

ISOMETRIC RIGIDITY OF WASSERSTEIN SPACES: THE GRAPH METRIC CASE

GERGELY KISS AND TAMÁS TITKOS

ABSTRACT. The aim of this paper is to prove that the p -Wasserstein space $\mathcal{W}_p(X)$ is isometrically rigid for all $p \geq 1$ whenever X is a countable graph metric space. As a consequence, we obtain that for every countable group H and any $p \geq 1$ there exists a p -Wasserstein space whose isometry group is isomorphic to H .

1. INTRODUCTION

Due to its deep impact on both pure and applied sciences, one of the most intensively studied metric spaces nowadays is the so-called p -Wasserstein space $\mathcal{W}_p(X)$: the collection of Borel probability measures on a complete separable metric space (X, ϱ) with finite p -th moment, endowed with a transport related metric d_p , which is calculated by means of optimal couplings and the p -th power of the underlying distance ϱ . We mention here only three comprehensive textbooks [1, 22, 24], more references and precise definitions will follow later. In this paper we consider those p -Wasserstein spaces whose underlying metric space (X, ϱ) is a graph metric space. This class contains many important metric spaces, just to mention a few: any countable set with the discrete metric; the set of natural numbers and the set of integers with the usual $|\cdot|$ -distance; d -dimensional lattices endowed with the l_1 -metric or the l_∞ -metric (for the relevance of these metrics in pattern recognition see e.g. [21]); finite strings with the Hamming distance (as it was mentioned in [5] in connection with the quantum 1-Wasserstein distance, the classical 1-Wasserstein distance with respect to the Hamming metric is called Ornstein's distance, and was first considered in [20]); and infinite regular trees (see the very recent manuscript [9]).

When working with a structure, the most fundamental and natural task is to explore its transformations and symmetries. In the case of metric spaces, such symmetries are *isometries*, that is, distance preserving bijections. In the recent past, many authors investigated isometries of various important metric spaces of probability measures [2, 6, 7, 11–14, 16, 19, 23, 25]. To mention a few, in [19] Molnár explored the structure of isometries of the space of distribution functions with respect to the Lévy distance. Later Gehér and the second author generalised his result to the Lévy-Prokhorov metric in [11]. Namely, it was shown that if the space $\mathcal{P}(X)$ of all Borel probability measures on a real Banach space $(X, \|\cdot\|)$ is endowed with the Lévy-Prokhorov metric d_{LP} , then the isometry group of $(\mathcal{P}(X), d_{LP})$ is isomorphic to the isometry group of the underlying space X . Bertrand and Kloeckner showed that a similar phenomenon occurs when one considers a 2-Wasserstein space built on a negatively curved metric space: each isometry

2020 *Mathematics Subject Classification*. Primary: 54E40; 46E27 Secondary: 54E70; 05C12.

Key words and phrases. Wasserstein space, graph metric space, isometry, isometric rigidity.

Corresponding author: Tamás Titkos, titkos.tamas@renyi.hu.

G. Kiss was supported by Premium Postdoctoral Fellowship of the Hungarian Academy of Sciences and by the Hungarian National Research, Development and Innovation Office - NKFIH (grant no. K124749). T. Titkos was supported by the Hungarian National Research, Development and Innovation Office - NKFIH (grant no. PD128374 and grant no. K134944), by the János Bolyai Research Scholarship and the Momentum Program No. LP2021-15/202 of the Hungarian Academy of Sciences, and by the ÚNKP-20-5-BGE-1 New National Excellence Program of the Ministry of Innovation and Technology.

of the space of measures is the push-forward of an isometry of the underlying space. This phenomenon is called *isometric rigidity*. Finally, we mention a very recent rigidity result, Santos-Rodríguez proved that p -Wasserstein spaces built on compact rank one symmetric spaces are all isometrically rigid [23] if $p > 1$.

Our main result is **Theorem 6**, where we prove that p -Wasserstein spaces over graph metric spaces are all isometrically rigid. As a consequence, in **Corollary 7** we will conclude that for every countable group G and for every $p \geq 1$ there exists a p -Wasserstein space whose isometry group is isomorphic to G .

Before going into the details, we make some comments on rigidity. It comes easy to say that these rigidity results are not surprising because of the intimate connection between d_p and ϱ . It is well known that if $p \geq 1$ then the distance between any two Dirac measures equals to the distance of their supporting points, and every measure can be approximated by convex combinations of Dirac measures. In other words, $\mathcal{W}_p(X)$ contains an isometric copy of X , and the convex hull of this copy is dense in $\mathcal{W}_p(X)$. Moreover, $\mathcal{W}_p(X)$ inherits many nice properties of X , e.g. completeness, compactness, existence of geodesics. So one may have the impression that although the p -Wasserstein space $\mathcal{W}_p(X)$ is much bigger than X (see e.g. [17] for many interesting results), the strong connection between the metrics does not allow $\mathcal{W}_p(X)$ to have more symmetries than X has. A possible sketch of proof is this:

- Step 1. Prove that an isometry $\Phi : \mathcal{W}_p(X) \rightarrow \mathcal{W}_p(X)$ leaves the set of Dirac masses invariant. Once it is done, one can conclude that the action on Dirac masses is generated by an isometry $f : X \rightarrow X$, that is, $\Phi(\delta_x) = \delta_{f(x)}$ for all $x \in X$.
- Step 2. Prove that this action extends to a set \mathcal{S} of finitely supported probability measures, that is, $\Phi\left(\sum_{j \in J} \lambda_j \delta_{x_j}\right) = \sum_{j \in J} \lambda_j \delta_{f(x_j)}$ for all $\sum_{j \in J} \lambda_j \delta_{x_j} \in \mathcal{S}$.
- Step 3. Show that \mathcal{S} is dense in $\mathcal{W}_p(X)$. Since Φ is continuous, Φ must be the push-forward of f^{-1} , where f is the above defined isometry.

The problem with this sketch is that it does not work in general. And even if it works, these seemingly easy steps can be nontrivial. For example, Step 1 fails if $p = 1$ and $X = [0, 1]$. In that case, there exists an isometry j (called *flip*) which is mass-splitting, i.e. which does not leave the set of Dirac masses invariant: $j(\delta_t) = t\delta_0 + (1-t)\delta_1$ for all $t \in [0, 1]$, for more details see [13, Section 2.1]. Step 2 can fail (even if Step 1 can be done) as it was shown by Kloeckner in [16]: if $p = 2$ and $X = \mathbb{R}$, then there exists a flow of strangely behaving isometries. These isometries leave all Dirac masses fixed, but they differ from the identity of $\mathcal{W}_2(\mathbb{R})$, for more details see [16, Section 5.1]. We mention that all these strange isometries disappear once we modify the value of p : it was proved in [13] that if $p \neq 1$, then $\mathcal{W}_p([0, 1])$ is isometrically rigid, and similarly, if $p \neq 2$, then $\mathcal{W}_p(\mathbb{R})$ is isometrically rigid. Furthermore, Gehér et al. showed in [14, Section 2] that for every $p \geq 1$ there exists a compact metric space X such that $\mathcal{W}_p(X)$ admits mass-splitting isometries.

Summarising the above comments, we can say that isometric rigidity of p -Wasserstein spaces is a quite regular phenomenon (only a few non-rigid example is known), but there is no general recipe which helps to decide whether a space is rigid or not.

Acknowledgements: We would like to thank the referee for the careful reading of the manuscript and the constructive comments that helped us to improve the presentation.

2. TECHNICAL PRELIMINARIES

First we fix the terminology. Given a metric space (Y, d) , we call a bijection $f : Y \rightarrow Y$ an *isometry* if $d(f(x), f(y)) = d(x, y)$ holds for all $x, y \in Y$. The *isometry group* of (Y, d) will be denoted by $\text{Isom}(Y, d)$. For two groups H_1, H_2 the symbol $H_1 \cong H_2$ means that they are isomorphic.

In this paper $G(X, E)$ always denotes a graph with a countable vertex set X and edge set E . Two different vertices $x, y \in X$ are called adjacent if there exists an edge $e \in E$ which joins them. A *path* between two vertices $x, y \in X$ is a finite sequence of distinct edges which joins a sequence $x_0 = x, x_1, \dots, x_k = y$ of distinct adjacent vertices. The length of such a path is the number of edges that the path contains. The graph $G(X, E)$ is called connected if every two different vertices can be connected by a path. The vertex set X of a connected graph can always be endowed with a metric: for $x, y \in X$ we say that the *shortest path distance* of x and y is the minimum number of k such that there exists a path between x and y of length k . Now we can define the central notion of this paper: we say that a countable metric space (X, ϱ) is a *graph metric space* if there exists a connected graph $G(X, E)$, such that the shortest path distance in the graph coincides with the distance ϱ in X . That is, $\varrho(x, y)$ equals to the minimum number of edges in a path in $G(X, E)$ between x and y . Since the existence of loops and multiple edges do not change the length of the shortest path, one can assume that the graph in question is simple.¹

Now we recall the notion of a p -Wasserstein space in the special case when the underlying metric space is a countable graph metric space (X, ϱ) . The symbol $\mathcal{M}_+(X)$ stands for the set of nonnegative Borel measures on X . In this setting each measure $\mu \in \mathcal{M}_+(X)$ is uniquely determined by its value on singletons:

$$(1) \quad \mu(A) = \sum_{x \in A} \mu(\{x\}) \quad \text{for all } A \subseteq X,$$

and therefore μ can be handled as a one-variable function on X . Such a function is often referred to as a *probability mass function*. For the sake of simplicity, we will write shortly $\mu(x)$ instead of $\mu(\{x\})$. For a given real number $p \geq 1$ we denote by $\mathcal{W}_p(X)$ the set of all probability measures such that

$$(2) \quad \sum_{x \in X} \varrho^p(x, \hat{x}) \mu(x) < \infty$$

for some (hence all) $\hat{x} \in X$. The *support* μ of a $\mu \in \mathcal{W}_p(X)$ in this setting equals to the set $\{x \in X \mid \mu(x) \neq 0\}$. A Borel probability measure π on $X \times X$ is said to be a *coupling* for μ and ν if the marginals of π are μ and ν , that is,

$$(3) \quad \sum_{x' \in X} \pi(x, x') = \mu(x) \quad \text{and} \quad \sum_{x \in X} \pi(x, x') = \nu(x').$$

The set of all couplings (which is never empty because the product measure is a coupling) is denoted by $\Pi(\mu, \nu)$. We will refer to couplings as transport plans, as $\pi(x, x')$ is the weight of mass that is transported from x to x' while μ is transported to ν along π . For a given measure μ we will denote by $\pi_\mu \in \Pi(\mu, \mu)$ the coupling which leaves μ undisturbed, that is, $\pi_\mu(x, x) := \mu(x)$ for all $x \in X$ and $\pi_\mu(x, y) := 0$ otherwise.

¹The following characterization of graph metric spaces was proved in [3]: a countable metric space (X, ϱ) is a graph metric space if and only if the distance between every two points of X is an integer, and if $a, b \in X$ and $\varrho(a, b) \geq 2$ then there exists a point $x \in M$ such that $\varrho(a, x) > 0$, $\varrho(x, b) > 0$, and x saturates the triangle inequality: $\varrho(a, b) = \varrho(a, x) + \varrho(x, b)$. It was assumed in [3] that the graph is finite, but the proof works in the countable case as well.

If the cost function on $X \times X$ is ϱ^p , then the optimal cost of transporting μ into ν is

$$(4) \quad d_p(\mu, \nu) := \left(\inf_{\pi \in \Pi(\mu, \nu)} \sum_{(x, x') \in X \times X} \varrho^p(x, x') \cdot \pi(x, x') \right)^{1/p}.$$

It is known (see e.g. Theorem 1.5 in [1] with $c = \varrho^p$) that the infimum in (4) is in fact a minimum. Those transport plans that minimise the transport cost are called *optimal transport plans*. We will refer to the metric space $(\mathcal{W}_p(X), d_p)$ as the p -Wasserstein space $\mathcal{W}_p(X)$. Let us denote the set of all *finitely supported probability measures* by $\mathcal{F}(X)$. A very important feature of p -Wasserstein spaces is that if $p \geq 1$, then X embeds into $\mathcal{W}_p(X)$ isometrically (that is, $d_p(\delta_x, \delta_y) = \varrho(x, y)$ for all $x, y \in X$) and that $\mathcal{F}(X)$ is dense in $\mathcal{W}_p(X)$ (see e.g. Example 6.3 and Theorem 6.16 in [24]). Although it is known that the isometry group of X embeds into the isometry group of $\mathcal{W}_p(X)$, we provide with a short proof here for the sake of completeness.

Proposition 1. Let (X, ϱ) be countable graph metric space and let $p \geq 1$ be fixed. Then the push-forward $\psi \mapsto \psi_{\#}$ defined by

$$(5) \quad (\psi_{\#}(\mu))(x) := \mu(\psi^{-1}(x)) \quad (x \in X)$$

induces an embedding, which is in fact a group homomorphism

$$(6) \quad \# : \text{Isom}(X, \varrho) \rightarrow \text{Isom}(\mathcal{W}_p(X), d_p).$$

Proof. Let us fix a $\psi \in \text{Isom}(X, \varrho)$, two measures $\mu, \nu \in \mathcal{W}_p(X)$, and an optimal transport plan $\tilde{\pi} \in \Pi(\mu, \nu)$. Since $\text{Isom}(X, \varrho)$ is a group, ψ^{-1} is a bijection such that $\varrho(\psi(x), \psi(y)) = \varrho(x, y) = \varrho(\psi^{-1}(x), \psi^{-1}(y))$ for all $x, y \in X$. Furthermore, $\pi'(x, y) := \tilde{\pi}(\psi^{-1}(x), \psi^{-1}(y))$ is a coupling for $\psi_{\#}(\mu)$ and $\psi_{\#}(\nu)$, because $\sum_{y \in X} \tilde{\pi}(\psi^{-1}(x), \psi^{-1}(y)) = \mu(\psi^{-1}(x)) = \psi_{\#}(\mu)(x)$ and $\sum_{x \in X} \tilde{\pi}(\psi^{-1}(x), \psi^{-1}(y)) = \nu(\psi^{-1}(y)) = \psi_{\#}(\nu)(y)$. Using that the p -Wasserstein distance is always smaller or equal to the cost of any coupling, the above observation implies

$$(7) \quad \begin{aligned} d_p^p(\mu, \nu) &= \sum_{(x, y) \in X \times X} \varrho^p(x, y) \cdot \tilde{\pi}(x, y) \\ &= \sum_{(x, y) \in X \times X} \varrho^p(\psi^{-1}(x), \psi^{-1}(y)) \cdot \tilde{\pi}(\psi^{-1}(x), \psi^{-1}(y)) \\ &= \sum_{(x, y) \in X \times X} \varrho^p(x, y) \cdot \tilde{\pi}(\psi^{-1}(x), \psi^{-1}(y)) \geq d_p^p(\psi_{\#}(\mu), \psi_{\#}(\nu)). \end{aligned}$$

The reverse inequality $d_p^p(\psi_{\#}(\mu), \psi_{\#}(\nu)) \geq d_p^p(\mu, \nu)$ can be proved along the same lines by using an optimal coupling $\hat{\pi} \in \Pi(\psi_{\#}(\mu), \psi_{\#}(\nu))$ and the observation that π'' defined by $\pi''(x, y) := \hat{\pi}(\psi(x), \psi(y))$ is a coupling for μ and ν .

Finally, we verify that the map $\psi \mapsto \psi_{\#}$ is indeed a group homomorphism. For all $\psi, \chi \in \text{Isom}(X, \varrho)$ and for all $\mu \in \mathcal{W}_p(X)$ and $x \in X$ we have

$$(8) \quad \begin{aligned} \left((\psi \circ \chi)_{\#}(\mu) \right)(x) &= \mu \left((\psi \circ \chi)^{-1}(x) \right) = \mu \left(\chi^{-1}(\psi^{-1}(x)) \right) \\ &= \left(\chi_{\#}(\mu) \right)(\psi^{-1}(x)) = \left(\psi_{\#}(\chi_{\#}(\mu)) \right)(x). \end{aligned}$$

□

Isometries of the form $\psi_{\#}$ are called *trivial isometries*. We say that $\mathcal{W}_p(X)$ is *isometrically rigid* if the map $\#$ is onto, i.e. $\text{Isom}(X, \varrho) \cong \text{Isom}(\mathcal{W}_p(X), d_p)$. In other words, if every isometry of $\mathcal{W}_p(X)$ is trivial.

In order to prove isometric rigidity, it would be useful to find properties which can be characterized by means of the metric, and thus are preserved by isometries. As we will see later in Proposition 4, the neighbouring property – which says that two probability mass functions differ at the end-points of a given edge and nowhere else (see Figure 1 below) – is one of such properties.

Definition 2. For a given $\alpha \in (0, 1]$ we say that two measures $\mu, \nu \in \mathcal{W}_p(X)$ are α -neighbouring, if there exists an $\eta \in \mathcal{M}_+(X)$ and $u, v \in X$ with $\varrho(u, v) = 1$ such that $\mu = \eta + \alpha\delta_u$ and $\nu = \eta + \alpha\delta_v$. In symbols, we write $\mu \equiv \nu [\alpha]$.

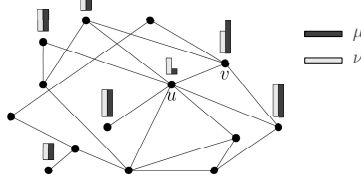


FIGURE 1. An illustration of α -neighbouring property.

In particular, $\mu \equiv \nu [1]$ means that $\mu = \delta_u$ and $\nu = \delta_v$, where u and v are different endpoints of an edge in the underlying graph. We note that \equiv is not an equivalence relation, as it is not reflexive and not transitive. In order to see the metric side of the neighbouring property, let us introduce the set $B_s(\mu, \nu)$ for $\mu, \nu \in \mathcal{W}_p(X)$

$$(9) \quad B_s(\mu, \nu) := \left\{ \xi \in \mathcal{W}_p(X) \mid d_p(\mu, \xi) \leq \sqrt[p]{s} d_p(\mu, \nu), \quad d_p(\xi, \nu) \leq \sqrt[p]{1-s} d_p(\mu, \nu) \right\}.$$

The following proposition says that $B_s(\mu, \nu)$ is always non-empty. For the sake of brevity, we will use the notation in the sequel $\xi_s^{\mu, \nu} := (1-s)\mu + s\nu$.

Proposition 3. Let (X, ϱ) be a countable graph metric space and let $p \geq 1$ and $s \in (0, 1)$ be fixed real numbers. Then for any $\mu, \nu \in \mathcal{W}_p(X)$ the measure $\xi_s^{\mu, \nu}$ belongs to $B_s(\mu, \nu)$. In particular, the statement that $B_s(\mu, \nu)$ is a singleton is equivalent to $B_s(\mu, \nu) = \{\xi_s^{\mu, \nu}\}$.

Proof. First we show that $\xi_s^{\mu, \nu}$ satisfies $d_p(\mu, \xi_s^{\mu, \nu}) \leq \sqrt[p]{s} d_p(\mu, \nu)$. Let us fix an optimal transport plan $\tilde{\pi} \in \Pi(\mu, \nu)$. Recall that π_μ is a coupling which leaves μ undisturbed: $\pi_\mu(x, x) = \mu(x)$ and $\pi_\mu(x, y) = 0$ otherwise. Then $\pi_s := (1-s)\pi_\mu + s\tilde{\pi} \in \Pi(\mu, \xi_s^{\mu, \nu})$. Indeed, $\sum_{y \in X} \left((1-s)\pi_\mu(x, y) + s\tilde{\pi}(x, y) \right) = (1-s)\mu(x) + s\mu(x) = \mu(x)$ for all $x \in X$ and

$\sum_{x \in X} \left((1-s)\pi_\mu(x, y) + s\tilde{\pi}(x, y) \right) = (1-s)\mu(y) + s\nu(y) = \xi_s^{\mu, \nu}(y)$ for all $y \in X$. Using the transport plan $\pi_s \in \Pi(\mu, \xi_s^{\mu, \nu})$ we can estimate $d_p(\mu, \xi_s^{\mu, \nu})$ as

$$(10) \quad \begin{aligned} d_p^p(\mu, \xi_s^{\mu, \nu}) &= \inf_{\pi \in \Pi(\mu, \xi_s^{\mu, \nu})} \sum_{(x, y) \in X \times X} \varrho^p(x, y) \cdot \pi(x, y) \leq \sum_{(x, y) \in X \times X} \varrho^p(x, y) \cdot \pi_s(x, y) \\ &= (1-s) \sum_{(x, y) \in X \times X} \varrho^p(x, y) \cdot \pi_\mu(x, y) + s \sum_{(x, y) \in X \times X} \varrho^p(x, y) \cdot \tilde{\pi}(x, y) \\ &= s d_p^p(\mu, \nu), \end{aligned}$$

where we used $\tilde{\pi}$ is optimal and that $\varrho^p(x, y) \cdot \pi_\mu(x, y) = 0$ for all $(x, y) \in X \times X$. The other inequality $d_p(\xi_s^{\mu, \nu}, \nu) \leq \sqrt[p]{1-s} d_p(\mu, \nu)$ can be proved in the same way using a similar combination of $\tilde{\pi}$ and π_ν (which leaves ν undisturbed). \square

In Proposition 3 we saw that $\xi_s^{\mu, \nu} \in B_s(\mu, \nu)$ for all $\mu, \nu \in \mathcal{W}_p(X)$. Our next aim is to find a metric characterization for those pairs such that $B_s(\mu, \nu) = \{\xi_s^{\mu, \nu}\}$.

Proposition 4. Let (X, ϱ) be a countable graph metric space and let $p \geq 1$ and $\alpha \in (0, 1]$ be fixed real numbers. Then the following statements are equivalent:

- (i) μ and ν are α -neighbouring, that is, there exists an $\eta \in \mathcal{M}_+(X)$ and $u, v \in X$ with $\varrho(u, v) = 1$ such that

$$(11) \quad \mu = \eta + \alpha\delta_u \quad \text{and} \quad \nu = \eta + \alpha\delta_v.$$

- (ii) $d_p(\mu, \nu) = \sqrt[p]{\alpha}$ and $B_s(\mu, \nu) = \{\xi_s^{\mu, \nu}\}$ for all $s \in (0, 1)$.

- (iii) $d_p(\mu, \nu) = \sqrt[p]{\alpha}$ and $B_{\frac{1}{2}}(\mu, \nu) = \{\xi_{\frac{1}{2}}^{\mu, \nu}\}$.

Proof. (i) \implies (ii): First we show that $d_p(\mu, \nu) \geq \sqrt[p]{|\nu(\hat{x}) - \mu(\hat{x})|}$ holds for all $\mu, \nu \in \mathcal{W}_p(X)$ and $\hat{x} \in X$. By symmetry, we can assume without loss of generality that $\mu(\hat{x}) \leq \nu(\hat{x})$. Then for any coupling $\pi \in \Pi(\mu, \nu)$ we have

$$(12) \quad \begin{aligned} \nu(\hat{x}) &= \sum_{x \in X} \pi(x, \hat{x}) = \pi(\hat{x}, \hat{x}) + \sum_{\substack{x \in X \\ x \neq \hat{x}}} \pi(x, \hat{x}) \\ &\leq \sum_{y \in X} \pi(\hat{x}, y) + \sum_{\substack{x \in X \\ x \neq \hat{x}}} \pi(x, \hat{x}) = \mu(\hat{x}) + \sum_{\substack{x \in X \\ x \neq \hat{x}}} \pi(x, \hat{x}). \end{aligned}$$

which implies

$$(13) \quad \nu(\hat{x}) - \mu(\hat{x}) \leq \sum_{\substack{x \in X \\ x \neq \hat{x}}} \pi(x, \hat{x}).$$

Since $x \neq y$ implies $\varrho^p(x, y) \geq 1$, we have the following lower bound for the cost of π

$$(14) \quad \nu(\hat{x}) - \mu(\hat{x}) \leq \sum_{\substack{x \in X \\ x \neq \hat{x}}} \pi(x, \hat{x}) \leq \sum_{x \in X} \varrho^p(x, \hat{x}) \cdot \pi(x, \hat{x}) \leq \sum_{(x, y) \in X \times X} \varrho^p(x, y) \cdot \pi(x, y).$$

By taking the minimum over $\Pi(\mu, \nu)$, one gets $\nu(\hat{x}) - \mu(\hat{x}) \leq d_p^p(\mu, \nu)$. In (11) we have $\nu(v) - \mu(v) = \alpha$, and therefore $d_p(\mu, \nu) \geq \sqrt[p]{\alpha}$ holds. To see the reverse inequality, observe first that the following $\pi \in \Pi(\mu, \nu)$ has cost α

$$\pi(x, y) = \begin{cases} \mu(x) & \text{if } x = y \text{ and } x \neq u, \\ \mu(x) - \alpha & \text{if } x = y = u, \\ \alpha & \text{if } x = u \text{ and } y = v, \\ 0 & \text{otherwise.} \end{cases}$$

And therefore, $d_p(\mu, \nu) \leq \sqrt[p]{\alpha}$. Now assume that $\xi \in B_s(\mu, \nu)$. We have to show that $\xi = \xi_s^{\delta_u, \delta_v} = \eta + (1-s)\alpha\delta_u + s\alpha\delta_v$. Let $\pi^* \in \Pi(\mu, \xi)$ be an optimal transport plan, i.e.

$$(15) \quad \alpha s \geq d_p^p(\mu, \xi) = \sum_{(x, y) \in X \times X} \varrho^p(x, y) \cdot \pi^*(x, y).$$

By adding $\pi^*(u, u) - \pi^*(u, u) = 0$, the right hand side in (15) can be written as

$$(16) \quad \sum_{\substack{(x, y) \in X \times X \\ x \neq u}} \varrho^p(x, y) \cdot \pi^*(x, y) + \left[\sum_{y \in X} \varrho^p(u, y) \cdot \pi^*(u, y) + \pi^*(u, u) \right] - \pi^*(u, u),$$

which gives

$$(17) \quad \alpha s \geq \mu(u) - \pi^*(u, u),$$

because

$$(18) \quad \sum_{\substack{(x,y) \in X \times X \\ x \neq y}} \varrho^p(x,y) \cdot \pi^*(x,y) \geq 0,$$

and if $u \neq y$, then $\varrho(u,y) \geq 1$, and thus

$$(19) \quad \sum_{y \in X} \varrho^p(u,y) \cdot \pi^*(u,y) + \pi^*(u,u) \geq \sum_{y \in X} \pi^*(u,y) = \mu(u).$$

Since $\mu(u) = \eta(u) + \alpha$ according to (11), we can rearrange (17) as

$$(20) \quad \pi^*(u,u) \geq \alpha(1-s) + \eta(u).$$

Using that $\pi^* \in \Pi(\mu, \xi)$ we get $\xi(u) = \sum_{x \in X} \pi^*(x,u) \geq \pi^*(u,u)$, and thus

$$(21) \quad \xi(u) \geq \pi^*(u,u) \geq \eta(u) + (1-s)\alpha\delta_u(u).$$

Combination of (15) and (21) asserts now that

$$(22) \quad s\alpha \geq d_p^p(\mu, \xi) \geq |\xi(u) - \mu(u)| \geq |\eta(u) + (1-s)\alpha - (\eta(u) + \alpha\delta_u(u))| = s\alpha.$$

Furthermore, a very similar calculation with ξ and ν gives

$$(23) \quad (1-s)\alpha \geq d_p^p(\nu, \xi) \geq |\xi(v) - \nu(v)| \geq |\eta(v) + s\alpha - (\eta(v) + \alpha\delta_v(v))| = (1-s)\alpha.$$

Since every inequality in (22) and (23) is actually an equality, we get

$$(24) \quad \xi(u) = \eta(u) + (1-s)\alpha\delta_u(u) \quad \text{and} \quad \xi(v) = \eta(v) + s\alpha\delta_v(v)$$

which together with $d_p^p(\mu, \xi) = |\mu(u) - \xi(u)|$ and $d_p^p(\xi, \nu) = |\xi(v) - \nu(v)|$ imply that

$$(25) \quad \xi = \eta + (1-s)\alpha\delta_u + s\alpha\delta_v = \xi_s^{\delta_u, \delta_v}.$$

This proves that (i) \implies (ii).

(ii) \implies (iii): This implication is straightforward.

(iii) \implies (i): We have to show that if $d_p(\mu, \nu) = \sqrt[p]{\alpha}$ and $B_{\frac{1}{2}}(\mu, \nu) = \{\xi_{\frac{1}{2}}^{\mu, \nu}\}$, then (11) holds. Let $\pi \in \Pi(\mu, \nu)$ be an optimal transport plan

$$(26) \quad \alpha = d_p^p(\mu, \nu) = \sum_{\substack{(x,y) \in \text{supp}(\pi) \\ x \neq y}} \varrho^p(x,y) \cdot \pi(x,y).$$

and recall that $\pi_{\frac{1}{2}} := \frac{1}{2}\pi_\mu + \frac{1}{2}\pi \in \Pi(\mu, \xi_{\frac{1}{2}}^{\mu, \nu})$. First assume indirectly that there exists an $(x', y') \in \text{supp}(\pi)$ for which $k := \varrho(x', y') > 1$. Let us choose a path of length k between x' and y' along the vertices $x_0 = x', x_1, \dots, x_{k-1}, x_k = y'$. Set $c = \frac{\pi(x', y')}{4}$ and modify $\xi_{\frac{1}{2}}^{\mu, \nu}$ along this path as

$$(27) \quad \xi := \xi_{\frac{1}{2}}^{\mu, \nu} - c\delta_{x_0} + c\delta_{x_1} + c\delta_{x_{k-1}} - c\delta_{x_k}.$$

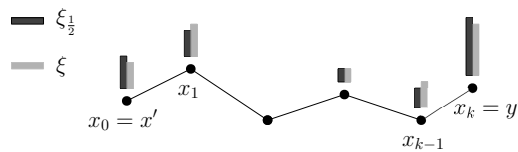


FIGURE 2. An illustration of how to modify $\xi_{\frac{1}{2}}^{\mu, \nu}$ in order to get ξ .

Since $k > 1$ and $p \geq 1$, we have $\xi \neq \xi_{\frac{1}{2}}^{\mu, \nu}$ and it follows from the construction that

$$(28) \quad d_p^p(\mu, \xi) \leq d_p^p(\mu, \xi_{\frac{1}{2}}^{\mu, \nu}) - k^p c + (k-1)^p c + 1^p c \leq d_p^p(\mu, \xi_{\frac{1}{2}}^{\mu, \nu}) \leq \frac{\alpha}{2}.$$

A similar calculation shows that $d_p^p(\xi, \nu) \leq \frac{\alpha}{2}$, and thus $\xi \in B_{\frac{1}{2}}(\mu, \nu)$, a contradiction. This contradiction means that $\varrho(x, y) = 1$ for all $(x, y) \in \text{supp}(\pi)$ in (26).

Now assume indirectly that $\text{supp}(\pi)$ has at least two different elements (x_1, y_1) and (x_2, y_2) . Set $\tilde{c} := \frac{\min\{\pi(x_1, y_1), \pi(x_2, y_2)\}}{4}$ and modify $\xi_{\frac{1}{2}}^{\mu, \nu}$ as follows

$$(29) \quad \xi := \xi_{\frac{1}{2}}^{\mu, \nu} + \tilde{c}\delta_{x_1} + \tilde{c}\delta_{y_1} - \tilde{c}\delta_{x_2} - \tilde{c}\delta_{y_2}.$$

Again, we have that $\xi \neq \xi_{\frac{1}{2}}^{\mu, \nu}$. In order to give an upper bound for $d_p^p(\mu, \xi)$, let us define the coupling $\tilde{\pi} \in \Pi(\mu, \xi)$ by modifying $\pi_{\frac{1}{2}} = \frac{1}{2}\pi_{\mu} + \frac{1}{2}\pi \in \Pi(\mu, \xi_{\frac{1}{2}}^{\mu, \nu})$ as follows

$$\tilde{\pi}(x, y) = \begin{cases} \pi_{\frac{1}{2}}(x_1, y_1) + \tilde{c} & \text{if } (x, y) = (x_1, y_1), \\ \pi_{\frac{1}{2}}(x_2, y_2) - \tilde{c} & \text{if } (x, y) = (x_2, y_2), \\ \pi_{\frac{1}{2}}(x, y) & \text{otherwise.} \end{cases}$$

Since $\varrho(x, y) = 1$ for all $(x, y) \in \text{supp}(\pi)$ and $\pi_{\mu}(x, y)\varrho(x, y) = 0$ for all $x, y \in X$, we have

$$(30) \quad d_p^p(\mu, \xi) \leq \sum_{\substack{(x, y) \in \text{supp}(\tilde{\pi}) \\ x \neq y}} \tilde{\pi}(x, y) = \sum_{\substack{(x, y) \in \text{supp}(\pi) \\ x \neq y}} \frac{\pi(x, y)}{2} = \frac{\alpha}{2}.$$

Similarly, $d_p^p(\xi, \nu) \leq \frac{\alpha}{2}$, and thus ξ belongs to $B_{\frac{1}{2}}(\mu, \nu)$, a contradiction. The only remaining possibility is that there exists $u, v \in X$ such that $\varrho(u, v) = 1$ and (26) can be written as $\alpha = d_p^p(\mu, \nu) = \varrho^p(u, v) \cdot \pi(u, v) = \pi(u, v)$, which means exactly that (11) holds. This proves (i) \iff (ii). \square

Corollary 5. Let (X, ϱ) be a countable graph metric space and let $p \geq 1$ and $\alpha \in (0, 1]$ be fixed real numbers. For any isometry $\Phi : \mathcal{W}_p(X) \rightarrow \mathcal{W}_p(X)$ and for any pair of measures $\mu, \nu \in \mathcal{W}_p(X)$ the following holds

$$(31) \quad \mu \equiv \nu [\alpha] \quad \iff \quad \Phi(\mu) \equiv \Phi(\nu) [\alpha].$$

Proof. Since Φ and Φ^{-1} are both distance preserving bijections, we have that $B_s(\mu, \nu)$ is a singleton if and only if $B_s(\Phi(\mu), \Phi(\nu))$ is a singleton. According to Proposition 4, this implies that $\mu \equiv \nu [\alpha]$ if and only if $\Phi(\mu) \equiv \Phi(\nu) [\alpha]$. \square

3. THE MAIN RESULT

Now we are ready to state and prove the main result of the paper.

Theorem 6. Let (X, ϱ) be a countable graph metric space and let $p \geq 1$ be fixed. Then the p -Wasserstein space $\mathcal{W}_p(X)$ is isometrically rigid, i.e., $\text{Isom}(\mathcal{W}_p(X), d_p) \cong \text{Isom}(X, \varrho)$.

Proof. We have seen in Proposition 1 that $\# : \text{Isom}(X, \varrho) \rightarrow \text{Isom}(\mathcal{W}_p(X), d_p)$ is a group homomorphism. Therefore, it is enough to prove that it is surjective, i.e., for any $\Phi \in \text{Isom}(\mathcal{W}_p(X), d_p)$ there exists a $\psi \in \text{Isom}(X, \varrho)$ such that $\Phi = \psi_{\#}$. The strategy of proof is similar to the sketch (Step 1–3.) mentioned in the introduction.

Step 1. First we prove that Φ maps the set of Dirac masses onto itself. Assume that $\mu = \delta_u$ is a Dirac measure and choose a $v \in X$ such that $\varrho(u, v) = 1$. Since $\delta_u \equiv \delta_v [1]$, we have $\Phi(\delta_u) \equiv \Phi(\delta_v) [1]$ according to Corollary 5. That is, $\Phi(\mu) = \delta_{\hat{u}}$ and $\Phi(\nu) = \delta_{\hat{v}}$ for some $\hat{u}, \hat{v} \in X$ with $\varrho(\hat{u}, \hat{v}) = 1$. In particular, $\Phi(\mu)$ is a Dirac measure. Since

$d_p(\delta_x, \delta_y) = \varrho(x, y)$ for all $x, y \in X$ and Φ^{-1} is an isometry as well, we see that Φ maps the set of Dirac masses bijectively onto itself, and the function $\psi : X \rightarrow X$ defined by

$$(32) \quad \Phi(\delta_x) := \delta_{\psi(x)} \quad (x \in X)$$

is an isometry. Let us consider the isometry $\tilde{\Phi} := \psi_{\#}^{-1} \circ \Phi$. On the one hand, $\tilde{\Phi}$ fixes all Dirac measures. On the other hand, if $\tilde{\Phi}(\mu) = \mu$ for all $\mu \in \mathcal{W}_p(X)$, then $\Phi = \psi_{\#}$. Therefore we can assume without loss of generality that $\psi(x) = x$, and hence $\Phi(\delta_x) = \delta_x$ for all $x \in X$.

Step 2. Our next task is to prove that Φ leaves a dense set of finitely supported measures fixed. We do this by induction with respect to the prescribed location of the supports in question. Since the underlying graph $G(X, E)$ is connected, starting with an arbitrary $x_1 \in X$ we can enumerate X in a way that for all $n \geq 2$ the vertex x_n is connected to the initial segment $X_{n-1} = \{x_1, \dots, x_{n-1}\}$ with at least one edge. If $\mu \in \mathcal{F}(X)$, then $\text{supp}(\mu) \subseteq X_N$ for some large enough $N \in \mathbb{N}$, and therefore it is enough to show that measures supported in X_n are fixed by Φ for all $n \in \mathbb{N}$.

If $n = 1$, then $\text{supp}(\mu) \subseteq X_1$ implies $\mu = \delta_{x_1}$, and $\Phi(\delta_{x_1}) = \delta_{x_1}$ according to Step 1.

If $n = 2$, then $\text{supp}(\mu) \subseteq X_2$ implies $\mu = (1-s)\delta_{x_1} + s\delta_{x_2}$ for some $s \in [0, 1]$. If $s = 0$ or $s = 1$ then μ is a Dirac measure and thus $\Phi(\mu) = \mu$. Assume now that $0 < s < 1$. In this case, $\mu = \xi_s^{\delta_{x_1}, \delta_{x_2}}$ and we know from Proposition 3 that $\xi_s^{\delta_{x_1}, \delta_{x_2}} \in B_s(\delta_{x_1}, \delta_{x_2})$. Moreover, $\Phi(\xi_s^{\delta_{x_1}, \delta_{x_2}}) \in B_s(\delta_{x_1}, \delta_{x_2})$ holds as well, because $\Phi(\delta_{x_i}) = \delta_{x_i}$ for $i = 1, 2$, and thus

$$(33) \quad d_p(\delta_{x_1}, \Phi(\mu)) = d_p(\Phi(\delta_{x_1}), \Phi(\mu)) = d_p(\delta_{x_1}, \mu) \leq \sqrt[p]{s} d_p(\delta_{x_1}, \delta_{x_2}),$$

and

$$(34) \quad d_p(\Phi(\xi_s^{\delta_{x_1}, \delta_{x_2}}), \delta_{x_2}) = d_p(\Phi(\xi_s^{\delta_{x_1}, \delta_{x_2}}), \Phi(\delta_{x_2})) = d_p(\xi_s^{\delta_{x_1}, \delta_{x_2}}, \delta_{x_2}) \leq \sqrt[p]{1-s} d_p(\delta_{x_1}, \delta_{x_2}).$$

But $\delta_{x_1} \equiv \delta_{x_2}$ [1], and thus $B_s(\delta_{x_1}, \delta_{x_2})$ is a singleton according to Proposition 4. This implies that $\Phi(\xi_s^{\delta_{x_1}, \delta_{x_2}}) = \xi_s^{\delta_{x_1}, \delta_{x_2}}$.

Assume now that $\Phi(\zeta) = \zeta$ holds whenever $\text{supp}(\zeta) \subseteq X_n = \{x_1, \dots, x_n\}$, and choose a finitely supported measure $\mu = \sum_{i=1}^{n+1} \mu(x_i) \delta_{x_i}$ which satisfies the following two properties:

$$(35) \quad \text{for all } x, y \in \text{supp}(\mu) : \quad x \neq y \text{ implies } \mu(x) \neq \mu(y),$$

and

$$(36) \quad \text{for all pairwise different elements } x, y, z \in \text{supp}(\mu) : \quad \mu(x) + \mu(y) \neq \mu(z).$$

If $\mu(x_{n+1}) = 0$, then $\text{supp}(\mu) \subseteq X_n$, and if $\mu(x_{n+1}) = 1$, then $\mu = \delta_{x_{n+1}}$. In both cases, $\Phi(\mu) = \mu$ according to the inductive hypothesis and Step 1. So it remains to deal with the case $0 < \mu(x_{n+1}) < 1$. According to the construction, there exists an $u \in X_n$ such that $\varrho(u, x_{n+1}) = 1$. Set $c := \mu(u) + \mu(x_{n+1})$ and define two (in a sense extremal) measures μ_* and μ^* as follows

$$(37) \quad \mu_* := \mu - \mu(u)\delta_u + \mu(u)\delta_{x_{n+1}} \quad \text{and} \quad \mu^* := \mu - \mu(x_{n+1})\delta_{x_{n+1}} + \mu(x_{n+1})\delta_u.$$

Observe that $\mu_* \equiv \mu^*$ [c], and $\text{supp}(\mu^*) \subseteq X_n$ implies $\Phi(\mu^*) = \mu^*$. Furthermore, we have

$$(38) \quad \mu_*(u) = \mu^*(x_{n+1}) = 0, \quad \mu_*(x_{n+1}) = \mu^*(u) = c$$

and

$$(39) \quad \mu_* \equiv \mu [\mu(u)], \quad \mu^* \equiv \mu [\mu(x_{n+1})].$$

Let us define a curve γ which connects μ^* and μ_* (see Figure 3 below)

$$(40) \quad \gamma : [0, c] \rightarrow \mathcal{W}_p(X); \quad \gamma(t) := \mu^* + t\delta_{x_{n+1}} - t\delta_u.$$

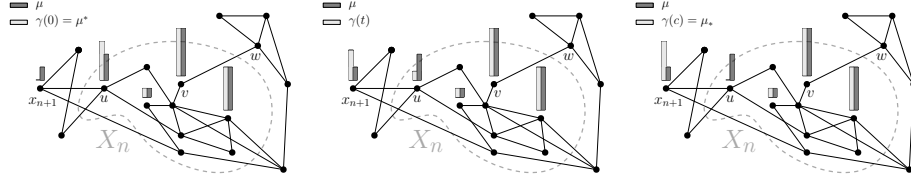


FIGURE 3. An illustration of teleporting μ^* to μ_* along the curve γ .

Observe that $\gamma(0) = \mu^*$, $\gamma(\mu(x_{n+1})) = \mu$, $\gamma(c) = \mu_*$, and that

$$(41) \quad \gamma(t) \equiv \mu [|t - \mu(x_{n+1})|] \quad \text{for all } t \in [0, c] \setminus \{x_{n+1}\}.$$

According to Corollary 5, $\mu_* \equiv \mu^* [c]$ implies $\Phi(\mu_*) \equiv \Phi(\mu^*) [c]$, and thus we can issue a curve Γ from $\Phi(\mu^*)$ to $\Phi(\mu_*)$ with a structure similar to that of γ . Recall that $\text{supp}(\mu^*) \subseteq X_n$ implies $\Phi(\mu^*) = \mu^*$, and therefore $\Phi(\mu_*) \equiv \Phi(\mu^*) [c]$ can be written as $\Phi(\mu_*) \equiv \mu^* [c]$. Proposition 4 indicates that there exist a $v \in \text{supp}(\mu^*)$ and a $w \in X$ with $\varrho(v, w) = 1$ such that $\Phi(\mu_*) = \mu^* - c\delta_v + c\delta_w$. Therefore the natural way to connect $\mu^* = \Phi(\mu^*)$ and $\Phi(\mu_*)$ with a curve is

$$(42) \quad \Gamma : [0, c] \rightarrow \mathcal{W}_p(X); \quad \Gamma(t) := \mu^* - t\delta_v + t\delta_w.$$

It follows from (31) and (39) that $\Phi(\mu) \equiv \Phi(\mu^*) [\mu(x_{n+1})]$ and $\Phi(\mu) \equiv \Phi(\mu_*) [\mu(u)]$. And thus, $\Phi(\mu)$ must have the form $\mu^* - t\delta_v + t\delta_w$ for some $t \in (0, c)$. In fact, $\Phi(\mu) = \Gamma(\mu(x_{n+1}))$. And similarly, it follows from (41) and Corollary 5 that the Φ -image of γ is Γ .

We need more, in fact we want to prove that $\gamma = \Gamma$, or equivalently, $x_{n+1} = w$ and $u = v$. To show that $u = v$, it is enough to prove that $\mu^*(u) = \mu^*(v)$. Indeed, assume indirectly that $\mu^*(u) = \mu^*(v)$, but $u \neq v$. First recall that $u \neq v$ implies $\mu^*(v) = \mu(v)$. If $\mu(u) = 0$, then

$$(43) \quad 0 < \mu(x_{n+1}) = \mu(x_{n+1}) + \mu(u) = \mu^*(u) = \mu^*(v) = \mu(v)$$

which contradicts (35), unless $v = x_{n+1}$. But $v = x_{n+1}$ is impossible because $v \in X_n$ and $x_{n+1} \notin X_n$. If $\mu(u) \neq 0$, then $\mu(v) = \mu^*(v) = \mu^*(u) > \mu(u) > 0$, and thus u, v, x_{n+1} are pairwise different elements of $\text{supp}(\mu)$ such that

$$(44) \quad \mu(v) = \mu^*(v) = \mu^*(u) = \mu(u) + \mu(x_{n+1}),$$

which contradicts (36). Now we know that if $\mu^*(u) = \mu^*(v)$, then $u = v$. In this case, $\Phi(\mu)$ can be written as $\Phi(\mu) = \mu^* - \mu(x_{n+1})\delta_u + \mu(x_{n+1})\delta_w$, and thus

$$(45) \quad d_p^p(\delta_{x_{n+1}}, \Phi(\mu)) = d_p^p(\delta_{x_{n+1}}, \mu^*) - \mu(x_{n+1}) + \varrho^p(x_{n+1}, w)\mu(x_{n+1}).$$

Moreover, we have that $d_p(\delta_{x_{n+1}}, \Phi(\mu)) = d_p(\Phi(\delta_{x_{n+1}}), \Phi(\mu)) = d_p(\delta_{x_{n+1}}, \mu)$, and that

$$(46) \quad d_p^p(\delta_{x_{n+1}}, \mu) = d_p^p(\delta_{x_{n+1}}, \mu^*) - \mu(x_{n+1}).$$

Now we can conclude from (45) and (46) that $\varrho^p(x_{n+1}, w)\mu(x_{n+1}) = 0$. Since $\mu(x_{n+1}) \neq 0$, we get $x_{n+1} = w$.

What remains to prove in this step is that $c = \mu^*(u) = \mu^*(v)$. On the one hand, we know from (38) that $\mu^*(u) = c$. On the other hand, we have $\mu^*(v) \geq c$, because

$$(47) \quad 0 \leq \Phi(\mu_*)(v) = \mu^*(v) - c\delta_v(v) + c\delta_w(v) = \mu^*(v) - c.$$

Assume indirectly that $\mu^*(v) > c$. In this case, we can extend Γ from $[0, c]$ to $[0, \mu^*(v)]$ by

$$(48) \quad \tilde{\Gamma} : [0, \mu^*(v)] \rightarrow \mathcal{W}_p(X); \quad \tilde{\Gamma}(t) := \mu^* - t\delta_v + t\delta_w.$$

This extension has the property that

$$(49) \quad \Phi(\mu) \equiv \tilde{\Gamma}(t) [|t - \mu(x_{n+1})|] \quad \text{for all } t \in [0, \mu^*(v)] \setminus \{\mu(x_{n+1})\}.$$

Since Φ^{-1} is an isometry as well, Corollary 5 says that $\Phi^{-1}(\tilde{\Gamma}(t)) \equiv \mu [|t - \mu(x_{n+1})|]$ holds for all $t \in [0, \mu^*(v)] \setminus \{\mu(x_{n+1})\}$, and thus γ can be extended through $\gamma(c) = \mu_*$ with measures which are all in neighbouring relation with μ , a contradiction. Indeed, in order to continue γ , we need to add more weight to x_{n+1} . But $\mu_*(u) = 0$, so we should teleport mass from a point $x \in X \setminus \{x_{n+1}, u\}$ which would ruin the neighbouring relation with μ . Summarising the above observations: $\mu^*(u) = \mu^*(v)$, and thus $\Phi(\mu) = \mu$.

Step 3. We saw that $\Phi(\mu) = \mu$ holds if μ satisfies (35) and (36). Since Φ is continuous and $\mathcal{F}(X)$ is dense in $\mathcal{W}_p(X)$, it is enough to show that every $\nu := \sum_{i=1}^L a_i \delta_{u_i} \in \mathcal{F}(X)$ can be approximated by such measures. Here we assume that $u_i \neq u_j$ if $i \neq j$ and that $a_i > 0$ for all $1 \leq i \leq L$. For an arbitrary $\varepsilon > 0$ we are going to construct a measure ν' with $\text{supp}(\nu) = \text{supp}(\nu')$ such that $d_p(\nu, \nu') < \varepsilon$.

If $L = 1$ then ν itself satisfies (35) and (36), so we can assume that $L \geq 2$. Set

$$(50) \quad K := \max\{\varrho(u_i, u_j) \mid 1 \leq i, j \leq L\},$$

and if necessary, choose a smaller $0 < \tilde{\varepsilon} \leq \varepsilon$ such that $0 < a_i - \frac{\tilde{\varepsilon}}{K^p L^2}$ holds for all $1 \leq i \leq L$. Using such an $\tilde{\varepsilon}$, the intersection of the cube

$$(51) \quad C = \prod_{i=1}^L \left[a_i - \frac{\tilde{\varepsilon}}{K^p L^2}, a_i + \frac{\tilde{\varepsilon}}{K^p L^2} \right] \subseteq \mathbb{R}^L$$

with the hyperplane $P = \{(c_1, \dots, c_L) \mid \sum_{i=1}^L c_i = 1\}$ contains only vectors (c_1, \dots, c_L) such that $\sum_{i=1}^L c_i \delta_{u_i} \in \mathcal{W}_p(X)$. The set of representing vectors of those measures which violate (35) or (36) can be covered by the union of finitely many lower dimensional linear subspaces in \mathbb{R}^L . Since none of these subspaces are identical with P , we can choose an uncovered $(c_1, \dots, c_L) \in C \cap P$ and set $\nu' = \sum_{i=1}^L c_i \delta_{u_i}$.

We claim that $d_p(\nu, \nu') < \varepsilon$. To see this, we construct a $\pi \in \Pi(\nu, \nu')$ which leaves all mass shared by ν and ν' undisturbed. Set $m_i = \min\{a_i, c_i\}$ ($1 \leq i \leq L$) and $M := \sum_{i=1}^L m_i$, and subtract $\sum_{i=1}^L m_i \delta_{u_i}$ from ν and ν' . Now we have $\zeta := \sum_{i=1}^L (a_i - m_i) \delta_{u_i}$ and $\zeta' := \sum_{i=1}^L (c_i - m_i) \delta_{u_i}$ with $\text{supp}(\zeta) \cap \text{supp}(\zeta') = \emptyset$. For the product measure $(\zeta \times \zeta')(u_i, u_j) := \zeta(u_i) \zeta'(u_j)$ we have $(\zeta \times \zeta')(u_i, u_i) = 0$ for all $1 \leq i \leq L$, and if $i \neq j$ then

$$(52) \quad (\zeta \times \zeta')(u_i, u_j) = \zeta(u_i) \zeta'(u_j) \leq \zeta(u_i) = a_i - m_i \leq |a_i - c_i| \leq \frac{\tilde{\varepsilon}}{K^p L^2}.$$

Now define $\pi \in \Pi(\nu, \nu')$ as follows: $\pi(u_i, u_j) := (\zeta \times \zeta')(u_i, u_j)$ if $i \neq j$, and $\pi(u_i, u_i) := m_i$ for $1 \leq i \leq L$. Since $\varrho^p(u_i, u_i) \pi(u_i, u_i) = 0$ for all $1 \leq i \leq L$, using (52) we have the following upper bound for $d_p(\nu, \nu')$

$$(53) \quad d_p(\nu, \nu') \leq \sqrt[p]{\sum_{1 \leq i, j \leq L} \varrho^p(u_i, u_j) \cdot \pi(u_i, u_j)} < \sqrt[p]{L^2 K^p \frac{\tilde{\varepsilon}}{K^p L}} = \tilde{\varepsilon} \leq \varepsilon.$$

□

4. WASSERSTEIN SPACES WITH PRESCRIBED ISOMETRY GROUP

A natural question was raised by Kónig in [18]: *which groups are isomorphic to the automorphism group of a graph?* We recall that an automorphism of a simple graph $G(X, E)$ is a permutation $f : X \rightarrow X$ such that for any two $x, y \in X$ the pair $\{x, y\}$ form

an edge (i.e., belongs to E) if and only if the pair $\{f(x), f(y)\}$ also form an edge. The group of automorphisms will be denoted by $\text{Aut}(G(X, E))$.

Of course, one can replace graphs with other mathematical structures, for example with Wasserstein spaces, and ask the same question. Since in the metric context, automorphisms are in particular isometries, the corresponding question reads as follows: which groups are isomorphic to the isometry group of a Wasserstein space? Using some famous results in graph theory, the answer for countable groups is a corollary of Theorem 6. (For analogous results for autohomeomorphism groups see [4, Theorem 7].)

Corollary 7. Let H be a countable group and $p \geq 1$ any real number. Then there exists a metric space (X, ϱ) such that $\text{Isom}(\mathcal{W}_p(X), d_p) \cong H$.

Proof. As an extension of Frucht's theorem [8], de Groot proved that every countable group H is isomorphic to the automorphism group of a countable simple graph $G(X, E)$ (see [4, comments on p.96]). Let (X, ϱ) be the metric space associated to $G(X, E)$ and consider the p -Wasserstein space $\mathcal{W}_p(X)$. According to Theorem 6, $\text{Isom}(\mathcal{W}_p(X), d_p) \cong \text{Isom}(X, \varrho)$, and therefore it is enough to show that $\text{Isom}(X, \varrho) \cong \text{Aut}(G(X, E))$. It will turn out that these groups are identical as a set with the same operation (composition), so the identity map is an isomorphism.

If $\psi \in \text{Isom}(X, \varrho)$ then for any pair $x, y \in X$ we have $\varrho(x, y) = 1$ if and only if $\varrho(\psi(x), \psi(y)) = 1$. Or equivalently, x and y are joined by an edge if and only if $\psi(x)$ and $\psi(y)$ are joined by an edge. Since ψ is a bijection, this means that $\psi \in \text{Aut}(G(X, E))$. On the other hand, every $\psi \in \text{Aut}(G(X, E))$ induces a length preserving bijection on the set of all paths as follows: if we have a path of length k along the sequence of distinct vertices $x_0 = x, x_1, \dots, x_k = y$, then the sequence $\psi(x_0), \dots, \psi(x_k)$ determines a path of length k between $\psi(x)$ and $\psi(y)$. And therefore, the shortest path distance of x and y must be the same as the shortest path distance of $\psi(x)$ and $\psi(y)$. Since an automorphism is a bijection by definition, we have that $\psi \in \text{Isom}(X, \varrho)$. \square

We remark that there is no uniqueness above, because there is no uniqueness in Frucht's and de Groot's theorems. In fact, Izbicki proved in [15] that there are uncountably many infinite graphs realizing any finite symmetry group. We also remark that although de Groot's theorem is valid for non-countable groups as well, we do not know the smallest possible order of the representing graph. However, it is important to note that if the cardinality of the vertex set is bigger than \aleph_0 , then our method of proof does not work, therefore the following question remains open.

Problem 8. Given an uncountable group G and a fixed number $p \geq 1$, does there exist a p -Wasserstein space whose isometry group is isomorphic to G ?

REFERENCES

- [1] L. AMBROSIO, N. GIGLI, *A user's guide to optimal transport. Modelling and optimisation of flows on networks*, Lecture Notes in Math., 2062, Fond. CIME/CIME Found. Subser., Springer, Heidelberg, 2013.
- [2] J. BERTRAND, B. KLOECKNER, *A geometric study of Wasserstein spaces: isometric rigidity in negative curvature*, Int. Math. Res. Notices **5** (2016), 1368–1386.
- [3] G. CHARTRAND, D. KAY, *A Characterization of Certain Ptolemaic Graphs*, Canadian Journal of Mathematics **17** (1965), 342–346.
- [4] J. DE GROOT, *Groups represented by homeomorphism groups I.*, Math. Ann. **138** (1959), 80–102.
- [5] G. DE PALMA, M. MARVIAN, D. TREVISAN AND S. LLOYD, *The quantum Wasserstein distance of order 1*, IEEE Transactions on Information Theory **67** (2021), 6627–6643.
- [6] G. DOLINAR, B. KUZMA, D. MITROVIC, *Isometries of probability measures with respect to the total variation distance* J. Math. Anal. Appl. (2021), Paper No. 125829, In Press.
- [7] G. DOLINAR, L. MOLNÁR, *Isometries of the space of distribution functions with respect to the Kolmogorov–Smirnov metric*, J. Math. Anal. Appl. **348** (2008), 494–498.
- [8] R. FRUCHT, *Herstellung von Graphen mit vorgegebener abstrakter Gruppe*, Compositio Math. **6** (1938), 239–250.
- [9] P. JIRADILOK, S. KAMTUE, *Transportation distance between probability measures on the infinite regular tree*, manuscript, arXiv:2107.09876, 2021.
- [10] GY.P. GEHÉR, *Surjective Kuiper isometries*, Houston J. Math. **44** (2018), 263–281.
- [11] GY.P. GEHÉR, T. TITKOS, *A characterisation of isometries with respect to the Lévy-Prokhorov metric*, Annali della Scuola Normale Superiore di Pisa - Classe di Scienze, Vol. **XIX** (2019), 655–677.
- [12] GY.P. GEHÉR, T. TITKOS, D. VIROSZTEK, *On isometric embeddings of Wasserstein spaces – the discrete case*, J. Math. Anal. Appl. **480** (2019), Paper No. 123435.
- [13] GY.P. GEHÉR, T. TITKOS, D. VIROSZTEK, *Isometric study of Wasserstein spaces – the real line*, Trans. Amer. Math. Soc. **373** (2020), 5855–5883.
- [14] GY. P. GEHÉR, T. TITKOS, D. VIROSZTEK, *The isometry group of Wasserstein spaces: the Hilbertian case*, manuscript, arXiv:2102.02037, 2021.
- [15] H. IZBICKI, *Unendliche Graphen endlichen Grades mit vorgegebenen Eigenschaften*, Monatshefte für Mathematik **63** (1959), 298–301.
- [16] B. KLOECKNER, *A geometric study of Wasserstein spaces: Euclidean spaces*, Annali della Scuola Normale Superiore di Pisa - Classe di Scienze **IX** (2010), 297–323.
- [17] B. KLOECKNER, *A generalization of Hausdorff dimension applied to Hilbert cubes and Wasserstein spaces*, J. Topol. Anal. **4** (2012), 203–235.
- [18] D. KÖNIG, *Theorie der endlichen und unendlichen Graphen*, Akad. Verlag, Leipzig 1936.
- [19] L. MOLNÁR, *Lévy isometries of the space of probability distribution functions*, J. Math. Anal. Appl. **380** (2011), 847–852.
- [20] D.S. ORNSTEIN, *An Application of Ergodic Theory to Probability Theory*, The Annals of Probability **1** (1973), 43–58.
- [21] F. RHODES, *Some characterizations of the chessboard metric and the city block metric*, Pattern Recognition Letters **11** (1990), 669–675.
- [22] F. SANTAMBROGIO, *Optimal Transport for Applied Mathematicians*, Progress in Nonlinear Differential Equations and Their Applications 87, Birkhauser Basel (2015).
- [23] J. SANTOS-RODRÍGUEZ, *Isometric rigidity of compact Wasserstein spaces*, manuscript, arXiv:2102.08725, 2021.
- [24] C. VILLANI, *Optimal Transport: Old and New*, Grundlehren der mathematischen Wissenschaften, Springer, 2009.
- [25] D. VIROSZTEK, *Maps on probability measures preserving certain distances — a survey and some new results.*, Acta Sci. Math. (Szeged) **84** (2018), 65–80.

GERGELY KISS, ALFRÉD RÉNYI INSTITUTE OF MATHEMATICS – EÖTVÖS LORÁND RESEARCH NETWORK, REÁLTANODA U. 13-15., BUDAPEST H-1053, HUNGARY.

Email address: `kiss.gergely@renyi.hu`

TAMÁS TITKOS, ALFRÉD RÉNYI INSTITUTE OF MATHEMATICS – EÖTVÖS LORÁND RESEARCH NETWORK, REÁLTANODA U. 13-15., BUDAPEST H-1053, HUNGARY, AND BBS UNIVERSITY OF APPLIED SCIENCES, ALKOTMÁNY U. 9., BUDAPEST H-1054, HUNGARY.

Email address: `titkos.tamas@renyi.hu`