The Hellinger-Kantorovich distance between positive measures and Optimal Entropy-Transport problems

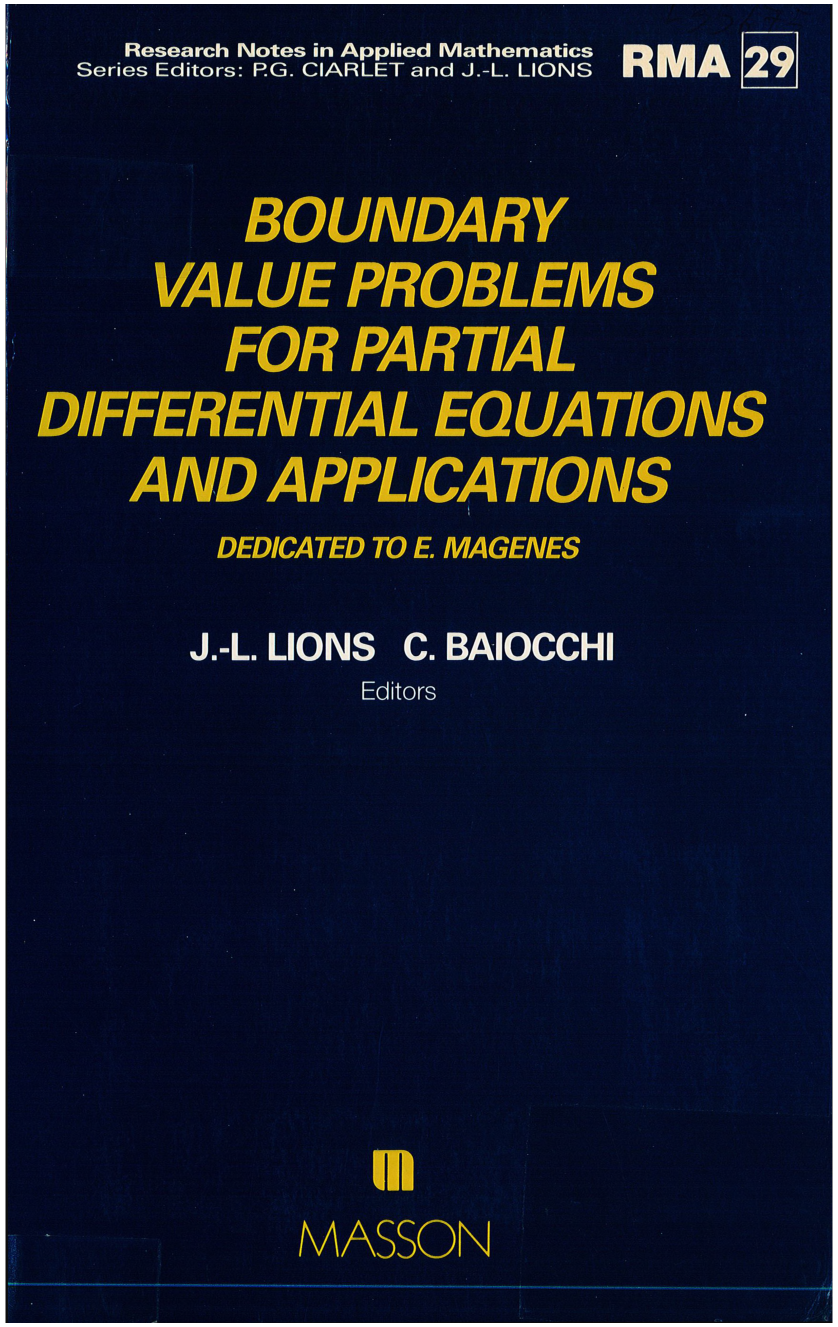
Giuseppe Savaré
http://www-dimat.unipv.it/∼savare

Dipartimento di Matematica, Università di Pavia

A Mathematical Tribute to Ennio De Giorgi, September 20, 2016

In collaboration with M. LIERO e A. MIELKE http://arxiv.org/abs/1508.07941

Standing on the shoulders of giants….



E. De Giorgi
New problems on Minimizing Movements
1993

New problems on minimizing movements

Ennio De Giorgi

Scuola Normale Superiore
Piazza dei Cavalieri 7
56126 Pisa

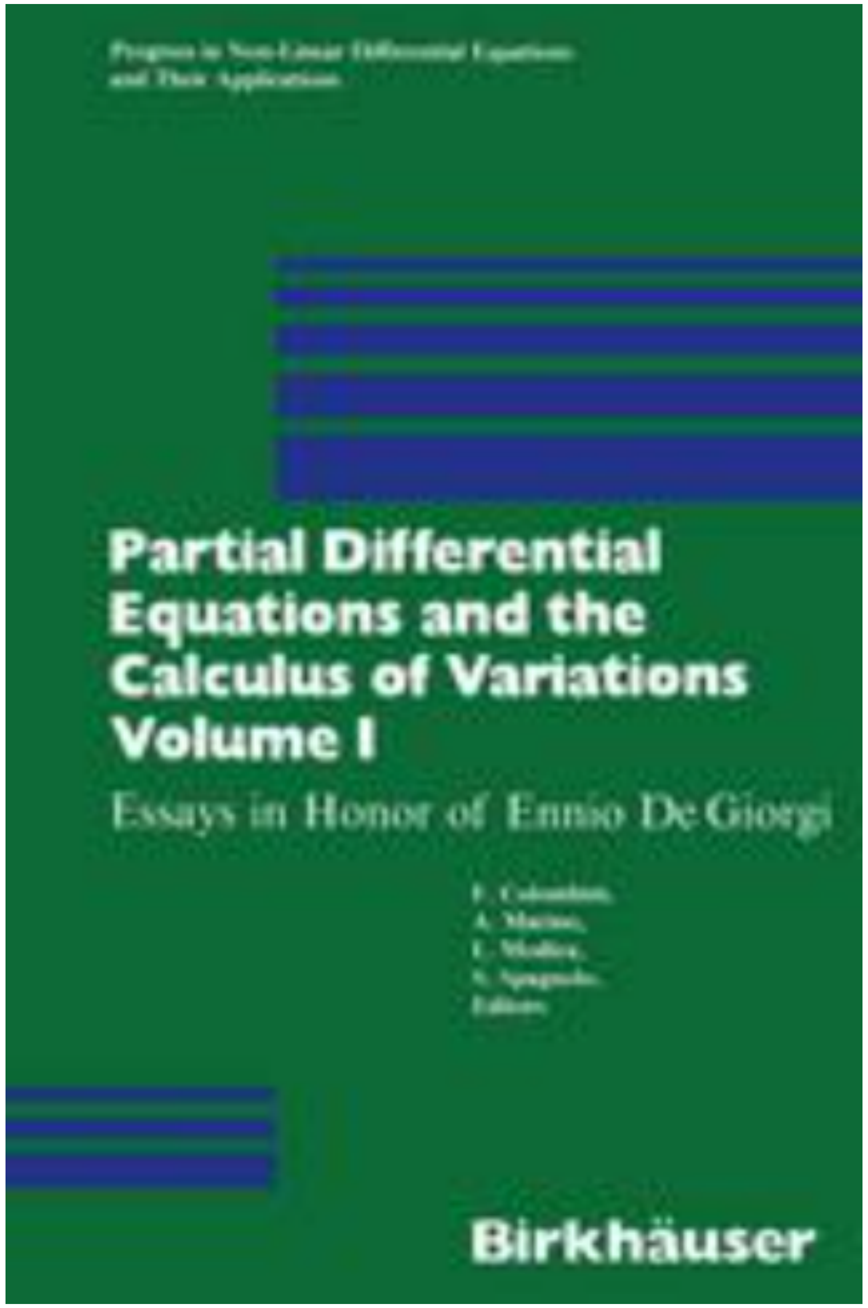
Dedicato ad Enrico Magenes per il suo settantesimo compleanno.

Introduction

In this paper I intend to deepen the idea of *minimizing movement* which has been presented in the conference [10]. Such idea seems to be suitable to unify many problems in calculus of variations, differential equations, geometric measure theory : among others, steepest descent methods, heat equation, mean curvature flow, monotone operators, various evolution problems, etc. This paper is completely self-contained and may be read independently of the papers quoted in the bibliography ; nevertheless, we remark that the main definitions of this paper may be considered as slight generalizations of the definitions given in [10] and that the paper [10] has been inspired mainly by the paper [1].

Minimizing movements are tied in various ways to penalty methods, Γ-convergence, singular perturbation, geometric measure theory, etc., hence the bibliographic indications will be unavoidably partial and far from being complete. In many cases the reader can surely find many other interesting references, as well as many interesting examples, problems, conjectures suggested by his own experience, which could be more interesting and expressive than those presented in this paper. One could think of finding general hypotheses on F and S such that the set of *minimizing movements* $MM(F,S)$ or the set of *generalized minimizing movements* $GMM(F,S)$ are nonvoid, or finite or such that their elements can be characterized by some differential equation, and/or some other meaningful condition.

Standing on the shoulders of giants….



C. Baiocchi
Discretization of
Evolution Variational Inequalities
1989

DISCRETIZATION OF EVOLUTION VARIATIONAL INEQUALITIES

CLAUDIO BAIOCCHI(*)

Dedicated to Ennio De Giorgi on his sixtieth birthday

1. Introduction. In the framework of the usual Hilbert triplet $\{V, H, V'\}$ (1) let K, A be given with:

(1.1) K is a closed convex nonempty subset of V

and

(*) This work was partly supported by CNR through the IAN and partly supported by MPI through 40% and 60% funds.

(1) Say H,V are real Hilbert spaces (norms $|\ |$ and $\|\ \|$ respectively) with $V \subset H$, V dense in H. We identify H with its dual H'; the dual space V' is the completion of H with respect to the dual norm $\|h\|_* = \sup\{(h,v) \mid \|v\| \le 1\}$ where $(\ ,\ )$ is the scalar product in H and can also be used for the pairing $V' - V$.

Outline

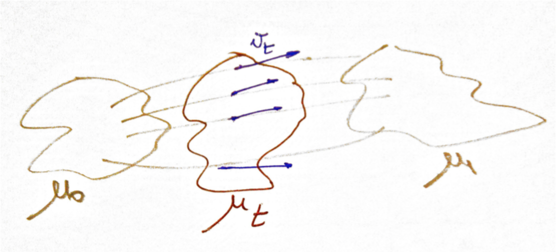
The Benamou-Brenier dynamic characterization

$\mu \in C^0([0,1]; \mathcal{M}(\mathbb{R}^d))$; $\boldsymbol{v} : \mathbb{R}^d \times (0,1) \to \mathbb{R}^d$ is a Borel vector field satisfying

$$\int_0^1 \int |\boldsymbol{v}_t(x)|^2 \, d\mu_t(x) \, dt < \infty$$

Continuity equation governed by the field $\boldsymbol{v}$

$$\partial_t \mu_t + \nabla \cdot (\boldsymbol{v}_t \mu_t) = 0 \quad \text{in the sense of distributions of } \mathbb{R}^d \times (0,1). \qquad \text{(CE)}$$



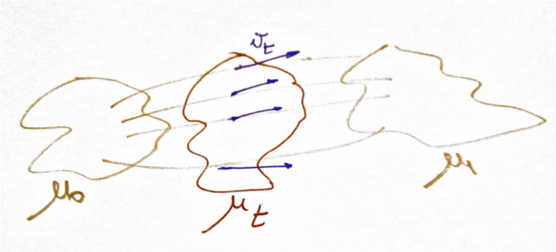
The Benamou-Brenier dynamic characterization

$\mu \in C^0([0,1];\mathcal{M}(\mathbb{R}^d))$; $\boldsymbol{v} : \mathbb{R}^d \times (0,1) \to \mathbb{R}^d$ is a Borel vector field satisfying

$$\int_0^1 \int |\boldsymbol{v}_t(x)|^2 \, d\mu_t(x) \, dt < \infty$$

Continuity equation governed by the field $\boldsymbol{v}$

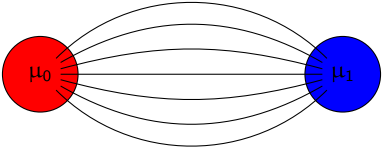
$$\partial_t \mu_t + \nabla \cdot (\boldsymbol{v}_t \mu_t) = 0 \quad \text{in the sense of distributions of } \mathbb{R}^d \times (0,1). \qquad \text{(CE)}$$



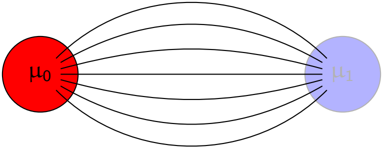
Benamou-Brenier dynamic formulation:

$$W^2(\mu_0,\mu_1) = \min\Big\{ \int_0^1 \int |\boldsymbol{v}_t|^2 \, d\mu_t \, dt : \mu \in C([0,1];\mathcal{M}(\mathbb{R}^d)), \qquad \text{(BB)}$$
$$\partial_t \mu_t + \nabla \cdot (\boldsymbol{v}_t \mu_t) = 0, \qquad \mu_{t=i} = \mu_i \Big\}.$$

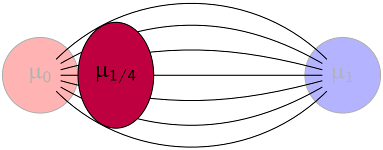
Interpolation of measures



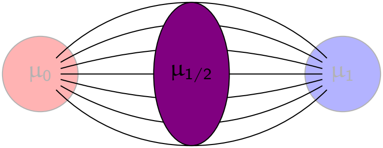
Interpolation of measures



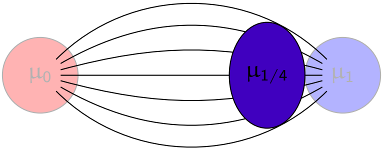
Interpolation of measures



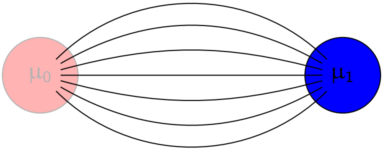
Interpolation of measures



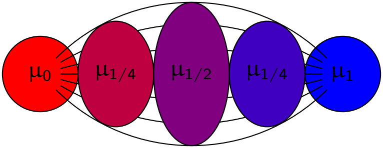
Interpolation of measures



Interpolation of measures



Interpolation of measures



Duality with the quadratic Hamilton-Jacobi equation

If ζ is a regular subsolution to the **Hamilton-Jacobi equation**

$$\partial_t \zeta_t + \frac{1}{2}|D\zeta_t|^2 \leqslant 0 \tag{HJ}$$

and

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

then

$$\int \zeta_1 \, d\mu_1 - \int \zeta_0 \, d\mu_0 \leqslant \frac{1}{2} \int_0^1 \int |\boldsymbol{v}_t|^2 \, d\mu_t \, dt.$$

Duality with the quadratic Hamilton-Jacobi equation

If ζ is a regular subsolution to the **Hamilton-Jacobi equation**

$$\partial_t \zeta_t + \frac{1}{2}|D\zeta_t|^2 \leqslant 0 \tag{HJ}$$

and

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

then

$$\int \zeta_1 \, d\mu_1 - \int \zeta_0 \, d\mu_0 \leqslant \frac{1}{2} \int_0^1 \int |\boldsymbol{v}_t|^2 \, d\mu_t \, dt.$$

HJ-duality (Otto-Villani, Bobkov-Ledoux)

$$\frac{1}{2} W^2(\mu_0, \mu_1) = \sup \Big\{ \int \zeta_1 \, d\mu_1 - \int \zeta_0 \, d\mu_0 : \zeta \in C^1([0,1]; \mathrm{Lip}_b(\mathbb{R}^d)) \\ \partial_t \zeta_t + \frac{1}{2}|D\zeta_t|^2 \leqslant 0 \Big\}.$$

Hopf-Lax formula and the dual Kantorovich formulation

Given $\zeta_0 \in \mathrm{Lip}_b(\mathbb{R}^d)$ the viscosity solution (or the maximal subsolution) of the Hamilton Jacobi equation

$$\partial_t \zeta_t + \frac{1}{2}|D\zeta_t|^2 = 0 \tag{HJ}$$

is given by the **Hopf-Lax semigroup**

$$\mathscr{Q}_t \zeta_0(x) := \inf_y \zeta_0(y) + \frac{1}{2t}|x-y|^2 \tag{Hopf-Lax}$$

Hopf-Lax formula and the dual Kantorovich formulation

Given $\zeta_0 \in \mathrm{Lip}_b(\mathbb{R}^d)$ the viscosity solution (or the maximal subsolution) of the Hamilton Jacobi equation

$$\partial_t \zeta_t + \frac{1}{2}|D\zeta_t|^2 = 0 \tag{HJ}$$

is given by the **Hopf-Lax semigroup**

$$\mathscr{Q}_t \zeta_0(x) := \inf_y \zeta_0(y) + \frac{1}{2t}|x-y|^2 \tag{Hopf-Lax}$$

Dual Kantorovich formulation

$$\begin{aligned}\frac{1}{2}W^2(\mu_0,\mu_1) &= \sup\left\{\int \zeta_1\,d\mu_1 - \int \zeta_0\,d\mu_0 : \zeta_1 = \mathscr{Q}_1\zeta_0\right\}\\ &= \sup\left\{\int \zeta_1\,d\mu_1 - \int \zeta_0\,d\mu_0 : \zeta_1(y)-\zeta_0(x) \leqslant \frac{1}{2}|x-y|^2\right\}\end{aligned}$$

The primal formulation: optimal transport

Kantorovich duality:

$$\sup\left\{\int \zeta_1\,\mathrm{d}\mu_1 - \int \zeta_0\,\mathrm{d}\mu_0 : \zeta_1(y)-\zeta_0(x) \leqslant \frac{1}{2}|x-y|^2\right\}$$
$$= \min\left\{\int \frac{1}{2}|x-y|^2\,\mathrm{d}\boldsymbol{\mu}(x,y) : \boldsymbol{\mu} \in \mathrm{Plan}(\mu_0,\mu_1)\right\}$$

Transport plans $\mathrm{Plan}(\mu_0,\mu_1)$: $\boldsymbol{\mu} \in \mathcal{M}(\mathbb{R}^d \times \mathbb{R}^d)$ with fixed marginals μ_0 and μ_1, i.e.

$$\boldsymbol{\mu}(A \times \mathbb{R}^d) = \mu_0(A), \quad \boldsymbol{\mu}(\mathbb{R}^d \times B) = \mu_1(B).$$

Compatibility condition: $\boxed{\mu_0(\mathbb{R}^d) = \mu_1(\mathbb{R}^d)}$

The primal formulation: optimal transport

Kantorovich duality:

$$\sup\left\{\int \zeta_1\, d\mu_1 - \int \zeta_0\, d\mu_0 : \zeta_1(y) - \zeta_0(x) \leqslant \frac{1}{2}|x-y|^2\right\}$$
$$= \min\left\{\int \frac{1}{2}|x-y|^2\, d\boldsymbol{\mu}(x,y) : \boldsymbol{\mu} \in \mathrm{Plan}(\mu_0,\mu_1)\right\}$$

Transport plans $\mathrm{Plan}(\mu_0,\mu_1)$: $\boldsymbol{\mu} \in \mathcal{M}(\mathbb{R}^d \times \mathbb{R}^d)$ with fixed marginals μ_0 and μ_1, i.e.

$$\boldsymbol{\mu}(A \times \mathbb{R}^d) = \mu_0(A), \quad \boldsymbol{\mu}(\mathbb{R}^d \times B) = \mu_1(B).$$

Compatibility condition: $\boxed{\mu_0(\mathbb{R}^d) = \mu_1(\mathbb{R}^d)}$

Primal formulation of the Kantorovich-Wasserstein distance via Optimal Transport

$$W^2(\mu_0,\mu_1) = \min\left\{\int |x-y|^2\, d\boldsymbol{\mu}(x,y) : \boldsymbol{\mu} \in \mathrm{Plan}(\mu_0,\mu_1)\right\}$$

Four equivalent formulations

$$\begin{tikzcd}[column sep=huge, row sep=large]
\textbf{Benamou Brenier} \arrow[r, Leftrightarrow, "\textit{Dynamic superposition}"] \arrow[d, Leftrightarrow, "\textit{Geodesic interpolation}"'] & \textbf{Hamilton Jacobi} \arrow[d, Leftrightarrow, "\textit{Hopf-Lax}"] \\
\textbf{Optimal Transport} \arrow[r, Leftrightarrow, "\textit{Linear programming}"'] & \textbf{Kantorovich duality}
\end{tikzcd}$$

Four equivalent formulations

$$
\begin{tikzcd}[column sep=huge, row sep=large]
\textbf{Benamou Brenier} \arrow[r, Leftrightarrow, "\textit{Dynamic superposition}"] \arrow[d, Updownarrow, "\textit{Geodesic interpolation}"'] & \textbf{Hamilton Jacobi} \arrow[d, Updownarrow, "\textit{Hopf-Lax}"] \\
\textbf{Optimal Transport} \arrow[r, Leftrightarrow, "\textit{Linear programming}"'] & \textbf{Kantorovich duality}
\end{tikzcd}
$$

$$
\begin{aligned}
W^2(\mu_0,\mu_1) &= \min\Big\{ \int_0^1 \int |\boldsymbol{v}_t|^2 \, d\mu_t \, dt : \mu \in C([0,1];\mathcal{M}(\mathbb{R}^d)), \\
&\qquad \partial_t \mu_t + \nabla \cdot (\boldsymbol{v}_t \mu_t) = 0, \quad \mu_{t=i} = \mu_i \Big\} && \text{(CE)} \\
&= 2 \sup \Big\{ \int \xi_1 \, d\mu_1 - \int \xi_0 \, d\mu_0 : \xi \in C^1([0,1]; \mathrm{Lip}_b(\mathbb{R}^d)) \\
&\qquad \partial_t \xi_t + \frac{1}{2} |D\xi_t|^2 \leqslant 0 \Big\} && \text{(HJ)} \\
&= 2 \sup \Big\{ \int \xi_1 \, d\mu_1 - \int \xi_0 \, d\mu_0 : \xi_1 = \mathscr{Q}_1 \xi_0 \Big\} && \text{(HL)} \\
&= \min \Big\{ \int |x-y|^2 \, d\boldsymbol{\mu}(x,y) : \boldsymbol{\mu} \in \mathrm{Plan}(\mu_0,\mu_1) \Big\} && \text{(OT)}
\end{aligned}
$$

Wasserstein distance, Entropy and Fokker-Planck equation

Consider the Guassian measure and the **Logarithmic Entropy Functional**

$$\mathfrak{m} := (2\pi/\lambda)^{-d/2}\,\mathrm{e}^{-\lambda|x|^2/2}\,\mathscr{L}^d, \qquad \mathscr{E}(\mu|\mathfrak{m}) := \int \mathsf{LE}\Big(\frac{\mathrm{d}\mu}{\mathrm{d}\mathfrak{m}}\Big)\,\mathrm{d}\mathfrak{m}, \quad \mathsf{LE}(s) := s\ln s - s + 1.$$

Wasserstein distance, Entropy and Fokker-Planck equation

Consider the Guassian measure and the **Logarithmic Entropy Functional**

$$\mathfrak{m} := (2\pi/\lambda)^{-d/2}\, e^{-\lambda|x|^2/2}\, \mathscr{L}^d, \qquad \mathscr{E}(\mu|\mathfrak{m}) := \int \mathsf{LE}\Big(\frac{d\mu}{d\mathfrak{m}}\Big)\, d\mathfrak{m}, \quad \mathsf{LE}(s) := s\ln s - s + 1.$$

[Jordan-Kinderlehrer-Otto '98] The Fokker-Planck equation

$$\boxed{\partial_t \mu - \Delta\mu + \lambda\nabla\cdot(x\mu) = 0}$$

can be interpreted as the **gradient flow of the Logarithmic Entropy functional with respect to the Wasserstein distance.**

Wasserstein distance, Entropy and Fokker-Planck equation

Consider the Guassian measure and the **Logarithmic Entropy Functional**

$$\mathfrak{m} := (2\pi/\lambda)^{-d/2}\, e^{-\lambda|x|^2/2}\, \mathscr{L}^d, \qquad \mathscr{E}(\mu|\mathfrak{m}) := \int \mathsf{LE}\Big(\frac{d\mu}{d\mathfrak{m}}\Big)\, d\mathfrak{m}, \quad \mathsf{LE}(s) := s\ln s - s + 1.$$

[Jordan-Kinderlehrer-Otto '98] The Fokker-Planck equation

$$\boxed{\partial_t \mu - \Delta\mu + \lambda\nabla\cdot(x\mu) = 0}$$

can be interpreted as the **gradient flow of the Logarithmic Entropy functional with respect to the Wasserstein distance.**

Minimizing Movement/JKO scheme: given $\tau > 0$ and $M_\tau^0 = \mu_0, \cdots, M_\tau^{n-1}$, recursively find $M_\tau^n \approx \mu_{n\tau}$ as the solution of

$$\boxed{\min \frac{1}{2\tau} W^2(M, M_\tau^{n-1}) + \mathscr{E}(M|\mathfrak{m})}$$

$$\mu_t = \lim_{N\to\infty} M_{t/N}^N \quad \text{solves the Fokker-Planck equation.}$$

Wasserstein distance, Entropy and Fokker-Planck equation

Consider the Guassian measure and the **Logarithmic Entropy Functional**

$$\mathfrak{m} := (2\pi/\lambda)^{-d/2}\, e^{-\lambda|x|^2/2}\, \mathscr{L}^d, \qquad \mathscr{E}(\mu|\mathfrak{m}) := \int \mathrm{LE}\Big(\frac{d\mu}{d\mathfrak{m}}\Big)\, d\mathfrak{m}, \quad \mathrm{LE}(s) := s\ln s - s + 1.$$

[Jordan-Kinderlehrer-Otto '98] The Fokker-Planck equation

$$\boxed{\partial_t \mu - \Delta\mu + \lambda\nabla\cdot(x\mu) = 0}$$

can be interpreted as the **gradient flow of the Logarithmic Entropy functional with respect to the Wasserstein distance.**

Minimizing Movement/JKO scheme: given $\tau > 0$ and $M_\tau^0 = \mu_0, \cdots, M_\tau^{n-1}$, recursively find $M_\tau^n \approx \mu_{n\tau}$ as the solution of

$$\boxed{\min \frac{1}{2\tau} W^2(M, M_\tau^{n-1}) + \mathscr{E}(M|\mathfrak{m})}$$

$$\mu_t = \lim_{N\to\infty} M^N_{t/N} \quad \text{solves the Fokker-Planck equation.}$$

Contraction and Evolution Variational Inequality: if μ_t, ν_t are solution to the Fokker-Planck equation

$$W(\mu_t, \nu_t) \leqslant e^{-\lambda t} W(\mu_0, \nu_0), \qquad \frac{1}{2}\frac{d}{dt} W^2(\mu_t, \nu) \leqslant \mathscr{E}(\mu_t, \mathfrak{m}) - \mathscr{E}(\nu|\mathfrak{m})$$

Outline

The dynamic formulation

Let $\mu \in C^0([0,1];\mathcal{M}(\mathbb{R}^d))$, $w : \mathbb{R}^d \times (0,1) \to \mathbb{R}$ be a Borel vector field satisfying

$$\int_0^1 \int w_t^2(x)\,d\mu_t(x)\,dt < \infty.$$

Pure reaction equation:

$$\partial_t \mu_t = w_t \mu_t \quad \text{in } \mathscr{D}'(\mathbb{R}^d \times (0,1))$$

The dynamic formulation

Let $\mu \in C^0([0,1];\mathcal{M}(\mathbb{R}^d))$, $w : \mathbb{R}^d \times (0,1) \to \mathbb{R}$ be a Borel vector field satisfying

$$\int_0^1 \int w_t^2(x)\,d\mu_t(x)\,dt < \infty.$$

Pure reaction equation:

$$\partial_t \mu_t = w_t \mu_t \quad \text{in } \mathscr{D}'(\mathbb{R}^d \times (0,1))$$

A reaction distance via dynamic interpolation

$$H^2(\mu_0,\mu_1) = \min\Big\{ \frac{1}{4}\int_0^1 \int |w_t|^2\,d\mu_t\,dt : \mu \in C([0,1];\mathcal{M}(\mathbb{R}^d)),\\ \partial_t\mu_t = w_t\mu_t, \quad \mu_{t=i} = \mu_i \Big\}.$$

ODE reaction and duality

If ζ is a regular subsolution of

$$\partial_t \zeta_t + \zeta_t^2 \leqslant 0$$

and

$$\partial_t \mu_t = w_t \mu_t$$

then

$$\int \zeta_1 \, d\mu_1 - \int \zeta_0 \, d\mu_0 \leqslant \frac{1}{4} \int_0^1 \int w_t^2 \, d\mu_t \, dt.$$

ODE reaction and duality

If ζ is a regular subsolution of

$$\partial_t \zeta_t + \zeta_t^2 \leqslant 0$$

and

$$\partial_t \mu_t = w_t \mu_t$$

then

$$\int \zeta_1 \, d\mu_1 - \int \zeta_0 \, d\mu_0 \leqslant \frac{1}{4} \int_0^1 \int w_t^2 \, d\mu_t \, dt.$$

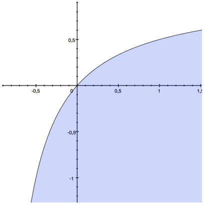
Duality

$$H^2(\mu_0, \mu_1) = \sup \Big\{ \int \zeta_1 \, d\mu_1 - \int \zeta_0 \, d\mu_0 : \zeta \in C^1([0,1]; B_b(\mathbb{R}^d)) \\ \partial_t \zeta_t + \zeta_t^2 \leqslant 0 \Big\}.$$

Static duality

Any finite subsolution of $\partial_t \zeta_t + \zeta_t^2 \leqslant 0$ for $t \in [0,1]$ should satisfy

$$\zeta_0 > -1, \quad \boxed{\zeta_1 < \frac{\zeta_0}{1+\zeta_0}} < 1, \quad (1-\zeta_1)(1+\zeta_0) > 1$$

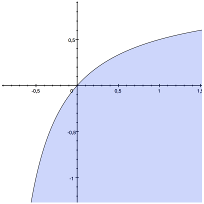


The convex set $K := \Big\{ (\zeta_0, \zeta_1) : (1-\zeta_1)(1+\zeta_0) > 1 \Big\}$

Static duality

Any finite subsolution of $\partial_t \zeta_t + \zeta_t^2 \leqslant 0$ for $t \in [0,1]$ should satisfy

$$\zeta_0 > -1, \quad \boxed{\zeta_1 < \frac{\zeta_0}{1+\zeta_0}} < 1, \quad (1-\zeta_1)(1+\zeta_0) > 1$$



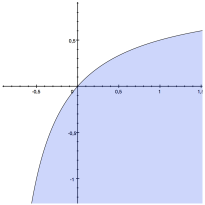
The convex set $K := \Big\{ (\zeta_0, \zeta_1) : (1-\zeta_1)(1+\zeta_0) > 1 \Big\}$

Static duality

$$H^2(\mu_0, \mu_1) = \sup \Big\{ \int \zeta_1 \, d\mu_1 - \int \zeta_0 \, d\mu_0 : (\zeta_0, \zeta_1) \in K \Big\}.$$

Primal formulation

We compute the **support function of K:**



$$H(r_0, r_1) := \sup \left\{ r_1\zeta_1 - r_0\zeta_0 : (\zeta_0, \zeta_1) \in K \right\}$$

obtaining

$$H(r_0, r_1) = \left(\sqrt{r_1} - \sqrt{r_0}\right)^2 \quad \text{if } r_0, r_1 \geqslant 0.$$

Hellinger-Kakutani distance

H is the Hellinger-Kakutani distance

IF $\mu_0, \mu_1 \in \mathcal{M}_+(\mathbb{R}^d)$ are positive finite Borel measures

$$\mathsf{H}^2(\mu_0, \mu_1) = \int \left(\sqrt{\rho_0} - \sqrt{\rho_1}\right)^2 \mathrm{d}\mu, \quad \mu_i = \rho_i \mu \ll \mu.$$

Hellinger-Kakutani distance

H is the Hellinger-Kakutani distance

IF $\mu_0, \mu_1 \in \mathcal{M}_+(\mathbb{R}^d)$ are positive finite Borel measures

$$H^2(\mu_0, \mu_1) = \int (\sqrt{\rho_0} - \sqrt{\rho_1})^2 \, d\mu, \quad \mu_i = \rho_i \mu \ll \mu.$$

Since the function

$$(r_0, r_1) \mapsto (\sqrt{r_0} - \sqrt{r_1})^2 = r_0 + r_1 - 2\sqrt{r_0 r_1}$$

is positively 1-homogeneous, the above definition does not depend on the choice of the dominating measure μ.

$$H^2(\mu_0, \mu_1) \leqslant \|\mu_0 - \mu_1\|_{TV} \leqslant H(\mu_0, \mu_1)\Big(2\mu_0(X) + 2\mu_1(X)\Big)^{1/2}$$

Hellinger-Kakutani distance

H is the Hellinger-Kakutani distance

IF $\mu_0, \mu_1 \in \mathcal{M}_+(\mathbb{R}^d)$ are positive finite Borel measures

$$H^2(\mu_0, \mu_1) = \int (\sqrt{\rho_0} - \sqrt{\rho_1})^2 \, d\mu, \quad \mu_i = \rho_i \mu \ll \mu.$$

Since the function

$$(r_0, r_1) \mapsto (\sqrt{r_0} - \sqrt{r_1})^2 = r_0 + r_1 - 2\sqrt{r_0 r_1}$$

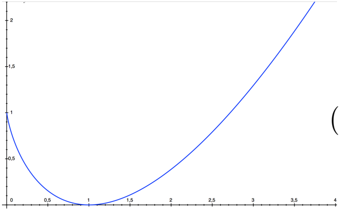
is positively 1-homogeneous, the above definition does not depend on the choice of the dominating measure μ.

$$H^2(\mu_0, \mu_1) \leqslant \|\mu_0 - \mu_1\|_{TV} \leqslant H(\mu_0, \mu_1)\Big(2\mu_0(X) + 2\mu_1(X)\Big)^{1/2}$$

Geodesics:

$$\mu_\vartheta := \Big((1-\vartheta)\sqrt{\rho_0} + \vartheta\sqrt{\rho_1}\Big)^2 \mu, \quad \vartheta \in [0, 1].$$

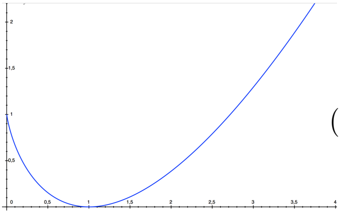
Hellinger distance and Logrithmic entropy



$$\mathsf{LE}(s) = s \ln s - s + 1$$

$$\left(\sqrt{r_1} - \sqrt{r_0}\right)^2 = \min_{s \geqslant 0}\; r_0 \mathsf{LE}(s/r_0) + r_1 \mathsf{LE}(s/r_1).$$

Hellinger distance and Logrithmic entropy



$$\mathsf{LE}(s) = s \ln s - s + 1$$

$$\left(\sqrt{r_1} - \sqrt{r_0}\right)^2 = \min_{s \geqslant 0} \; r_0 \mathsf{LE}(s/r_0) + r_1 \mathsf{LE}(s/r_1).$$

A variational characterization in terms of the logarithmic entropy

$$\mathsf{H}^2(\mu_0, \mu_1) = \min_{\gamma} \Big(\mathscr{E}(\gamma|\mu_0) + \mathscr{E}(\gamma|\mu_1) \Big)$$

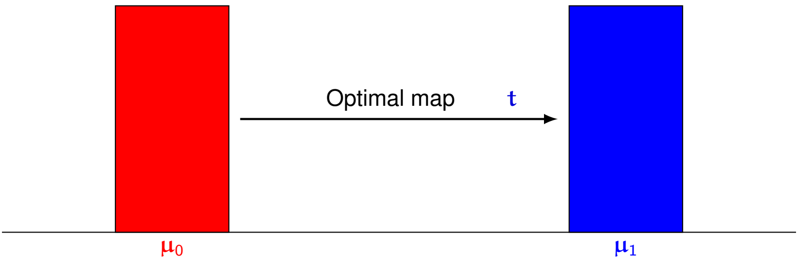
where

$$\mathscr{E}(\gamma|\mu) = \int \mathsf{LE}\Big(\frac{d\gamma}{d\mu}\Big)\, d\mu = \int \frac{d\gamma}{d\mu} \log\Big(\frac{d\gamma}{d\mu}\Big)\, d\mu - \gamma(\mathbb{R}^d) + \mu(\mathbb{R}^d).$$

If μ is a dominating measure, the optimal γ is given by

$$\gamma = \sqrt{\rho_0\,\rho_1}\,\mu, \quad \mu_i = \rho_i \mu \ll \mu$$

Displacement vs Hellinger interpolation



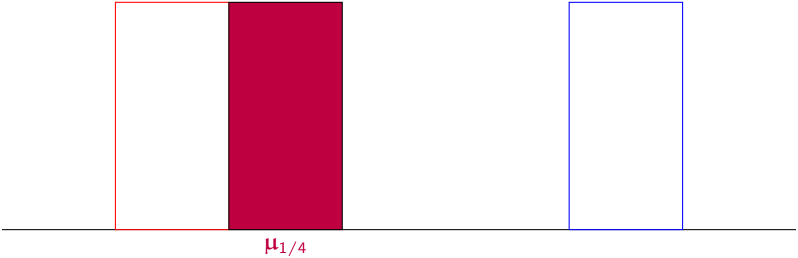
Initial configuration

Displacement vs Hellinger interpolation

Displacement interpolation: “points are moving along optimal rays”

$$\mu_\vartheta = \big((1-\vartheta)x + \vartheta \mathbf{t}(x)\big)_\sharp \mu_0$$

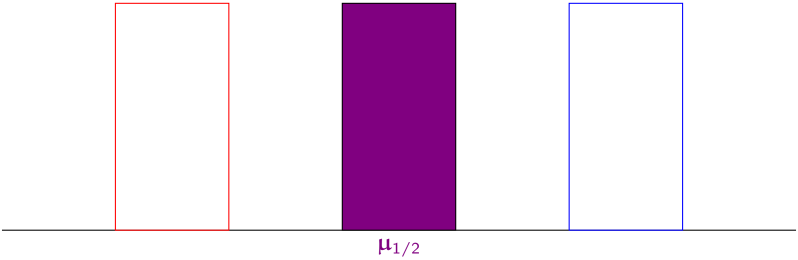
Displacement vs Hellinger interpolation



Displacement interpolation: "points are moving along optimal rays"

$$\mu_\vartheta = \big((1-\vartheta)x + \vartheta \mathbf{t}(x)\big)_\sharp \mu_0$$

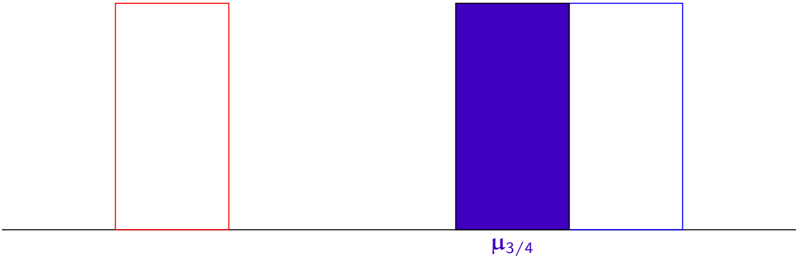
Displacement vs Hellinger interpolation



Displacement interpolation: "points are moving along optimal rays"

$$\mu_\vartheta = \big((1-\vartheta)x + \vartheta \mathbf{t}(x)\big)_\sharp \mu_0$$

Displacement vs Hellinger interpolation



Displacement interpolation: "points are moving along optimal rays"

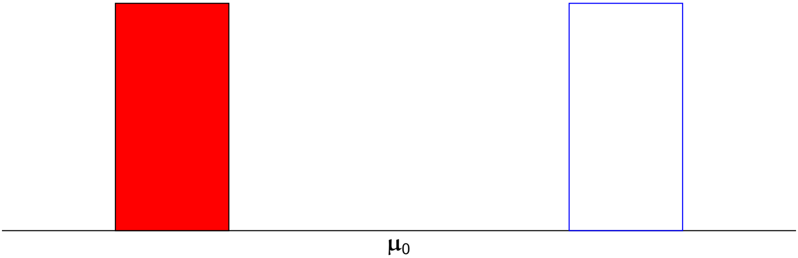
$$\mu_\vartheta = \big((1-\vartheta)x + \vartheta \mathbf{t}(x)\big)_\sharp \mu_0$$

Displacement vs Hellinger interpolation

Displacement interpolation: “points are moving along optimal rays”

$$\mu_\vartheta = \big((1-\vartheta)x + \vartheta \mathbf{t}(x)\big)_\sharp \mu_0$$

Displacement vs Hellinger interpolation



The Hellinger interpolation: "points remain fixed"

$$\mu_\vartheta = \Big((1-\vartheta)\sqrt{f_0} + \vartheta\sqrt{f_1}\Big)^2$$

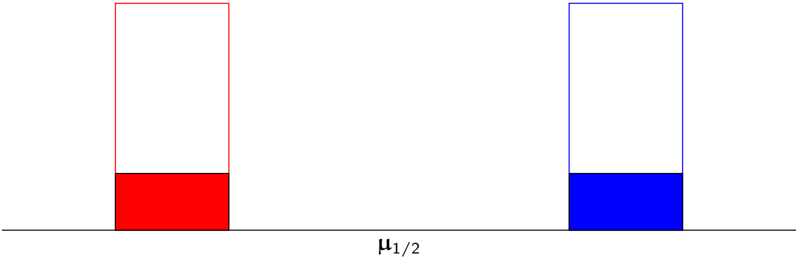
Displacement vs Hellinger interpolation

$\mu_{1/4}$

The Hellinger interpolation: “points remain fixed”

$$\mu_\vartheta = \Big((1-\vartheta)\sqrt{f_0} + \vartheta\sqrt{f_1} \Big)^2$$

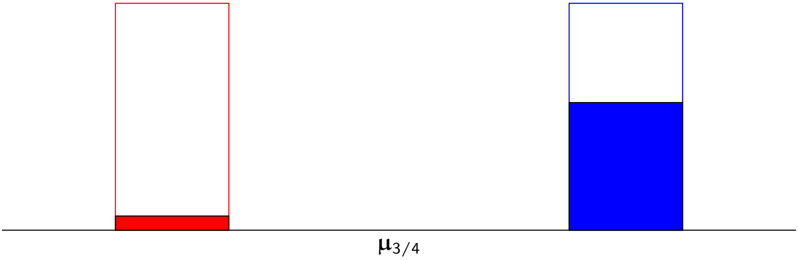
Displacement vs Hellinger interpolation



The Hellinger interpolation: “points remain fixed”

$$\mu_\vartheta = \left((1-\vartheta)\sqrt{f_0} + \vartheta\sqrt{f_1}\right)^2$$

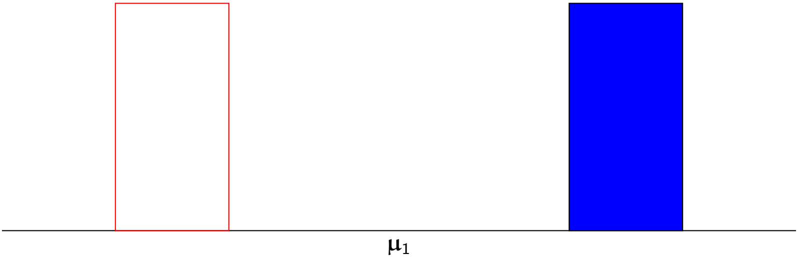
Displacement vs Hellinger interpolation



The Hellinger interpolation: "points remain fixed"

$$\mu_\vartheta = \Big((1-\vartheta)\sqrt{f_0} + \vartheta\sqrt{f_1}\Big)^2$$

Displacement vs Hellinger interpolation



The Hellinger interpolation: “points remain fixed”

$$\mu_\vartheta = \left((1-\vartheta)\sqrt{f_0} + \vartheta\sqrt{f_1}\right)^2$$

Outline

Starting point (I): the dynamic formulation of HK

Let $\mu \in C^0([0,1];\mathcal{M}(\mathbb{R}^d))$, $(\boldsymbol{v}, w) : \mathbb{R}^d \times (0,1) \to \mathbb{R}^{d+1}$ be a Borel vector field satisfying

$$\int_0^1 \int \Big(|\boldsymbol{v}_t(x)|^2 + w_t^2(x) \Big) \, d\mu_t(x) \, dt < \infty.$$

Continuity equation with reaction governed by the field $(\boldsymbol{v}, w)$ if

$$\partial_t \mu_t + \nabla \cdot (\boldsymbol{v}_t \mu_t) = w_t \mu_t \quad \text{in } \mathscr{D}'(\mathbb{R}^d \times (0,1)) \qquad \text{(CER)}$$

Starting point (I): the dynamic formulation of HK

Let $\mu \in C^0([0,1];\mathcal{M}(\mathbb{R}^d))$, $(\boldsymbol{v}, w) : \mathbb{R}^d \times (0,1) \to \mathbb{R}^{d+1}$ be a Borel vector field satisfying

$$\int_0^1 \int \left(|\boldsymbol{v}_t(x)|^2 + w_t^2(x) \right) d\mu_t(x)\, dt < \infty.$$

Continuity equation with reaction governed by the field $(\boldsymbol{v}, w)$ if

$$\partial_t \mu_t + \nabla \cdot (\boldsymbol{v}_t \mu_t) = w_t \mu_t \quad \text{in } \mathscr{D}'(\mathbb{R}^d \times (0,1)) \tag{CER}$$

The Hellinger-Kantorovich distance via dynamic interpolation

$$\mathsf{HK}^2(\mu_0, \mu_1) = \min \left\{ \int_0^1 \int \left(|\boldsymbol{v}_t|^2 + \frac{1}{4} |w_t|^2 \right) d\mu_t\, dt : \mu \in C([0,1];\mathcal{M}(\mathbb{R}^d)), \right.$$
$$\left. \partial_t \mu_t + \nabla \cdot (\boldsymbol{v}_t \mu_t) = w_t \mu_t, \quad \mu_{t=i} = \mu_i \right\}.$$

Starting point (I): the dynamic formulation of HK

Let $\mu \in C^0([0,1];\mathcal{M}(\mathbb{R}^d))$, $(\boldsymbol{v}, w) : \mathbb{R}^d \times (0,1) \to \mathbb{R}^{d+1}$ be a Borel vector field satisfying

$$\int_0^1 \int \Big(|\boldsymbol{v}_t(x)|^2 + w_t^2(x)\Big)\,\mathrm{d}\mu_t(x)\,\mathrm{d}t < \infty.$$

Continuity equation with reaction governed by the field $(\boldsymbol{v}, w)$ if

$$\partial_t \mu_t + \nabla \cdot (\boldsymbol{v}_t \mu_t) = w_t \mu_t \quad \text{in } \mathscr{D}'(\mathbb{R}^d \times (0,1)) \tag{CER}$$

The Hellinger-Kantorovich distance via dynamic interpolation

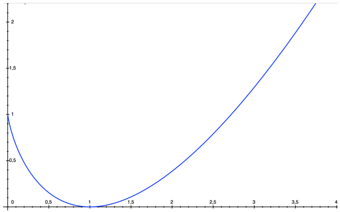
$$\mathsf{HK}^2(\mu_0, \mu_1) = \min\Big\{ \int_0^1 \int \Big(|\boldsymbol{v}_t|^2 + \frac{1}{4}|w_t|^2\Big)\,\mathrm{d}\mu_t\,\mathrm{d}t : \mu \in C([0,1];\mathcal{M}(\mathbb{R}^d)),$$
$$\partial_t \mu_t + \nabla \cdot (\boldsymbol{v}_t \mu_t) = w_t \mu_t, \quad \mu_{t=i} = \mu_i \Big\}.$$

HK is a **convex and subadditive functional** (cf. DOLBEAULT-NAZARET-S.).
A similar approach has been independently proposed by KONDRATIEV, MONSAINGEON, VOROTNIKOV and CHIZAT, PEYRÉ, VIALARD, SCHMITZER.

Starting point (II): combining Entropy and Transport

$$W^2(\mu_0, \mu_1) = \min\left\{ \int |x-y|^2 \, d\boldsymbol{\gamma}(x,y) : \boldsymbol{\gamma} \in \mathrm{Plan}(\mu_0, \mu_1) \right\}$$

$$H^2(\mu_0, \mu_1) = \min_{\gamma} \Big(\mathscr{E}(\gamma|\mu_0) + \mathscr{E}(\gamma|\mu_1) \Big)$$



Starting point (II): combining Entropy and Transport

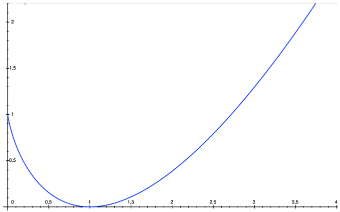
$$W^2(\mu_0,\mu_1) = \min\Big\{\int |x-y|^2\,d\boldsymbol{\gamma}(x,y) : \boldsymbol{\gamma} \in \mathsf{Plan}(\mu_0,\mu_1)\Big\}$$

$$H^2(\mu_0,\mu_1) = \min_{\gamma}\Big(\mathscr{E}(\gamma|\mu_0) + \mathscr{E}(\gamma|\mu_1)\Big)$$

Gaussian Entropy Transport problem

$$\mathsf{GET}(\mu_0,\mu_1) := \min_{\boldsymbol{\gamma}}\ \mathscr{E}(\gamma_0|\mu_0) + \mathscr{E}(\gamma_1|\mu_1) + \int |x-y|^2\,d\boldsymbol{\gamma}(x,y), \qquad \text{(GET)}$$

where γ_i are the marginals of $\boldsymbol{\gamma} \in \mathcal{M}(\mathbb{R}^d \times \mathbb{R}^d)$.



Starting point (II): combining Entropy and Transport

$$W^2(\mu_0, \mu_1) = \min \left\{ \int |x-y|^2 \, d\boldsymbol{\gamma}(x,y) : \boldsymbol{\gamma} \in \mathrm{Plan}(\mu_0, \mu_1) \right\}$$

$$H^2(\mu_0, \mu_1) = \min_{\gamma} \Big(\mathscr{E}(\gamma|\mu_0) + \mathscr{E}(\gamma|\mu_1) \Big)$$

Gaussian Entropy Transport problem

$$\mathrm{GET}(\mu_0, \mu_1) := \min_{\boldsymbol{\gamma}} \; \mathscr{E}(\gamma_0|\mu_0) + \mathscr{E}(\gamma_1|\mu_1) + \int |x-y|^2 \, d\boldsymbol{\gamma}(x,y), \qquad \text{(GET)}$$

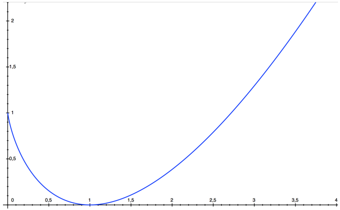
where γ_i are the marginals of $\boldsymbol{\gamma} \in \mathcal{M}(\mathbb{R}^d \times \mathbb{R}^d)$.

(GET) can be considered as a sort of **relaxation of the Optimal Transport problem**, where the constraint

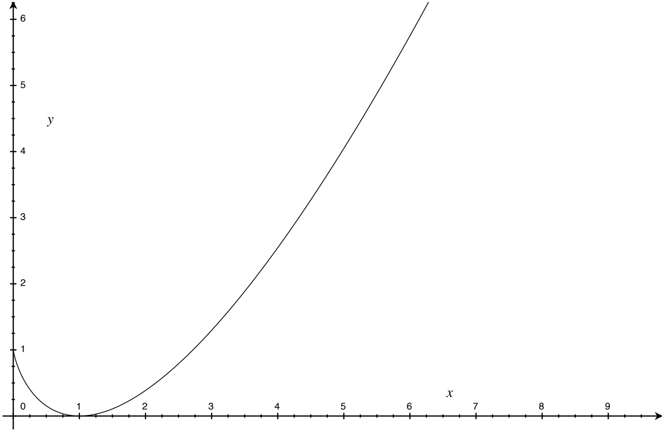
$$\boldsymbol{\gamma} \in \mathrm{Plan}(\mu_0, \mu_1) \quad \text{i.e. the } i\text{-marginal } \gamma_i \text{ of } \boldsymbol{\gamma} \text{ is } \mu_i$$

has been substituted by the penalizing entropy term

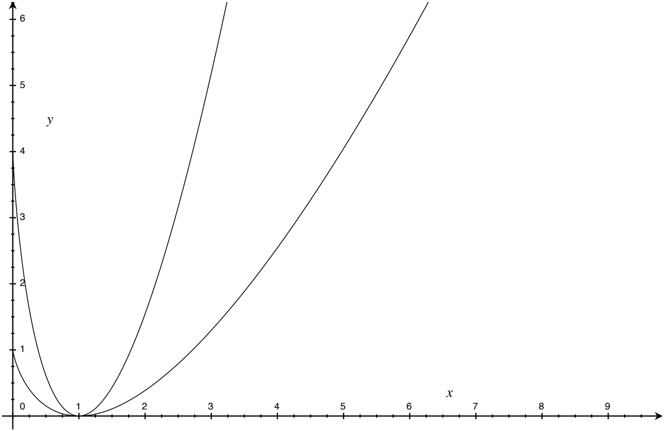
$$\mathscr{E}(\gamma_0|\mu_0) + \mathscr{E}(\gamma_1|\mu_1).$$



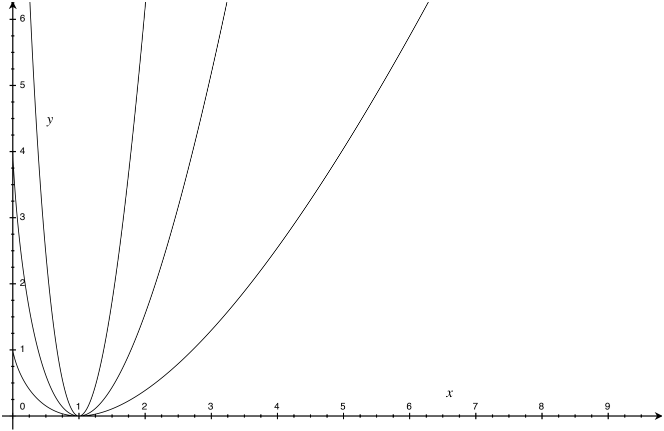
Relaxation/penalization of the density constraint $\rho = 1$



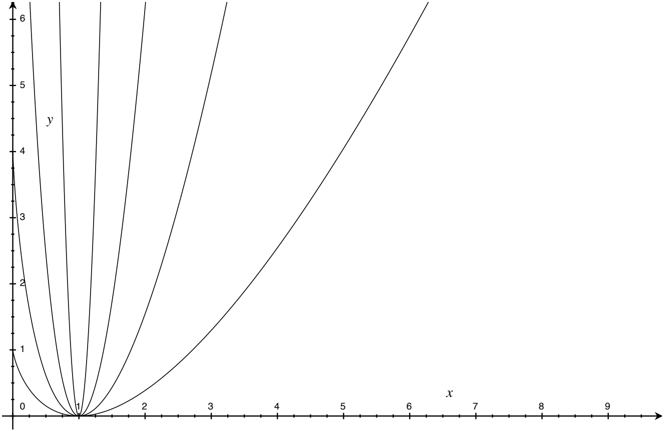
Relaxation/penalization of the density constraint $\rho = 1$



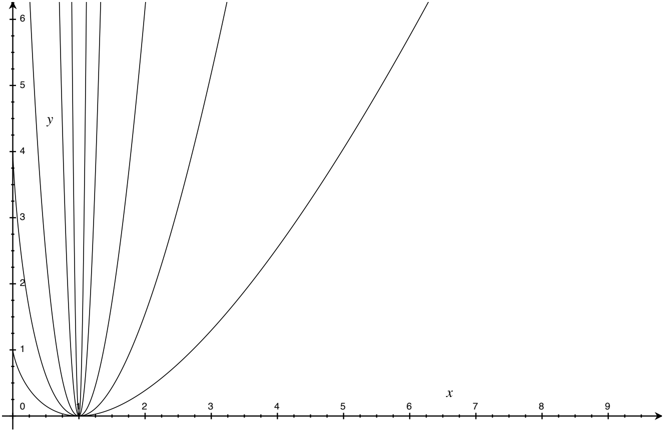
Relaxation/penalization of the density constraint $\rho = 1$



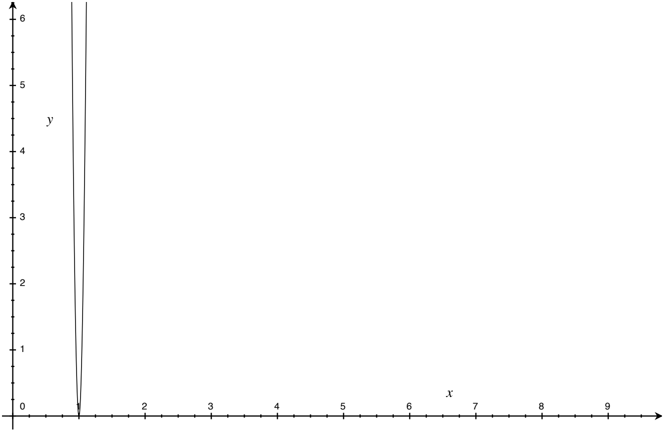
Relaxation/penalization of the density constraint $\rho = 1$



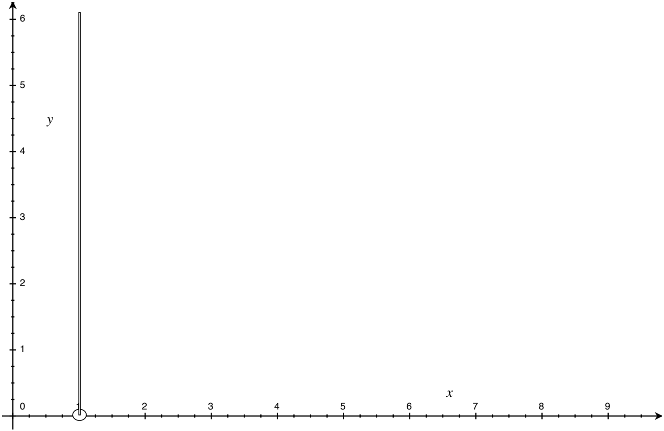
Relaxation/penalization of the density constraint $\rho = 1$



Relaxation/penalization of the density constraint $\rho = 1$



Relaxation/penalization of the density constraint $\rho = 1$



HK and GET are always finite and dominated by the Hellinger and the Kantorovich-Wasserstein distance

Continuity equation with reaction governed by the field $(\boldsymbol{v}, w)$

$$\partial_t \mu_t + \nabla \cdot (\boldsymbol{v}_t \mu_t) = w_t \mu_t \quad \text{in } \mathscr{D}'(\mathbb{R}^d \times (0,1)) \tag{CER}$$

HK and GET are always finite and dominated by the Hellinger and the Kantorovich-Wasserstein distance

Continuity equation with reaction governed by the field $(\boldsymbol{v}, w)$

$$\partial_t \mu_t + \nabla \cdot (\boldsymbol{v}_t \mu_t) = w_t \mu_t \quad \text{in } \mathscr{D}'(\mathbb{R}^d \times (0,1)) \tag{CER}$$

Estimate of the pure reaction part, $\boldsymbol{v} \equiv 0$.

$$\boxed{\mathsf{HK}^2(\mu_0, \mu_1) \leqslant \mathsf{H}^2(\mu_0, \mu_1)} \leqslant \mu_0(\mathbb{R}^d) + \mu_1(\mathbb{R}^d).$$

HK and GET are always finite and dominated by the Hellinger and the Kantorovich-Wasserstein distance

Continuity equation with reaction governed by the field $(\boldsymbol{v}, w)$

$$\partial_t \mu_t + \nabla \cdot (\boldsymbol{v}_t \mu_t) = w_t \mu_t \quad \text{in } \mathscr{D}'(\mathbb{R}^d \times (0,1)) \tag{CER}$$

Estimate of the pure reaction part, $\boldsymbol{v} \equiv 0$.

$$\boxed{\mathsf{HK}^2(\mu_0, \mu_1) \leqslant \mathsf{H}^2(\mu_0, \mu_1)} \leqslant \mu_0(\mathbb{R}^d) + \mu_1(\mathbb{R}^d).$$

Estimate of the pure transport part, $\boldsymbol{w} = 0$: in the case when $\mu_0(\mathbb{R}^d) = \mu_1(\mathbb{R}^d)$

$$\boxed{\mathsf{HK}^2(\mu_0, \mu_1) \leqslant \mathsf{W}^2(\mu_0, \mu_1)}.$$

The distances between two Dirac masses

Suppose that $\mu_i = r_i^2 \delta_{x_i}$ and denote $|x|_\alpha = |x| \wedge \alpha$. We can compute

$$\boxed{\mathrm{HK}^2(r_0^2\delta_{x_0}, r_1^2\delta_{x_1}) = r_0^2 + r_1^2 - 2r_0 r_1 \cos(|x_1 - x_0|_{\pi/2})}$$

The distances between two Dirac masses

Suppose that $\mu_i = r_i^2 \delta_{x_i}$ and denote $|x|_\alpha = |x| \wedge \alpha$. We can compute

$$\boxed{\mathsf{HK}^2(r_0^2\delta_{x_0}, r_1^2\delta_{x_1}) = r_0^2 + r_1^2 - 2r_0r_1\cos(|x_1 - x_0|_{\pi/2})}$$

Truncation effect: when $|x_0 - x_1| \geqslant \pi/2$ we have

$$\mathsf{HK}^2(r_0^2\delta_{x_0}, r_1^2\delta_{x_1}) = r_0^2 + r_1^2.$$

The distances between two Dirac masses

Suppose that $\mu_i = r_i^2 \delta_{x_i}$ and denote $|x|_\alpha = |x| \wedge \alpha$. We can compute

$$\boxed{\mathsf{HK}^2(r_0^2\delta_{x_0}, r_1^2\delta_{x_1}) = r_0^2 + r_1^2 - 2r_0 r_1 \cos(|x_1 - x_0|_{\pi/2})}$$

Truncation effect: when $|x_0 - x_1| \geqslant \pi/2$ we have

$$\mathsf{HK}^2(r_0^2\delta_{x_0}, r_1^2\delta_{x_1}) = r_0^2 + r_1^2.$$

For the Gaussian Entropy Transport problem we get a different value:

$$\boxed{\mathsf{GET}(r_0^2\delta_{x_0}, r_1^2\delta_{x_1}) = r_0^2 + r_1^2 - 2r_0 r_1 e^{-|x_1 - x_0|^2/2}}$$

The distances between two Dirac masses

Suppose that $\mu_i = r_i^2\delta_{x_i}$ and denote $|x|_\alpha = |x| \wedge \alpha$. We can compute

$$\boxed{\mathsf{HK}^2(r_0^2\delta_{x_0}, r_1^2\delta_{x_1}) = r_0^2 + r_1^2 - 2r_0r_1\cos(|x_1 - x_0|_{\pi/2})}$$

Truncation effect: when $|x_0 - x_1| \geqslant \pi/2$ we have

$$\mathsf{HK}^2(r_0^2\delta_{x_0}, r_1^2\delta_{x_1}) = r_0^2 + r_1^2.$$

For the Gaussian Entropy Transport problem we get a different value:

$$\boxed{\mathsf{GET}(r_0^2\delta_{x_0}, r_1^2\delta_{x_1}) = r_0^2 + r_1^2 - 2r_0r_1e^{-|x_1-x_0|^2/2}}$$

HK is associated to the cone distance:

$$d_{\mathfrak{C}}^2((x_0, r_0), (x_1, r_1)) = r_0^2 + r_1^2 - 2r_0r_1\cos(|x_1 - x_0|_\pi)$$

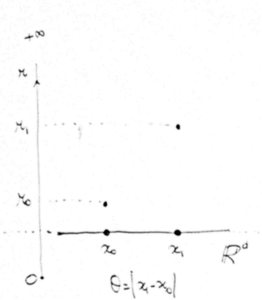
$d_{\mathfrak{C}}((x_0, r_0), (x_1, r_1))$ is a **length distance**.

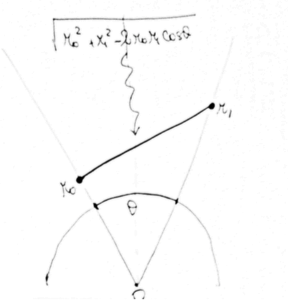
The function $r_0^2 + r_1^2 - 2r_0r_1e^{-|x_1-x_0|^2/2}$ does not satisfy the length property.

The Cone space

HK is associated to the cone distance:

$$d^2_{\mathfrak{C}}((x_0, r_0), (x_1, r_1)) = r_0^2 + r_1^2 - 2r_0 r_1 \cos(|x_1 - x_0|_\pi)$$





The Cone space

HK is associated to the cone distance:

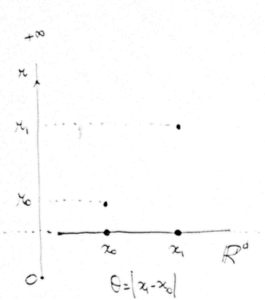
$$d^2_{\mathfrak{C}}((x_0, r_0), (x_1, r_1)) = r_0^2 + r_1^2 - 2r_0 r_1 \cos(|x_1 - x_0|_\pi)$$

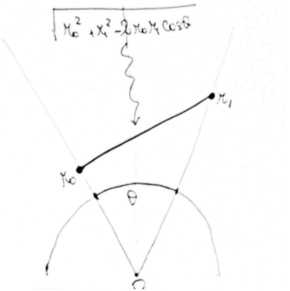
Cone space: identify all the points $(x, 0)$ with the vertex $\mathfrak{o}$.

$$\mathfrak{C} := \big(\mathbb{R}^d \times [0, \infty)\big)/\sim, \quad (x', r') \sim (x'', r'') \quad \Leftrightarrow \quad \begin{cases} x' = x'',\ r' = r'' \neq 0, \\ r' = r'' = 0 \end{cases}$$

$\mathfrak{C} \setminus \{\mathfrak{o}\}$ can be considered as a **smooth Riemannian manifold** with metric

$$\mathfrak{g}(dx, dr) := r^2|dx|^2 + |dr|^2$$





Duality with the conical Hamilton-Jacobi equation

If

$$\partial_t \xi_t + \frac{1}{2}|D\xi_t|^2 + 2\xi_t^2(x) \leqslant 0 \tag{CHJ}$$

and

$$\partial_t \mu_t + \nabla \cdot (\boldsymbol{v}_t \mu_t) = w_t \mu_t$$

then

$$\int \xi_1 \, d\mu_1 - \int \xi_0 \, d\mu_0 \leqslant \frac{1}{2} \int_0^1 \int \Big(|\boldsymbol{v}_t|^2 + \frac{1}{4} w_t^2 \Big) \, d\mu_t \, dt.$$

Duality with the conical Hamilton-Jacobi equation

If

$$\partial_t \xi_t + \frac{1}{2}|D\xi_t|^2 + 2\xi_t^2(x) \leqslant 0 \tag{CHJ}$$

and

$$\partial_t \mu_t + \nabla \cdot (\boldsymbol{v}_t \mu_t) = w_t \mu_t$$

then

$$\int \xi_1 \, d\mu_1 - \int \xi_0 \, d\mu_0 \leqslant \frac{1}{2} \int_0^1 \int \left(|\boldsymbol{v}_t|^2 + \frac{1}{4} w_t^2 \right) d\mu_t \, dt.$$

HK in duality with subsolutions to the conical Hamilton-Jacobi equations

$$\frac{1}{2}\mathsf{HK}^2(\mu_0, \mu_1) = \sup \Big\{ \int \xi_1 \, d\mu_1 - \int \xi_0 \, d\mu_0 : \xi \in C^1([0,1]; \mathrm{Lip}_b(\mathbb{R}^d))$$

$$\partial_t \xi_t + \frac{1}{2}|D\xi_t|^2 + 2\xi_t^2 \leqslant 0 \Big\}.$$

Conical Hopf-Lax representation formula

Given $\xi_0 \in \mathrm{Lip}_b(\mathbb{R}^d)$ with $\xi_0 > -1/2$, the viscosity solution (or the maximal subsolution) of the conical Hamilton Jacobi equation

$$\partial_t \xi_t + \frac{1}{2}|D\xi_t|^2 + 2\xi_t^2 = 0 \tag{CHJ}$$

is given by the **conical Hopf-Lax semigroup** (cf. Barron-Jensen-Liu for different representation formulae)

$$\boxed{\mathscr{P}_t \xi(x) := \inf_y \frac{1}{2t}\left[1 - \frac{\cos^2(|y-x|_{\pi/2})}{1+2t\xi(x)}\right]} \tag{CHL}$$

where

$$|x-y|_{\pi/2} := |x-y| \wedge \pi/2.$$

Conical Hopf-Lax representation formula

Given $\xi_0 \in \mathrm{Lip}_b(\mathbb{R}^d)$ with $\xi_0 > -1/2$, the viscosity solution (or the maximal subsolution) of the conical Hamilton Jacobi equation

$$\partial_t \xi_t + \frac{1}{2}|D\xi_t|^2 + 2\xi_t^2 = 0 \tag{CHJ}$$

is given by the **conical Hopf-Lax semigroup** (cf. BARRON-JENSEN-LIU for different representation formulae)

$$\boxed{\mathscr{P}_t \xi(x) := \inf_y \frac{1}{2t}\left[1 - \frac{\cos^2(|y-x|_{\pi/2})}{1+2t\xi(x)}\right]} \tag{CHL}$$

where

$$|x-y|_{\pi/2} := |x-y| \wedge \pi/2.$$

Conical Hopf-Lax representation for HK

$$\frac{1}{2}\mathsf{HK}^2(\mu_0,\mu_1) = \sup\left\{\int \xi_1\,d\mu_1 - \int \xi_0\,d\mu_0 : \xi_1 = \mathscr{P}_1\xi_0\right\}$$

Conical lift of the Hopf-Lax formula

Formally, if ξ is a solution of

$$\partial_t \xi_t + \frac{1}{2}|D\xi_t|^2 + 2\xi_t^2 \leqslant 0 \tag{CHJ}$$

then $\boxed{\zeta_t(x,r) := \xi_t(x) r^2}$ is a solution of

$$\partial_t \zeta_t + \frac{1}{2}|D_{\mathfrak{C}}\xi_t|^2 \leqslant 0 \tag{HJ}$$

since

$$\frac{1}{2}|D_{\mathfrak{C}}\zeta|^2 = \frac{1}{2}\mathfrak{g}^*(D_x\zeta, \partial_r\zeta) = \frac{1}{2}\Big(\frac{1}{r^2}|D_x\zeta|^2 + (\partial_r\zeta)^2\Big) = \Big(\frac{1}{2}|D\xi_t|^2 r^2 + 2\xi_t^2\Big) r^2$$

Conical lift of the Hopf-Lax formula

Formally, if ξ is a solution of

$$\partial_t \xi_t + \frac{1}{2}|D\xi_t|^2 + 2\xi_t^2 \leqslant 0 \tag{CHJ}$$

then $\boxed{\zeta_t(x,r) := \xi_t(x) r^2}$ is a solution of

$$\partial_t \zeta_t + \frac{1}{2}|D_{\mathfrak{C}} \xi_t|^2 \leqslant 0 \tag{HJ}$$

since

$$\frac{1}{2}|D_{\mathfrak{C}}\zeta|^2 = \frac{1}{2}\mathfrak{g}^*(D_x\zeta, \partial_r\zeta) = \frac{1}{2}\Big(\frac{1}{r^2}|D_x\zeta|^2 + (\partial_r\zeta)^2\Big) = \Big(\frac{1}{2}|D\xi_t|^2 r^2 + 2\xi_t^2\Big) r^2$$

The Hopf-Lax semigroup in $\mathfrak{C}$

$$\begin{aligned}\mathscr{Q}_t^{\mathfrak{C}}\zeta(x,r) &= \min_{y,s} \zeta(y,s) + \frac{1}{2t} d_{\mathfrak{C}}^2((x,r),(y,s)) \\ &= \min_{y,s} \xi(y)s^2 + \frac{1}{2t}\Big(r^2 + s^2 - 2rs\cos(|x-y|_\pi)\Big)\end{aligned}$$

yields

$$\mathscr{Q}_t^{\mathfrak{C}}\zeta(x,r) = \xi_t(x) r^2, \quad \xi_t = \mathscr{P}_t \xi.$$

Conical Hopf-Lax and dual Kantorovich formulation

Setting $\psi_i := 2\xi_i$ in the conical Hopf-Lax formula

$$\boxed{\mathscr{P}_1\xi(x) := \inf_y \frac{1}{2}\left[1 - \frac{\cos^2(|y-x|_{\pi/2})}{1+2\xi(x)}\right]} \tag{CHL}$$

Dual Kantorovich formulation (I)

$$\mathsf{HK}^2(\mu_0,\mu_1) = \sup\Big\{\int \psi_1\,\mathrm{d}\mu_1 - \int \psi_0\,\mathrm{d}\mu_0 : \ \psi_0 > -1,\ \psi_1 < 1$$
$$(1-\psi_1(y))(1+\psi_0(x)) \geqslant \cos^2(|y-x|_{\pi/2})\Big\}$$

Conical Hopf-Lax and dual Kantorovich formulation

Setting $\psi_i := 2\xi_i$ in the conical Hopf-Lax formula

$$\boxed{\mathscr{P}_1\xi(x) := \inf_y \frac{1}{2}\left[1 - \frac{\cos^2(|y-x|_{\pi/2})}{1+2\xi(x)}\right]} \tag{CHL}$$

Dual Kantorovich formulation (I)

$$\mathsf{HK}^2(\mu_0,\mu_1) = \sup\Big\{\int \psi_1\,d\mu_1 - \int \psi_0\,d\mu_0 : \ \psi_0 > -1,\ \psi_1 < 1$$
$$(1-\psi_1(y))(1+\psi_0(x)) \geqslant \cos^2(|y-x|_{\pi/2})\Big\}$$

Compare with the corresponding definition of the Hellinger distance

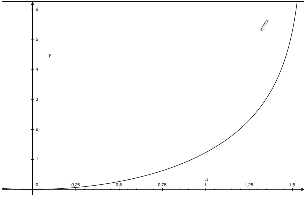
$$\mathsf{H}^2(\mu_0,\mu_1) = \sup\Big\{\int \psi_1\,d\mu_1 - \int \psi_0\,d\mu_0 : \ \psi_0 > -1,\ \psi_1 < 1$$
$$(1-\psi_1(x))(1+\psi_0(x)) \geqslant 1\Big\}$$

Dual formulation (II)

Change of variable: $\phi_1 := -\log(1-\psi_1)$, $\phi_0 := \log(1+\psi_0)$

$$(1-\psi_1(y))(1+\psi_0(x)) \geqslant \cos^2(|y-x|_{\pi/2}) \quad \Leftrightarrow \quad \phi_1(y)-\phi_0(x) \leqslant \ell(x,y),$$

$$\ell(x,y) = -\log\left(\cos^2(|y-x|_{\pi/2})\right) = \log\left(1+\tan^2(|y-x|_{\pi/2})\right)$$

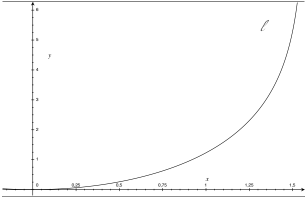


Dual formulation (II)

Change of variable: $\phi_1 := -\log(1-\psi_1)$, $\phi_0 := \log(1+\psi_0)$

$$(1-\psi_1(y))(1+\psi_0(x)) \geqslant \cos^2(|y-x|_{\pi/2}) \quad \Leftrightarrow \quad \phi_1(y)-\phi_0(x) \leqslant \ell(x,y),$$

$$\ell(x,y) = -\log\left(\cos^2(|y-x|_{\pi/2})\right) = \log\left(1+\tan^2(|y-x|_{\pi/2})\right)$$



Dual Kantorovich formulation (II)

$$\mathsf{HK}^2(\mu_0,\mu_1) = \sup\left\{ \int (1-e^{-\phi_1})\,d\mu_1 - \int (e^{\phi_0}-1)\,d\mu_0 : \right.$$
$$\left. \phi_1(y)-\phi_0(x) \leqslant \ell(x,y) \right\}$$

Primal formulation: Logarithmic Entropy-Transport problem

The **Legendre conjugate** of $G(\phi) := e^{\phi} - 1$ is

$$\mathrm{LE}(s) := s \log s - (s - 1)$$

Primal formulation: Logarithmic Entropy-Transport problem

The **Legendre conjugate** of $G(\phi) := e^{\phi} - 1$ is

$$\mathsf{LE}(s) := s \log s - (s - 1)$$

When $\boldsymbol{\gamma}$ is a plan in $\mathcal{M}(\mathbb{R}^d \times \mathbb{R}^d)$ with marginals γ_i we find

Logarithmic Entropy-Transport (LET) formulation

$$\mathsf{LET}(\mu_0, \mu_1) = \min_{\boldsymbol{\gamma} \in \mathcal{M}(\mathbb{R}^d \times \mathbb{R}^d)} \Big(\mathscr{E}(\gamma_0 | \mu_0) + \mathscr{E}(\gamma_1 | \mu_1) + \int \ell(x, y) \, d\boldsymbol{\gamma}(x, y) \Big)$$

where $\ell(x, y) = \log(1 + \tan^2(|x - y|_{\pi/2}))$.

Primal formulation: Logarithmic Entropy-Transport problem

The **Legendre conjugate** of $G(\phi) := e^{\phi} - 1$ is

$$\mathrm{LE}(s) := s \log s - (s-1)$$

When γ is a plan in $\mathcal{M}(\mathbb{R}^d \times \mathbb{R}^d)$ with marginals γ_i we find

Logarithmic Entropy-Transport (LET) formulation

$$\mathrm{LET}(\mu_0, \mu_1) = \min_{\gamma \in \mathcal{M}(\mathbb{R}^d \times \mathbb{R}^d)} \Big(\mathscr{E}(\gamma_0|\mu_0) + \mathscr{E}(\gamma_1|\mu_1) + \int \ell(x,y)\, \mathrm{d}\gamma(x,y) \Big)$$

where $\ell(x,y) = \log(1 + \tan^2(|x-y|_{\pi/2}))$.

$$\mathrm{HK}^2(\mu_0,\mu_1) = \mathrm{LET}(\mu_0,\mu_1)$$

Four equivalent formulations for HK

$$
\begin{tikzcd}[column sep=large, row sep=large]
\textbf{Dynamic formulation} \arrow[r, leftrightarrow, "\textit{Duality}"] \arrow[d, leftrightarrow] & \textbf{Conical Hamilton Jacobi} \arrow[d, leftrightarrow, "\textit{Conical Hopf-Lax}"] \\
\textbf{Optimal Entropy-Transport} \arrow[r, leftrightarrow, "\textit{Convex duality}"'] & \textbf{Kantorovich duality}
\end{tikzcd}
$$

Four equivalent formulations for HK

Dynamic formulation	$\overset{\textit{Duality}}{\Longleftrightarrow}$	Conical Hamilton Jacobi
$\Updownarrow$		$\Updownarrow$ *Conical Hopf-Lax*
Optimal Entropy-Transport	$\underset{\textit{Convex duality}}{\Longleftrightarrow}$	Kantorovich duality

$$
\begin{aligned}
\mathsf{HK}^2(\mu_0,\mu_1) &= \min\Big\{\int_0^1\int\Big(|v_t|^2+\frac{1}{4}|w_t|^2\Big)\,d\mu_t\,dt : \mu\in C([0,1];\mathcal{M}(\mathbb{R}^d)),\\
&\qquad \partial_t\mu_t+\nabla\cdot(v_t\mu_t)=w_t\mu_t,\quad \mu_{t=i}=\mu_i\Big\} && \text{(CER)}\\
&= 2\sup\Big\{\int\xi_1\,d\mu_1-\int\xi_0\,d\mu_0 : \xi\in C^1([0,1];\mathrm{Lip}_b(\mathbb{R}^d))\\
&\qquad \partial_t\xi_t+\frac{1}{2}|D\xi_t|^2+2\xi_t^2\leqslant 0\Big\} && \text{(CHJ)}\\
&= 2\sup\Big\{\int\xi_1\,d\mu_1-\int\xi_0\,d\mu_0 : \xi_1=\mathscr{P}_1\xi_0\Big\} && \text{(CHL)}\\
&= \min_{\gamma}\mathscr{E}(\gamma_0|\mu_0)+\mathscr{E}(\gamma_1|\mu_1)+\int\ell(x,y)\,d\boldsymbol{\gamma}(x,y). && \text{(LET)}
\end{aligned}
$$

Homogeneous and Monge formulation

Homogeneous formulation by scaling invariance

$$\mathsf{HK}^2(\mu_0,\mu_1) = \min_{\gamma} \int \mathrm{H}(r_0^2(x), r_1^2(y), \mathsf{c}(x,y))\, \mathrm{d}\boldsymbol{\gamma}(x,y)$$

where $\mu_i = r_i^2 \gamma_i = r_i^2 \pi^i_\sharp \boldsymbol{\gamma}$ and

$$\mathrm{H}(r_0, r_1, c) := \inf_{\vartheta > 0} \Big(r_0 \mathsf{LE}(\vartheta/r_0) + r_1 \mathsf{LE}(\vartheta/r_1) + c\,\vartheta \Big)$$

A simple calculation yields

$$\boxed{\mathrm{H}(r_0^2, r_1^2, c) = r_0^2 + r_1^2 - 2 r_0 r_1 \mathrm{e}^{-c/2}} = r_0^2 + r_1^2 - 2 r_0 r_1 \cos(|x-y|_{\pi/2}).$$

Homogeneous and Monge formulation

Homogeneous formulation by scaling invariance

$$\mathsf{HK}^2(\mu_0, \mu_1) = \min_{\gamma} \int H(r_0^2(x), r_1^2(y), \mathsf{c}(x, y))\, d\boldsymbol{\gamma}(x, y)$$

where $\mu_i = r_i^2 \gamma_i = r_i^2 \pi^i_\sharp \boldsymbol{\gamma}$ and

$$H(r_0, r_1, c) := \inf_{\vartheta > 0} \Big(r_0 \mathsf{LE}(\vartheta / r_0) + r_1 \mathsf{LE}(\vartheta / r_1) + c\,\vartheta \Big)$$

A simple calculation yields

$$\boxed{H(r_0^2, r_1^2, c) = r_0^2 + r_1^2 - 2 r_0 r_1 e^{-c/2}} = r_0^2 + r_1^2 - 2 r_0 r_1 \cos(|x - y|_{\pi/2}).$$

Homogeneous and Monge formulation

Homogeneous formulation by scaling invariance

$$\mathsf{HK}^2(\mu_0,\mu_1) = \min_{\gamma} \int \mathsf{H}(r_0^2(x), r_1^2(y), c(x,y))\, d\boldsymbol{\gamma}(x,y)$$

where $\mu_i = r_i^2 \gamma_i = r_i^2 \pi^i_\sharp \boldsymbol{\gamma}$ and

$$\mathsf{H}(r_0, r_1, c) := \inf_{\vartheta > 0} \Big(r_0 \mathsf{LE}(\vartheta/r_0) + r_1 \mathsf{LE}(\vartheta/r_1) + c\,\vartheta \Big)$$

A simple calculation yields

$$\boxed{\mathsf{H}(r_0^2, r_1^2, c) = r_0^2 + r_1^2 - 2 r_0 r_1 e^{-c/2}} = r_0^2 + r_1^2 - 2 r_0 r_1 \cos(|x-y|_{\pi/2}).$$

Dilation-transport couples and Monge formulation If $\mu_0 \ll \mathscr{L}^d$, then

$$\begin{aligned}\mathsf{HK}^2(\mu_0,\mu_1) =& \mu_0(\mathbb{R}^d) + \mu_1(\mathbb{R}^d) \\ & - 2\max\Big\{ \int q(x) \cos(|x - \mathbf{t}(x)|_{\pi/2}) \Big)\, d\mu_0(x) : \ q \geqslant 0,\ \mathbf{t}_\sharp(q^2\mu_0) \leqslant \mu_1 \Big\}.\end{aligned}$$

Homogeneous marginals and Kantorovich-Wasserstein distance on $\mathfrak{C}$

Any measure $\mu \in \mathcal{M}(\mathbb{R}^d)$ can be easily **lifted** to a measure in the cone $\mathfrak{C}$ by $\alpha := \mu \otimes \delta_1$.

Conversely, any measure α in the cone $\mathfrak{C}$ can be "projected" on $\mathbb{R}^d$ by taking the **homogeneous marginal:**

$$\boxed{\mu = \mathfrak{h}\alpha = \pi^x_\sharp\big(r^2 \cdot \alpha\big), \quad \int \xi(x)\,\mathrm{d}\mu(x) = \int_{\mathfrak{C}} \xi(x) r^2\,\mathrm{d}\alpha(x,r).}$$

Homogeneous marginals and Kantorovich-Wasserstein distance on $\mathfrak{C}$

Any measure $\mu \in \mathcal{M}(\mathbb{R}^d)$ can be easily **lifted** to a measure in the cone $\mathfrak{C}$ by $\alpha := \mu \otimes \delta_1$.

Conversely, any measure α in the cone $\mathfrak{C}$ can be "projected" on $\mathbb{R}^d$ by taking the **homogeneous marginal:**

$$\boxed{\mu = \mathfrak{h}\alpha = \pi^x_\sharp\big(r^2 \cdot \alpha\big), \quad \int \xi(x)\, d\mu(x) = \int_{\mathfrak{C}} \xi(x) r^2\, d\alpha(x,r).}$$

HK and $W_{d_{\mathfrak{C}}}$ via homogeneous marginals

$$\mathsf{HK}^2(\mu_0,\mu_1) = \min\Big\{W^2_{d_{\mathfrak{C}}}(\alpha_0,\alpha_1) : \alpha_i \in \mathcal{M}(\mathfrak{C}),\ \mathfrak{h}\alpha_i = \mu_i\Big\}$$

A list of useful properties

- HK can be defined starting from **any metric space** (X, d).

A list of useful properties

- HK can be defined starting from **any metric space** (X, d).
- $(\mathcal{M}(X), \mathsf{HK})$ is a **complete and separable metric space** if X is complete and separable; the induced topology coincides with the topology of weak convergence (no bounds on moments are required). If X is compact then bounded sets in $\mathcal{M}(X)$ are relatively compact.

A list of useful properties

- HK can be defined starting from **any metric space** (X, d).
- $(\mathcal{M}(X), \mathsf{HK})$ is a **complete and separable metric space** if X is complete and separable; the induced topology coincides with the topology of weak convergence (no bounds on moments are required). If X is compact then bounded sets in $\mathcal{M}(X)$ are relatively compact.
- If (X, d) is length (resp. geodesic) then $(\mathcal{M}(X), \mathsf{HK})$ is **length (resp. geodesic).**

A list of useful properties

- HK can be defined starting from **any metric space** (X, d).
- $(\mathcal{M}(X), \mathsf{HK})$ is a **complete and separable metric space** if X is complete and separable; the induced topology coincides with the topology of weak convergence (no bounds on moments are required). If X is compact then bounded sets in $\mathcal{M}(X)$ are relatively compact.
- If (X, d) is length (resp. geodesic) then $(\mathcal{M}(X), \mathsf{HK})$ is **length (resp. geodesic).**
- If $X = \mathbb{R}^d$ and $\mu_0 \ll \mathscr{L}^d$ then **there exists a unique geodesic** connecting μ_0 to μ_1 and a **unique optimal plan** $\boldsymbol{\gamma}$ minimizing $\mathsf{LET}(\mu_0, \mu_1)$.

A list of useful properties

- HK can be defined starting from **any metric space** (X, d).
- $(\mathcal{M}(X), \mathsf{HK})$ is a **complete and separable metric space** if X is complete and separable; the induced topology coincides with the topology of weak convergence (no bounds on moments are required). If X is compact then bounded sets in $\mathcal{M}(X)$ are relatively compact.
- If (X, d) is length (resp. geodesic) then $(\mathcal{M}(X), \mathsf{HK})$ is **length (resp. geodesic).**
- If $X = \mathbb{R}^d$ and $\mu_0 \ll \mathscr{L}^d$ then **there exists a unique geodesic** connecting μ_0 to μ_1 and a **unique optimal plan** $\boldsymbol{\gamma}$ minimizing $\mathsf{LET}(\mu_0, \mu_1)$.
- Power-like entropies $\int \rho^\alpha \, dx$, $\mu = \rho\mathscr{L}^d$, are **geodesically convex** if $\alpha > 1$ (reinforced McCann condition).

Contraction and regularizing effects for the Heat/Fokker-Planck flow

Solutions to the **Fokker-Planck equation**

$$\partial_t \mu - \Delta \mu + \nabla \cdot (x\mu) = 0, \quad \lambda \geqslant 0$$

satisfy the **contraction** property [Luise-S.]

$$\boxed{\mathsf{H}(\mu_t, \nu_t) \leqslant \mathsf{H}(\mu_0, \nu_0), \quad \mathsf{HK}(\mu_t, \nu_t) \leqslant \mathsf{HK}(\mu_0, \nu_0)}$$

Contraction and regularizing effects for the Heat/Fokker-Planck flow

Solutions to the **Fokker-Planck equation**

$$\partial_t \mu - \Delta \mu + \nabla \cdot (x\mu) = 0, \quad \lambda \geqslant 0$$

satisfy the **contraction** property [Luise-S.]

$$\boxed{\mathsf{H}(\mu_t, \nu_t) \leqslant \mathsf{H}(\mu_0, \nu_0), \quad \mathsf{HK}(\mu_t, \nu_t) \leqslant \mathsf{HK}(\mu_0, \nu_0)}$$

with the **regularazing effect**

$$\mathsf{H}^2(\mu_t, \nu_t) \leqslant \frac{1}{c(t)} \mathsf{HK}^2(\mu_0, \nu_0), \quad c(t) \sim t \quad \text{as } t \downarrow 0.$$

Contraction and regularizing effects for the Heat/Fokker-Planck flow

Solutions to the **Fokker-Planck equation**

$$\partial_t \mu - \Delta\mu + \nabla \cdot (x\mu) = 0, \quad \lambda \geqslant 0$$

satisfy the **contraction** property [Luise-S.]

$$\boxed{H(\mu_t, \nu_t) \leqslant H(\mu_0, \nu_0), \quad HK(\mu_t, \nu_t) \leqslant HK(\mu_0, \nu_0)}$$

with the **regularazing effect**

$$H^2(\mu_t, \nu_t) \leqslant \frac{1}{c(t)} HK^2(\mu_0, \nu_0), \quad c(t) \sim t \quad \text{as } t \downarrow 0.$$

The Hellinger contraction property holds for **general sub-Markov semigroups** generated by Dirichlet forms in some $L^2(X, \mathfrak{m})$.

The Hellinger-Kantorovich contraction and the regularizing effect depend on **lower weighted Ricci curvature bounds** (the curvature condition $\mathrm{CD}(K, \infty)$).

Open problems and future directions

- Study **Optimal Entropy-Transport problems** for different classes of cost and entropies: they provide interesting Transport versions of well known distances in statistics and information theory, as the *Jensen-Shannon divergence*, the *triangular discrimination* and the *total variation distance* [Piccoli-Rossi]
- Apply these techniques to other dynamic distances and gradient flows (cf. DOLBEAULT-NAZARET-S., MIELKE, MAAS,...)

Combining Entropy and Transport

- Two **convex entropy functions** $F_i : [0,\infty) \to [0,\infty]$ with the corresponding **Entropy functionals** $\mathscr{F}_i$

$$\boxed{\mathscr{F}_i(\gamma_i|\mu_i) := \int F_i(\rho_i)\,d\mu_i + F_i'(\infty)\gamma_i^\perp(\mathbb{R}^d)} \quad \gamma_i = \rho_i\mu_i + \gamma_i^\perp$$

Combining Entropy and Transport

- Two **convex entropy functions** $F_i : [0, \infty) \to [0, \infty]$ with the corresponding **Entropy functionals** $\mathscr{F}_i$

$$\boxed{\mathscr{F}_i(\gamma_i|\mu_i) := \int F_i(\rho_i)\,d\mu_i + F_i'(\infty)\gamma_i^\perp(\mathbb{R}^d)} \quad \gamma_i = \rho_i\mu_i + \gamma_i^\perp$$

- c is a general nonnegative lower semicontinuous cost.

Combining Entropy and Transport

- Two **convex entropy functions** $F_i : [0, \infty) \to [0, \infty]$ with the corresponding **Entropy functionals** $\mathscr{F}_i$

$$\boxed{\mathscr{F}_i(\gamma_i|\mu_i) := \int F_i(\rho_i)\,d\mu_i + F_i'(\infty)\gamma_i^\perp(\mathbb{R}^d)} \quad \gamma_i = \rho_i\mu_i + \gamma_i^\perp$$

- c is a general nonnegative lower semicontinuous cost.

Optimal Entropy Transport problem

$$\mathrm{ET}(\mu_0, \mu_1) := \min_{\boldsymbol{\gamma}} \ \mathscr{F}_0(\gamma_0|\mu_0) + \mathscr{F}_1(\gamma_1|\mu_1) + \int c(x, y)\,d\boldsymbol{\gamma}(x, y). \qquad \text{(ET)}$$

where γ_i are the marginals of $\boldsymbol{\gamma} \in \mathcal{M}(\mathbb{R}^d \times \mathbb{R}^d)$.

Combining Entropy and Transport

- Two **convex entropy functions** $F_i : [0,\infty) \to [0,\infty]$ with the corresponding **Entropy functionals** $\mathscr{F}_i$

$$\boxed{\mathscr{F}_i(\gamma_i|\mu_i) := \int F_i(\rho_i)\,d\mu_i + F_i'(\infty)\gamma_i^\perp(\mathbb{R}^d)} \quad \gamma_i = \rho_i\mu_i + \gamma_i^\perp$$

- c is a general nonnegative lower semicontinuous cost.

Optimal Entropy Transport problem

$$\mathrm{ET}(\mu_0,\mu_1) := \min_{\boldsymbol{\gamma}} \; \mathscr{F}_0(\gamma_0|\mu_0) + \mathscr{F}_1(\gamma_1|\mu_1) + \int c(x,y)\,d\boldsymbol{\gamma}(x,y). \tag{ET}$$

where γ_i are the marginals of $\boldsymbol{\gamma} \in \mathcal{M}(\mathbb{R}^d \times \mathbb{R}^d)$.

(ET) can be considered as a sort of **relaxation of the Optimal Transport problem**, where the constraints "i-marginal of $\boldsymbol{\gamma}$=μ_i" has been substituted by the penalizing entropies $\mathscr{F}_i$.