x) d\rho(y)$$

on  $\mathcal{P}_{2,0}(\mathbb{R}^d)$ .

# Convex Functionals on Probability Measures

Potential energy:

$\phi$ -uniform convexity of  $V$  on  $\mathbb{R}^d$  implies  $\phi$ -uniform convexity of

$$\mathcal{V}(\rho) = \int_{\mathbb{R}^d} V(x) d\rho(x)$$

on  $(\mathcal{P}_2(\mathbb{R}^d), W_2)$ .

Interaction energy:

$\sqrt{2}\phi(\cdot/\sqrt{2})$ -uniform convexity of  $W$  on  $\mathbb{R}^d$  implies  $\phi$ -uniform convexity of

$$\mathcal{W}(\rho) = \frac{1}{2} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} W(x-y) d\rho(x) d\rho(y)$$

on  $\mathcal{P}_{2,0}(\mathbb{R}^d)$ .

# Convex Functionals on Probability Measures

Internal energy:

Condition (HDC)  $\frac{P(\rho)}{\rho^{\frac{d-1}{d}}}$  non-decreasing  $\Leftrightarrow P(\rho) \leq \frac{d}{d-1} \rho P'(\rho)$ , implies that the functional

$$\mathcal{U}(\rho) = \int_{\mathbb{R}^d} U(\rho(x)) dx$$

is displacement convex on  $\mathcal{P}_2^{ac}(\mathbb{R}^d)$ .<sup>a</sup> Remember  $P(\rho) = \rho U'(\rho) - U(\rho)$ .

---

<sup>a</sup>R. J. McCann, Adv. Math. (1997).

# Consequences of Geodesic Convexity

## Characterization by first Derivatives

For all minimal geodesics  $s \in [0, 1] \rightarrow u_s \in M$  whose endpoints have finite energy:  $E(u_s)$  is continuous on  $[0, 1]$ , its distributional derivative belongs to  $BV_{loc}(0, 1)$ , and the left and right derivatives, when they exist, satisfy

$$\left. \frac{d}{ds} E(u_s) \right|_{1-} - \left. \frac{d}{ds} E(u_s) \right|_{0+} \geq \phi(\text{dist}(u_0, u_1)) \text{dist}(u_0, u_1).$$

## First Derivatives of Probability Measures Functionals

**Potential Energy Functional:** Given any geodesic  $\rho_s = [id + s\nabla\varphi]_{\#}\rho$ , we have

$$\mathcal{V}(\rho_s) = \int_{\mathbb{R}^d} V(x) d\rho_s(x) = \int_{\mathbb{R}^d} V(x + s\nabla\varphi(x)) d\rho(x)$$

and thus

$$\left. \frac{d}{ds} \mathcal{V}(\rho_s) \right|_{0+} = \int_{\mathbb{R}^d} \nabla V(x) \cdot \nabla\varphi(x) d\rho(x).$$

# Consequences of Geodesic Convexity

## Characterization by first Derivatives

For all minimal geodesics  $s \in [0, 1] \rightarrow u_s \in M$  whose endpoints have finite energy:  $E(u_s)$  is continuous on  $[0, 1]$ , its distributional derivative belongs to  $BV_{loc}(0, 1)$ , and the left and right derivatives, when they exist, satisfy

$$\left. \frac{d}{ds} E(u_s) \right|_{1-} - \left. \frac{d}{ds} E(u_s) \right|_{0+} \geq \phi(\text{dist}(u_0, u_1)) \text{dist}(u_0, u_1).$$

## First Derivatives of Probability Measures Functionals

**Potential Energy Functional:** Given any geodesic  $\rho_s = [id + s\nabla\varphi]_{\#}\rho$ , we have

$$\mathcal{V}(\rho_s) = \int_{\mathbb{R}^d} V(x) d\rho_s(x) = \int_{\mathbb{R}^d} V(x + s\nabla\varphi(x)) d\rho(x)$$

and thus

$$\left. \frac{d}{ds} \mathcal{V}(\rho_s) \right|_{0+} = \int_{\mathbb{R}^d} \nabla V(x) \cdot \nabla\varphi(x) d\rho(x).$$

# Consequences of Geodesic Convexity

## First Derivatives of Probability Measures Functionals

**Interaction Energy Functional:** Given any geodesic  $\rho_s = [id + s\nabla\varphi]_{\#}\rho$ , we have

$$\begin{aligned}\mathcal{W}(\rho_s) &= \frac{1}{2} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} W(x-y) d\rho_s(y) d\rho_s(x) \\ &= \frac{1}{2} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} W(x + s\nabla\varphi(x) - y - s\nabla\varphi(y)) d\rho(y) d\rho(x)\end{aligned}$$

and thus

$$\begin{aligned}\left. \frac{d}{ds} \mathcal{W}(\rho_s) \right|_{0+} &= \frac{1}{2} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \nabla W(x-y) \cdot (\nabla\varphi(x) - \nabla\varphi(y)) d\rho(y) d\rho(x) \\ &= \int_{\mathbb{R}^d} \left( \int_{\mathbb{R}^d} \nabla W(x-y) d\rho(y) \right) \cdot \nabla\varphi(x) d\rho(x)\end{aligned}$$

# Consequences of Geodesic Convexity

## First Derivatives of Probability Measures Functionals

**Internal Energy Functional:** Given any geodesic  $\rho_s = [id + s\nabla\varphi]_{\#}\rho$ , we have

$$\mathcal{U}(\rho_s) = \int_{\mathbb{R}^d} U \left( \frac{\rho(x)}{\det(I + sD^2\varphi(x))} \right) \det(I + sD^2\varphi(x)) dx,$$

with  $D^2\varphi$  being understood in the sense of Aleksandrov. Now, the derivative is

$$\begin{aligned} \lim_{s \rightarrow 0} \frac{\mathcal{U}(\rho_s) - \mathcal{U}(\rho)}{s} &= - \int_{\mathbb{R}^d} P(\rho) \operatorname{tr}[D^2\varphi(x)] dx \geq \int_{\mathbb{R}^d} \nabla P(\rho) \cdot \nabla\varphi(x) dx \\ &= \int_{\mathbb{R}^d} \nabla U'(\rho) \cdot \nabla\varphi(x) d\rho(x). \end{aligned}$$

The inequality follows from the fact that  $P$  is nonnegative and the Aleksandrov (i.e. pointwise a.e.) Laplacian  $\operatorname{tr}D^2\varphi$  of a convex function is always less than the distributional Laplacian  $\Delta\varphi$ .

# Consequences of Geodesic Convexity

## First Derivatives of Probability Measures Functionals

**Summary:** The derivative along a geodesic  $\rho_s = [id + s\nabla\varphi]_{\#}\rho$  of  $E(\rho) = \mathcal{U}(\rho) + \mathcal{V}(\rho) + \mathcal{W}(\rho)$  can be estimated as

$$\begin{aligned} \left. \frac{d}{ds} E(\rho_s) \right|_{0+} &\geq \int_{\mathbb{R}^d} \left( \nabla V(x) + \int_{\mathbb{R}^d} \nabla W(x-y) d\rho(y) + \nabla U'(\rho) \right) \cdot \nabla \varphi(x) d\rho(x) \\ &= \int_{\mathbb{R}^d} \left( \nabla \frac{\delta E}{\delta \rho}(\rho(x)) \right) \cdot \nabla \varphi(x) d\rho(x) \end{aligned}$$

# Outline

- 1 Schematic Cartoon of Rate Arguments
  - Finite Dimensional Gradient Flows
  - First Example of Gradient Flow
- 2 Gradient Flow Structure
  - Gradient Flow
  - Convexity
- 3 Decay Rates
  - **Conclusion**
  - Theorems

# Main Theorem for Probability Measures

## Recollection of Facts:

Let us consider two probability solutions  $\rho_1(t)$  and  $\rho_2(t)$  of the equation

$$\frac{\partial \rho}{\partial t} = \operatorname{div} \left( \rho \nabla \frac{\delta E}{\delta \rho} \right), \quad (x \in \mathbb{R}^d, t > 0).$$

The gradient flow structure leads to

$$\frac{d}{dt} \Big|_{t_0^+} W_2^2(\rho_1(t), \rho_2(t))/2 \leq - \int_{\mathbb{R}^d \times \mathbb{R}^d} (x-y) \cdot \left( \nabla \frac{\delta E}{\delta \rho}(\rho_1(x)) - \nabla \frac{\delta E}{\delta \rho}(\rho_2(y)) \right) d\gamma_o(x, y)$$

where  $\gamma_o(x, y)$  is the optimal transference plan between  $\rho_1(t_0)$  and  $\rho_2(t_0)$ . Take the geodesic  $\rho_s = [id + s\nabla\varphi]_{\#}\rho_1(t_0)$  joining  $\rho_1(t_0)$  and  $\rho_2(t_0)$ . Then

$$\frac{d}{dt} \Big|_{t_0^+} W_2^2(\rho_1(t), \rho_2(t))/2 \leq \int_{\mathbb{R}^d} \nabla\varphi(x) \cdot \left( \nabla \frac{\delta E}{\delta \rho}(\rho_1(x)) - \nabla \frac{\delta E}{\delta \rho}(\rho_2(x + \nabla\varphi(x))) \right) d\rho_1(x)$$

# Main Theorem for Probability Measures

## Recollection of Facts:

Let us consider two probability solutions  $\rho_1(t)$  and  $\rho_2(t)$  of the equation

$$\frac{\partial \rho}{\partial t} = \operatorname{div} \left( \rho \nabla \frac{\delta E}{\delta \rho} \right), \quad (x \in \mathbb{R}^d, t > 0).$$

The gradient flow structure leads to

$$\frac{d}{dt} \Big|_{t_0^+} W_2^2(\rho_1(t), \rho_2(t))/2 \leq - \int_{\mathbb{R}^d \times \mathbb{R}^d} (x-y) \cdot \left( \nabla \frac{\delta E}{\delta \rho}(\rho_1(x)) - \nabla \frac{\delta E}{\delta \rho}(\rho_2(y)) \right) d\gamma_o(x, y)$$

where  $\gamma_o(x, y)$  is the optimal transference plan between  $\rho_1(t_0)$  and  $\rho_2(t_0)$ . Take the geodesic  $\rho_s = [id + s\nabla\varphi]_{\#}\rho_1(t_0)$  joining  $\rho_1(t_0)$  and  $\rho_2(t_0)$ . Then

$$\frac{d}{dt} \Big|_{t_0^+} W_2^2(\rho_1(t), \rho_2(t))/2 \leq \int_{\mathbb{R}^d} \nabla\varphi(x) \cdot \left( \nabla \frac{\delta E}{\delta \rho}(\rho_1(x)) - \nabla \frac{\delta E}{\delta \rho}(\rho_2(x + \nabla\varphi(x))) \right) d\rho_1(x)$$

# Main Theorem for Probability Measures

## Recollection of Facts:

Let us consider two probability solutions  $\rho_1(t)$  and  $\rho_2(t)$  of the equation

$$\frac{\partial \rho}{\partial t} = \operatorname{div} \left( \rho \nabla \frac{\delta E}{\delta \rho} \right), \quad (x \in \mathbb{R}^d, t > 0).$$

The gradient flow structure leads to

$$\frac{d}{dt} \Big|_{t_0^+} W_2^2(\rho_1(t), \rho_2(t))/2 \leq - \int_{\mathbb{R}^d \times \mathbb{R}^d} (x-y) \cdot \left( \nabla \frac{\delta E}{\delta \rho}(\rho_1(x)) - \nabla \frac{\delta E}{\delta \rho}(\rho_2(y)) \right) d\gamma_o(x, y)$$

where  $\gamma_o(x, y)$  is the optimal transference plan between  $\rho_1(t_0)$  and  $\rho_2(t_0)$ . Take the geodesic  $\rho_s = [id + s\nabla\varphi]_{\#}\rho_1(t_0)$  joining  $\rho_1(t_0)$  and  $\rho_2(t_0)$ . Then

$$\frac{d}{dt} \Big|_{t_0^+} W_2^2(\rho_1(t), \rho_2(t))/2 \leq \int_{\mathbb{R}^d} \nabla\varphi(x) \cdot \left( \nabla \frac{\delta E}{\delta \rho}(\rho_1(x)) - \nabla \frac{\delta E}{\delta \rho}(\rho_2(x + \nabla\varphi(x))) \right) d\rho_1(x)$$

# Main Theorem for Probability Measures

## Recollection of Facts:

The computation of first derivatives tell us

$$\left. \frac{d}{ds} E(\rho_s) \right|_{0+} - \left. \frac{d}{ds} E(\rho_s) \right|_{1-} \geq \int_{\mathbb{R}^d} \nabla \varphi(x) \cdot \left( \nabla \frac{\delta E}{\delta \rho}(\rho_1(x)) - \nabla \frac{\delta E}{\delta \rho}(\rho_2(x + \nabla \varphi(x))) \right) d\rho_1(x)$$

Therefore, we have shown

$$\left. \frac{d}{dt} \right|_{t_0^+} W_2^2(\rho_1(t), \rho_2(t))/2 \leq - \left( \left. \frac{d}{ds} E(\rho_s) \right|_{1-} - \left. \frac{d}{ds} E(\rho_s) \right|_{0+} \right)$$

Finally, using the convexity characterization, we conclude

$$\left. \frac{d}{dt} \right|_{t_0^+} W_2^2(\rho_1(t), \rho_2(t))/2 \leq -\phi(W_2(\rho_1(t_0), \rho_2(t_0))) W_2(\rho_1(t_0), \rho_2(t_0))$$

# Main Theorem for Probability Measures

## Recollection of Facts:

The computation of first derivatives tell us

$$\left. \frac{d}{ds} E(\rho_s) \right|_{0+} - \left. \frac{d}{ds} E(\rho_s) \right|_{1-} \geq \int_{\mathbb{R}^d} \nabla \varphi(x) \cdot \left( \nabla \frac{\delta E}{\delta \rho}(\rho_1(x)) - \nabla \frac{\delta E}{\delta \rho}(\rho_2(x + \nabla \varphi(x))) \right) d\rho_1(x)$$

Therefore, we have shown

$$\left. \frac{d}{dt} \right|_{t_0^+} W_2^2(\rho_1(t), \rho_2(t))/2 \leq - \left( \left. \frac{d}{ds} E(\rho_s) \right|_{1-} - \left. \frac{d}{ds} E(\rho_s) \right|_{0+} \right)$$

Finally, using the convexity characterization, we conclude

$$\left. \frac{d}{dt} \right|_{t_0^+} W_2^2(\rho_1(t), \rho_2(t))/2 \leq -\phi(W_2(\rho_1(t_0), \rho_2(t_0))) W_2(\rho_1(t_0), \rho_2(t_0))$$

# Main Theorem for Probability Measures

## Recollection of Facts:

The computation of first derivatives tell us

$$\left. \frac{d}{ds} E(\rho_s) \right|_{0^+} - \left. \frac{d}{ds} E(\rho_s) \right|_{1^-} \geq \int_{\mathbb{R}^d} \nabla \varphi(x) \cdot \left( \nabla \frac{\delta E}{\delta \rho}(\rho_1(x)) - \nabla \frac{\delta E}{\delta \rho}(\rho_2(x + \nabla \varphi(x))) \right) d\rho_1(x)$$

Therefore, we have shown

$$\left. \frac{d}{dt} \right|_{t_0^+} W_2^2(\rho_1(t), \rho_2(t))/2 \leq - \left( \left. \frac{d}{ds} E(\rho_s) \right|_{1^-} - \left. \frac{d}{ds} E(\rho_s) \right|_{0^+} \right)$$

Finally, using the convexity characterization, we conclude

$$\left. \frac{d}{dt} \right|_{t_0^+} W_2^2(\rho_1(t), \rho_2(t))/2 \leq -\phi(W_2(\rho_1(t_0), \rho_2(t_0))) W_2(\rho_1(t_0), \rho_2(t_0))$$

# Outline

- 1 Schematic Cartoon of Rate Arguments
  - Finite Dimensional Gradient Flows
  - First Example of Gradient Flow
- 2 Gradient Flow Structure
  - Gradient Flow
  - Convexity
- 3 Decay Rates
  - Conclusion
  - Theorems

# Decay Rates

Exponential contraction / expansion rates for gradient flows:

If  $V : \mathbb{R}^d \rightarrow \mathbb{R}$  and  $W : \mathbb{R}^d \rightarrow \mathbb{R}$  are (semi)convex, say  $D^2V(x) \geq \beta I$  and  $D^2W(x) \geq \gamma I$  for a.e.  $x \in \mathbb{R}^d$ , some  $\beta \in \mathbb{R}$  and  $\gamma \leq 0$ , then

$$W_2(\rho_1(t), \rho_2(t)) \leq e^{-(\beta+\gamma)t} W_2(\rho_1(0), \rho_2(0))$$

holds for all  $t \geq 0$ .

# Decay Rates

Algebraic contraction by gradient flow:

Let  $\phi(s) = (k/r)s^{r+1}$ ,  $k, r > 0$ , and assume that two convex functions  $V : \mathbb{R}^d \rightarrow \mathbb{R}$  and  $W : \mathbb{R}^d \rightarrow \mathbb{R}$  satisfy one of the following conditions:

- (i)  $V(x)$  is  $\phi$ -uniformly convex on  $\mathbb{R}^d$ , or
- (ii)  $W(x)$  is  $\phi$ -uniformly convex on  $\mathbb{R}^d$ , and the solutions verify  $\langle x \rangle_{\rho_1(t)} = \langle x \rangle_{\rho_2(t)} = 0$  for all  $t \geq 0$ .

Then for all  $t \geq 0$

$$W_2^2(\rho_1(t), \rho_2(t)) \leq \frac{W_2^2(\rho_1(0), \rho_2(0))}{(1 + tkW_2^r(\rho_1(0), \rho_2(0)))^{2/r}}.$$

# Contractivity versus Entropy dissipation

## Advantages:

- Gives a decay of the distance between any two solutions and thus, it can be easily used to extend flow maps to initial measures.
- It proves the existence and uniqueness of steady states by simple fixed point arguments, and thus, the decay rates towards them. These ideas are suitable for being generalized to situations in which the existence of an entropy functional is not known.

## Drawbacks:

- Gives a weaker information on the asymptotic decay. Although the HWI inequalities imply  $L^1$  convergence.
- It cannot deal with cases where  $\frac{P(\rho)}{\rho^{\frac{d-1}{d}}}$  non-decreasing is not satisfied.<sup>5</sup>

<sup>5</sup> J.A. Carrillo, J.L. Vázquez, Comm. PDE (2003); R.J. McCann, D. Slepcev, preprint (2005); R.J. McCann, Y.-M. Jiang, J. Math. Pures Appl. (2005).

# Contractivity versus Entropy dissipation

## Advantages:

- Gives a decay of the distance between any two solutions and thus, it can be easily used to extend flow maps to initial measures.
- It proves the existence and uniqueness of steady states by simple fixed point arguments, and thus, the decay rates towards them. These ideas are suitable for being generalized to situations in which the existence of an entropy functional is not known.

## Drawbacks:

- Gives a weaker information on the asymptotic decay. Although the HWI inequalities imply  $L^1$  convergence.
- It cannot deal with cases where  $\frac{P(\rho)}{\rho^{\frac{d-1}{d}}}$  non-decreasing is not satisfied.<sup>5</sup>

<sup>5</sup> J.A. Carrillo, J.L. Vázquez, Comm. PDE (2003); R.J. McCann, D. Slepcev, preprint (2005); R.J. McCann, Y.-G. Ma, J. Ma, Pure Appl. Math. (2005).

# Contractivity versus Entropy dissipation

## Advantages:

- Gives a decay of the distance between any two solutions and thus, it can be easily used to extend flow maps to initial measures.
- It proves the existence and uniqueness of steady states by simple fixed point arguments, and thus, the decay rates towards them. These ideas are suitable for being generalized to situations in which the existence of an entropy functional is not known.

## Drawbacks:

- Gives a weaker information on the asymptotic decay. Although the HWI inequalities imply  $L^1$  convergence.
- It cannot deal with cases where  $\frac{P(\rho)}{\rho^{\frac{d-1}{d}}}$  non-decreasing is not satisfied.<sup>5</sup>

<sup>5</sup> J.A. Carrillo, J.L. Vázquez, Comm. PDE (2003); R.J. McCann, D. Slepcev, preprint (2005); R.J. McCann, Y. Ma, J. Ma, Pure Appl. Math. (2005)

# Contractivity versus Entropy dissipation

## Advantages:

- Gives a decay of the distance between any two solutions and thus, it can be easily used to extend flow maps to initial measures.
- It proves the existence and uniqueness of steady states by simple fixed point arguments, and thus, the decay rates towards them. These ideas are suitable for being generalized to situations in which the existence of an entropy functional is not known.

## Drawbacks:

- Gives a weaker information on the asymptotic decay. Although the HWI inequalities imply  $L^1$  convergence.
- It cannot deal with cases where  $\frac{P(\rho)}{\rho^{\frac{d-1}{d}}}$  non-decreasing is not satisfied.<sup>5</sup>

<sup>5</sup> J.A. Carrillo, J.L. Vázquez, Comm. PDE (2003); R.J. McCann, D. Slepcev, preprint (2005); R.J. McCann, Y.J. Kim, J. Math. Pures Appl. (2005).

# Some open problems

- Degenerate convex potentials in the presence of linear diffusion are known to produce exponential convergence by entropy dissipation, we do not know when looking at Wasserstein metrics decay.
- A proof of the Holley-Stroock lemma by any of these methods is missing.
- Contraction in  $W_p$  in more dimensions for nonlinear diffusions or friction equations is not known. Very nice counterexample of Vázquez for the porous medium equation.<sup>6</sup>

---

<sup>6</sup>J.L. Vázquez, Gaeta (2005).

# Some open problems

- Degenerate convex potentials in the presence of linear diffusion are known to produce exponential convergence by entropy dissipation, we do not know when looking at Wasserstein metrics decay.
- A proof of the Holley-Stroock lemma by any of these methods is missing.
- Contraction in  $W_p$  in more dimensions for nonlinear diffusions or friction equations is not known. Very nice counterexample of Vázquez for the porous medium equation.<sup>6</sup>

---

<sup>6</sup>J.L. Vázquez, Gaeta (2005).

# Some open problems

- Degenerate convex potentials in the presence of linear diffusion are known to produce exponential convergence by entropy dissipation, we do not know when looking at Wasserstein metrics decay.
- A proof of the Holley-Stroock lemma by any of these methods is missing.
- Contraction in  $W_p$  in more dimensions for nonlinear diffusions or friction equations is not known. Very nice counterexample of Vázquez for the porous medium equation.<sup>6</sup>

---

<sup>6</sup>J.L. Vázquez, Gaeta (2005).