The intrinsic dynamics of optimal transport

Ludovic Rifford

Université Nice Sophia Antipolis
&
Institut Universitaire de France

CMC conference
"Analysis, Geometry, and Optimal Transport"
KIAS, Seoul, June 2016

Monge vs. Kantorovich

Gaspard Monge

(1746-1818)

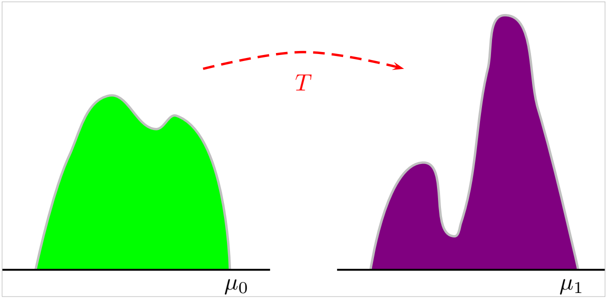
Leonid Kantorovich

(1912-1986)

Transport maps

Let μ_0 and μ_1 be **probability measures** on M. We call **transport map** from μ_0 to μ_1 any measurable map $T : M \to M$ such that $T_\sharp \mu_0 = \mu_1$, that is

$$\mu_1(B) = \mu_0\big(T^{-1}(B)\big), \quad \forall B \text{ measurable } \subset M.$$



The original Monge Problem

Let $M = \mathbb{R}^n$, given two probabilities measures μ_0, μ_1 on M, find a **transport map** $T : M \to M$ from μ_0 to μ_1 which **minimizes** the transportation cost

$$\int_M \|T(x) - x\| \, d\mu_0(x),$$

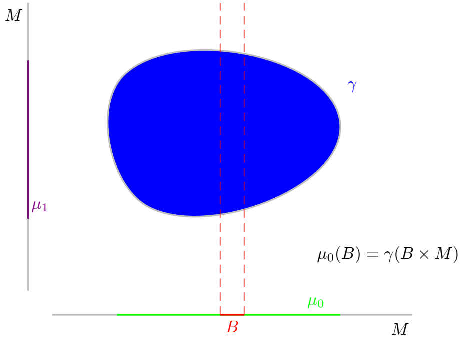
among all transport maps from μ_0 to μ_1.

Transport plans

Let μ_0 and μ_1 be **probability measures** on M. We call **transport plan** between μ_0 and μ_1 any probability measure γ on $M \times M$ having **marginals** μ_0 and μ_1, i.e.

$$(\pi_1)_\sharp \gamma = \mu_0 \qquad \text{and} \qquad (\pi_2)_\sharp \gamma = \mu_1,$$

(where $\pi_1 : M \times M \to M$ and $\pi_2 : M \times M \to M$ are the canonical projections),



The Kantorovich Optimal Transport Problem

Given M, a cost $c : M \times M \to \mathbb{R}$ and two probability measures μ_0, μ_1 on M, we want to find a transport plan γ on $M \times M$ between μ_0 and μ_1 which minimizes the transportation cost

$$\int_{M\times M} c(x,y)d\gamma(x,y),$$

among all transport maps from μ_0 to μ_1.

Monge vs. Kantorovitch

Let M be a smooth compact manifold, $c : M \times M \to \mathbb{R}$ be a continuous cost function, and μ_0, μ_1 two probability measures on M.

Monge's Problem

Minimize

$$\int_M c(x, T(x)) d\mu_0(x)$$

among all transport maps T, that is $T_\sharp \mu_0 = \mu_1$.

Kantorovitch's Problem

Minimize

$$\int_M c(x, y) d\gamma(x, y)$$

among all transport plans γ, that is $(\pi_1)_\sharp \gamma = \mu_0, (\pi_2)_\sharp \gamma = \mu_1$.

Kantorovitch's Duality

Theorem

There are two continuous function $\psi_1, \psi_2 : M \to \mathbb{R}$ *satisfying*

$$\psi_1(x) = \max_{y \in M} \{\psi_2(y) - c(x,y)\} \qquad \forall x \in M,$$

$$\psi_2(y) = \min_{x \in M} \{\psi_1(x) + c(x,y)\} \qquad \forall y \in M.$$

such that the following holds: a transport plan γ *is optimal if and only if one has*

$$\psi_2(y) - \psi_1(x) = c(x,y) \qquad \text{for } \gamma \text{ a.e. } (x,y) \in M \times M.$$

As a consequence, to obtain that an optimal transport plan corresponds to a Monge's optimal transport map, we have to show that γ is concentrated on a graph.

Kantorovich $\rightsquigarrow$ Monge

Let $\psi_1, \psi_2 : M \to \mathbb{R}$ be a pair of **Kantorovich potentials** given by the previous result. A way to get existence and uniqueness for Monge is to proceed as follows:

- Show that ψ_1 admits a super-differential for μ_0-almost every point.
- Let $\bar{x} \in \mathrm{supp}(\mu_0)$ be such that ψ_1 admits a super-differential $d_{\bar{x}}f$ at $\bar{x}$ and let $\bar{y}$ be such that

$$\psi_1(\bar{x}) = \psi_2(\bar{y}) - c(\bar{x}, \bar{y}).$$

Then we have for every $x \in M$

$$c(x, \bar{y}) \geq \psi_2(\bar{y}) - \psi_1(x) \geq \psi_2(\bar{y}) - f(x).$$

$\rightsquigarrow$ The function $x \mapsto c(x, \bar{y})$ admits $-d_{\bar{x}}f$ as a sub-differential at $\bar{x}$.

School matching around a lake

Find a **transport map** ($T_\sharp \mu_X = \mu_Y$)

$$T : X = \{pupils\} = \mathbb{S}^1 \longrightarrow Y = \{schools\} = \mathbb{S}^1$$

which minimizes the **transportation cost**

$$\int_X c\,(x, T(x))\; d\mu_X(x)$$

for some **cost** $c : \mathbb{S}^1 \times \mathbb{S}^1 \to [0, \infty)$.

Geodesic cost

$$c(x,y) = d_g(x,y)^2$$

Kingsley lake, FL

Euclidean cost

$$c(x,y) = |y-x|^2$$

The (quadratic) geodesic cost

Let (M, g) be a smooth compact Riemannian manifold, given two probabilities measures μ_0, μ_1 on M, find a transport map $T : M \to M$ from μ_0 to μ_1 which minimizes

$$\int_M d_g^2(x, T(x)) d\mu_0(x).$$

Theorem (McCann '01)

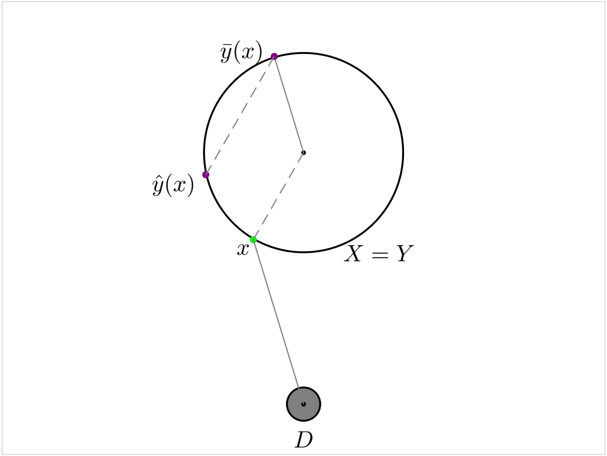
If μ_0 is absolutely continuous w.r.t. Lebesgue, then **there exists a unique** *optimal transport map T from μ_0 to μ_1.*

Comments:

- Sub-TWIST $(D_x^- c(\cdot, y_1) \cap D_x^- c(\cdot, y_2) = \emptyset \quad \forall y_1 \neq y_2, \forall x)$ $\Longrightarrow$ existence and uniqueness
- **No smooth costs satisfy Sub-TWIST**

The (quadratic) Euclidean cost

Let $X = Y = \mathbb{S}^1 \subset \mathbb{R}^2$ and $c(x,y) = |y-x|^2$.



The (quadratic) Euclidean cost

Let $\tilde{\psi}$ be the distance function to the disc D, then we define the pair of potentials $\psi_1, \psi_2 : \mathbb{S}^1 \to \mathbb{R}$ by

$$\psi_1(x) := \tilde{\psi}(x) - \frac{1}{2}|x|^2,$$

$$\psi_2(y) := \min_x \left\{ \psi_1(x) + c(x,y) \right\}.$$

For x close to the south pole, we define $\bar{y}(x), \hat{y}(x)$ in $\mathbb{S}^1$ by

$$\begin{cases} \bar{y}(x) & := & \nabla_x \tilde{\psi}, \\ \hat{y}(x) & := & \nabla_x \tilde{\psi} + \lambda(x)\, x \text{ with } \lambda(x) \geq 0. \end{cases}$$

The (quadratic) Euclidean cost

By convexity of $\tilde{\psi}$,

$$\langle \bar{y}(x), x' - x \rangle \leq \tilde{\psi}(x') - \tilde{\psi}(x) \qquad \forall x'.$$

$$\rightsquigarrow \bar{y}(x) \in \partial_c \psi_1(x) := \{(x,y) \,|\, c(x,y) = \psi_2(y) - \psi_1(x)\} \,.$$

We also have for any x',

$$\begin{aligned} \langle \hat{y}(x), x' - x \rangle &= \langle \bar{y}(x), x' - x \rangle + \lambda(x) \langle x, x' - x \rangle \\ &\leq \langle \bar{y}(x), x' - x \rangle \\ &\leq \tilde{\psi}(x') - \tilde{\psi}(x). \end{aligned}$$

$$\rightsquigarrow \hat{y}(x) \in \partial_c \psi_1(x) := \{(x,y) \,|\, c(x,y) = \psi_2(y) - \psi_1(x)\} \,.$$

In consequence, for x close to the south pole, we have

$$\partial_c \psi_1(x) = \{\bar{y}(x), \hat{y}(x)\}.$$

The (quadratic) Euclidean cost

Let us consider an absolutely continuous probability measure μ_0 on $X = \mathbb{S}^1$ whose support is close to the south pole. Then define the measures $\bar{\nu}, \hat{\nu}$ on N by

$$\bar{\nu} := \frac{1}{2}\bar{y}_\sharp\mu_0, \ \hat{\nu} := \frac{1}{2}\hat{y}_\sharp\mu_0, \text{ and set } \mu_1 := \bar{\nu} + \hat{\nu}.$$

Any plan γ with marginals μ_0 and μ_1 satisfies

$$\begin{aligned}
\int_{X\times Y} c(x,y)\, d\gamma(x,y) &\geq \int_{X\times Y} [\psi_2(y) - \psi_1(x)]\, d\gamma(x,y) \\
&= \int_Y \psi_2(y)\, d\mu_1(y) - \int_X \psi_1(x)\, d\mu_0(x) \\
&= \int_{X\times Y} c(x,y)\, d\bar{\gamma}(x,y),
\end{aligned}$$

with equality in the first inequality if and only if $\gamma = \bar{\gamma}$ with $\bar{\gamma} := \frac{1}{2}\left(Id, \bar{y}\right)_\sharp \mu_0 + \frac{1}{2}\left(Id, \hat{y}\right)_\sharp \mu_0$.

Non-genericity of twist

Theorem (McCann, LR)

Let M, N be smooth compact manifolds of dimensions $n \geq 1$ and $c : M \times N \to [0, \infty)$ a cost function of class C^2. Assume that

$$\exists (\bar{x}, \bar{y}) \in M \times N \quad \text{such that} \quad \frac{\partial^2 c}{\partial x \partial y}(\bar{x}, \bar{y}) \quad \text{is invertible.} \qquad (1)$$

Then there is a pair μ_0, μ_1 of probability measures respectively on M and N which are both absolutely continuous w.r.t. Lebesgue for which **there is a unique optimal transport plan** *and such that this plan is* **not supported on a graph**. **The set of costs c satisfying (1) is open and dense in $C^2(M \times N; \mathbb{R})$.**

$\rightsquigarrow$ I do not know if assumption (1) is necessary.

Purpose of the talk

- Study sufficient conditions for smooth costs that insure uniqueness of Kantorovitch optimizers (minimizing transport plans).
- Exhibit such costs on arbitrary manifolds.
- Study the size of the set of such costs (genericity for some topology)

References

- K. Hestir and S. C. Williams. Supports of doubly stochastic measures (1995)
- S. Bianchini and L. Caravenna. On the extremality, uniqueness, and optimality of transference plans (2009)
- W. Gangbo and R. J. McCann. Shape recognition via Wasserstein distance (2000)
- P.-A. Chiappori, R. McCann and L. Nesheim. Hedonic price equilibria, stable matching, and optimal transport: equivalence, topology, and uniqueness (2010)
- R. J. McCann and LR. The intrinsic dynamics of optimal transport (2015)

Setting

- M, N be smooth compact manifolds of dimensions ≥ 1.
- $c : M \times N \to [0, \infty)$ of class C^1.
- Given two probabilities measures μ_0, μ_1 on M, denote by $\Pi(\mu_0, \mu_1)$, the set of probability measures on $M \times N$ having marginals μ_0 and ν_0.
- A transport plan $\gamma \in \Pi(\mu_0, \mu_1)$ is called optimal if it minimizes the transportation cost

$$\int_{M \times N} c(x, y) d\gamma(x, y).$$

Observation

If the measures μ_0, μ_1 are two Borel probability measures on M and N, then

Theorem (Folklore)

For each $k \in \mathbb{N} \cup \{\infty\}$, there exists a residual set $\mathcal{C} \subset C^k(M \times N; \mathbb{R})$ such that for every $c \in \mathcal{C}$, there is a unique optimal transport plan between μ_0 and μ_1.

We want to find sufficient conditions depending only upon the cost, such that we have uniqueness of an optimal transport plan for any datas!!!!

Alternant chains

Definition

We call *L-chain* in S ($L \geq 1$) any ordered family of pairs

$$\Big((x_1, y_1), \ldots, (x_L, y_L)\Big) \in (M \times N)^L$$

such that:

- The set $\{(x_1, y_1), \ldots (x_L, y_L)\}$ is c-**cyclically monotone**.
- For every $l = 1, \ldots, L-1$ odd,

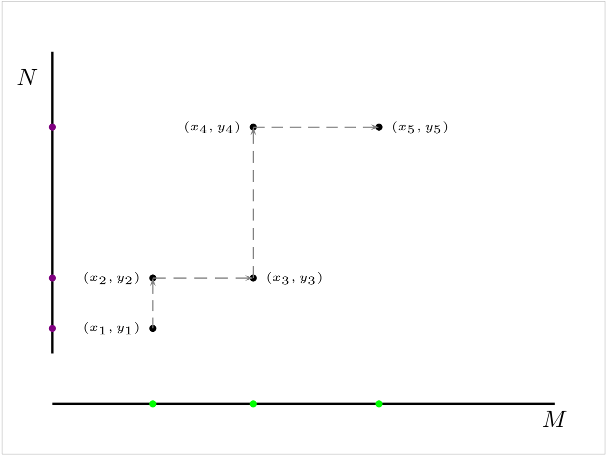
$$x_l = x_{l+1},\ y_l \neq y_{l+1},\ \frac{\partial c}{\partial x}(x_l, y_l) = \frac{\partial c}{\partial x}(x_l, y_{l+1}),$$

- For every $l = 1, \ldots, L-1$ even,

$$y_l = y_{l+1},\ x_l \neq x_{l+1},\ \frac{\partial c}{\partial y}(x_l, y_l) = \frac{\partial c}{\partial y}(x_{l+1}, y_l).$$

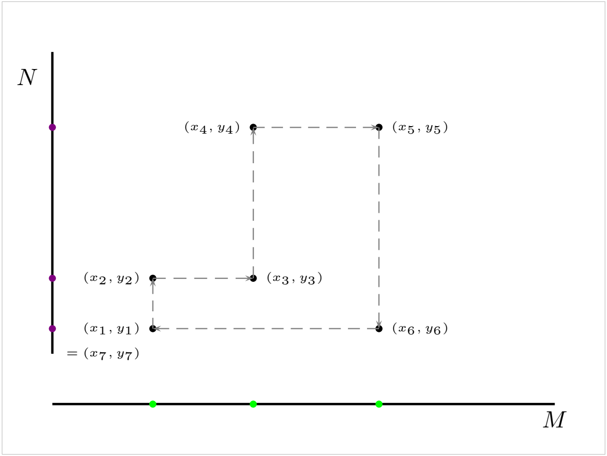
Alternant chains (picture)

A 5-chain



Alternant chains (picture)

Cyclic chains $\rightsquigarrow$ infinite chains



Optimal transport is unique if long chains are rare

Theorem

Let μ_0, μ_1 be probability measures respectively on M and N which are both absolutely continuous w.r.t. Lebesgue. Denote by $\mathcal{S}^\infty$ the set of points in $M \times N$ which occur in L-chains for arbitrarily large L and assume that $\mu_0\left(\pi^M(\mathcal{S}^\infty)\right) = 0$ or $\mu_1\left(\pi^N(\mathcal{S}^\infty)\right) = 0$. Then there is a **unique** *optimal transport plan.*

Comments:

- The theorem applies if there is a uniform bound on the length of all chains in $M \times N$.
- The theorem does not apply if there are cyclic chains on a set of positive measure.

Sketch of proof

Given μ_0, μ_1, there is a c-**cyclically monotone** set $\mathcal{S}$ and Lipschitz potentials $\psi : M \to \mathbb{R}$ and $\phi : M \to \mathbb{R}$ which satisfy

$$\psi(x) = \max_y \{\phi(y) - c(x,y)\}, \ \phi(y) = \min_x \{\psi(x) + c(x,y)\},$$

$$\mathcal{S} \subset \partial_c \psi := \Big\{ (x,y) \in M \times N \,|\, c(x,y) = \phi(y) - \psi(x) \Big\},$$

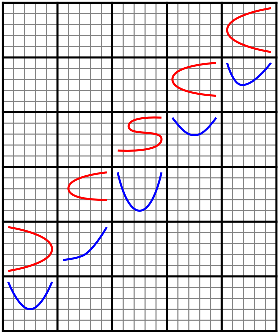
such that $\gamma \in \Pi(\mu, \nu)$ is optimal if and only if $\mathrm{Supp}(\gamma) \subset \mathcal{S}$.

Observation:
If ψ is differentiable at x, then

$$y \in \partial_c \psi(x) \implies \frac{\partial c}{\partial x}(x,y) = -d_x \psi.$$

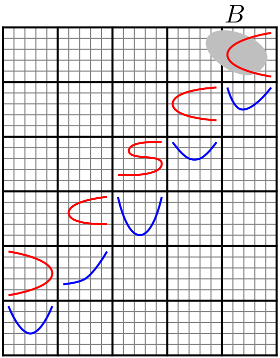
Sketch of proof

The previous observation allows to decompose $\mathcal{S}$ into a **numbered limb system** consisting of Borel graph and antigraphs (apart from a set of measure zero).



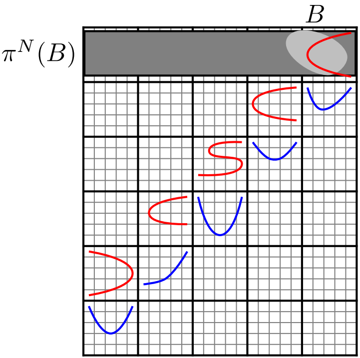
Sketch of proof

Then the result follows from uniqueness of transport plans in $\Pi(\mu_0, \mu_1)$ concentrated on the numbered limb system.



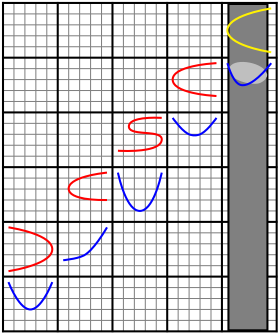
Sktech of proof

Then the result follows from uniqueness of transport plans in $\Pi(\mu_0, \mu_1)$ concentrated on the numbered limb system.



Sketch of proof

Then the result follows from uniqueness of transport plans in $\Pi(\mu_0, \mu_1)$ concentrated on the numbered limb system.

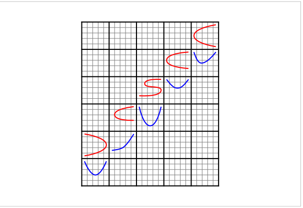


A remark (after Hestir and Williams)

Given a set $\mathcal{S} \subset M \times N$, define the equivalence relation $\sim_S$ on S by saying that $(x, y) \sim_S (x', y')$ if there is an alternating chain from (x, y) to (x', y').

Theorem

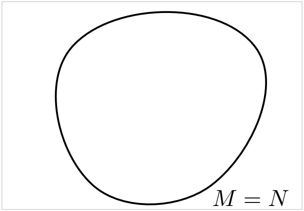
If the orbits of $\sim_S$ do not admit cycles, then $\mathcal{S}$ can be decomposed into a countable numbered limb system.



$\rightsquigarrow$ This can of formal result is not sufficient to get uniqueness of optimal plans.

Examples: Strictly convex sets

Setting: $M = N =$ smooth strictly convex compact hypersurface in $\mathbb{R}^n$, $c(x, y) = |y - x|^2$.



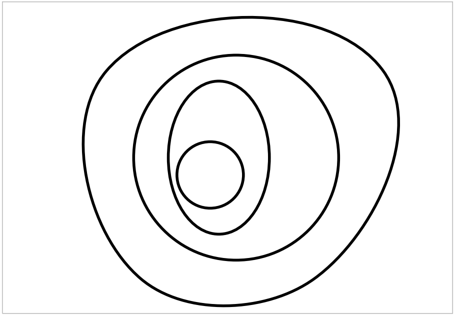
Lemma

There is no chain of length ≥ 4.

$\rightsquigarrow$ Uniqueness of optimal transport plans

Examples: Nested strictly convex sets

Setting: $M = N = \cup_{k=1}^{K} \mathcal{C}_k$ nested family of smooth strictly convex compact hypersurfaces in $\mathbb{R}^n$, $c(x,y) = |y-x|^2$.



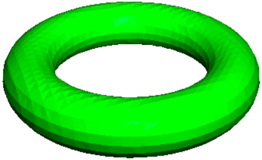
Lemma

There is no chain of length $\geq 4K+1$.

$\rightsquigarrow$ Uniqueness of optimal transport plans

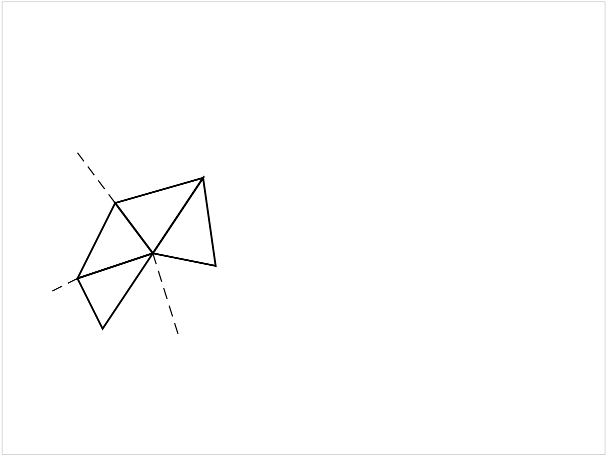
Examples: On arbitrary manifold

Setting: $M = N$ smooth compact manifold of dimension n

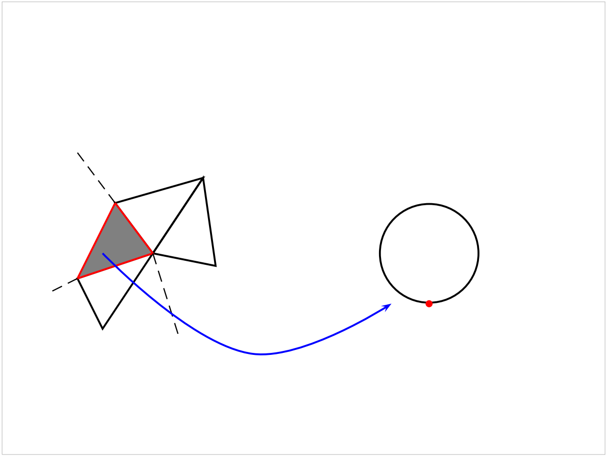


Let us consider a triangulation of the manifold.

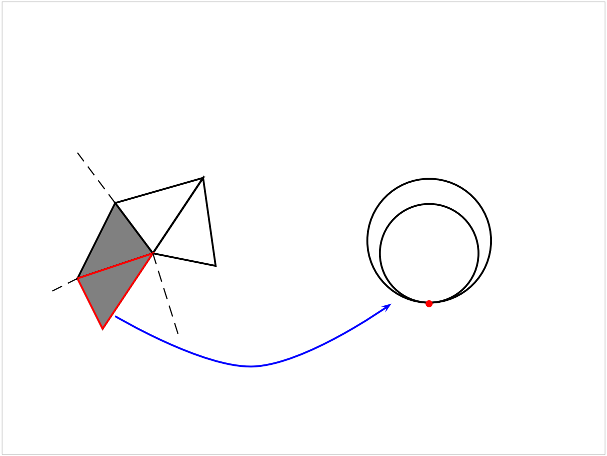
Examples: On arbitrary manifold



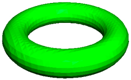
Examples: On arbitrary manifold



Examples: On arbitrary manifold



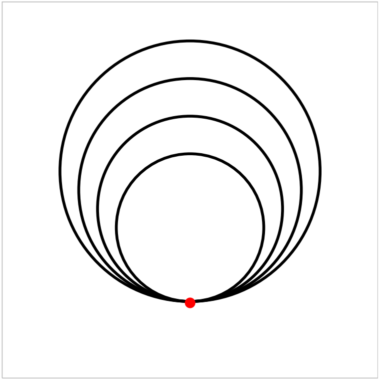
Examples: On arbitrary manifold



F

$\longrightarrow$

smooth



Then we define

$$c(x, y) = |F(y) - F(x)|^2$$

$\rightsquigarrow$ Uniqueness of optimal transport plans

Final comments

Open question

Given $k \in \mathbb{N} \cup \{\infty\}$, is the set of costs for which we have uniqueness (of optimal transport plans between absolutely continuous measures w.r.t. Lebesgue) dense in the C^k topology ?

Open question

The dynamics of the results presented previously are always the same. Can we find other examples with more involved dynamics ?

Thank you for your attention !!