Unbalanced Optimal Transport: Convex Relaxation and Dynamic Perspectives

Giuseppe Savaré
Department of Decision Sciences, Bocconi University, Milan, Italy
In collaboration with Matthias Liero, Alexander Mielke (WIAS, Berlin), and Giacomo Sodini (Vienna)

Kantorovich Initiative Seminar
November 30, 2023
1 Unbalanced Optimal Transport: a relaxation viewpoint

2 The Hellinger-Kantorovich metric between positive measures of arbitrary mass

3 Geodesics and geodesic convexity

4 Regularity of solutions to the Conical Hopf-Lax semigroup
Unbalanced Optimal Transport starting from Dirac masses

$X_i$ Polish topological spaces (the topology is induced by a separable and complete metric).

$\mathcal{M}(X)$ is the space of nonnegative Borel measures $\mu$ on $X$ with finite mass $\mu(X) < \infty$. 

Cone space: identify all the points $(x, 0)$ with the vertex $o$ (they correspond to the null measure $0 \delta_x = 0$).
**Unbalanced Optimal Transport starting from Dirac masses**

$X_i$ Polish topological spaces (the topology is induced by a separable and complete metric).

$\mathcal{M}(X)$ is the space of nonnegative Borel measures $\mu$ on $X$ with finite mass $\mu(X) < \infty$.

We introduce a function $h : (X_0 \times \mathbb{R}_+) \times (X_1 \times \mathbb{R}_+) \rightarrow [0, +\infty]$ which characterizes the cost of connecting two Dirac measures with possibly different mass:

$$h(x_0, r_0; x_1, r_1) := \text{UOT}_{\text{Dirac}}(r_0 \delta_{x_0}, r_1 \delta_{x_1}) \quad x_i \in X_i, \; r_i \geq 0$$
Unbalanced Optimal Transport starting from Dirac masses

$X_i$ Polish topological spaces (the topology is induced by a separable and complete metric).

$\mathcal{M}(X)$ is the space of nonnegative Borel measures $\mu$ on $X$ with finite mass $\mu(X) < \infty$.

We introduce a function $h : (X_0 \times \mathbb{R}_+) \times (X_1 \times \mathbb{R}_+) \to [0, +\infty]$ which characterizes the cost of connecting two Dirac measures with possibly different mass:

$$h(x_0, r_0; x_1, r_1) := \text{UOT}_{\text{Dirac}}(r_0 \delta_{x_0}, r_1 \delta_{x_1}) \quad x_i \in X_i, \ r_i \geq 0$$

Simplifying assumptions: for every $x_0, x_1$

$$\begin{cases} 
  h(x_0, r_0; x_1, 0) \text{ is independent of } x_1, & h(x_0, 0; x_1, r_1) \text{ is independent of } x_0. \\
  (r_0, r_1) \mapsto h(x_0, r_0; x_1, r_1) \text{ is positively } 1\text{-homogeneous and convex}
\end{cases}$$
Unbalanced Optimal Transport starting from Dirac masses

$X_i$ Polish topological spaces (the topology is induced by a separable and complete metric).

$\mathcal{M}(X)$ is the space of nonnegative Borel measures $\mu$ on $X$ with finite mass $\mu(X) < \infty$.

We introduce a function $h : (X_0 \times \mathbb{R}_+) \times (X_1 \times \mathbb{R}_+) \rightarrow [0, +\infty]$ which characterizes the cost of connecting two Dirac measures with possibly different mass:

$$h(x_0, r_0; x_1, r_1) := \text{UOT}_{\text{Dirac}}(r_0 \delta_{x_0}, r_1 \delta_{x_1}) \quad x_i \in X_i, \ r_i \geq 0$$

Simplifying assumptions: for every $x_0, x_1$

$$\begin{cases}
    h(x_0, r_0; x_1, 0) \text{ is independent of } x_1, \\
    h(x_0, 0; x_1, r_1) \text{ is independent of } x_0.
\end{cases}$$

$(r_0, r_1) \mapsto h(x_0, r_0; x_1, r_1)$ is positively 1-homogeneous and convex

Cone space: identify all the points $(x, 0)$ with the vertex $o$ (they correspond to the null measure $0\delta_x = 0$)

$$\mathcal{C}[X] := (X \times [0, \infty)) / \sim, \quad (x', r') \sim (x'', r'') \iff \begin{cases}
    x' = x'', \\
    r' = r'' \neq 0, \\
    r' = r'' = 0
\end{cases}$$
The Balanced OT case: $c : X_0 \times X_1 \rightarrow \mathbb{R}$ is a cost function,

$$h(x_0, r_0; x_1, r_1) = \begin{cases} 
rc(x_0, x_1) & \text{if } r_0 = r_1 = r; \\
+\infty & \text{if } r_0 \neq r_1
\end{cases}$$
Examples

The Balanced OT case: $c : X_0 \times X_1 \to \mathbb{R}$ is a cost function,

$$h(x_0, r_0; x_1, r_1) = \begin{cases} 
rc(x_0, x_1) & \text{if } r_0 = r_1 = r; \\
+\infty & \text{if } r_0 \neq r_1
\end{cases}$$

The Hellinger-Kakutani case:

$$h(x_0, r_0; x_1, r_1) = \begin{cases} 
(r_0 - r_1)^2 = r_0 + r_1 - 2\sqrt{r_0r_1} & \text{if } x_0 = x_1 \\
+\infty & \text{if } x_0 \neq x_1
\end{cases}$$
Examples

The Balanced OT case: $c : X_0 \times X_1 \rightarrow \mathbb{R}$ is a cost function,

$$h(x_0, r_0; x_1, r_1) = \begin{cases} rc(x_0, x_1) & \text{if } r_0 = r_1 = r; \\ +\infty & \text{if } r_0 \neq r_1 \end{cases}$$

The Hellinger-Kakutani case:

$$h(x_0, r_0; x_1, r_1) = \begin{cases} (\sqrt{r_0} - \sqrt{r_1})^2 = r_0 + r_1 - 2\sqrt{r_0r_1} & \text{if } x_0 = x_1 \\ +\infty & \text{if } x_0 \neq x_1 \end{cases}$$

The Entropic Unbalanced Cost

$$h(x_0, r_0; x_1, r_1) = r_0 + r_1 - 2\sqrt{r_0r_1}e^{-c(x_0, x_1)}$$

$$= (\sqrt{r_0} - \sqrt{r_1})^2 + 2\sqrt{r_0r_1}(1 - e^{-c(x_0, x_1)}).$$
What is the most natural way (from the convex analysis viewpoint) to extend $UOT_{\text{Dirac}}$

$$h(x_0, r_0; x_1, r_1) := UOT_{\text{Dirac}}(r_0 \delta_{x_0}, r_1 \delta_{x_1}) \quad \forall x_i \in X_i, r_i \geq 0$$

to a function in $\mathcal{M}(X_0) \times \mathcal{M}(X_1)$?
Unbalanced Optimal Transport as convex envelope

What is the most natural way (from the convex analysis viewpoint) to extend $\text{UOT}_{\text{Dirac}}$

$$h(x_0, r_0; x_1, r_1) := \text{UOT}_{\text{Dirac}}(r_0\delta_{x_0}, r_1\delta_{x_1}) \quad x_i \in X_i, \ r_i \geq 0$$

to a function in $\mathcal{M}(X_0) \times \mathcal{M}(X_1)$?

**$\Gamma$-relaxation of $\text{UOT}_{\text{Dirac}}$: the largest convex and l.s.c. functional**

$\Gamma$-$\text{UOT}_{\text{Dirac}} : \mathcal{M}(X_0) \times \mathcal{M}(X_1) \to [0, +\infty]$

dominated by $\text{UOT}_{\text{Dirac}}$:

$$\begin{cases} 
\Gamma$-$\text{UOT}_{\text{Dirac}}(r_0\delta_{x_0}, r_1\delta_{x_1}) \leq \text{UOT}_{\text{Dirac}}(r_0\delta_{x_0}, r_1\delta_{x_1}) & \text{for every } r_i \geq 0, \ x_i \in X_i \\
\hat{\text{UOT}} \text{ convex, l.s.c., } \hat{\text{UOT}} \leq \text{UOT}_{\text{Dirac}} \Rightarrow \hat{\text{UOT}} \leq \Gamma$-$\text{UOT}_{\text{Dirac}}.
\end{cases}$$
If $\mathcal{F} : V \to (-\infty, +\infty]$ is a given function, defined in a vector space $V$ in duality with $V'$, its $\Gamma$-regularization can be characterized in two equivalent ways:

- Using the **Legendre transform thanks to Fenchel-Moreau Theorem**:

  $$\mathcal{F}^*(\phi) := \sup_{v \in V} \langle \phi, v \rangle - \mathcal{F}(v), \quad \phi \in V'$$

  $$\Gamma-\mathcal{F}(v) = \mathcal{F}^{**}(v) := \sup_{\phi \in V'} \langle \phi, v \rangle - \mathcal{F}^*(\phi)$$
Two equivalent constructions

If $\mathcal{F}: V \to (-\infty, +\infty]$ is a given function, defined in a vector space $V$ in duality with $V'$, its $\Gamma$-regularization can be characterized in two equivalent ways:

- Using the **Legendre transform thanks to Fenchel-Moreau Theorem:**

  $$\mathcal{F}^*((\phi)) := \sup_{v \in V} \langle \phi, v \rangle - \mathcal{F}(v), \quad \phi \in V'$$

  $$\Gamma-\mathcal{F}(v) = \mathcal{F}^{**}(v) := \sup_{\phi \in V'} \langle \phi, v \rangle - \mathcal{F}^*(\phi)$$

- Computing the **convex envelope:**

  $$\text{co.}\mathcal{F}(v) := \inf \left\{ \sum_i \alpha_i \mathcal{F}(v_i) : \alpha_i \geq 0, \sum_i \alpha_i = 1, \sum_i \alpha_i v_i = v \right\}$$

  and then taking the **l.s.c. relaxation** of $\text{co.}\mathcal{F}(v)$. If $\mathcal{F}$ is coercive we have the integral description

  $$\Gamma-\mathcal{F}(v) = \min \left\{ \int_V \mathcal{F}(w) \, d\alpha(w) : \alpha \in \mathcal{P}(V), \int_V w \, d\alpha(w) = v \right\}.$$
Convex duality

Γ-UOT\textsubscript{Dirac} can be computed by **Legendre transform thanks to Fenchel-Moreau Theorem**, using the duality between $\mathcal{M}(X)$ and $C_b(X)$.

$$\text{UOT}^*\text{Dirac}(\phi_0, \phi_1) = \sup \left\{ \tau_0 \phi_0(x_0) + \tau_1 \phi_1(x_1) - h(x_0, \tau_0; x_1, \tau_1) : \tau_i \geq 0, \ x_i \in X_i \right\}$$

$$= \begin{cases} 0 & \text{if } \tau_0 \phi_0(x_0) + \tau_1 \phi_1(x_1) \leq h(x_0, \tau_0; x_1, \tau_1), \\ +\infty & \text{otherwise} \end{cases}$$
Convex duality

\( \Gamma - \text{UOT}_{\text{Dirac}} \) can be computed by Legendre transform thanks to Fenchel-Moreau Theorem, using the duality between \( \mathcal{M}(X) \) and \( C_b(X) \).

\[
\text{UOT}_{\text{Dirac}}^*(\phi_0, \phi_1) = \sup \left\{ r_0 \phi_0(x_0) + r_1 \phi_1(x_1) - h(x_0, r_0; x_1, r_1) : r_i \geq 0, x_i \in X_i \right\} \\
= \begin{cases} 
0 & \text{if } r_0 \phi_0(x_0) + r_1 \phi_1(x_1) \leq h(x_0, r_0; x_1, r_1), \\
+\infty & \text{otherwise}
\end{cases}
\]

\( \text{UOT}_{\text{Dirac}}^* \) is just the indicator function of a convex set \( K[h] \) of admissible Kantorovich potentials \((\phi_0, \phi_1) \in C_b(X_0) \times C_b(X_1)\).

The dual Kantorovich formulation of Unbalanced Optimal Transport

\[
\Gamma - \text{UOT}_{\text{Dirac}}(\mu_0, \mu_1) = \text{UOT}_{\text{Dirac}}^{**}(\mu_0, \mu_1) = \\
= \sup \left\{ \int \phi_0 \, d\mu_0 + \int \phi_1 \, d\mu_1 : (\phi_0, \phi_1) \in K[h] \right\}.
\]
Primal formulation

How to represent convex combinations of pair of Dirac masses?

Given $\alpha_k \geq 0$, $\sum_k \alpha_k = 1$, we may consider

$$(\mu_0, \mu_1) = \sum_k \alpha_k (r_{0,k} \delta_{x_{0,k}}, r_{1,k} \delta_{x_{1,k}})$$

$$\leadsto \Gamma\text{-UOT}_{\text{Dirac}}(\mu_0, \mu_1) \leq \sum_k \alpha_k h(x_{0,k}, r_{0,k}; x_{1,k}, r_{1,k}) = \int h \, d\alpha$$

$$\leadsto \alpha = \sum_k \alpha_k \delta(x_{0,k}, r_{0,k}; x_{1,k}, r_{1,k}) \in \mathcal{P}(X_0 \times \mathbb{R}_+ \times X_1 \times \mathbb{R}_+)$$
Primal formulation

How to represent convex combinations of pair of Dirac masses?

Given $\alpha_k \geq 0$, $\sum_k \alpha_k = 1$, we may consider

$$(\mu_0, \mu_1) = \sum_k \alpha_k (r_{0,k} \delta_{x_{0,k}}, r_{1,k} \delta_{x_{1,k}})$$

$$\Rightarrow \Gamma\text{-UOT}_{\text{Dirac}}(\mu_0, \mu_1) \leq \sum_k \alpha_k h(x_{0,k}, r_{0,k}; x_{1,k}, r_{1,k}) = \int h \, d\alpha$$

$$\Rightarrow \alpha = \sum_k \alpha_k \delta_{(x_{0,k}, r_{0,k}; x_{1,k}, r_{1,k})} \in \mathcal{P}(X_0 \times \mathbb{R}_+ \times X_1 \times \mathbb{R}_+)$$

Constraints:

$$\mu_0(A) = \sum_k \alpha_k r_{0,k} \delta_{x_{0,k}}(A) = \int_{A \times \mathbb{R}_+ \times X_1 \times \mathbb{R}_+} r_0 \, d\alpha(x_0, r_0; x_1, r_1)$$

$$= h_0 \alpha(A)$$

$$\mu_1(B) = \sum_k \alpha_k r_{1,k} \delta_{x_{1,k}}(B) = \int_{X_0 \times \mathbb{R}_+ \times B \times \mathbb{R}_+} r_1 \, d\alpha(x_0, r_0; x_1, r_1)$$

$$= h_1 \alpha(B)$$

$$\mu_0 = h_0 \alpha = \pi_{\mu_0}^{X_0}(r_0 \alpha), \quad \mu_1 = h_1 \alpha = \pi_{\mu_1}^{X_1}(r_1 \alpha) \quad 1\text{-homogeneous marginals of } \alpha$$
We introduce the set of plans with homogeneous marginals $\mu_0, \mu_1$:

$$\mathcal{H}(\mu_0, \mu_1) := \left\{ \alpha \in \mathcal{P}(\mathbb{R}_+^n \times \mathbb{R}_+^m) : \begin{align*}
    h_0 \alpha &= \pi_{\mu_0}^{X_0}(r_0 \alpha) = \mu_0, \\
    h_1 \alpha &= \pi_{\mu_1}^{X_1}(r_1 \alpha) = \mu_1
\end{align*} \right\}$$
We introduce the set of plans with homogeneous marginals $\mu_0, \mu_1$:

$$\mathcal{H}(\mu_0, \mu_1) := \left\{ \alpha \in \mathcal{P}(X_0 \times \mathbb{R}_+ \times X_1 \times \mathbb{R}_+) : h_0 \alpha = \pi_{\mu_0}^{X_0}(r_0 \alpha) = \mu_0, \ h_1 \alpha = \pi_{\mu_1}^{X_1}(r_1 \alpha) = \mu_1 \right\}$$

**Primal formulation**

$$\text{UOT}(\mu_0, \mu_1) = \min \left\{ \int h(x_0, r_0; x_1, r_1) \, d\alpha : \alpha \in \mathcal{H}(\mu_0, \mu_1) \right\}$$
We introduce the set of plans with homogeneous marginals $\mu_0, \mu_1$:

$$\mathcal{H}(\mu_0, \mu_1) := \left\{ \alpha \in \mathcal{P}(X_0 \times \mathbb{R}_+ \times X_1 \times \mathbb{R}_+) : \right.$$

$$ h_0 \alpha = \pi^X_0(r_0 \alpha) = \mu_0, \ h_1 \alpha = \pi^X_1(r_1 \alpha) = \mu_1 \left\} $$

Primal formulation

$$\text{UOT}(\mu_0, \mu_1) = \min \left\{ \int h(x_0, r_0; x_1, r_1) \, d\alpha : \alpha \in \mathcal{H}(\mu_0, \mu_1) \right\}$$

It is possible to check that $\text{UOT}$ is convex, l.s.c., and it is dominated by $\text{UOT}_{\text{Dirac}}$, so that

$$\text{UOT}(\mu_0, \mu_1) \leq \text{UOT}^{**}_{\text{Dirac}}(\mu_0, \mu_1)$$

On the other hand it is also immediate to check that

$$\text{UOT}(\mu_0, \mu_1) \geq \text{UOT}^{**}_{\text{Dirac}}(\mu_0, \mu_1),$$

Primal-dual equivalence of Unbalanced Optimal Transport

$$\text{UOT}(\mu_0, \mu_1) = \text{UOT}^{**}_{\text{Dirac}}(\mu_0, \mu_1) = \sup \left\{ \int \phi_0 \, d\mu_0 + \int \phi_1 \, d\mu_1 : (\phi_0, \phi_1) \in K[h] \right\},$$

$$K[h] = \left\{ (\phi_0, \phi_1) \in C_b(X_0) \times C_b(X_1) : r_0 \phi_0(x_0) + r_1 \phi_1(x_1) \leq h(x_0, r_0; x_1, r_1) \right\}.$$
Consider the space $\mathcal{C}[X_i] = (X_i \times \mathbb{R}_+)/\sim$ and the cost functional $h$. It induces the OT problem

$$\text{OT}_h(\alpha_0, \alpha_1) = \min \left\{ \int h \, d\alpha : \alpha \in \Gamma(\alpha_1, \alpha_2) \right\}.$$ 

We have

**Optimal transport formulation via homogeneous marginals**

$$\text{UOT}(\mu_0, \mu_1) = \min \left\{ \text{OT}_h(\alpha_0, \alpha_1) : \alpha_i \in \mathcal{P}(\mathcal{C}[X_i]), h\alpha_i = \mu_i \right\}$$

where

$$h\alpha_i = \pi^X_i(r_i \alpha_i).$$
Consider the space $\mathcal{C}[X_i] = (X_i \times \mathbb{R}_+)/\sim$ and the cost functional $h$. It induces the OT problem

$$\text{OT}_h(\alpha_0, \alpha_1) = \min \left\{ \int h \, d\alpha : \alpha \in \Gamma(\alpha_1, \alpha_2) \right\}.$$ 

We have

**Optimal transport formulation via homogeneous marginals**

$$\text{UOT}(\mu_0, \mu_1) = \min \left\{ \text{OT}_h(\alpha_0, \alpha_1) : \alpha_i \in \mathcal{P}(\mathcal{C}[X_i]), h\alpha_i = \mu_i \right\}$$

where

$$h\alpha_i = \pi_X^i (r_i \alpha_i).$$

**How to choose interesting costs $h$?** We discuss the particular case of the hellinger-Kantorovich metric, induced by the natural cone distance on $\mathcal{C}[\mathbb{R}^d]$. 

1 Unbalanced Optimal Transport: a relaxation viewpoint

2 The Hellinger-Kantorovich metric between positive measures of arbitrary mass

3 Geodesics and geodesic convexity

4 Regularity of solutions to the Conical Hopf-Lax semigroup
Let $\mu \in C^0([0, 1]; \mathcal{M}(\mathbb{R}^d))$, $(\nu, w) : \mathbb{R}^d \times (0, 1) \to \mathbb{R}^{d+1}$ be a Borel vector field satisfying
\[
\int_0^1 \int \left( |\nu_t(x)|^2 + w_t^2(x) \right) d\mu_t(x) dt < \infty.
\]

**Continuity equation with reaction** governed by the field $(\nu, w)$ if
\[
\partial_t \mu_t + \nabla \cdot (\nu_t \mu_t) = 2w_t \mu_t \quad \text{in } \mathcal{D}'(\mathbb{R}^d \times (0, 1)) \quad \text{(CER)}
\]
The dynamic perspective

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

\[
\int_0^1 \int (|\nu_t(x)|^2 + w_t^2(x)) \, d\mu_t(x) \, dt < \infty.
\]

Continuity equation with reaction governed by the field \((\nu, w)\) if

\[
\partial_t \mu_t + \nabla \cdot (\nu_t \mu_t) = 2w_t \mu_t \quad \text{in} \ \mathcal{D}'(\mathbb{R}^d \times (0, 1)) \quad \text{(CER)}
\]

The Hellinger-Kantorovich distance via dynamic interpolation

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

This approach has been independently proposed by

KONDRAFIEV, MONSAingeON, VOROTNIKOV and CHIZAT, PEYRÉ, VIALARD, SCHMITZER.
Suppose that \( \mu_i = r_i^2 \delta_{x_i} \); if we look for \( \mu_t := r^2(t) \delta_{x(t)} \)

\[
\partial_t \mu_t + \nabla \cdot (\mu_t \nu_t) = 2 w_t \mu_t, \quad \nu_t(x(t)) = \dot{x}(t), \quad w_t(x(t)) = \dot{r}(t)/r(t)
\]

We can compute

\[
HK^2(r_0^2 \delta_{x_0}, r_1^2 \delta_{x_1}) = \min \left\{ \int_0^1 \left( r^2(t)|\dot{x}(t)|^2 + |\dot{r}(t)|^2 \right) dt : \right.
\]

\[
(x, r) : [0, 1] \rightarrow \mathbb{R}^d \times \mathbb{R}^+, (x(i), r(i)) = (x_i, r_i) i = 0, 1 \right\}
\]
The distances between two Dirac masses

Suppose that \( \mu_i = r_i^2 \delta_{x_i} \); if we look for \( \mu_t := r^2(t) \delta_{x(t)} \)

\[
\partial_t \mu_t + \nabla \cdot (\mu_t \mathbf{v}_t) = 2w_t \mu_t, \quad \mathbf{v}_t(x(t)) = \dot{x}(t), \quad w_t(x(t)) = \dot{r}(t)/r(t)
\]

We can compute

\[
\mathcal{H}\mathcal{K}^2(r_0^2 \delta_{x_0}, r_1^2 \delta_{x_1}) = \min \left\{ \int_0^1 \left( r^2(t)|\dot{x}(t)|^2 + |\dot{r}(t)|^2 \right) \, dt : (x, r) : [0, 1] \to \mathbb{R}^d \times \mathbb{R}_+, (x(i), r(i)) = (x_i, r_i) \text{ } i = 0, 1 \right\}
\]

\( \mathcal{H}\mathcal{K} \) is associated to the cone distance:

\[
d_c^2((x_0, r_0), (x_1, r_1)) = r_0^2 + r_1^2 - 2r_0r_1 \cos \alpha(|x_1 - x_0|)
\]

where \( \cos \alpha(r) = \cos(r \wedge \alpha) \). \( d_c((x_0, r_0), (x_1, r_1)) \) is a length distance.
The distances between two Dirac masses

Suppose that \( \mu_i = r_i^2 \delta_{x_i} \); if we look for \( \mu_t := r^2(t) \delta_{x(t)} \)

\[
\partial_t \mu_t + \nabla \cdot (\mu_t \nu_t) = 2w_t \mu_t, \quad \nu_t(x(t)) = \dot{x}(t), \quad w_t(x(t)) = \dot{r}(t)/r(t)
\]

We can compute

\[
\mathcal{H} \mathcal{K}^2(r_0^2 \delta_{x_0}, r_1^2 \delta_{x_1}) = \min \left\{ \int_0^1 \left( r^2(t)|\dot{x}(t)|^2 + |\dot{r}(t)|^2 \right) \, dt : \right. \\
(x, r) : [0, 1] \to \mathbb{R}^d \times \mathbb{R}_+, (x(i), r(i)) = (x_i, r_i) \ i = 0, 1 \left. \right\}
\]

\( \mathcal{H} \mathcal{K} \) is associated to the cone distance:

\[
d_C^2((x_0, r_0), (x_1, r_1)) = r_0^2 + r_1^2 - 2r_0r_1 \cos(\alpha/2)(|x_1 - x_0|)
\]

where \( \cos(\alpha)(r) = \cos(r \wedge \alpha) \). \( d_C((x_0, r_0), (x_1, r_1)) \) is a length distance.

**Truncation effect:** when \( |x_0 - x_1| \geq \pi/2 \) a better competitor is provided by \( \mu_t := [(1 - t)r_0]^2 \delta_{x_0} + (tr_1)^2 \delta_{x_1} \) and we have

\[
\mathcal{H} \mathcal{K}^2(r_0^2 \delta_{x_0}, r_1^2 \delta_{x_1}) = r_0^2 + r_1^2.
\]
The distances between two Dirac masses

Suppose that $\mu_i = r_i^2 \delta_{x_i}$; if we look for $\mu_t := r^2(t)\delta_{x(t)}$

\[
\partial_t \mu_t + \nabla \cdot (\mu_t \nu_t) = 2w_t \mu_t, \quad \nu_t(x(t)) = \dot{x}(t), \quad w_t(x(t)) = \dot{r}(t)/r(t)
\]

We can compute

\[
HK^2(r_0^2 \delta_{x_0}, r_1^2 \delta_{x_1}) = \min \left\{ \int_0^1 \left( r^2(t)|\dot{x}(t)|^2 + |\dot{r}(t)|^2 \right) \, dt : 
\begin{array}{l}
(x, r) : [0, 1] \to \mathbb{R}^d \times \mathbb{R}_+, (x(i), r(i)) = (x_i, r_i) i = 0, 1
\end{array}\right\}
\]

$HK$ is associated to the cone distance:

\[
d^2_c((x_0, r_0), (x_1, r_1)) = r_0^2 + r_1^2 - 2r_0 r_1 \cos_{\pi/2}(|x_1 - x_0|)
\]

where $\cos_\alpha(r) = \cos(r \wedge \alpha)$.

$d_c((x_0, r_0), (x_1, r_1))$ is a length distance.

**Truncation effect:** when $|x_0 - x_1| \geq \pi/2$ a better competitor is provided by $\mu_t := [(1 - t)r_0]^2\delta_{x_0} + (tr_1)^2\delta_{x_1}$ and we have

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

\[
HK^2(r_0^2 \delta_{x_0}, r_1^2 \delta_{x_1}) = r_0^2 + r_1^2 - 2r_0 r_1 \cos_{\pi/2}(|x_1 - x_0|)
\]
The Cone space

**Cone metric:** \[ d^2_C((x_0, r_0), (x_1, r_1)) = r_0^2 + r_1^2 - 2r_0r_1 \cos(\pi| x_1 - x_0 |) \]

**Cone space:** identify all the points \((x, 0)\) with the vertex \(o\).

\[ C := (\mathbb{R}^d \times [0, \infty)) / \sim, \quad (x', r') \sim (x'', r'') \iff \begin{cases} x' = x'', \ r' = r'' \neq 0, \\ r' = r'' = 0 \end{cases} \]
The Cone space

**Cone metric:** \[ d_C^2((x_0, r_0), (x_1, r_1)) = r_0^2 + r_1^2 - 2r_0r_1 \cos\pi(|x_1 - x_0|) \]

**Cone space:** identify all the points \((x, 0)\) with the vertex \(o\).

\[ \mathcal{C} := (\mathbb{R}^d \times [0, \infty)) / \sim, \quad (x', r') \sim (x'', r'') \iff \begin{cases} x' = x'', & r' = r'' \neq 0, \\ r' = r'' = 0 \end{cases} \]

\(\mathcal{C} \setminus \{o\}\) can be considered as a **smooth Riemannian manifold** with metric

\[ g(dx, dr) = r^2|dx|^2 + |dr|^2 \]
HK^2 is a convex and subadditive functional

We introduce a function \( h : (\mathbb{R}^d \times \mathbb{R}_+)^2 \to [0, +\infty) \) which characterizes the cost of connecting two Dirac measures with possibly different mass:

\[
h(x_0, r_0; x_1, r_1) := HK^2(r_0 \delta_{x_0}, r_1 \delta_{x_1}) = d_c^2((x_0, \sqrt{r_0}), (x_1, \sqrt{r_1})) \\
= r_0 + r_1 - 2\sqrt{r_0 r_1} \cos \pi/2(|x_1 - x_0|) \quad x_i \in X_i, \ r_i \geq 0
\]
$\mathcal{H}^2$ is a **convex and subadditive functional**

We introduce a function $h : (\mathbb{R}^d \times \mathbb{R}_+)^2 \rightarrow [0, +\infty)$ which characterizes the cost of connecting two Dirac measures with possibly different mass:

$$h(x_0, r_0; x_1, r_1) := \mathcal{H}^2(r_0 \delta_{x_0}, r_1 \delta_{x_1}) = d_c^2((x_0, \sqrt{r_0}), (x_1, \sqrt{r_1}))$$

$$= r_0 + r_1 - 2\sqrt{r_0 r_1} \cos\frac{\pi}{2}(|x_1 - x_0|) \quad x_i \in X_i, \ r_i \geqslant 0$$

$$(r_0, r_1) \mapsto h(x_0, r_0; x_1, r_1)$$ is **positively 1-homogeneous and convex**

thanks to the truncation $(- \cos\frac{\pi}{2} \leqslant 0)$. Define $\text{UOT}_{\text{Dirac}}(\mu_0, \mu_1) := \mathcal{H}^2(\mu_0, \mu_1)$ if $\mu_i = r_i \delta_{x_i}$, $+\infty$ otherwise.
HK^2 is a convex and subadditive functional

We introduce a function \( h : (\mathbb{R}^d \times \mathbb{R}_+)^2 \to [0, +\infty) \) which characterizes the cost of connecting two Dirac measures with possibly different mass:

\[
\begin{align*}
  h(x_0, r_0; x_1, r_1) := HK^2(r_0 \delta_{x_0}, r_1 \delta_{x_1}) &= d^2_Q((x_0, \sqrt{r_0}), (x_1, \sqrt{r_1})) \\
  &= r_0 + r_1 - 2 \sqrt{r_0 r_1} \cos \frac{\pi}{2} |x_1 - x_0| \\
  x_i &\in X_i, \quad r_i \geq 0
\end{align*}
\]

\((r_0, r_1) \mapsto h(x_0, r_0; x_1, r_1)\) is positively 1-homogeneous and convex thanks to the truncation \((- \cos \frac{\pi}{2} \leq 0\). Define \( \hat{UOT}_{\text{Dirac}}(\mu_0, \mu_1) := HK^2(\mu_0, \mu_1) \) if \( \mu_i = r_i \delta_{x_i}, +\infty \) otherwise.

**Theorem**

HK^2 is the \( \Gamma \)-relaxation of \( \hat{UOT}_{\text{Dirac}} \): the largest convex and lower semicontinuous functional defined in \( \mathcal{M}(\mathbb{R}^d) \times \mathcal{M}(\mathbb{R}^d) \to [0, +\infty] \) dominated by \( \hat{UOT}_{\text{Dirac}} \):

\[
\hat{UOT} \text{ convex, l.s.c., } \hat{UOT} \leq UOT_{\text{Dirac}} \Rightarrow \hat{UOT} \leq HK^2.
\]
\[ H(\mu_0, \mu_1) := \left\{ \alpha \in \mathcal{P}(X_0 \times \mathbb{R}_+ \times X_1 \times \mathbb{R}_+) : h^0 \alpha = \pi^{X_0}_\mu (r_0^2 \alpha) = \mu_0, \ h^1 \alpha = \pi^{X_1}_\mu (r_1^2 \alpha) = \mu_1 \right\} \]
$\mathcal{H}(\mu_0, \mu_1) := \left\{ \alpha \in \mathcal{P}(X_0 \times \mathbb{R}_+ \times X_1 \times \mathbb{R}_+) : \right.$

$h^0 \alpha = \pi_{X_0}^{r_0}(r_0^2 \alpha) = \mu_0, \ h^1 \alpha = \pi_{X_1}^{r_1}(r_1^2 \alpha) = \mu_1 \right\}$

**Primal formulation**

$H^K(\mu_0, \mu_1) = \min \left\{ \int h(x_0, r_0^2; x_1, r_1^2) \, d\alpha : \alpha \in \mathcal{H}(\mu_0, \mu_1) \right\}$

$= \min \left\{ \int d^2_c((x_0, r_0), (x_1, r_1)) \, d\alpha : \alpha \in \mathcal{H}(\mu_0, \mu_1) \right\}$
Transport-growth pairs

We can represent $\alpha \in H(\mu_0, \mu_1)$ as $\alpha = ((T_0, q_0), (T_1, q_1)) \# \lambda$ where $\lambda \in \mathcal{M}(Y)$, $Y$ is some Polish space, and $(T_i, q_i) : Y \rightarrow \mathbb{R}^d \times \mathbb{R}_+$ with $q_i \in L^2(\lambda)$. 

Problem (Monge formulation of $H K$)

Given $\mu_0, \mu_1 \in \mathcal{M}(\mathbb{R}^d)$ find an optimal transport-growth pair $(T, q) : \mathbb{R}^d \rightarrow \mathbb{R}^d \times \mathbb{R}_+$ minimizing the cost

$$M(T, q; \mu_0) := \int (1 + q_2(x) - 2q(x) \cos \pi/2 |T(x) - x|) \, d\mu_0(x) \quad (1)$$

among all the transport-growth maps satisfying $(T, q) \# \mu_0 = \mu_1$. 


Transport-growth pairs

We can represent $\alpha \in \mathcal{H}(\mu_0, \mu_1)$ as $\alpha = ((T_0, q_0), (T_1, q_1)) \# \lambda$ where $\lambda \in \mathcal{M}(Y)$, $Y$ is some Polish space, and $(T_i, q_i) : Y \to \mathbb{R}^d \times \mathbb{R}_+$ with $q_i \in L^2(\lambda)$.

We say that $(T_i, q_i)$ is a transport-growth pair. $(T, q)$ acts on $\lambda$ according to this rule:

$$(T, q)_* \lambda := T_#(q^2 \lambda) = \mathcal{H}((T, q)_# \lambda),$$

$$\mathcal{H}^2(\mu_0, \mu_1) = \min \left\{ \int_{Y \times Y} \left( q_0^2 + q_1^2 - 2q_0 q_1 \cos \frac{\pi}{2} |T_0 - T_1| \right) \, d\lambda \right\} \lambda \in \mathcal{M}(Y),$$

$Y$ Polish, $(T_i, q_i) : Y \to \mathbb{R}^d \times \mathbb{R}_+$, $\mu_i := (T_i, q_i)_* \lambda$;

moreover, it is not restrictive to choose $Y = C[\mathbb{R}^d] \times C[\mathbb{R}^d]$. 

We can represent \( \alpha \in \mathcal{H}(\mu_0, \mu_1) \) as \( \alpha = ((T_0, q_0), (T_1, q_1)) \# \lambda \) where \( \lambda \in \mathcal{M}(Y) \), \( Y \) is some Polish space, and (\( T_i, q_i \)) : \( Y \rightarrow \mathbb{R}^d \times \mathbb{R}^+ \) with \( q_i \in L^2(\lambda) \).

We say that (\( T_i, q_i \)) is a **transport-growth pair**. (\( T, q \)) acts on \( \lambda \) according to this rule:

\[
(T, q)_* \lambda := T \# (q^2 \lambda) = \mathfrak{h}(T, q)_* \lambda,
\]

\[
\mathcal{H}^2(\mu_0, \mu_1) = \min \left\{ \int_{Y \times Y} \left( q_0^2 + q_1^2 - 2q_0q_1 \cos \frac{\pi}{2} |T_0 - T_1| \right) d\lambda \mid \lambda \in \mathcal{M}(Y), Y \text{ Polish}, (T_i, q_i) : Y \rightarrow \mathbb{R}^d \times \mathbb{R}^+, \mu_i := (T_i, q_i)_* \lambda \right\};
\]

moreover, it is not restrictive to choose \( Y = \mathcal{C}[\mathbb{R}^d] \times \mathcal{C}[\mathbb{R}^d] \).

**Problem (Monge formulation of \( \mathcal{H}^2 \))**

Given \( \mu_0, \mu_1 \in \mathcal{M}(\mathbb{R}^d) \) find an optimal transport-growth pair (\( T, q : \mathbb{R}^d \rightarrow \mathbb{R}^d \times \mathbb{R}^+ \)) minimizing the cost

\[
\mathcal{M}(T, q; \mu_0) := \int \left( 1 + q^2(x) - 2q(x) \cos \frac{\pi}{2} |T(x) - x| \right) d\mu_0(x)
\]

among all the transport-growth maps satisfying (\( T, q \)) \( \# \mu_0 = \mu_1 \).
If
\[ \partial_t \xi_t + \frac{1}{2} |D\xi_t|^2 + 2\xi_t^2(x) \leq 0 \]  
(CHJ)

and
\[ \partial_t \mu_t + \nabla \cdot (\nu_t \mu_t) = 2w_t \mu_t \]

then
\[ \int \xi_1 \, d\mu_1 - \int \xi_0 \, d\mu_0 \leq \frac{1}{2} \int_0^1 \int (|v_t|^2 + w_t^2) \, d\mu_t \, dt. \]
If
\[ \partial_t \xi_t + \frac{1}{2} |D \xi_t|^2 + 2 \xi_t^2(x) \leq 0 \]  
(CHJ)
and
\[ \partial_t \mu_t + \nabla \cdot (v_t \mu_t) = 2w_t \mu_t \]
then
\[ \int \xi_1 \, d\mu_1 - \int \xi_0 \, d\mu_0 \leq \frac{1}{2} \int_0^1 \left( |v_t|^2 + w_t^2 \right) \, d\mu_t \, dt. \]

\[ \mathbb{H} \text{ in duality with subsolutions to the conical Hamilton-Jacobi equations} \]

\[ \frac{1}{2} \mathbb{H}^2(\mu_0, \mu_1) = \sup \left\{ \int \xi_1 \, d\mu_1 - \int \xi_0 \, d\mu_0 : \xi \in C^1([0, 1]; \text{Lip}_b(\mathbb{R}^d)) \right\} \]
\[ \partial_t \xi_t + \frac{1}{2} |D \xi_t|^2 + 2 \xi_t^2 \leq 0 \].
Given $\xi_0 \in \text{Lip}_b(\mathbb{R}^d)$ with $\xi_0 > -1/2$, the viscosity solution (or the maximal subsolution) of the conical Hamilton Jacobi equation

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

(CHJ)

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

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

(CHL)
Conical Hopf-Lax representation formula

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

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

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

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

Conical Hopf-Lax representation for $H_K$

$$\frac{1}{2} H_K^2(\mu_0, \mu_1) = \sup \left\{ \int \xi_1 \, d\mu_1 - \int \xi_0 \, d\mu_0 : \xi_1 = \mathcal{P}_1 \xi_0 \right\}$$
Formally, if $\xi$ is a solution of

$$\partial_t \xi_t + \frac{1}{2} |D\xi_t|^2 + 2\xi_t^2 \leq 0 \quad \text{(CHJ)}$$

then \[\zeta_t(x, r) := \xi_t(x)r^2\] is a solution of

$$\partial_t \zeta_t + \frac{1}{2} |D\zeta_t|^2 \leq 0 \quad \text{(HJ)}$$

since

$$\frac{1}{2} |D\zeta_t|^2 = \frac{1}{2}g^*(D_x \zeta, \partial_r \zeta) = \frac{1}{2} \left( \frac{1}{r^2} |D_x \zeta|^2 + (\partial_r \zeta)^2 \right) = \left( \frac{1}{2} |D\xi_t|^2 + 2\xi_t^2 \right) r^2$$
Conical lift of the Hopf-Lax formula

Formally, if $\xi$ is a solution of

$$\partial_t \xi_t + \frac{1}{2} |D\xi_t|^2 + 2\xi_t^2 \leq 0$$  \hspace{1cm} (CHJ)$$

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

$$\partial_t \zeta_t + \frac{1}{2} |D\zeta_t|^2 \leq 0$$  \hspace{1cm} (HJ)$$

since

$$\frac{1}{2} |D\zeta_t|^2 = \frac{1}{2} g^*(D_x \zeta, \partial_r \zeta) = \frac{1}{2} \left( \frac{1}{r^2} |D_x \zeta|^2 + (\partial_r \zeta)^2 \right) = \left( \frac{1}{2} |D\xi_t|^2 + 2\xi_t^2 \right) r^2$$

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

$$\mathcal{D}_t^e \zeta(x, r) = \min_{y, s} \zeta(y, s) + \frac{1}{2t} d_e^2((x, r), (y, s)) = \min_{y, s} \xi(y)s^2 + \frac{1}{2t} \left( r^2 + s^2 - 2rs \cos(|x - y|) \pi \right)$$

yields

$$\mathcal{D}_t^e \zeta(x, r) = \xi_t(x)r^2, \quad \xi_t = \mathcal{P}_t \xi.$$
Dual formulation (II)

Change of variable: \( \varphi_1 := -\frac{1}{2} \log(1 - 2\xi_1), \varphi_0 := \frac{1}{2} \log(1 + 2\xi_0) \)

\[
2\xi_1(y) \leq 1 - \frac{\cos^2\frac{\pi}{2}(|y - x|)}{1 + 2\xi_0(x)} \iff \varphi_1(y) - \varphi_0(x) \leq \ell(y - x),
\]

\[
\ell(r) = -\frac{1}{2} \log\left(\cos^2\frac{\pi}{2} |r|\right) = \frac{1}{2} \log\left(1 + \tan^2\frac{\pi}{2} |r|\right), \quad D\ell(r) = \tan(r)
\]
Dual formulation (II)

Change of variable: \( \varphi_1 := -\frac{1}{2} \log(1 - 2\xi_1), \varphi_0 := \frac{1}{2} \log(1 + 2\xi_0) \)

\[
2\xi_1(y) \leq 1 - \frac{\cos^2(\pi/2 \, |y - x|)}{1 + 2\xi_0(x)} \quad \Leftrightarrow \quad \varphi_1(y) - \varphi_0(x) \leq \ell(y - x),
\]

\[
\ell(r) = -\frac{1}{2} \log \left( \cos^2(\pi/2 \, |r|) \right) = \frac{1}{2} \log \left( 1 + \tan^2(\pi/2 \, |r|) \right), \quad D\ell(r) = \tan(r)
\]

Dual Kantorovich formulation

\[
\frac{1}{2} \mathcal{H}^2(\mu_0, \mu_1) = \sup \left\{ \int \frac{1}{2}(1 - e^{-2\varphi_1}) \, d\mu_1 - \int \frac{1}{2}(e^{2\varphi_0} - 1) \, d\mu_0 : \varphi_1(y) - \varphi_0(x) \leq \ell(y - x) \right\}
\]
The **Legendre conjugate** of \( G(\varphi) := \frac{1}{2} \left( e^{2\varphi} - 1 \right) \) is

\[
G^*(s) = \frac{1}{2} \mathcal{L}(s), \quad \mathcal{L}(s) := s \log s - (s - 1)
\]
The **Legendre conjugate** of $G(\phi) := \frac{1}{2} \left( e^{2\phi} - 1 \right)$ is

$$G^*(s) = \frac{1}{2} \mathcal{LE}(s), \quad \mathcal{LE}(s) := s \log s - (s - 1)$$

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

### Logarithmic Entropy-Transport (LET) formulation

$$\text{LET}(\mu_0, \mu_1) = \min_{\gamma \in \mathcal{M}(\mathbb{R}^d \times \mathbb{R}^d)} \left( \mathcal{E}(\gamma_0|\mu_0) + \mathcal{E}(\gamma_1|\mu_1) + 2 \int \ell(y - x) \, d\gamma(x, y) \right)$$

where $\ell(r) = \frac{1}{2} \log(1 + \tan^2_{\pi/2}(|r|))$. 
The **Legendre conjugate** of $G(\varphi) := \frac{1}{2} \left( e^{2\varphi} - 1 \right)$ is

$$G^*(s) = \frac{1}{2} \mathcal{LE}(s), \quad \mathcal{LE}(s) := s \log s - (s - 1)$$

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

**Logarithmic Entropy-Transport (LET) formulation**

$$\text{LET}(\mu_0, \mu_1) = \min_{\gamma \in \mathcal{M}(\mathbb{R}^d \times \mathbb{R}^d)} \left( \mathcal{E}(\gamma_0|\mu_0) + \mathcal{E}(\gamma_1|\mu_1) + 2 \int \ell(y - x) d\gamma(x, y) \right)$$

where $\ell(r) = \frac{1}{2} \log(1 + \tan^2 \frac{\pi}{2}(|r|))$.  

$$\mathcal{HK}^2(\mu_0, \mu_1) = \text{LET}(\mu_0, \mu_1)$$
Four equivalent formulations for $HK$

\[
H_K(\mu_0, \mu_1) = \min \left\{ \int_0^1 \left( |v_t|^2 + |w_t|^2 \right) d\mu_t dt : \mu \in C([0,1]; M(\mathbb{R}^d)) \right\}
\]

\[
\partial_t \mu_t + \nabla \cdot (v_t \mu_t) = 2w_t \mu_t,
\]

\[
\mu_t = i = \mu_i
\]

\[(CER)\]

\[
\sup \left\{ \int_\xi_1 d\mu_1 - \int_\xi_0 d\mu_0 : \xi \in C^1([0,1]; \text{Lip}_b(\mathbb{R}^d)) \right\}
\]

\[
\partial_t \xi_t + \frac{1}{2} |D\xi_t|^2 + 2\xi_t^2 \leq 0
\]

\[(CHJ)\]

\[
\sup \left\{ \int_\xi_1 d\mu_1 - \int_\xi_0 d\mu_0 : \xi_1 = P_1 \xi_0 \right\}
\]

\[(CHL)\]

\[
\min \gamma E(\gamma|\mu_0) + E(\gamma|\mu_1) + 2\int \ell(x,y) d\gamma(x,y).
\]

\[(LET)\]
Four equivalent formulations for $HK^2$:

$$HK^2(\mu_0, \mu_1) = \min \left\{ \int_0^1 \left( |v_t|^2 + |w_t|^2 \right) d\mu_t \, dt : \mu \in C([0, 1]; \mathcal{M}(\mathbb{R}^d)) \right\}$$

\[ \begin{align*}
\partial_t \mu_t + \nabla \cdot (v_t \mu_t) &= 2w_t \mu_t, \quad \mu_{t=i} = \mu_i \\
= 2 \sup \left\{ \int \xi_1 \, d\mu_1 - \int \xi_0 \, d\mu_0 : \xi \in C^1([0, 1]; \text{Lip}_b(\mathbb{R}^d)) \right\} \\
&= 2 \sup \left\{ \int \xi_1 \, d\mu_1 - \int \xi_0 \, d\mu_0 : \xi_1 = \mathcal{P}_1 \xi_0 \right\} \\
&= \min_{\gamma} \mathcal{E}(\gamma_0 | \mu_0) + \mathcal{E}(\gamma_1 | \mu_1) + 2 \int \ell(x, y) \, d\gamma(x, y). \tag{LET}
\end{align*} \]
1. Unbalanced Optimal Transport: a relaxation viewpoint

2. The Hellinger-Kantorovich metric between positive measures of arbitrary mass

3. Geodesics and geodesic convexity

4. Regularity of solutions to the Conical Hopf-Lax semigroup
Important properties

- \((\mathcal{M}(\mathbb{R}^d), \mathcal{H})\) is a complete and separable metric space if \(\mathbb{X}\) is complete and separable; the induced topology coincides with the topology of weak convergence (no bounds on moments are required).

Problem
Characterize geodesics and study the convexity properties of integral functionals.

In particular, we want to prove that power-like entropies
\[
E_\alpha(\mu) := \int c^\alpha \, d\mu,
\]
\(\mu = cL^d\) are geodesically convex if \(\alpha \geq 1\) (reinforced McCann condition).
Important properties

• $(\mathcal{M}(\mathbb{R}^d), HK)$ 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).

• $(\mathcal{M}(\mathbb{R}^d), HK)$ is **geodesic**
Important properties

- \((\mathcal{M}(\mathbb{R}^d), HK)\) 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).

- \((\mathcal{M}(\mathbb{R}^d), HK)\) is geodesic

- If \(\mu_0 \ll \mathcal{L}^d\) then there exists a unique geodesic connecting \(\mu_0\) to \(\mu_1\) and a unique optimal plan \(\gamma\) minimizing \(\text{LET}(\mu_0, \mu_1)\).
Important properties

- $(\mathcal{M}(\mathbb{R}^d), HK)$ 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).
- $(\mathcal{M}(\mathbb{R}^d), HK)$ is **geodesic**
- If $\mu_0 \ll \mathcal{L}^d$ then there exists a unique geodesic connecting $\mu_0$ to $\mu_1$ and a **unique optimal plan** $\gamma$ minimizing $\text{LET}(\mu_0, \mu_1)$.

**Problem**

*Characterize geodesics and study the convexity properties of integral functionals.*

In particular, we want to prove that power-like entropies

$$E_\alpha(\mu) := \int c^\alpha \, dx, \quad \mu = c\mathcal{L}^d$$

are **geodesically convex** if $\alpha \geq 1$ (**reinforced McCann condition**).
The $\frac{\pi}{2}$ threshold and $HK$ geodesics between Dirac masses

$\mu_0 = r_0^2 \delta_{x_0}$, $\mu_1 = r_1^2 \delta_{x_1}$, $|x_1 - x_0| \in [0, \pi]$, $\mu_t := r_t \delta_{x_t}$ geodesic.

Initial velocities $(u, v) \in \mathbb{R} \times \mathbb{R}^d$

\[
u := \frac{r_1}{r_0} \sin(x_1 - x_0), \quad \sin(w) := \sin(|w|) \frac{w}{|w|}
\]

curve:

\[
r_t := r_0 \left( (1+tu)^2 + t^2|v|^2 \right)^{1/2}, \quad x_t := x_0 + \arctan\left( \frac{tv}{1+tu} \right)
\]
Theorem

For every $\mu_0, \mu_1 \in \mathcal{M}(\mathbb{R}^d)$ there exists a pair of optimal potentials $(\varphi_0, \varphi_1)$ such that

$$\varphi_1(y) - \varphi_0(x) \leq 2\ell(y - x) \text{ and }$$

$$H K^2(\mu_0, \mu_1) = \int (1 - e^{-2\varphi_1}) \, d\mu_1 - \int (e^{2\varphi_0} - 1) \, d\mu_0.$$
Regularity of optimal potentials for the LET formulation

**Theorem**

For every $\mu_0, \mu_1 \in M(\mathbb{R}^d)$ there exists a pair of optimal potentials $(\varphi_0, \varphi_1)$ such that $\varphi_1(y) - \varphi_0(x) \leq 2 \ell(y - x)$ and

$$HK^2(\mu_0, \mu_1) = \int (1 - e^{-2\varphi_1}) \, d\mu_1 - \int (e^{2\varphi_0} - 1) \, d\mu_0.$$ 

The optimal potential $\varphi_0$ is locally semiconcave outside a closed $(d - 1)$-rectifiable set.
Regularity of optimal potentials for the LET formulation

**Theorem**

For every \( \mu_0, \mu_1 \in \mathcal{M}(\mathbb{R}^d) \) there exists a pair of optimal potentials \((\varphi_0, \varphi_1)\) such that \(\varphi_1(y) - \varphi_0(x) \leq 2\ell(y - x)\) and

\[
HK^2(\mu_0, \mu_1) = \int (1 - e^{-2\varphi_1}) \, d\mu_1 - \int (e^{2\varphi_0} - 1) \, d\mu_0.
\]

The Optimal potential \(\varphi_0\) is locally semiconcave outside a closed \((d - 1)\)-rectifiable set.

When \(\mu_0 \ll \mathcal{L}^d\) and \(\mu_1\left\{y \in \mathbb{R}^d : d(y, \text{supp}(\mu_0)) \geq \pi/2\right\} = 0\), then Monge formulation has a unique solution \((T, q)\) such that \((T, q)_*\mu_0 = \mu_1\) and

\[
\tan(T(x) - x) = \nabla \varphi_0(x), \quad q^2(x) = (e^{2\varphi_0(x)})^2 + \frac{1}{4} |\nabla e^{2\varphi_0(x)}|^2
\]
Regularity of optimal potentials for the LET formulation

**Theorem**

For every $\mu_0, \mu_1 \in \mathcal{M}(\mathbb{R}^d)$ there exists a pair of optimal potentials $(\varphi_0, \varphi_1)$ such that

$$\varphi_1(y) - \varphi_0(x) \leq 2\ell(y - x)$$

and

$$HK^2(\mu_0, \mu_1) = \int (1 - e^{-2\varphi_1}) \, d\mu_1 - \int (e^{2\varphi_0} - 1) \, d\mu_0.$$

The optimal potential $\varphi_0$ is locally semiconcave outside a closed $(d - 1)$-rectifiable set.

When $\mu_0 \ll \mathcal{L}^d$ and $\mu_1 \left\{ y \in \mathbb{R}^d : d(y, \text{supp}(\mu_0)) \geq \pi/2 \right\} = 0$, then Monge formulation has a unique solution $(T, q)$ such that $\left( T, q \right)_* \mu_0 = \mu_1$ and

$$\tan(T(x) - x) = \nabla \varphi_0(x), \quad q^2(x) = \left( e^{2\varphi_0(x)} \right)^2 + \frac{1}{4} \left| \nabla e^{2\varphi_0}(x) \right|^2.$$

After the transformation $\xi_0 := \frac{1}{2} \left( e^{2\varphi_0} - 1 \right)$ we can identify

$$T(x) = x + \arctan \left( \frac{\nabla \xi_0}{1 + 2\xi_0} \right), \quad q^2 = (1 + 2\xi_0)^2 + \left| \nabla \xi_0 \right|^2.$$
Regularity of optimal potentials for the LET formulation

Theorem

For every $\mu_0, \mu_1 \in \mathcal{M}(\mathbb{R}^d)$ there exists a pair of optimal potentials $(\varphi_0, \varphi_1)$ such that

$$\varphi_1(y) - \varphi_0(x) \leq 2\ell(y - x)$$

and

$$\mathcal{H}_K^2(\mu_0, \mu_1) = \int (1 - e^{-2\varphi_1}) \, d\mu_1 - \int (e^{2\varphi_0} - 1) \, d\mu_0.$$

The optimal potential $\varphi_0$ is locally semiconcave outside a closed $(d - 1)$-rectifiable set.

When $\mu_0 \ll \mathcal{L}^d$ and $\mu_1 \{ y \in \mathbb{R}^d : d(y, \text{supp}(\mu_0)) \geq \pi/2 \} = 0$, then Monge formulation has a unique solution $(T, q)$ such that $(T, q)_* \mu_0 = \mu_1$ and

$$\tan(T(x) - x) = \nabla \varphi_0(x), \quad q^2(x) = (e^{2\varphi_0(x)})^2 + \frac{1}{4} |\nabla e^{2\varphi_0(x)}|^2$$

After the transformation $\xi_0 := \frac{1}{2} (e^{2\varphi_0} - 1)$ we can identify

$$T(x) = x + \arctan\left( \frac{\nabla \xi_0}{1 + 2\xi_0} \right), \quad q^2 = (1 + 2\xi_0)^2 + |\nabla \xi_0|^2$$

$$\mathcal{H}_K^2(\mu_0, \mu_1) = \mathcal{M}(T, q; \mu_0) = \int_{\mathbb{R}^d} \left( 4\xi_0^2 + |\nabla \xi_0|^2 \right) \, d\mu_0.$$

Tangent space: $\text{Tan}_{\mu_0} \mathcal{M}(\mathbb{R}^d) = H^{1,2}(\mathbb{R}^d, \mu_0)$. 
Recalling

\[ T(x) = x + \arctan\left(\frac{\nabla \xi_0}{1 + 2 \xi_0}\right), \quad q^2 = (1 + 2 \xi_0)^2 + |\nabla \xi_0|^2 \]

the geodesic interpolations can be obtained by rescaling \( \xi_0 \mapsto t \xi_0, \ t \in [0, 1] \):

\[
\begin{align*}
T_{0 \to t}(x) &:= x + \arctan\left(\frac{t \nabla \xi_0}{1 + 2 t \xi_0(x)}\right), \\
q^2_{0 \to t}(x) &:= (1 + 2 t \xi_0(x))^2 + t^2 |\nabla \xi_0(x)|^2 
\end{align*}
\]

They provide an explicit characterization of the unique HK geodesic connecting \( \mu_0 \) to \( \mu_1 \):

\[
\mu_t = (T_{0 \to t}, q_{0 \to t}^{-1})_* \mu_0, \quad \mu_t = c_t \mathcal{L}^d, \quad c_t(T_{0 \to t}(x)) = c_0(x) \frac{q^2_{0 \to t}(x)}{\det D T_{0 \to t}(x)}
\]
Recalling

\[ T(x) = x + \arctan \left( \frac{\nabla \xi_0}{1+2\xi_0} \right), \quad q^2 = (1+2\xi_0)^2 + |\nabla \xi_0|^2 \]

the geodesic interpolations can be obtained by rescaling \( \xi_0 \sim t\xi_0, \ t \in [0, 1] \):

\[
T_{0 \to t}(x) := x + \arctan \left( \frac{t \nabla \xi_0}{1 + 2t \xi_0(x)} \right), \quad q_{0 \to t}^2(x) := (1 + 2t \xi_0(x))^2 + t^2 |\nabla \xi_0(x)|^2
\]

They provide an explicit characterization of the unique HK geodesic connecting \( \mu_0 \) to \( \mu_1 \):

\[
\mu_t = (T_{0 \to t}, q_{0 \to t})_* \mu_0, \quad \mu_t = c_t \mathcal{L}^d, \quad c_t (T_{0 \to t}(x)) = c_0(x) \frac{q_{0 \to t}^2(x)}{\det DT_{0 \to t}(x)}
\]

**Simplifying assumption:** \( \mu_0, \mu_1 \) have compact support, \( \text{supp}(\mu_1) \subset B_{\pi/2}(\text{supp}(\mu_0)), \text{supp}(\mu_0) \subset B_{\pi/2}(\text{supp}(\mu_1)) \).

**Optimal potentials** \( \phi_0 \) and \( \xi_0 \) are semiconvex, \( \phi_1 \) and \( \xi_1 \) are semiconcave, all the functions are globally Lipschitz and for suitable constants \( a, b \in \mathbb{R} \)

\[-\frac{1}{2} < -a \leq \xi_0(x) \leq b, \quad -b \leq \xi_1(y) \leq a < \frac{1}{2}.\]
Dynamic optimality conditions for geodesics

Theorem (Formal)

A continuous curve $(\mu_t)_{t \in [0,1]}$ is a geodesic if and only if there exists a curve $(\xi_t)_{t \in [0,1]}$ such that

$$\begin{align*}
\partial_t \mu_t + \nabla \cdot (\mu_t \nu_t) &= 2w_t \mu_t \\
\partial_t \xi_t + \frac{1}{2} |\nabla \xi_t|^2 + 2\xi_t^2 &\leq 0 \\
\partial_t \xi_t + \frac{1}{2} |\nabla \xi_t|^2 + 2\xi_t^2 &= 0 \quad \text{on the support of } \mu,
\end{align*}$$

where 

- $\nu_t = \nabla \xi_t$
- $w_t = 2\xi_t$
Dynamic optimality conditions for geodesics

**Theorem (Formal)**

A continuous curve \((\mu_t)_{t\in[0,1]}\) is a geodesic if and only if there exists a curve \((\xi_t)_{t\in[0,1]}\) such that

\[
\begin{align*}
\partial_t \mu_t + \nabla \cdot (\mu_t v_t) &= 2w_t \mu_t \\
\partial_t \xi_t + \frac{1}{2} |\nabla \xi_t|^2 + 2\xi_t^2 &\leq 0 \\
\partial_t \xi_t + \frac{1}{2} |\nabla \xi_t|^2 + 2\xi_t^2 &= 0 \quad \text{on the support of } \mu, \\
v_t &= \nabla \xi_t \\
w_t &= 2\xi_t
\end{align*}
\]

**Characteristic flow:** fix \(s \in (0, 1)\) \(T(t, \cdot) := T_{s \to t} (\cdot), q(t, \cdot) := q_{s \to t} (\cdot),\)

\[
\begin{align*}
\dot{T}(t, x) &= \nabla \xi_t (T(t, x)) \\
\dot{q}(t, x) &= 4\xi_t (T(t, x))q(t, x) \\
T(s, x) &= x, \\
q(s, x) &= 1.
\end{align*}
\]
Formal computations

\[ \partial_t \xi_t + \frac{1}{2} |\nabla \xi_t|^2 + 2 \xi_t^2 = 0. \]

**Characteristic flow:**  \( T(t, \cdot) := T_{s \to t}(\cdot), \ q(t, \cdot) := q_{s \to t}(\cdot), \)

\[
\begin{align*}
\dot{T}(t, x) &= \nabla \xi_t(T(t, x)) \\
\dot{q}(t, x) &= 4 \xi_t(T(t, x))q(t, x)
\end{align*}
\]

\( B(t, \cdot) := DT(t, \cdot), \ \delta(t, \cdot) := \det B(t, \cdot) \)
Formal computations

\[ \partial_t \xi_t + \frac{1}{2} |\nabla \xi_t|^2 + 2 \xi_t^2 = 0. \]

**Characteristic flow:**

\[ T(t, \cdot) := T_{s \to t}(\cdot), \quad q(t, \cdot) := q_{s \to t}(\cdot), \]

\[
\begin{align*}
\dot{T}(t, x) &= \nabla \xi_t(T(t, x)) \\
\dot{q}(t, x) &= 4\xi_t(T(t, x))q(t, x)
\end{align*}
\]

\[ B(t, \cdot) := DT(t, \cdot), \quad \delta(t, \cdot) := \det B(t, \cdot) \]

\[ \ddot{T}(t) = \partial_t \nabla \xi_t(T(t)) + D^2 \xi_t \nabla \xi_t(T(t)), \]

\[ \partial_t \nabla \xi_t = -D^2 \xi_t \nabla \xi_t + 4 \xi_t \nabla \xi_t \]
Formal computations

$$\partial_t \xi_t + \frac{1}{2} |\nabla \xi_t|^2 + 2 \xi_t^2 = 0.$$  

**Characteristic flow:** \( T(t, \cdot) := T_{s \to t}(\cdot), \ q(t, \cdot) := q_{s \to t}(\cdot), \)

\[
\begin{cases}
    \dot{\xi}(t, x) = \nabla \xi_t(T(t, x)) \\
    \dot{q}(t, x) = 4 \xi_t(T(t, x))q(t, x)
\end{cases}
\]

\( B(t, \cdot) := DT(t, \cdot), \ \delta(t, \cdot) := \text{det } B(t, \cdot) \)

\[
\ddot{T}(t) = \partial_t \nabla \xi_t(T(t)) + D^2 \xi_t \nabla \xi_t(T(t)), \\
\dot{q}(t, x) = |\nabla \xi_t(T(t))|^2 q(t)
\]

**Second order relations**

\[
\ddot{T}(t) = 4 \xi_t \nabla \xi_t(T(t)) \\
\ddot{q}(t, x) = |\nabla \xi_t(T(t))|^2 q(t)
\]

\[
\dot{B}(t) = -4 \left( \nabla \xi_t \otimes \nabla \xi_t + \xi_t D^2 \xi_t \right) \circ T(t) \cdot B(t) \\
\delta(t) = \left( (\Delta \xi_t)^2 - |D^2 \xi_t|^2 - 4|\nabla \xi_t|^2 - 4 \xi_t \Delta \xi_t \right) \circ T(t) \cdot \delta(t).
\]
Structural second order estimates for the densities

\[ \mu_t = c(t, \cdot) \mathcal{L}^d \] with

\[ c(t) = \frac{q^2(t)}{\delta(t)} = \frac{q^{d+2}(t)}{q^d(t) \delta(t)} = \frac{q^{d+2}(t)}{\rho^d(t)}, \]

\[ \rho(t) := q(t) \delta^{1/d}(t) \]
Structural second order estimates for the densities

\[\mu_t = c(t, \cdot)\mathcal{L}^d \text{ with} \]

\[c(t) = \frac{q^2(t)}{\delta(t)} = \frac{q^{d+2}(t)}{q^d(t)\delta(t)} = \frac{q^{d+2}(t)}{\rho^d(t)}, \quad \rho(t) := q(t)\delta^{1/d}(t)\]

**Structural estimates**

since the previous identities yield

\[\frac{\ddot{q}(t)}{q(t)} \geq \frac{|\nabla \xi_t|^2}{q(t)}, \quad \frac{\ddot{\rho}(t)}{\rho(t)} \leq \left(1 - \frac{4}{d}\right)\frac{\ddot{q}(t)}{q(t)}.

Theorem

The density \(c(t, \cdot)\mathcal{L}^d\) is convex along the characteristics:

\[\dddot{c}c \geq 6\dddot{q}q \geq 0.\]
Structural second order estimates for the densities

\[ \mu_t = c(t, \cdot) \mathcal{L}^d \]  

with

\[ c(t) = \frac{q^2(t)}{\delta(t)} = \frac{q^{d+2}(t)}{q^d(t)\delta(t)} = \frac{q^{d+2}(t)}{\rho^d(t)}, \quad \rho(t) := q(t)\delta^{1/d}(t) \]

**Structural estimates**

\[ \frac{\ddot{q}(t)}{q(t)} \geq 0, \quad \frac{\ddot{\rho}(t)}{\rho(t)} \leq \left(1 - \frac{4}{d}\right) \frac{\ddot{q}(t)}{q(t)}. \]

since the previous identities yield

\[ \begin{aligned}
\frac{\ddot{q}(t)}{q(t)} &= |\nabla \xi_t|^2 \\
\frac{\ddot{\rho}(t)}{\rho(t)} &= \frac{1}{d^2} \left( (\Delta \xi_t)^2 - d|D^2 \xi_t|^2 \right) + \left(1 - \frac{4}{d}\right)|\nabla \xi_t|^2
\end{aligned} \]

**Theorem**

The density \( c(t, \cdot) \) is **convex along the characteristics**: 

\[ \frac{\dddot{c}}{c} \geq 6 \frac{\ddot{q}}{q} \geq 0. \]

The functional \( \mu \mapsto \|d\mu/d\mathcal{L}^d\|_{L^\infty} \) is **geodesically convex**.
Consider a functional

\[ \mathcal{E}(\mu) := \int E(c(x)) \, dx, \quad c = \frac{d\mu}{d\mathcal{L}^d} \]

where \( E \) is convex (smooth).
Consider a functional 

$$E(\mu) := \int E(c(x)) \, dx, \quad c = \frac{d\mu}{d\mathcal{L}^d}$$

where $E$ is convex (smooth).

The case $E(c) = c$ corresponds to the total mass of $\mu$: it is quadratic.
Application: geodesic convexity of integral functionals

Consider a functional

$$E(\mu) := \int E(c(x)) \, dx, \quad c = \frac{d\mu}{d\mathcal{L}^d}$$

where $E$ is convex (smooth).

The case $E(c) = c$ corresponds to the total mass of $\mu$: it is quadratic.

More generally, we set $\varepsilon_0(c) := E(c)$, $\varepsilon_1(c) := cE'(c)$, $\varepsilon_2(c) := c^2E''(c)$.

McCann condition:

$$\varepsilon_2(c) \geq \left(1 - \frac{1}{d}\right)(\varepsilon_1(c) - \varepsilon_0(c)) \geq 0 \iff r^dE(r^{-d}) \text{ convex, nonincreasing.}$$
Consider a functional
\[ E(\mu) := \int E(c(x)) \, dx, \quad c = \frac{d\mu}{d\mathcal{L}^d} \]
where \( E \) is convex (smooth).

The case \( E(c) = c \) corresponds to the total mass of \( \mu \): it is quadratic.

More generally, we set \( \varepsilon_0(c) := E(c), \varepsilon_1(c) := cE'(c), \varepsilon_2(c) := c^2E''(c) \).

**McCann condition:**
\[ \varepsilon_2(c) \geq \left(1 - \frac{1}{d}\right)(\varepsilon_1(c) - \varepsilon_0(c)) \geq 0 \iff r^dE(r^{-d}) \text{ convex, nonincreasing}. \]

**Convexity with respect to the Hellinger-Kakutani distance:**
\[ \varepsilon_2(c) + \frac{1}{2}\varepsilon_0(c) \geq 0 \iff r \mapsto E(r^2) \text{ convex}. \]
Consider a functional

$$E(\mu) := \int E(c(x)) \, dx, \quad c = \frac{d\mu}{d\mathcal{L}^d}$$

where $E$ is convex (smooth).

The case $E(c) = c$ corresponds to the total mass of $\mu$: it is quadratic.

More generally, we set $\varepsilon_0(c) := E(c)$, $\varepsilon_1(c) := cE'(c)$, $\varepsilon_2(c) := c^2E''(c)$.

**McCann condition:**

$$\varepsilon_2(c) \geq \left(1 - \frac{1}{d}\right)(\varepsilon_1(c) - \varepsilon_0(c)) \geq 0 \iff r^d E(r^{-d}) \text{ convex, nonincreasing.}$$

**Convexity with respect to the Hellinger-Kakutani distance:**

$$\varepsilon_2(c) + \frac{1}{2}\varepsilon_0(c) \geq 0 \iff r \mapsto E(r^2) \text{ convex.}$$

**Theorem**

$E$ is geodesically convex w.r.t. HK if and only if

$$G(c) := \begin{pmatrix}
\varepsilon_2(c) - \frac{d-1}{d}(\varepsilon_1(c)-\varepsilon_0(c)) & \varepsilon_2(c) - \frac{1}{2}(\varepsilon_1(c)-\varepsilon_0(c)) \\
\varepsilon_2(c) - \frac{1}{2}(\varepsilon_1(c)-\varepsilon_0(c)) & \varepsilon_2(c) + \frac{1}{2}\varepsilon_1(c)
\end{pmatrix} \succeq 0, \quad \varepsilon_1 \geq \varepsilon_0.$$
An equivalent condition

$$G(c) := \begin{pmatrix} \varepsilon_2(c) - \frac{d-1}{d}(\varepsilon_1(c) - \varepsilon_0(c)) & \varepsilon_2(c) - \frac{1}{2}(\varepsilon_1(c) - \varepsilon_0(c)) \\ \varepsilon_2(c) - \frac{1}{2}(\varepsilon_1(c) - \varepsilon_0(c)) & \varepsilon_2(c) + \frac{1}{2}\varepsilon_1(c) \end{pmatrix} \geq 0, \quad \varepsilon_1 \geq \varepsilon_0.$$ 

**Theorem**

Define

$$N(\rho, q) := \left( \frac{\rho}{q} \right)^d E \left( \frac{q^{d+2}}{\rho^d} \right)$$

$\mathcal{E}$ is geodesically convex if and only if

$N_E$ is jointly convex and nonincreasing w.r.t. $\rho$. 

Main examples:

- the power functions $E(c) := c^p$ are convex if $p \geq 1$.
- In dimension $d = 2$ also $E(c) = -\sqrt{c}$ is convex.
- In dimension $d = 1$ all the power functions $E(c) = -c^p$, $p \in \left[\frac{1}{3}, \frac{1}{2}\right]$ induces convex functionals.
An equivalent condition

\[ G(c) := \begin{pmatrix} \varepsilon_2(c) - \frac{d-1}{d} (\varepsilon_1(c) - \varepsilon_0(c)) & \varepsilon_2(c) - \frac{1}{2} (\varepsilon_1(c) - \varepsilon_0(c)) \\ \varepsilon_2(c) - \frac{1}{2} (\varepsilon_1(c) - \varepsilon_0(c)) & \varepsilon_2(c) + \frac{1}{2} \varepsilon_1(c) \end{pmatrix} \geq 0, \quad \varepsilon_1 \geq \varepsilon_0. \]

Theorem

Define

\[ N(\rho, q) := \left( \frac{\rho}{q} \right)^d E \left( \frac{q^{d+2}}{\rho^d} \right) \]

\( E \) is geodesically convex if and only if

\[ N_E \text{ is jointly convex and nonincreasing w.r.t. } \rho. \]

Main examples: the power functions \( E(c) := c^p \) are convex if \( p \geq 1 \).

In dimension \( d = 2 \) also \( E(c) = -\sqrt{c} \) is convex.

In dimension \( d = 1 \) all the power functions \( E(c) = -c^p, \ p \in [1/3, 1/2] \) induces convex functionals.
1 Unbalanced Optimal Transport: a relaxation viewpoint

2 The Hellinger-Kantorovich metric between positive measures of arbitrary mass

3 Geodesics and geodesic convexity

4 Regularity of solutions to the Conical Hopf-Lax semigroup
Conical Hopf-Lax representation formula:

\[ P_t \xi(x) := \inf_{y} \frac{1}{2t} \left[ 1 - \frac{\cos^2(\pi/2)(|y - x|)}{1 + 2t \xi(x)} \right] \quad \text{(CHL)} \]

It is useful to introduce the reverse evolution (Villani '09)

\[ R_t \bar{\xi}(x) := -P_{1-t}(-\bar{\xi})(x) = \sup_{y} \frac{1}{2(1-t)} \left[ \frac{\cos^2(\pi/2)(|y - x|)}{1 - 2(1-t) \bar{\xi}(x)} - 1 \right] \quad \text{(RCHL)} \]

If \( \bar{\xi}_1 : \mathbb{R}^d \to [-b, a] \) with \(-\infty < -b < a < 1/2\) then the functions \( \bar{\xi}_t := R_t \bar{\xi}_1(x) \) are globally bounded, Lipschitz and semiconvex, \( \xi_t = P_{t-s} \xi_s, \xi_1 < 1/2 \).
Conical Hopf-Lax representation formula:

\[ \mathcal{P}_t \xi(x) := \inf_{y} \frac{1}{2t} \left\{ 1 - \frac{\cos^2(\pi/2 \cdot |y - x|)}{1 + 2t \xi(x)} \right\} \]  

(CHL)

It is useful to introduce the reverse evolution (Villani '09)

\[ \mathcal{R}_t \bar{\xi}(x) := -\mathcal{P}_{1-t}(-\bar{\xi})(x) = \sup_{y} \frac{1}{2(1-t)} \left[ \frac{\cos^2(\pi/2 \cdot |y - x|)}{1 - 2(1-t) \bar{\xi}(x)} - 1 \right] \]  

(RCHL)

Theorem

If \( \xi_0 : \mathbb{R}^d \to [-a, b] \) with \(-1/2 < -a < b < \infty\) then the functions \( \xi_t := \mathcal{P}_t \xi_0(x) \) are globally bounded, Lipschitz and semiconcave \( \xi_t = \mathcal{P}_{t-s} \xi_s, \xi_1 < 1/2 \).
Conical Hopf-Lax representation formula:

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

It is useful to introduce the reverse evolution (Villani '09)

\[ \mathcal{R}_t \bar{\xi}(x) := -\mathcal{P}_{1-t}(-\bar{\xi})(x) = \sup_y \frac{1}{2(1-t)} \left[ \frac{\cos^2(\pi/2)(|y - x|)}{1 - 2(1-t)\bar{\xi}(x)} - 1 \right] \] (RCHL)

**Theorem**

If \( \xi_0 : \mathbb{R}^d \to [-a, b] \) with \(-1/2 < -a < b < \infty\) then the functions \( \xi_t := \mathcal{P}_t \xi_0(x) \) are globally bounded, Lipschitz and semiconcave \( \xi_t = \mathcal{P}_{t-s} \xi_s, \xi_1 < 1/2 \).

If \( \bar{\xi}_1 : \mathbb{R}^d \to [-b, a] \) with \(-\infty < -b < a < 1/2\) then the functions \( \bar{\xi}_t := \mathcal{R}_t \bar{\xi}_1(x) \) are globally bounded, Lipschitz and semiconvex, \( \xi_t = \mathcal{P}_{t-s} \xi_s, \xi_1 < 1/2 \).
Regularity of CHL (II)

Theorem (Differentiability on the contact set)

If $\bar{\xi}_1 = \xi_1 = \mathcal{P}_1(\xi_0)$, $\xi_0 = \mathcal{R}_1 \xi_1$ then $\xi_t \geq \bar{\xi}_t$ and the contact set $\Xi_t := \{x : \bar{\xi}_t(x) = \xi_t(x)\}$ is closed and contains $\text{supp}(\mu_t)$. 

\[ q_{t_1} \rightarrow t_2 \circ T_{t_0} \rightarrow t_1 = T_{t_0} \rightarrow t_2 \]
Theorem (Differentiability on the contact set)

If $\bar{\xi}_1 = \xi_1 = \mathcal{P}_1(\xi_0)$, $\xi_0 = \mathcal{R}_1 \xi_1$ then $\xi_t \geq \bar{\xi}_t$ and the contact set

$$\Xi_t := \{ x : \bar{\xi}_t(x) = \xi_t(x) \}$$

is closed and contains $\text{supp}(\mu_t)$.

$\xi_t$, $\bar{\xi}_t$ are differentiable in $\Xi_t$ with gradient $g_t$. 

q
Theorem (Differentiability on the contact set)

If \( \bar{\xi}_1 = \xi_1 = \mathcal{P}_1(\xi_0) \), \( \xi_0 = \mathcal{R}_1 \xi_1 \) then \( \xi_t \geq \bar{\xi}_t \) and the **contact set**

\[
\Xi_t := \left\{ x : \bar{\xi}_t(x) = \xi_t(x) \right\}
\]

is closed and contains \( \text{supp}(\mu_t) \).

\( \xi_t, \bar{\xi}_t \) are **differentiable** in \( \Xi_t \) with gradient \( g_t \).

\( g_t \) is **Lipschitz** in \( \Xi_t \) and for every \( s, t \in [0, 1] \) defining

\[
T_{s \to t}(x) := x + \arctan \left( \frac{(t - s)g_s(x)}{1 + 2(t - s)g_s(x)} \right)
\]

the map \( T_{s \to t} \) is **Lipschitz**, it satisfies \( T_{s \to t}(\Xi_s) = \Xi_t \) and

the **concatenation property**

\[
T_{t_1 \to t_2} \circ T_{t_0 \to t_1} = T_{t_0 \to t_2}
\]
Regularity of CHL (II)

**Theorem (Differentiability on the contact set)**

If $\overline{\xi}_1 = \xi_1 = \mathcal{P}_1(\xi_0)$, $\xi_0 = \mathcal{P}_1\xi_1$ then $\xi_t \geq \overline{\xi}_t$ and the **contact set**

$$\Xi_t := \left\{ x : \overline{\xi}_t(x) = \xi_t(x) \right\}$$

is closed and contains $\text{supp}(\mu_t)$.

$\xi_t$, $\overline{\xi}_t$ are **differentiable** in $\Xi_t$ with gradient $g_t$.

$g_t$ is **Lipschitz** in $\Xi_t$ and for every $s, t \in [0, 1]$ defining

$$T_{s \rightarrow t}(x) := x + \arctan\left(\frac{(t-s)g_s(x)}{1 + 2(t-s)g_s(x)}\right)$$

the map $T_{s \rightarrow t}$ is **Lipschitz**, it satisfies $T_{s \rightarrow t}(\Xi_s) = \Xi_t$ and

**the concatenation property**

$$T_{t_1 \rightarrow t_2} \circ T_{t_0 \rightarrow t_1} = T_{t_0 \rightarrow t_2}$$

Setting

$$q^2_{s \rightarrow t}(x) := (1 + 2(t-s)\xi_s(x))^2 + (t-s)^2|g_s(x)|^2$$

we have $q_{t_1 \rightarrow t_2} \circ T_{t_0 \rightarrow t_1} \cdot q_{t_0 \rightarrow t_1} = q_{t_0 \rightarrow t_2}$
Theorem

For every $s \in (0, 1)$ and $t \in [0, 1]$ the transport-growth pair $(T_{s \to t}, q_{s \to t})$ is the unique solution to the Monge formulation for the HK problem between $\mu_s$ and $\mu_t$. 

If $\mu_s \ll L_d$ then $\mu_t \ll L_d$ for every $t \in (0, 1)$.

If $\text{supp}(\nu_s) \subset \text{supp}(\mu_s)$ then $\nu_t := (T_{s \to t}, q_{s \to t})^* \nu_s$ is a geodesic.
For every $s \in (0, 1)$ and $t \in [0, 1]$ the transport-growth pair $(T_{s \to t}, q_{s \to t})$ is the unique solution to the Monge formulation for the $\mathcal{H}K$ problem between $\mu_s$ and $\mu_t$.

In particular, if for given $\mu_0, \mu_1, \mu_s$

$$\mathcal{H}K(\mu_0, \mu_s) = s\mathcal{H}K(\mu_0, \mu_1), \quad \mathcal{H}K(\mu_s, \mu_1) = (1 - s)\mathcal{H}K(\mu_0, \mu_1)$$

then there exists a unique geodesic $\mu : [0, 1] \to \mathcal{M}(\mathbb{R}^d)$ connecting $\mu_0$ to $\mu_1$ such that $\mu(s) = \mu_s$. 


Nonbranching and restrictions

**Theorem**

For every \( s \in (0, 1) \) and \( t \in [0, 1] \) the transport-growth pair \((T_{s \rightarrow t}, q_{s \rightarrow t})\) is the unique solution to the Monge formulation for the HK problem between \( \mu_s \) and \( \mu_t \).

In particular, if for given \( \mu_0, \mu_1, \mu_s \)

\[
HK(\mu_0, \mu_s) = sHK(\mu_0, \mu_1), \quad HK(\mu_s, \mu_1) = (1 - s)HK(\mu_0, \mu_1)
\]

then there exists a unique geodesic \( \mu : [0, 1] \rightarrow \mathcal{M}(\mathbb{R}^d) \) connecting \( \mu_0 \) to \( \mu_1 \) such that \( \mu(s) = \mu_s \).

If \( \mu_s \ll \mathcal{L}^d \) then \( \mu_t \ll \mathcal{L}^d \) for every \( t \in (0, 1) \).
Theorem

For every $s \in (0, 1)$ and $t \in [0, 1]$ the transport-growth pair $(T_{s \to t}, q_{s \to t})$ is the unique solution to the Monge formulation for the HK problem between $\mu_s$ and $\mu_t$.

In particular, if for given $\mu_0, \mu_1, \mu_s$

$$\text{HK}(\mu_0, \mu_s) = s \text{HK}(\mu_0, \mu_1), \quad \text{HK}(\mu_s, \mu_1) = (1 - s) \text{HK}(\mu_0, \mu_1)$$

then there exists a unique geodesic $\mu : [0, 1] \to \mathcal{M}(\mathbb{R}^d)$ connecting $\mu_0$ to $\mu_1$ such that $\mu(s) = \mu_s$.

If $\mu_s \ll \mathcal{L}^d$ then $\mu_t \ll \mathcal{L}^d$ for every $t \in (0, 1)$.

If $\text{supp}(\nu_s) \subset \text{supp}(\mu_s)$ then $\nu_t := (T_{s \to t}, q_{s \to t})_* \nu_s$ is a geodesic.
Let $\mathcal{D}_s \subset \Xi_s$ the set of points of density 1 where $g_s$ is differentiable.

**Theorem**

$A_s := Dg_s$ is symmetric. $\xi_s$ has a second order Taylor expansion in terms of $g_s$ and $A_s$. We thus can set $g_s = \nabla \xi_s, B_s = D\nabla \xi_s = D^2 \xi_s$ in $\mathcal{D}_s$.

\[
\begin{align*}
\ddot{T}(t) &= 4\xi_t \nabla \xi_t(T(t)) \\
\ddot{q}(t) &= |\nabla \xi_t(T(t))|^2 q(t) \\
\ddot{B}(t) &= -4 \left( \nabla \xi_t \otimes \nabla \xi_t + \xi_t D^2 \xi_t \right) \circ T(t) \cdot B(t) \\
\ddot{\delta}(t) &= \left( (\Delta \xi_t)^2 - |D^2 \xi_t|^2 - 4|\nabla \xi_t|^2 - 4\xi_t \Delta \xi_t \right) \circ T(t) \cdot \delta(t).
\end{align*}
\]
Second order regularity of CHL (III)

Let $\mathcal{D}_s \subset \Xi_s$ the **set of points of density 1** where $g_s$ is differentiable.

**Theorem**

$A_s := Dg_s$ is **symmetric**. $\xi_s$ has a **second order Taylor expansion** in terms of $g_s$ and $A_s$. We thus can set $g_s = \nabla \xi_s$, $B_s = D\nabla \xi_s = D^2 \xi_s$ in $\mathcal{D}_s$.

If $\mu_s \ll L^d$ then $\mu_s(\Xi_s \setminus \mathcal{D}_s) = 0$.

\[
\begin{align*}
\ddot{T}(t) &= 4\xi_t \nabla \xi_t(T(t)) \\
\ddot{q}(t) &= |\nabla \xi_t(T(t))|^2 q(t) \\
\ddot{B}(t) &= -4 \left( \nabla \xi_t \otimes \nabla \xi_t + \xi_t D^2 \xi_t \right) \circ T(t) \cdot B(t) \\
\ddot{\delta}(t) &= \left( (\Delta \xi_t)^2 - |D^2 \xi_t|^2 - 4|\nabla \xi_t|^2 - 4\xi_t \Delta \xi_t \right) \circ T(t) \cdot \delta(t).
\end{align*}
\]
Let $\mathcal{D}_s \subset \Xi_s$ the set of points of density 1 where $g_s$ is differentiable.

**Theorem**

$A_s := Dg_s$ is symmetric. $\xi_s$ has a second order Taylor expansion in terms of $g_s$ and $A_s$. We thus can set $g_s = \nabla \xi_s$, $B_s = D\nabla \xi_s = D^2 \xi_s$ in $\mathcal{D}_s$.

If $\mu_s \ll \mathcal{L}^d$ then $\mu_s(\Xi_s \setminus \mathcal{D}_s) = 0$.

$T_{s\to t}$ is differentiable in $\mathcal{D}_s$ and $T_{s\to t}(\mathcal{D}_s) = \mathcal{D}_t$.

The maps $T(t) := T_{s\to t}$, $B(t, \cdot) := DT(t, \cdot)$, $\delta(t, \cdot) := \det B(t, \cdot)$ are analytic in time and satisfy the characteristic systems of ODE.

\[
\begin{align*}
\ddot{T}(t) &= 4\xi_t \nabla \xi_t (T(t)) \\
\ddot{q}(t) &= |\nabla \xi_t (T(t))|^2 q(t) \\
\ddot{B}(t) &= -4 \left( \nabla \xi_t \otimes \nabla \xi_t + \xi_t D^2 \xi_t \right) \circ T(t) \cdot B(t) \\
\ddot{\delta}(t) &= \left( (\Delta \xi_t)^2 - |D^2 \xi_t|^2 - 4|\nabla \xi_t|^2 - 4\xi_t \Delta \xi_t \right) \circ T(t) \cdot \delta(t).
\end{align*}
\]


