Isometries of Wasserstein spaces

Dániel Virosztek
IST Austria

8th European Congress of Mathematics, 24 June 2021

Joint work with G. P. Gehér (U Reading) and T. Titkos (Rényi)

MARIE CURIE ACTIONS

European Commission
Horizon 2020
European Union funding
for Research & Innovation

IST AUSTRIA
Institute of Science and Technology

What is a Wasserstein space?

- roughly speaking: the space of sufficiently concentrated probability measures endowed with a metric which is calculated by means of *optimal transport* (Monge: 1781, Kantorovich: WW II, Villani: 2010, Figalli: 2018)
- more precisely: for a Polish space (X, ρ) and a parameter $0 < p < \infty$ the p-*Wasserstein space* is

$$\mathcal{W}_p(X) = \left\{ \mu \in \mathcal{P}(X) \,\middle|\, \int_X \rho(x, \hat{x})^p \, \mathrm{d}\mu(x) < \infty \text{ for some } \hat{x} \in X \right\}$$

endowed with the p-*Wasserstein distance*

$$d_{\mathcal{W}_p}(\mu, \nu) := \left(\inf_{\pi \in \Pi(\mu,\nu)} \int_{X^2} \rho(x, y)^p \, \mathrm{d}\pi(x, y) \right)^{\min\left\{\frac{1}{p}, 1\right\}}.$$

where $\Pi(\mu, \nu) = \left\{ \pi \in \mathcal{P}\left(X^2\right) \,\middle|\, \pi_1 = \mu, \pi_2 = \nu \right\}$ is the collection of all *transport plans* between μ and ν

Probabilistic and dynamical interpretations

- a probabilistic reformulation of the p-Wasserstein distance:

$$d_{\mathcal{W}_p}(\mu,\nu) = \left(\inf_{(\mathrm{x},\mathrm{y}):\, \mathrm{x}\sim\mu,\, \mathrm{y}\sim\nu} \mathrm{E}\left(d(\mathrm{x},\mathrm{y})^p\right) \right)^{\min\left\{\frac{1}{p},1\right\}}.$$

- a dynamical interpretation: Benamou-Brenier formula (fluid mechanics or "Eulerian" formulation):

$$d^2_{\mathcal{W}_2}(\mu,\nu) = \inf_{(\rho,v)\in V(\mu,\nu)} \int_0^1 \int_{R^n} \rho_t(x) \left\| v_t(x) \right\|^2 \mathrm{d}x\mathrm{d}t,$$

where $\{\rho_t\}_{t\in[0,1]}$ is the weak solution of the linear transport equation

$$\frac{\partial \rho_t}{\partial t} + \nabla_x \cdot (\rho_t v_t) = 0$$

with initial and final conditions $\rho_0 = \mu, \rho_1 = \nu$

Motivation and basic notions

- when working in a metric setting, a natural question arises: *how do isometries look like?*
- classical results:
  - the Banach–Lamperti theorem describing all linear non-surjective isometries of L^p spaces
  - the Banach–Stone theorem on the group of all linear isometries of commutative C^*-algebras ($C(K)$)
- Bertrand and Kloeckner: a series of papers on isometries of quadratic Wasserstein spaces over various metric spaces
- Kloeckner:[1] a description of the isometry group of $\mathcal{W}_2(E)$ for $\dim(E) < \infty$
- main question: is $\mathcal{W}_p(X)$ more symmetric than X?
- isometric rigidity: $\mathrm{Isom}(\mathcal{W}_p(X)) = \mathrm{Isom}(X)$

[1] B. Kloeckner, *A geometric study of Wasserstein spaces: Euclidean spaces,* Annali della Scuola Normale Superiore di Pisa - Classe di Scienze **IX** (2010), 297–323.

Basic notions, notation

Definition (Push-forward)

For a measurable map $g\colon X \to X$ the induced *push-forward map* $g_\#\colon \mathcal{P}(X) \to \mathcal{P}(X)$ is defined by

$$\left(g_\#(\mu)\right)(A) = \mu(g^{-1}[A]) \qquad (A \subseteq X \text{ Borel set}, \ \mu \in \mathcal{P}(X))$$

where $g^{-1}[A] = \{x \in X \,|\, g(x) \in A\}$. We call $g_\#(\mu)$ the *push-forward* of μ with g. If $\psi \in \mathrm{Isom}(\mathbb{R})$, then the push-forward map $\psi_\#$ is an isometry of $\mathcal{W}_p(X)$, and the embedding

$$\#\colon\ \mathrm{Isom}(X) \to \mathrm{Isom}\left(\mathcal{W}_p(X)\right), \qquad \psi \mapsto \psi_\#$$

is a group homomorphism. Isometries of the form $\psi_\#$ are called *trivial isometries*.

Basic notions, notation

- p-Wasserstein distance expressed by cumulative distribution and quantile functions:
  - Vallender:[2]

$$d_{\mathcal{W}_1}(\mu,\nu) = \int_{-\infty}^{\infty} |F_\mu(x) - F_\nu(x)| \, \mathrm{d}x = \int_0^1 \left|F_\mu^{-1}(x) - F_\nu^{-1}(x)\right| \, \mathrm{d}x$$

  - this can be generalized:[3]

$$d_{\mathcal{W}_p}(\mu,\nu) = \left(\int_0^1 \left|F_\mu^{-1} - F_\nu^{-1}\right|^p \mathrm{d}t\right)^{\frac{1}{p}} \qquad (p > 1, \ \mu,\nu \in \mathcal{W}_p(\mathbb{R})),$$

[2] S. S. Vallender, *Calculation of the Wasserstein distance between probability distributions on the line,* Theory Probab. Appl. 18 (1973), 784–786.

[3] C. Villani, *Topics in optimal transportation,* Graduate studies in Mathematics vol. 58, American Mathemtical Society, Providence, RI, 2003.

The discrete case

Theorem (Gehér, Titkos, V., *J. Math. Anal. Appl.* **480** (2019), 123435)

For $p \in (0, \infty)$ and $f : \mathcal{W}_p(\mathcal{X}) \to \mathcal{W}_p(\mathcal{X})$ an isometric embedding, there exists a unique family of measures

$$\Phi := \left(\varphi_{x,t}\right)_{x \in \mathcal{X}, t \in (0,1]} \in \mathcal{M}(\mathcal{X})^{\mathcal{X} \times (0,1]}$$

that satisfies the properties

(a) *for all $x \neq y$: $S_{\varphi_{x,1}} \cap S_{\varphi_{y,1}} = \emptyset$,*

(b) *for all $x \in \mathcal{X}$ and $0 < t \leq 1$: $\varphi_{x,t}(\mathcal{X}) = t$,*

(c) *for all $x \in \mathcal{X}$ and $0 < s < t \leq 1$: $\varphi_{x,s} \leq \varphi_{x,t}$,*

and that generates f in the sense that $f(\mu) = \sum_{x \in S_\mu} \varphi_{x, \mu(\{x\})}$.
Conversely, every $\mathcal{X} \times (0, 1]$-indexed family of measures satisfying (a)–(c) *generates an isometric embedding.*

The discrete case

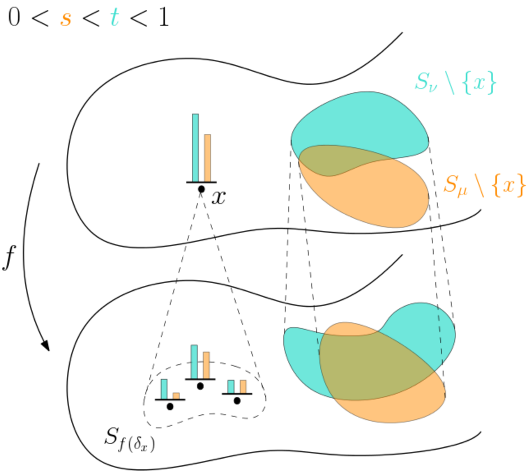


Figure: $f(\mu)|_{S_{f(\delta_x)}} \leq f(\nu)|_{S_{f(\delta_x)}}$.

Unit interval, real line, and Euclidean spaces

Theorem (Gehér, Titkos, V., *Trans. Amer. Math. Soc.* (2020) & G-T-V, arXiv:2102.02037 (2021))

For $0 < p < 1$, *the Wasserstein space* $\mathcal{W}_p(X)$ *is isometrically rigid* **for every Polish space** X. *Moreover, the isometry group of* $\mathcal{W}_p(X)$ *depending on* $X \in \{[0,1], \mathbb{R}, E\}$ *and* $p \geq 1$ *is isomorphic to:*

$\mathrm{Isom}\left(\mathcal{W}_p(X)\right)$	$X = [0,1]$	$X = \mathbb{R}$	$X = E$
$p = 1$	$C_2 \times C_2$	$\mathrm{Isom}(\mathbb{R})$	$\mathrm{Isom}(E)$
$p > 1,\ p \neq 2$	C_2	$\mathrm{Isom}(\mathbb{R})$	$\mathrm{Isom}(E)$
$p = 2$	C_2	$\mathrm{Isom}(\mathbb{R}) \ltimes \mathrm{Isom}(\mathbb{R})$	$\mathrm{Isom}(E) \ltimes O(E)$

E stands for a separable real Hilbert space. Color code: Kloeckner (2010), G-T-V (2020-21), Kloeckner: $\dim(E) < \infty$, G-T-V: $\dim(E) = \infty$.

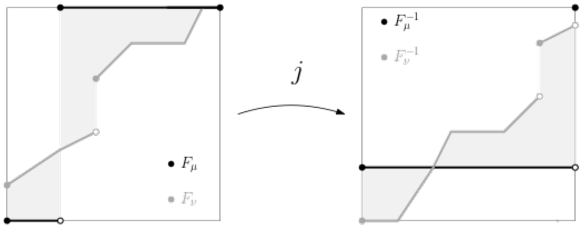
The unit interval

Theorem (G-T-V, TAMS (2020))

Every isometric embedding of $\mathcal{W}_1([0,1])$ *is surjective, and*

$$\mathrm{Isom}(\mathcal{W}_1([0,1])) = \{\mathrm{id}, r_\#, j, r_\# j\} \simeq C_2 \times C_2$$

(Klein group), where r *is the reflection of* $[0,1]$, *and* j *is the flip operation defined by* $F_{j(\mu)} = F_\mu^{-1}$



The action of isometries of $\mathcal{W}_1([0,1])$

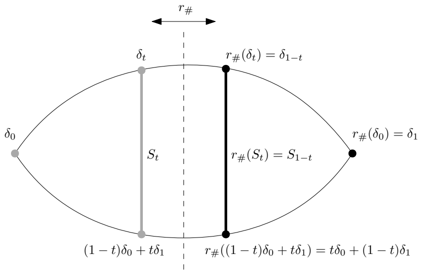
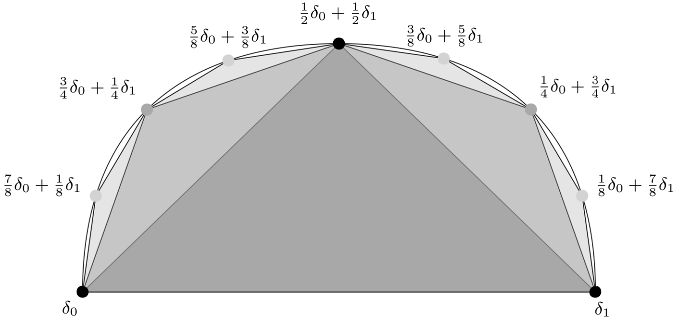


Figure: The action of $r_\#$ and j on $\mathcal{W}_1([0,1])$

The shape of $\mathcal{W}_2([0,1])$



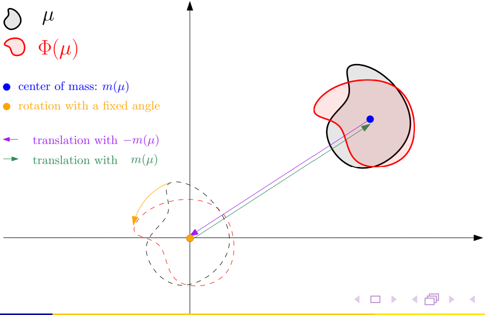
Wasserstein-Hilbert spaces (G-T-V, arXiv:2102.02037)

(a) If $p \neq 2$, then Φ is necessarily a push-forward of an isometry ψ of E, that is, $\Phi(\mu) = \psi_{\#}\mu$.

(b) If $p = 2$ and E is infinite dimensional then

$$\Phi(\mu) = \left(\psi \circ t_{m(\mu)} \circ R \circ t_{m(\mu)}^{-1}\right)_{\#} \mu \qquad (\mu \in \mathcal{W}_2(E)),$$

where $\psi \colon E \to E$ is an affine, $R \colon E \to E$ is a linear isometry, and $t_{m(\mu)} \colon E \to E$ is the translation on E by the barycenter $m(\mu)$ of μ



Wasserstein-Hilbert spaces (G-T-V, arXiv:2102.02037)

- main tool: the Wasserstein potential defined by

$$\mathcal{T}_\mu^p \colon E \to \mathbb{R}, \quad x \mapsto d_{\mathcal{W}_p}^p(\mu, \delta_x) = \int_E \|x-y\|^p \,\mathrm{d}\mu(y)$$

  from which *atoms* can be recovered if $1 \leq p < \infty$, $2 \nmid p$ by

$$\lim_{h\to 0} \frac{\sum_{j=0}^{2k} \binom{2k}{j}(-1)^j \mathcal{T}_\mu^p\left(x+(k-j)h\right)}{\left(\sum_{j=0}^{2k} \binom{2k}{j}(-1)^j |k-j|^p\right)\|h\|^p} = \mu(\{x\})$$

- for $2 \mid p$ the potential *does not contain enough information about the measure*, still, we proved rigidity by different techniques

Future plan

Describe the isometric structure of p-Wasserstein spaces over:

- graphs with the shortest path distance
- Banach spaces, length spaces, non-branching spaces: no inner product structure $\Rightarrow$ genuinely new methods are needed
- spheres, tori: Fourier method combined with the Wasserstein potential technique
- projective spaces: represent the sets of pure states in quantum mechanics

Thank you for your kind attention!