A geometric study of Wasserstein spaces: Hadamard spaces

Jérôme Bertrand, Benoît Kloeckner

A GEOMETRIC STUDY OF WASSERSTEIN SPACES: HADAMARD SPACES

by

Jérôme Bertrand & Benoît Kloeckner

Abstract. — Optimal transport enables one to construct a metric on the set of (sufficiently small at infinity) probability measures on any (not too wild) metric space X, called its Wasserstein space $\mathscr{W}_2(X)$.

In this paper we investigate the geometry of $\mathscr{W}_2(X)$ when X is a Hadamard space, by which we mean that X has globally non-positive sectional curvature and is locally compact. Although it is known that –except in the case of the line– $\mathscr{W}_2(X)$ is not non-positively curved, our results show that $\mathscr{W}_2(X)$ have large-scale properties reminiscent of that of X. In particular we define a geodesic boundary for $\mathscr{W}_2(X)$ that enables us to prove a non-embeddablity result: if X has the visibility property, then the Euclidean plane does not admit any isometric embedding in $\mathscr{W}_2(X)$.

1. Introduction

The goal of this paper is to contribute to the understanding of the geometry of Wasserstein spaces. Given a metric space X, the theory of optimal transport (with quadratic cost) gives birth to a new metric space, made of probability measures on X, often called its Wasserstein space and denoted here by $\mathscr{W}_2(X)$ (precise definitions are recalled in the first part of this paper). One can use this theory to study X, for example by defining lower Ricci curvature bounds as in the celebrated works of Lott-Villani [**LV09**] and Sturm [**Stu06**]. Conversely, here we assume some understanding of X and try to use it to study geometric properties of $\mathscr{W}_2(X)$. A similar philosophy underlines the works of Lott in [**Lot08**] and Takatsu and Yokota in [**TY12**].

In a previous paper [**Klo10**], the second named author studied the case when X is a Euclidean space. Here we are interested in the far broader class of *Hadamard spaces* which are roughly the globally non-positively curved spaces. The first part of the paper gives the classical definitions and property we need both on optimal transport and Hadamard spaces; in particular the precise hypotheses under which we shall work are given there (Definition 2.1, see also examples 2.2). Let us stress that we allow X to be branching; trees, product involving trees, some buildings are in particular treated in the same framework than, for example, symmetric spaces of non-compact type.

While non-negative curvature is an assumption that is inherited by Wasserstein spaces, it is well-known that non-positive curvature is not (an argument is recalled in Section 2.2.3). We shall however show that some features of Hadamard spaces still hold in their Wasserstein spaces. Let us now describe the main results of the article.

A Hadamard space admits a well-known geometric compactification, obtained by adding a boundary at infinity made of asymptote classes of geodesic rays. In sections 3 and 4, we study the geodesic rays of $\mathscr{W}_2(X)$. Using a displacement interpolation procedure (Proposition 3.2), we associate to each ray its *asymptotic measure* which lies in a subset $\mathscr{P}_1(c\partial X)$ of probability measures on the cone $c\partial X$ over the boundary of X (Definition 4.1). It encodes the asymptotic distribution of the direction and speed of a measure running along the ray. Our first main result is the *asymptotic formula* (Theorem 4.2) which enables one to compute the asymptotic behavior of the distance between two rays in terms of the Wasserstein distance of the asymptotic measures, with respect to the angular cone distance on $c\partial X$. This asymptotic distance is either bounded or asymptotically linear, so that the boundary $\partial\mathscr{W}_2(X)$ of the Wasserstein space, defined as the set of asymptote classes of unit geodesic rays, inherits an angular metric, just like X does. A striking consequence of the asymptotic formula concerns the rank of $\mathscr{W}_2(X)$, and partially answers a question raised in the previous paper cited above.

Theorem 1.1. — *If X is a visibility space (e.g. if it has curvature bounded from above by a negative constant), then it is not possible to embed the Euclidean plane isometrically in $\mathscr{W}_2(X)$.*

In other words, when X is strongly negatively curved –which implies that its has rank 1–, then although $\mathscr{W}_2(X)$ is not negatively curved it has rank 1 too. Note that our large-scale method in fact implies more general non-embedding results, see Proposition 5.3 and the discussion below. It is important to stress that asking for *isometric* embedding is the right regularity: more flexible conditions are easily dealt with, see example 5.5. Besides this property on the rank, the Wasserstein space over a visibility space can also be differentiated from the one over an Euclidean space through its isometry group $\operatorname{Isom}\mathscr{W}_2(X)$. Indeed, contrary to the Euclidean case where the isometry group $\operatorname{Isom}\mathscr{W}_2(\mathbb{R}^n)$ is larger than $\operatorname{Isom}\mathbb{R}^n$ [**Klo10**], negatively curved spaces seem to have isometrically rigid Wasserstein spaces in the sense that $\operatorname{Isom}\mathscr{W}_2(X) = \operatorname{Isom} X$. This holds at least in the case of manifolds and trees, as proved in a previous version of this paper [**BK10**]; this result uses different methods from the ones developped here and we aim at extending it, it shall therefore appear in a subsequent article.

Section 6 is devoted to the definition of a so-called *cone topology* on $\partial\mathscr{W}_2(X)$ and $\overline{\mathscr{W}_2(X)} = \mathscr{W}_2(X) \cup \partial\mathscr{W}_2(X)$, see Proposition 6.1. Note that the angular metric alluded to above, however useful and meaningful, does not define a satisfactory topology (just as in ∂X, where the angular metric is usually not separable and can even be discrete). The point is that many monotony properties used in the case of Hadamard spaces hold when one restricts to angles based at a Dirac mass. This enables us to carry out the construction of this topology despite the presence of positive curvature. The main result of this part is the following, restated as Theorem 7.2.

Theorem 1.2. — *The asymptotic measure map defined from the boundary $\partial\mathscr{W}_2(X)$ to the set of measures $\mathscr{P}_1(c\partial X)$ is a homeomorphism.*

Note that the two natural topologies on $\partial\mathscr{W}_2(X)$, namely the cone topology and the quotient topology of the topology of uniform convergence on compact sets, coincide. The set $\mathscr{P}_1(c\partial X)$ is simply endowed with the weak topology (where the topology on $c\partial X$ is induced by the cone topology of ∂X).

The possibility to identify $\partial\mathscr{W}_2(X)$ to $\mathscr{P}_1(c\partial X)$ should be thought of as an interversion result, similar to displacement interpolation. The latter says that "a geodesic in the set of measures is a measure on the set of geodesics", while the former can be roughly restated as "a boundary point of the set of measures is a measure on the (cone over the) set of boundary points". Note that $\overline{\mathscr{W}_2(X)}$ is not compact; this is quite inevitable since $\mathscr{W}_2(X)$ is not locally compact.

2. Reminders and notations

As its title indicates, this part contains nothing new. We chose to give quite a lot of recalls, so that the reader familiar with non-positively curved spaces can get a crash-course on Wasserstein spaces, and the reader familiar with optimal transport can be introduced to Hadamard spaces.

2.1. Hadamard spaces.

— Most properties of Hadamard spaces stated here are proved in [**Bal95**]. Another more extensive reference is [**BH99**].

2.1.1. Geodesics. — Let us first fix some conventions for any metric space Y (this letter shall be used to design arbitrary spaces, while X shall be reserved to the (Hadamard) space under study).

A *geodesic* in Y is a curve $\gamma : I \to Y$ defined on some interval I, such that there is a constant v that makes the following hold for all times t, t':

$$d(\gamma_t, \gamma_{t'}) = v|t - t'|.$$

In particular, all geodesics are assumed to be *globally* minimizing and to have constant, non necessarily unitary speed. A metric space is *geodesic* if any pair of points can be linked by a geodesic.

When $v = 0$ we say that the geodesic is constant and it will be necessary to consider this case. We denote by $\mathscr{G}^{T,T'}(Y)$ the set of geodesics defined on the interval $[T, T']$. A *geodesic ray* (or *ray*, or *complete ray*) is a geodesic defined on the interval $[0, +\infty)$. A *complete geodesic* is a geodesic defined on $\mathbb{R}$. The set of rays is denoted by $\mathscr{R}(Y)$, the set of unit speed rays by $\mathscr{R}_1(Y)$ and the set of non-constant rays by $\mathscr{R}_{>0}(Y)$. We shall also denote by $\mathscr{G}^{\mathbb{R}}(Y)$ the set of complete geodesics, and by $\mathscr{G}_1^{\mathbb{R}}(Y)$ the set of unit-speed complete geodesics.

2.1.2. Non-positive curvature. — A triangle in a geodesic space Y is the datum of three points (x, y, z) together with three geodesics parametrized on $[0, 1]$ linking x to y, y to z and z to x. Given a triangle, one defines its *comparison triangle* $(\tilde{x}, \tilde{y}, \tilde{z})$ as any triangle of the Euclidean plane $\mathbb{R}^2$ that has the same side lengths: $d(x, y) = d(\tilde{x}, \tilde{y})$, $d(x, z) = d(\tilde{x}, \tilde{z})$ and $d(y, z) = d(\tilde{y}, \tilde{z})$. The comparison triangle is defined up to congruence.

A triangle with vertices (x, y, z) is said to satisfy the CAT(0) inequality along its $[yz]$ side (parametrized by a geodesic $\gamma \in \mathscr{G}^{0,1}(Y)$) if for all $t \in [0, 1]$, the following inequality holds:

$$d(x, \gamma(t)) \leqslant d(\tilde{x}, (1 - t)\tilde{y} + t\tilde{z}) \tag{1}$$

see figure 1. A geodesic space is said to be locally CAT(0) if every point admits a neighborhood where all triangles satisfy the CAT(0) inequality (along all there sides). When Y is a Riemannian manifold, this is equivalent to ask that Y has non-positive sectional curvature. A geodesic space is said to be globally CAT(0) if all its triangles satisfy the CAT(0) inequality. Globally CAT(0) is equivalent to simply connected plus locally CAT(0). We shall simply say CAT(0) for "globally CAT(0)", but this is not a universal convention.

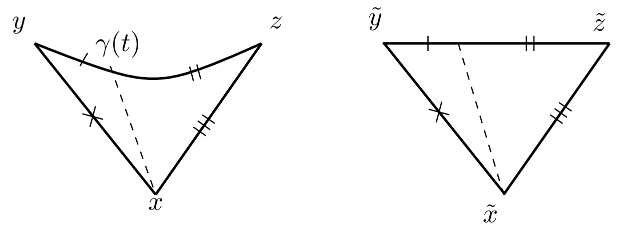

FIGURE 1. The CAT(0) inequality: the dashed segment is shorter in the original triangle on the left than in the comparison triangle on the right.

2.1.3. Angles. — The CAT(0) condition can be translated in terms of angles as follows. Given any geodesic triangle, choose any of its vertices, say x, and assume that the sides containing x are parametrized by two geodesics $\sigma, \gamma \in \mathscr{G}^{0,1}(Y)$. If Y is CAT(0), then the Euclidean angle $\widetilde{\angle}_{\tilde{\gamma}_s \tilde{x} \tilde{\sigma}_t}$ at $\tilde{x}$ is a nondecreasing function of s and t.

One then defines in Y the angle $\angle_{\gamma_1 x \sigma_1}$ at x as the limit, when s and t go to zero, of $\widetilde{\angle}_{\tilde{\gamma}_s \tilde{x} \tilde{\sigma}_t}$. As a consequence, one gets that for any geodesic triangle with vertices (x, y, z) in a CAT(0) space, angle and comparison angle satisfy

$$\angle_{xyz} \leqslant \widetilde{\angle}_{\tilde{x}\tilde{y}\tilde{z}}.$$

2.1.4. Distance convexity. — In a CAT(0) space, given two geodesics γ and β, the distance function $t \longmapsto d(\gamma_t, \beta_t)$ is convex. This important property shall be kept in mind since it will be used very often in the sequel.

2.1.5. Hadamard spaces: definition and examples. — We can now introduce the class of spaces we are interested in.

Definition 2.1. — A metric space is a *Hadamard space* if it is:

– Polish (i.e. complete and separable),
– locally compact,
– geodesic,
– CAT(0), implying that it is simply connected.

In all what follows, we consider a Hadamard space X. The Hadamard assumption may not always be made explicit, but the use of the letter X shall always implicitely imply it. Not all authors assume Hadamard spaces to be locally compact, and this assumption excludes for example real trees.

Example 2.2. — There are many important examples of Hadamard spaces. Let us give some of them:

– the Euclidean space $\mathbb{R}^n$,
– the real hyperbolic space $\mathbb{R}\mathrm{H}^n$,
– the other hyperbolic spaces $\mathbb{C}\mathrm{H}^n$, $\mathbb{H}\mathrm{H}^n$, $\mathbb{O}\mathrm{H}^2$,
– more generally the symmetric spaces of non-compact type, like the quotient $\mathrm{SL}(n;\mathbb{R})/\mathrm{SO}(n)$ endowed with the metric induced by the Killing form of $\mathrm{SL}(n;\mathbb{R})$,
– more generally any simply connected Riemannian manifold whose sectional curvature is non-positive,
– trees,
– any product of Hadamard spaces,
– some buildings, like product of trees having unit edges and no leaf or I_{pq} buildings (see [**Bou97, BP99**]),

– the gluing of any two Hadamard spaces along isometric, convex subsets; for example any Hadamard space with an additional geodesic ray glued at some point, or three hyperbolic half-planes glued along their limiting geodesics, etc.

2.1.6. Geodesic boundary. — The construction of the geodesic boundary that we will shortly describe seems to date back to [**EO73**], but note that [**Bus55**] is at the origin of many related ideas.

Two rays of X are *asymptotic* if they stay at bounded distance when $t \to +\infty$, and this relation is denoted by $\sim$. The asymptote class of a ray γ is often denoted by $\gamma(\infty)$ or γ_∞, and is called the *endpoint* or *boundary point* of γ.

The *geodesic* (or *Hadamard*) *boundary* of X is defined as the set

$$\partial X = \mathscr{R}_1(X)/\sim .$$

Using the convexity of distance along geodesics, one can for example prove that, given points $x \in X$ and $\zeta \in \partial X$, there is a unique unit ray starting at x and ending at ζ.

The union $\bar{X} = X \cup \partial X$ can be endowed with its so-called cone topology, which makes $\bar{X}$ and ∂X compact. Without entering into the details, let us say that this topology induces the original topology on X, and that given a base point x_0 a basic neighborhood of a point $\zeta = \gamma(\infty) \in \partial X$ (where γ starts at x_0) is the union, over all rays σ starting at x_0 such that $d(\sigma_t, \gamma_t) < \varepsilon$ for all $t < R$, of $\sigma([R, +\infty])$ (see figure 2).

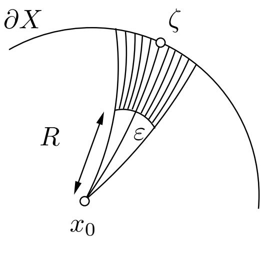

FIGURE 2. A basic neighborhood of a point $\zeta \in \partial X$ in the cone topology.

Consistently whith the cone topology, all previously defined sets of geodesics, as well as the larger sets $C(I, X)$ of continuous curves defined on an interval I with values in X, are endowed with the topology of uniform convergence on compact sets. With this topology, since X is Hadamard, a geodesic segment is uniquely and continuously defined by its endpoints and a geodesic ray is uniquely and continuously defined by its starting point, its endpoint in the boundary and its speed. As a consequence, there are natural homeomorphisms

$$\mathscr{G}^{T,T'}(X) \simeq X^2, \qquad \mathscr{R}(X) \simeq X \times c\partial X$$

where $c\partial X$ is the cone over ∂X, that is the quotient of $\partial X \times [0,+\infty)$ by the relation $(x,0) \sim (y,0)$ for all $x, y \in \partial X$. We usually use the same notation (x,s) for a couple and its equivalence class under this relation; here s shall be considered as a speed. In particular one has $(\mathscr{R}(X)/\sim) \simeq c\partial X$.

Note that in view of our assumptions on X, all these spaces are locally compact Polish topological spaces (that is, metrizable by a metric that is separable and complete). It ensures that all finite measures on these spaces are Radon.

2.1.7. Possible additional assumptions. — At some points, we shall make explicitely additional hypotheses on X. One says that a space is:

– *geodesically complete* if every geodesic can be extended to a complete geodesic,

– *non-branching* if two geodesics that coincide on an open set of parameters coincide everywhere on their common definition interval,

– CAT(κ) fo some $\kappa < 0$ if its triangles satisfy the (1) inequality when the comparison triangle is taken in $\mathbb{R}\mathrm{H}^2_\kappa$, the hyperbolic plane of curvature κ, instead of $\mathbb{R}^2$.

– a *visibility space* if for all pairs of distinct points $\alpha, \omega \in \partial X$ there is a complete geodesic γ such that $\gamma(-\infty) = \alpha$ and $\gamma(+\infty) = \omega$.

Note that for all $\kappa < 0$, the condition CAT(κ) implies that X is a visibility space.

Geodesic completeness is quite mild (avoiding examples as trees with leafs), while the other are strong assumptions (for example non-branching rules out trees and visibility rules out products). Another possible assumption, that we shall not use directly, is for X to have *rank one*, meaning that it admits no isometric embedding of the Euclidean plane. It is a weaker condition than visibility. More generally the rank of a space Y is the maximal dimension of an isometrically embedded Eulidean space, and it has been proved a very important invariant in the study of symmetric spaces. For example the hyperbolic spaces $\mathbb{R}\mathrm{H}^n$, $\mathbb{C}\mathrm{H}^n$, $\mathbb{H}\mathrm{H}^n$ and $\mathbb{O}\mathrm{H}^2$ are the only rank one symmetric spaces of non-compact type.

2.1.8. Asymptotic distance. — Given two rays γ, σ, one defines their *asymptotic distance* by

$$d_\infty(\gamma,\beta) = \lim_{t\to+\infty} \frac{d(\gamma_t,\beta_t)}{t}.$$

This limit always exists because of the convexity of the distance function along geodesics. Moreover d_∞ defines a metric on $c\partial X$ and, by restriction, on ∂X (in particular two rays whose distance grows sub-linearly must be asymptotic). It can be proved that d_∞ is the cone metric over ∂X endowed with the angular metric. Namely, for any $(\xi, s), (\xi', t) \in c\partial X$,

$$d_\infty^2((\xi, s), (\xi', t)) = s^2 + t^2 - 2st\cos\angle(\xi, \xi') \tag{2}$$

where $\angle(\xi, \xi') = \sup_{x\in X} \angle_x(\xi, \xi')$ is the supremum of angles between rays issuing from x and asympotic to ξ and ξ' respectively (we refer to [**BH99**, Section II.9] for more details and proof).

It is most important to keep in mind that the metric d_∞ *does not induce the cone topology*, but a much finer topology. The most extreme case is that of visibility spaces, where $d_\infty(\gamma, \sigma)$ is 0 if $\gamma \sim \sigma$ and the sum of the speeds of γ and σ otherwise: the topology induced on ∂X is discrete. However, the function d_∞ is lower semi-continuous with respect to the cone topology [**BH99**, Proposition II.9.5], so that it is a mesurable function.

In higher rank spaces, it can be useful to turn d_∞ into a length metric, called the Tits metric, but we shall not use it so we refer the interested reader to the books cited above. Let us just note that d_∞ resembles in some aspects the chordal metric on a sphere. In particular, it is naturally isometric to this metric when X is a Euclidean space.

2.2. The Wasserstein space. — In this section, we recall the definition of Wasserstein space and some of its main properties. For more details, we refer to the books [**Vil03**] and [**Vil09**].

2.2.1. Optimal transport. — Let us start with the concept of optimal transport which is the theory aimed at studying the *Monge-Kantorovich* problem.

Standard data for this problem are the following. We are given a Polish metric space (Y, d), a lower semicontinuous and nonnegative function $c : Y \times Y \to \mathbb{R}^+$ called the cost function and two Borel probability measures μ, ν defined on Y. A *transport plan* between μ and ν is a measure on $Y \times Y$ whose marginals are μ and ν. One should think of a transport plan as a specification of how the mass in Y, distributed according to μ, is moved so as to be distributed according to ν. We denote by $\Gamma(\mu, \nu)$ the set of transport plans which is never empty (it contains $\mu\otimes\nu$) and most of the time not reduced to one element. The Monge-Kantorovich problem is now

$$\min_{\Pi\in\Gamma(\mu,\nu)} \int_{Y\times Y} c(x, y)\, \Pi(dxdy)$$

where a minimizer is called an *optimal transport plan.* The set of optimal transport plans is denoted by $\Gamma_o(\mu, \sigma)$.

Let us make a few comments on this problem. First, note that under these assumptions, the cost function is measurable (see, for instance, [**Vil03**, p. 26]). Now, existence of minimizers follows readily from the lower semicontinuity of the cost function together with the following compactness result which will be used throughout this paper. We refer to [**Bil99**] for a proof.

***Theorem 2.3* (Prokhorov's Theorem).** — *Given a Polish space Y, a set $P \subset \mathscr{P}(Y)$ is totally bounded (that is, has compact closure) for the weak topology if and only if it is tight, namely for any $\varepsilon > 0$, there exists a compact set K_ε such that $\mu(Y \setminus K_\varepsilon) \leqslant \varepsilon$ for any $\mu \in P$.*

For example, the set $\Gamma(\mu, \nu)$ is always compact.

We also mention that, compared to the existence problem, the issue of the *uniqueness* of minimizers is considerably harder and requires, in general, additional assumptions. To conclude this introduction, we state a useful criterium to detect optimal transport plan among other plans, named cyclical monotonicity.

***Definition 2.4* (cyclical monotonocity).** — Given a cost function $c : Y \times Y \longrightarrow \mathbb{R}^+$, a set $\Gamma \subset Y \times Y$ is called c-cyclically monotone if for any finite family of pairs $(x_1, y_1), \cdots, (x_m, y_m)$ in Γ, the following inequality holds

$$\sum_{i=1}^{m} c(x_i, y_{i+1}) \geqslant \sum_{i=1}^{m} c(x_i, y_i) \tag{3}$$

where $y_{m+1} = y_1$.

In other words, a c-cyclically monotone set does not contain cycles of pairs (starting point, ending point) along which a shift in the ending points would reduce the total cost.

Theorem 2.5. — *Let (Y, d) be a Polish space and $c : Y \times Y \longrightarrow \mathbb{R}^+$ be a lower semi-continuous cost function. Then, a transport plan is optimal relatively to c if and only if it is concentrated on a c-cyclically monotone set.*

If c is continuous, this is equivalent to its support being c-cyclically monotone.

Under these assumptions, this result is due to Schachermayer and Teichmann [**ST09**]; see also [**Vil09**] for a proof.

2.2.2. Wasserstein space. — Wasserstein spaces arise in a particular variant of the setting above.

***Definition 2.6* (Wasserstein space).** — Given a Polish metric space Y, its (quadratic) Wasserstein space $\mathscr{W}_2(Y)$ is the set of Borel probability measures μ on Y with finite second moment, that is such that

$$\int_Y d(x_0, x)^2 \, \mu(dx) < +\infty \qquad \text{for some, hence all } x_0 \in Y,$$

endowed with the Wasserstein metric defined by

$$\mathrm{W}^2(\mu_0, \mu_1) = \min_{\Pi \in \Gamma(\mu_0, \mu_1)} \int_{Y \times Y} d^2(x, y) \, \Pi(dx, dy).$$

From now on, the cost c will therefore be $c = d^2$.

The fact that W is indeed a metric follows from the so-called "gluing lemma" which enables one to propagate the triangular inequality, see e.g. [**Vil09**].

Remark 2.7. — In this paper, we will also consider the Wasserstein space over the cone $c\partial X$ relative to the cost d_∞^2. Since d_∞ is lower semi-continuous, the usual theory of optimal transport applies, and we do get a metric on the suitable space of measures. We shall denote by W_∞ the Wasserstein metric derived from d_∞.

The Wasserstein space has several nice properties: it is Polish; it is compact as soon as Y is, in which case the Wasserstein metric metrizes the weak topology; but if Y is not compact, then $\mathscr{W}_2(Y)$ is not even locally compact and the Wasserstein metric induces a topology stronger than the weak one (more precisely, convergence in Wasserstein distance is equivalent to weak convergence plus convergence of the second moment). A very important property is that $\mathscr{W}_2(Y)$ is geodesic as soon as Y is; let us give some details.

2.2.3. Displacement interpolation. — The proof of what we explain now can be found for example in chapter 7 of [**Vil09**], see in particular corollary 7.22 and Theorem 7.30. Note that the concept of displacement interpolation has been introduced by McCann in [**McC97**]. We write this section in the case of a Hadamard space X, but most of it stays true for any Polish geodesic space.

Definition 2.8. — Define a *dynamical transport plan* between two measures $\mu_0, \mu_1 \in \mathscr{W}_2(X)$ as a probability measure μ on $C([t_0, t_1]; X)$ such that for $i = 0, 1$ the law at time t_i of a random curve drawn with law μ is μ_i. In other words we ask $e_{t_i\#}\mu = \mu_i$ where e_t is the map $C([t_0, t_1]; X) \to X$ defined by $e_t(\gamma) = \gamma_t$.

The *cost* of μ is then
$$|\mu|^2 = \int \ell(\gamma)^2 \mu(d\gamma)$$
where $\ell(\gamma)$ is the length of the curve γ (possibly $+\infty$). A dynamical transport plan is *optimal* if it minimizes the cost over all dynamical transport plans

It is known that a dynamical transport plan exists. Morevover, if μ is an optimal dynamical transport plan then:

(1) the law $(e_{t_0}, e_{t_1})_\#\mu$ of the couple $(\gamma_{t_0}, \gamma_{t_1})$ where γ is a random curve drawn with law μ, is an optimal transport plan between μ_0 and μ_1,

(2) μ-almost all $\gamma \in C([t_0, t_1]; X)$ are geodesics.

Conversely, if Π is a (non-dynamical) optimal transport plan, then one can construct for any $t_0 < t_1$ an optimal dynamical transport plan by the following construction. Let $F : X^2 \to C([t_0, t_1]; X)$ be the map that sends a couple (x, y) of points to the unique geodesic parametrized on $[t_0, t_1]$ that starts at x and ends at y. Then $\mu = F_{\#}\Pi$ is an optimal dynamical transport plan, whose associated optimal transport plan is obviously Π.

Given a dynamical plan $\mu \in \mathscr{P}(\mathscr{G}^I(X))$ where I is an arbitrary interval, μ is said to be *c-cyclically monotone* if for any $s, t \in I$, the support of the plan $(e_s, e_t)_{\#}\mu$ is c-cyclically monotone (note that here the cost is continuous). As soon as μ_t have finite second moments, this is equivalent to μ being optimal, but cyclical monotonicity has the advantage of being well-defined without any integrability assumption.

The main use of optimal dynamical transport plans is that they define geodesic segments. Indeed, let as before e_t be the map $\gamma \mapsto \gamma_t$ defined on the set of continuous curves. If μ is an optimal dynamical transport plan, consider the law $\mu_t = e_{t\#}\mu$ at time t of a random geodesic drawn with law μ: then $(\mu_t)_{t_0 \leqslant t \leqslant t_1}$ is a geodesic of $\mathscr{W}_2(X)$. Displacement interpolation is the converse to this principle.

***Proposition 2.9* (Displacement interpolation).** — *Given $(\mu_t)_{t_0 \leqslant t \leqslant t_1}$ any geodesic segment in $\mathscr{W}_2(X)$, there is a probability measure μ on $\mathscr{G}^{t_0,t_1}(X)$ such that for all t, $\mu_t = e_{t\#}\mu$.*

If μ is a dynamical transport plan on I, for all $t_0, t_1 \in I$ define the *time restriction* of μ to $[t_0, t_1]$ as $\mu^{t_0,t_1} = r_{t_0,t_1\#}\mu$ where $r_{t_0,t_1}(\gamma)$ is the restriction of the curve γ to the interval $[t_0, t_1]$.

Let μ be an optimal dynamical transport plan on $[0,1]$. For all $t_0, t_1 \in [0,1]$ the following holds (see [**Vil09**, Theorem 7.30]):

(1) μ^{t_0,t_1} is an optimal dynamical transport plan,

(2) if X is non-branching and $(t_0, t_1) \neq (0,1)$, then μ^{t_0,t_1} is the *unique* (up to parametrization) optimal dynamical transport plan between μ_{t_0} and μ_{t_1}.

***Remark 2.10* (The Wasserstein space is not non-positively curved)**

It is well known that the non-positive curvature assumption on X is not inherited by $\mathscr{W}_2(X)$ except if X is the real line or a subset of it. Let us give a brief explanation of this fact.

The affine structure makes the Wasserstein space contractible, so it cannot be locally CAT(0) without being uniquely geodesic. But as soon as X is not reduced to a geodesic, there exists four distinct points x, x', y, z such that $d(x,y) = d(x,z)$ and $d(x',y) = d(x',z)$. Between the measures $\mu = \frac{1}{2}\delta_x + \frac{1}{2}\delta_{x'}$ (where δ's are Dirac masses) and $\nu = \frac{1}{2}\delta_y + \frac{1}{2}\delta_z$, all transport plans are optimal. Each one of them defines a geodesic in $\mathscr{W}_2(Y)$ from μ to ν, therefore $\mathscr{W}_2(Y)$ is very far from being uniquely geodesic, and is in particular not non-positively curved.

3. Displacement interpolation for rays

When X is a Hadamard space, the *geodesic boundary* of $\mathscr{W}_2(X)$ is simply defined as the set of asymptote classes of unitary geodesic rays:

$$\partial \mathscr{W}_2(X) = \mathscr{R}_1(\mathscr{W}_2(X))/\sim .$$

To study its structure, we need a good understanding of the geodesic rays in the Wasserstein space, starting with the *displacement interpolation* of rays. There is not much, but some work needed to extend the case of geodesic segment; the crucial point being to handle the branching case. Note that the case of complete geodesic is not different than that of the rays.

***Proposition 3.1* (Displacement interpolation for rays in the non-branching case)**

If X is non-branching, any geodesic ray (μ_t) of $\mathscr{W}_2(X)$ admits a unique displacement interpolation, that is a probability measure μ on $\mathscr{R}(X)$ such that μ_t is the law of the time t of a random ray having law μ (in other words, such that $\mu_t = (e_t)_\#\mu$).

Note that since (μ_t) is a geodesic, the time restriction of μ to any segment is an optimal dynamical transport plan.

Proof. — Since X is non-branching and (μ_t) is defined for all positive times, we know that there is a unique optimal dynamical transport plan $\tilde{\mu}(T)$ from μ_0 to μ_T, and that $\tilde{\mu}(T) = \tilde{\mu}(T')^{0,T}$ whenever $T < T'$ (see the discussion after Proposition 2.9). The fact that X is non-branching also implies that for any two points $x, y \in X$ there is a unique maximal geodesic ray $F(x,y)$ starting at x and passing at y at time 1; moreover this ray depends continuously and therefore measurably on (x,y). In other words, $(r_{0,1})_\#$ identifies $\mathscr{P}(\mathscr{G}^{0,T}(X))$ and $\mathscr{P}(\mathscr{G}^{0,1}(X))$. It follows that all $\tilde{\mu}(T)$ are uniquely defined by $\tilde{\mu}(1)$ and that the probability measure $\mu = F_\#(\Pi)$, where Π is the optimal transport plan associated to $\tilde{\mu}(1)$, has the required property. □

When X is branching, the previous proof fails for two reasons. The first one is that we cannot determine $\tilde{\mu}(T)$ and define μ from $\tilde{\mu}(1)$ alone; although Prokhorov's theorem will do the trick. The second problem is that there may exist several optimal dynamical transport plans corresponding to the same geodesic; but the set of these transports is always compact and a diagonal process will solve the problem. However, we lose uniqueness in the process and it would be interesting to single out one of the dynamical transport plan obtained.

***Proposition 3.2* (Displacement interpolation for rays)**

Any geodesic ray of $\mathscr{W}_2(X)$ admits a displacement interpolation.

Proof. — Let $(\mu_t)_{t\geqslant 0}$ be a geodesic ray in $\mathscr{W}_2(X)$ and for all $T>0$, let $M(T)$ be the set of all optimal dynamical transport plans parametrized on $[0,T]$ that induce the geodesic segment $(\mu_t)_{0\leqslant t\leqslant T}$. It is a compact set according to [**Vil09**], Corollary 7.22. For all $T\in\mathbb{N}$, choose $\tilde\mu(T)$ in $M(T)$ and for all integer $0\leqslant T'\leqslant T$ define $\tilde\mu(T'|T)$ as the restriction $\tilde\mu(T)^{0,T'}$.

By a diagonal process, one can extract an increasing sequence T_k of integers such that for all $T'\in\mathbb{N}$, $\tilde\mu(T'|T_k)$ has a limit $\bar\mu(T')\in M(T')$ when $k\to+\infty$. Since for all $T'<T$ the restriction maps

$$p^{T,T'}:\mathscr{G}^{0,T}\to\mathscr{G}^{0,T'}$$

are continuous, we get that $p^{T,T'}(\bar\mu(T))=\bar\mu(T')$. We therefore have a projective system of measures; the projection map

$$p^T:\mathscr{R}(X)\to\mathscr{G}^{0,T}(X)$$

commutes with the $p^{T,T'}$ thus, according to a variant of Prokhorov's theorem 2.3 in the setting of projective system of measures [**Sch70**], if we prove tightness, *i.e.* that for all $\varepsilon>0$ there is a compact $K\subset\mathscr{R}(X)$ such that for all T, $\bar\mu(T)(p^TK)\geqslant 1-\varepsilon$, then we can conclude that there is a unique measure μ on $\mathscr{R}(X)$ such that $p^T_\#\mu=\bar\mu(T)$ for all T. This measure will have the required property since $\bar\mu(T)\in M(T)$.

Fix any $\varepsilon>0$. Let K_0,K_1 be compact subsets of X such that $\mu_i(K_i)\geqslant 1-\varepsilon/2$ for $i=0,1$. Let K be the compact subset of $\mathscr{R}(X)$ consisting in all geodesic rays starting in K_0 and whose time 1 is in K_1. Then for all $T>1$, $\bar\mu(T)(p^T(K))\geqslant 1-\varepsilon$, as needed. □

The following result shall make displacement interpolation particularly useful.

***Lemma 3.3* (lifting).** — *Let μ, σ be probability measures on $\mathscr{R}(X)$ (or similarly $\mathscr{G}^{T,T'}(X)$, ...) and denote by $\mu_t=(e_t)_\#\mu$ and $\sigma_t=(e_t)_\#\sigma$ their time t.*

Any transport plan $\Pi_t\in\Gamma(\mu_t,\sigma_t)$ admits a lift*, that is a transport plan $\Pi\in\Gamma(\mu,\sigma)$ such that $\Pi_t=(e_t,e_t)_\#\Pi$.*

Note that, as shall be apparent in the proof, the same holds for example with the map (e_s,e_t) when $s\neq t$.

Proof. — Disintegrate μ along μ_t: there is a family $(\zeta_x)_{x\in X}$ of probability measures on $\mathscr{R}(X)$, each one supported on the set $e_t^{-1}(x)$ of geodesic rays passing at x at time t, such that $\mu=\int\zeta_x\,\mu_t(dx)$ in the sense that

$$\mu(A)=\int_{\mathscr{R}(X)}\zeta_x(A)\mu_t(dx)$$

for all measurable A. Similarly, write $\sigma = \int \xi_y \, \sigma_t(dy)$ the disintegration of σ along σ_t.

Define then

$$\Pi(A \times B) = \int_{\mathscr{R}(X)^2} \zeta_x(A)\xi_y(B)\Pi_t(dxdy).$$

It is a probability measure on $\mathscr{R}(X)^2$, and for any measurable sets A, B in X we have $\Pi(e_t^{-1}(A), e_t^{-1}(B)) = \Pi_t(A \times B)$ because $\zeta_x\left(e_t^{-1}(A)\right)$ is 1 if $x \in A$, 0 otherwise (and similarly for $\xi_y\left(e_t^{-1}(B)\right)$). A similar computation gives that Π has marginals μ and σ. □

Note that we gave the proof for the sake of completness, but one could simply apply twice the gluing lemma, after noticing that the projection e_t from $\mathscr{R}(X)$ to X gives deterministic transport plans in $\Gamma(\mu, \mu_t)$ and $\Gamma(\sigma, \sigma_t)$.

The lift of Π_t need not be unique; the one constructed in the proof is very peculiar, and can be called the *most independent lift* of Π_t. It is well defined since other disintegration families $(\zeta'_x)_x$ and $(\xi'_y)_y$ must coincide with $(\zeta_x)_x$ and $(\xi_y)_y$ for μ_t-almost all x and σ_t-almost all y respectively.

The lifting lemma shall be used to translate the optimal transport problems between μ_t and σ_t, where these measures move (usually along geodesics) to transport problems between the fixed μ and σ, where it is the cost that moves. In other words, we have just shown that minimizing $\int c(x, y) \, \Pi_t$ over the $\Pi_t \in \Gamma(\mu_t, \sigma_t)$ is the same than minimizing $\int c(\gamma_t, \beta_t) \, \Pi(d\gamma d\beta)$ over the $\Pi \in \Gamma(\mu, \sigma)$.

4. Asymptotic measures

Let us denote by e_∞ the map defined by the formula

$$\begin{array}{rccc} e_\infty : & \mathscr{R}(X) & \longrightarrow & c\partial X \\ & \gamma & \longmapsto & ([\gamma^1], s(\gamma)) \end{array} \tag{4}$$

where γ^1 is the unitary reparametrization of γ, $[\gamma^1]$ is its asymptote class and $s(\gamma)$ is the speed of γ. It is to be understood that whenever $s(\gamma) = 0$, $[\gamma^1]$ can be taken arbitrarily in ∂X and this choice does not matter.

Definition 4.1 **(asymptotic measure)**. — Let $(\mu_t)_{t \geqslant 0}$ be a geodesic ray in $\mathscr{W}_2(X)$ and μ be a displacement interpolation (so that $\mu_t = e_{t\#}\mu$). We define the *asymptotic measure* of the ray by

$$\mu_\infty := e_{\infty\#}\mu.$$

We denote by $\mathscr{P}_1(c\partial X)$ the set of probability measures ν on $c\partial X$ such that $\int v^2 \nu(dv) = 1$.

In the branching case, the dynamical optimal transport plan is not unique in general. Therefore, the asymptotic measure depends *a priori* on the choice of the dynamical optimal transport plan. We will see soon that it is not the case.

Note that the speed of the geodesic (μ_t) is

$$\left(\int s^2(\gamma)\,\mu(d\gamma)\right)^{1/2} = \left(\int v^2\, s_\#\mu(dv)\right)^{1/2}$$

and we denote it by $s(\mu)$. In particular, $\mathscr{P}_1(c\partial X)$ is the set of measures that correspond to unit speed geodesics. We shall use that the speed function s defined in $\mathscr{R}(X)$ is in $L^2(\mu)$ several times.

The main result of this section is the following.

***Theorem 4.2* (asymptotic formula).** — *Consider two geodesic rays $(\mu_t)_{t\geqslant 0}$ and $(\sigma_t)_{t\geqslant 0}$, let μ and σ be any of their displacement interpolations and μ_∞, σ_∞ be the corresponding asymptotic measures. Then (μ_t) and (σ_t) are asymptotic if and only if $\mu_\infty = \sigma_\infty$, and we have*

$$\lim_{t\to\infty} \frac{\mathrm{W}(\mu_t,\sigma_t)}{t} = \mathrm{W}_\infty(\mu_\infty,\sigma_\infty).$$

Therefore: as in X itself, the distance between two rays is either bounded or of linear growth, and two displacement interpolations of the same ray define the same asymptotic measure.

The rôle of the asymptotic formula goes far beyond justifying Definition 4.1 in the branching case: it gives us a very good control on geodesic rays of $\mathscr{W}_2(X)$ on which several of our results rely. To cite one, the asymptotic formula is the main ingredient of Theorem 1.1 on the rank of $\mathscr{W}_2(X)$.

For every $t \geqslant 0$, let d_t be the function defined on $\mathscr{R}(X)\times\mathscr{R}(X)$ by $d_t(\gamma,\beta) = d(\gamma_t,\beta_t)$. We start with an implementation of a classical principle.

***Lemma 4.3*.** — *The function d_t is in $L^2(\Gamma(\mu,\sigma))$, by which we mean that there is a constant $C = C(\mu,\sigma)$ such that for all $\Pi\in\Gamma(\mu,\sigma)$, $\int d_t^2\,\Pi \leqslant C$.*

In the following, it will be of primary importance that C does not depend on Π.

Proof. — Denoting by x any base point in X we have

$$\begin{aligned}\int d_t^2\,\Pi &\leqslant \int (d(\gamma_t,x)+d(x,\beta_t))^2\;\Pi(d\gamma d\beta)\\ &\leqslant 2\int \left(d^2(\gamma_t,x)+d^2(x,\beta_t)\right)\;\Pi(d\gamma d\beta)\\ &= 2\int d^2(\gamma_t,x)\,\mu(d\gamma)+2\int d^2(x,\beta_t)\,\sigma(d\beta)\\ &= 2\,\mathrm{W}^2(\mu_t,\delta_x)+2\,\mathrm{W}^2(\sigma_t,\delta_x).\end{aligned}$$

□

Proof of the asymptotic formula. — We shall use several times the following translation of the convexity of the distance function: given any geodesic rays γ,β of X, the function

$$f_t(\gamma,\beta):=\frac{d(\gamma_t,\beta_t)-d(\gamma_0,\beta_0)}{t}$$

is nondecreasing in t and has limit $d_\infty(\gamma,\beta)$.

Assume first that $\mu_\infty=\sigma_\infty$, and let us prove that $\mathrm{W}(\mu_t,\sigma_t)$ is bounded. The lifting lemma gives us a transport plan $\Pi\in\Gamma(\mu,\sigma)$ such that for all (γ,β) in its support, $\gamma_\infty=\beta_\infty$ (simply observe that the lifting Lemma applies equally well to $t=\infty$, and lift the trivial transport $(\mathrm{Id}\times\mathrm{Id})_\#\mu_\infty$). If $f_t(\gamma,\beta)$ is positive for some t, then $d_\infty(\gamma,\beta)>0$. It follows that on $\operatorname{supp}\Pi$, $d_t\leqslant d_0$. Therefore:

$$\mathrm{W}(\mu_t,\sigma_t)\leqslant\int d_t^2\,\Pi\leqslant\int d_0^2\,\Pi$$

which is bounded by the previous lemma.

Let now Π_∞ be a transport plan from μ_∞ to σ_∞ that is optimal with respect to d_∞. Such a minimizer exists since d_∞ is non-negative and lower-semicontinuous with respect to the cone topology of $c\partial X$; note that taking an almost minimizer would be sufficient anyway. Denote by $\tilde\Pi$ a lift of Π_∞ to $\Gamma(\mu,\sigma)$; then

$$\frac{\mathrm{W}^2(\mu_t,\sigma_t)}{t^2}\leqslant\int\frac{d_t^2}{t^2}\,\tilde\Pi.$$

We have $2\geqslant d_\infty\geqslant f_t\geqslant f_1\geqslant -d_0$ for all $t\geqslant 1$, so that $t^{-1}d_t$ is bounded by $2+d_0$ and $-d_0$. We can thus apply the dominated convergence theorem, which gives

$$\limsup\frac{\mathrm{W}^2(\mu_t,\sigma_t)}{t^2}\leqslant\int d_\infty^2\,\tilde\Pi=\mathrm{W}_\infty^2(\mu_\infty,\sigma_\infty).$$

To prove the other inequality, we introduce $g_t:=\max(0,f_t)$. It is a nondecreasing, nonnegative function with d_∞ as limit, and it satisfies $t^2g_t^2\leqslant d_t^2$.

Let Π_t be an optimal transport plan between (μ_t) and (σ_t) and $\tilde{\Pi}_t$ be a lift to $\Gamma(\mu,\sigma)$, which by Prokhorov Theorem is compact in the weak topology. Let $(t_k)_k$ be an increasing sequence such that

$$\lim_k \mathrm{W}(\mu_{t_k},\sigma_{t_k}) = \liminf_t \mathrm{W}(\mu_t,\sigma_t)$$

and $\tilde{\Pi}_{t_k}$ weakly converges to some $\tilde{\Pi}_\infty$.

For all $k' < k$, we have

$$\begin{aligned} \frac{\mathrm{W}^2(\mu_{t_k},\sigma_{t_k})}{t_k^2} &= \frac{1}{t_k^2}\int d_{t_k}^2\, \tilde{\Pi}_{t_k} \\ &\geqslant \int g_{t_k}^2\, \tilde{\Pi}_{t_k} \\ &\geqslant \int g_{t_{k'}}^2\, \tilde{\Pi}_{t_k}. \end{aligned}$$

Letting $k \to \infty$ and using that the g_t are continuous, we obtain

$$\liminf_t \frac{\mathrm{W}^2(\mu_t,\sigma_t)}{t^2} \geqslant \int g_{t_{k'}}^2\, \tilde{\Pi}_\infty$$

for all k'. But $g_{t_{k'}} \leqslant 2$ and the dominated convergence theorem enables us to let $k' \to \infty$:

$$\liminf_t \frac{\mathrm{W}^2(\mu_t,\sigma_t)}{t^2} \geqslant \int d_\infty^2\, \tilde{\Pi}_\infty \geqslant \mathrm{W}^2(\mu_\infty,\sigma_\infty).$$

This ends the proof of the asymptotic formula, and shows that if (μ_t) and (σ_t) stay at bounded distance then $\mu_\infty = \sigma_\infty$. □

5. Complete geodesics and the rank

In this section we study complete geodesics in $\mathscr{W}_2(X)$, in particular to understand its rank. Recall that the *rank* of a metric space is the highest dimension of a Euclidean space that embeds isometrically in it.

The main result of this section is the following.

Theorem 5.1. — *If X is a visibility space, in particular if it is* $\mathrm{CAT}(\kappa)$ *with $\kappa < 0$, then $\mathscr{W}_2(X)$ has rank* 1.

We expect more generally that for most, if not all Hadamard space X, the rank of $\mathscr{W}_2(X)$ is equal to the rank of X. Theorem 5.1 is a first step in this direction. Note that the fact that we ask the embedding to be nothing weaker than an isometry and to be global is important, as shown by example 5.5 at the end of the section.

To prove Theorem 5.1, let us first see that the asymptotic measure of a complete geodesic is much more constrained than that of a mere ray.

Proposition 5.2. — *Let $(\mu_t)_{t\in\mathbb{R}}$ be a complete unit speed geodesic of $\mathscr{W}_2(X)$, and μ be one of its displacement interpolations. Then μ is concentrated on the set of* unit speed *geodesics of X.*

Note that here we do not use any assumption on X (besides the existence of displacement interpolations).

Proof. — Let γ, β be two geodesics in the support of μ and let $a = s(\gamma)$ and $b = s(\beta)$. Fix some point $x \in X$. Then we have the equivalents $d(x, \gamma_t) \sim at$ and $d(x, \beta_t) \sim bt$ when $t \to \pm\infty$. In particular, we get that

$$\begin{aligned} d^2(\gamma_t, \beta_{-t}) &\leqslant (d(\gamma_t, x) + d(x, \beta_{-t}))^2 \\ &\leqslant (a+b)^2 t^2 + o(t^2) \end{aligned}$$

and similarly $d^2(\beta_t, \gamma_{-t}) \leqslant (a+b)^2 t^2 + o(t^2)$. We also have $d^2(\gamma_t, \gamma_{-t}) = 4a^2t^2$ and $d^2(\beta_t, \beta_{-t}) = 4b^2t^2$. But since (μ_t) is a geodesic, the transport plan $\Pi^t \in \Gamma(\mu_t, \mu_{-t})$ induced by μ must respect the cyclical monotonicity. In particular we have

$$d^2(\gamma_t, \gamma_{-t}) + d^2(\beta_t, \beta_{-t}) \leqslant d^2(\gamma_t, \beta_{-t}) + d^2(\beta_t, \gamma_{-t}).$$

From this and letting $t \to \infty$ we get $4a^2 + 4b^2 \leqslant 2a^2 + 2b^2 + 4ab$, which is only possible when $a = b$.

We proved that the speed of geodesics in the support of μ is constant, and since their square integrates (with respect to μ) to 1 we get the desired conclusion. □

As a consequence, the asymptotic measure of a unit speed ray that can be extended to a complete geodesic lies in the subset $\mathscr{P}(\partial X)$ of $\mathscr{P}_1(c\partial X)$ (where we identify a space Y with the level $Y \times \{1\}$ in its cone).

Proposition 5.3. — *If X is a visibility space, the space $\mathscr{P}(\partial X)$ endowed with the metric W_∞ (where a ray is identified with its asymptote class) contains no non-constant rectifiable curve.*

Proof. — First, the visibility assumption implies that whenever γ and β are non asymptotic, unit speed geodesic rays of X, we have $d_\infty(\gamma, \beta) = 2$. Let us prove that for all displacement interpolation μ, σ of rays in $\mathscr{W}_2(X)$, assumed to be concentrated on $\mathscr{R}_1(X)$, we have

$$\mathrm{W}_\infty(\mu_\infty, \sigma_\infty) = 2|\mu_\infty - \sigma_\infty|_v^{1/2} \tag{5}$$

where $|\cdot|_v$ is the total variation norm. Let us recall that, using the Jordan measure decomposition of $\mu_\infty - \sigma_\infty$, one can find positive (not probability) measures μ'_∞, σ'_∞ and ν such that μ'_∞ and σ'_∞ are mutually singular, $\mu_\infty = \nu + \mu'_\infty$ and $\sigma_\infty = \nu + \sigma'_\infty$. By definition, $|\mu_\infty - \sigma_\infty|_v$ is the total mass of μ'_∞ (or, equivalently, of σ'_∞). To find an optimal transport plan, one can simply leave the common mass in place and move arbitrarily what is left, for example taking

$$\Pi = (\mathrm{Id}, \mathrm{Id})_{\#}\nu + \mu'_\infty \otimes \sigma'_\infty$$

where $\otimes$ denotes the product measure *normalized to have the same mass as each factor*. Since Π moves a mass $|\mu_\infty - \sigma_\infty|_v$ by a distance of 2, it has cost $4|\mu_\infty - \sigma_\infty|_v$. More generally, due to the behavior of d_∞, any transport plan Π' has cost

$$4\Pi'(\{(\zeta, \xi) \,|\, \zeta \neq \xi\})$$

which cannot be smaller than the cost of Π.

The proposition now results from the more general following lemma, which is well-known at least in the case of Euclidean space.

***Lemma 5.4* (Snowflaked metrics).** — *Let (Y, d) be any metric space and $\alpha < 1$ be a positive number. Then (Y, d^α) is a metric space not containing any non-constant rectifiable curve.*

The fact that d^α is a metric comes from the inequality $a + b \leqslant (a^\alpha + b^\alpha)^{1/\alpha}$ for positive a, b.

Let $c : I \to Y$ be a non-constant curve. Up to restriction and reparametrization, we can assume that $I = [0, 1]$ and $c_0 \neq c_1$. Take any positive integer n; since d is continuous, by the intermediate value theorem there are numbers $t_1 = 0 < t_2 < \cdots < t_n < 1$ such that $d(c_{t_{i-1}}, c_{t_i}) = d(c_0, c_1)/n$ and $d(c_{t_n}, c_1) \geqslant d(c_0, c_1)/n$. Denoting by ℓ the length according to the "snowflaked" metric d^α, we get that

$$\ell(c) \geqslant n \left(\frac{d(c_0, c_1)}{n}\right)^\alpha \geqslant d^\alpha(c_0, c_1) n^{1-\alpha}.$$

Since this holds for all n, $\ell(c) = \infty$ and c is not rectifiable. □

Proof of Theorem 5.1. — Assume that there is an isometric embedding $\varphi : \mathbb{R}^2 \to \mathscr{W}_2(X)$. Let r^θ be the ray starting at the origin and making an angle θ with some fixed direction. Then, since r^θ extends to a complete geodesic, so does $\varphi \circ r^\theta$. The displacement interpolation μ^θ of this ray of $\mathscr{W}_2(X)$ must be concentrated on $\mathscr{R}_1(X)$ by Proposition 5.2, so that $\mu^\theta_\infty \in \mathscr{P}(\partial X)$. But φ being isometric, the map $\theta \to \mu^\theta_\infty$ should be an isometric embedding from the boundary of $\mathbb{R}^2$ (that is, the unit circle endowed with the chordal metric) to $(\mathscr{P}(\partial X), \mathrm{W}_\infty)$. In particular its image would be a non-constant rectifiable curve, in contradiction with Proposition 5.3. □

Note that the same method yields more general results: we can rule out the isometric embedding of Minkowski planes ($\mathbb{R}^2$ endowed with any norm), and of their cones of the form $\{x^2 < \varepsilon y^2\}$ for any $\varepsilon > 0$. This contrasts with above-mentionned fact that even when X is reduced to a line, some Euclidean *half*-cones of arbitrary dimension embeds isometrically in $\mathscr{W}_2(X)$.

Example 5.5. — Let us remind an example of [**Klo10**] showing that there are plenty of weaker Euclidean embedding in most Wasserstein spaces. Consider the set $\mathbb{R}^n_<$ of increasingly ordered real n-tuples. The map

$$\begin{array}{rcl} f:\mathbb{R}^n_< & \to & \mathscr{W}_2(\mathbb{R}) \\ (x_1,\dots,x_n) & \mapsto & \frac{1}{n}\delta_{\sqrt{n}x_1}+\cdots+\frac{1}{n}\delta_{\sqrt{n}x_n} \end{array}$$

is an isometric embedding. Assume X contains a complete geodesic; then its Wasserstein space contains a copy of $\mathscr{W}_2(\mathbb{R})$ so that $\mathbb{R}^n_<$ embeds isometrically into $\mathscr{W}_2(X)$. But $\mathbb{R}^n_<$ is an open half-cone invariant under a 1-parameter group of translations so that:

– $\mathscr{W}_2(X)$ contains round Euclidean half-cones of arbitrary dimension, in particular $\mathbb{R}^{n-1}$ admits bi-Lipschitz emdedding in $\mathscr{W}_2(X)$ for arbitrary n,

– the cylinder $\mathbb{R}\times B^n(r)$ where $B^n(r)$ is a Euclidean ball of arbitrary radius r embeds isometrically in $\mathscr{W}_2(X)$ for all n.

We see that under a very mild assumption, most weak ranks of $\mathscr{W}_2(X)$ are infinite (examples of weak ranks include the largest dimension of a bi-Lipshitz embedded Euclidean space and the largest rank of a quasi-isometrically embedded $\mathbb{Z}^n$).

Another customary definition of rank for Hadamard manifolds uses Jacobi fields; one could try to extend the notion of Jacobi fields in the setting of Wasserstein spaces, but our example seems to indicate that with such a definition even $\mathscr{W}_2(\mathbb{R})$ should have infinite rank. Let us give a precise result showing this.

Proposition 5.6. — *Given any geodesic (μ_t) in $\mathscr{W}_2(\mathbb{R})$ and any $n\in\mathbb{N}$, there are independent unit vectors $v_1,\dots,v_n$ in $\mathbb{R}^n$ and a positive ε such that the set*

$$C=\{tv_1+s_2v_2+\cdots+s_nv_n\,|\,t\in\mathbb{R}, s_i\in[0,\varepsilon)\}$$

embeds isometrically in $\mathscr{W}_2(\mathbb{R})$, with tv_1 sent to μ_t for all t.

In particular, for any sensible metric definition of Jacobi fields, a geodesic admits arbitrarily many independent pairwise commuting Jacobi fields.

Proof. — It is easy to see and proved in [**Klo10**] that there is some measure μ on $\mathbb{R}$ such that $\mu_t = T_{t\#}\mu$ for all t, where T_t is the translation $x\mapsto x+t$.

If μ has finite support, then an embedding similar to the one described in example 5.5 gives the conclusion. Assume that the support of μ contains at least $n-1$ points $x_2,\dots,x_n$ and choose disjoint neighborhoods $U_2,\dots,U_n$ of these points. Let $X_2,\dots,X_n$ be smooth vector fields on $\mathbb{R}$, each X_i having support in U_i, define $\Phi(t,s_2,\dots,s_n)(x) = x+t+\sum_i s_i X_i(x)$ and consider the map

$$\begin{array}{rcl} f:\mathbb{R}\times[0,+\infty)^{n-1} & \to & \mathscr{W}_2(\mathbb{R}) \\ (t,s_2,\dots,s_n) & \mapsto & \Phi(t,s_2,\dots,s_n)_\#\mu \end{array}$$

Using convexity of the cost and cyclical monotonocity, one sees that optimal transport plans in $\mathscr{W}_2(\mathbb{R})$ are exactly those where no inversion of mass occurs (i.e. whose support does not contain pairs (x,y) and (x',y') such that $x>x'$ and $y<y'$). In particular, the push forward by the map $\Phi(t,s_2,\dots,s_n)$ defines an optimal transport plan between a measure and its image as soon as s_i's are small enough. It follows that for small enough s_i's we have

$$\begin{aligned} \mathrm{W}\left(f(t,(s_i)),f(t',(s_i'))\right)^2 =& (t-t')^2+\sum_i (s_i-s_i')^2\left(\int_{U_i} X_i^2(x)\,\mu(dx)\right) \\ &+2(t-t')\sum_i (s_i-s_i')\int_{U_i} X_i(x)\,\mu(dx) \end{aligned}$$

which is a quadratic expression in $(t-t',s_2-s_2',\dots,s_n-s_n')$. Modifying f by a linear change of coordinates, we get the desired embedding. □

Note that when the support of μ contains at least n points, we can in fact construct a n dimensional uniformly large Euclidean neighborhood of (μ_t). However, we only get the announced "corner" when the support of μ is too small, in particular when μ is a Dirac mass.

As a last remark, let us point out that the Jacobi fields constructed here have well-defined flows only for small times. The study (and definition) of fully integrable Jacobi fields on $\mathscr{W}_2(X)$ could lead to an understanding of its flats even when X has higher rank.

6. The geodesic boundary and its cone topology

In this section, we adapt to $\mathscr{W}_2(X)$ the classical construction of the cone topology on the geodesic compactification of Hadamard space, see for instance [**Bal95**]. We introduce this topology on $\overline{\mathscr{W}_2(X)} = \mathscr{W}_2(X)\cup\partial\mathscr{W}_2(X)$. We shall prove in Proposition 6.1 that the cone topology turns $\overline{\mathscr{W}_2(X)}$ into a first-countable Hausdorff space and that the topology induced on $\mathscr{W}_2(X)$ coincides with the topology derived from the Wasserstein metric.

In the next section, we shall show in Theorem 7.2 that $\partial\mathscr{W}_2(X)$ is homeomorphic to

$$\mathscr{P}_1(c\partial X) = \left\{\zeta\in\mathscr{P}(c\partial X);\int s^2\,\zeta(d\xi,ds)=1\right\}$$

endowed with the weak topology. In Corollary 7.3, we rewrite the above result in terms of Wasserstein space over $c\partial X$.

From now on, we will use the following notations. Let (Y, d) be a geodesic space and $y \in Y$. We set $\mathscr{R}_y(Y)$ (respectively $\mathscr{R}_{y,1}(Y)$) the set of geodesic rays in Y starting at y (respectively the set of unitary geodesic rays starting at y). These sets are closed subsets of $\mathscr{R}(Y)$ endowed with the topology of uniform convergence on compact subsets.

The cone topology on $\overline{\mathscr{W}_2(X)} = \mathscr{W}_2(X) \cup \partial\mathscr{W}_2(X)$ is defined by using as a basis the open sets of $\mathscr{W}_2(X)$ together with

$$U(x, \xi, R, \varepsilon) = \left\{\theta \in \overline{\mathscr{W}_2(X)}; \theta \notin \overline{B}(\delta_x, R), \mathrm{W}((\mu_{\delta_x,\theta})_R, (\mu_{\delta_x,\xi})_R) < \varepsilon\right\}$$

where $x \in X$ is a fixed point, ξ runs over $\partial\mathscr{W}_2(X)$, R and ε run over $(0, +\infty)$ and $\mu_{\delta_x,\theta}$ is the unitary geodesic between δ_x and θ (existence and uniqueness follow from Lemma 6.3).

The main properties of the cone topology on $\overline{\mathscr{W}_2(X)}$ are gathered together in the following proposition.

Proposition 6.1. — *The cone topology on $\overline{\mathscr{W}_2(X)}$ is well-defined and is independent of the choice of the basepoint δ_x. Moreover, endowed with this topology, $\overline{\mathscr{W}_2(X)}$ is a first-countable Hausdorff space. By definition, the topology induced on $\mathscr{W}_2(X)$ coincides with the topology derived from the Wasserstein metric.*

Remark 6.2. — We emphasize that the topology induced on $\partial\mathscr{W}_2(X)$ by the cone topology coincides with the quotient topology induced by the topology of uniform convergence on compact subsets on the set of unitary rays in $\mathscr{W}_2(X)$.

Moreover, since $\partial\mathscr{W}_2(X)$ endowed with the cone topology is first-countable, continuity and sequential continuity are equivalent in this topological space.

The scheme of proof is the same as in the nonpositively curved case. However, to get the result, we first need to generalize to our setting some properties related to nonpositive curvature.

Lemma 6.3. — *Given $x \in X$, the set of unitary rays in $\mathscr{W}_2(X)$ starting at δ_x is in one-to-one correspondence with the set $\mathscr{P}_1(c\partial X)$. Moreover, for any $\xi \in \partial\mathscr{W}_2(X)$, there exists a unique unitary ray starting at δ_x and belonging to ξ.*

Proof. — Recall that there exists a unique transport plan between a Dirac mass and any measure in $\mathscr{W}_2(X)$. Since there is a unique geodesic between two given points in X, the same property remains true for dynamical transportation plans. Using the previous remarks, we get that any $\mu \in \mathscr{P}(\mathscr{R}_x(X))$ such that $\int s^2(\gamma)\mu(d\gamma) < +\infty$ induces a ray starting at δ_x. Moreover, since displacement interpolation always exists (see Proposition 3.2), the set $\mathscr{R}_{\delta_x,1}(\mathscr{W}_2(X))$

is in one-to-one correspondence with the unitary dynamical transportation plans starting at δ_x, namely with the measures $\mu \in \mathscr{P}(\mathscr{R}_x(X))$ such that $\int s^2(\gamma)\mu(d\gamma) = 1$. Now, since X is a Hadamard space, we recall that $\mathscr{R}(X)$ is homeomorphic to $X \times c\partial X$ (where the left coordinate is the initial location of the ray). Therefore, for any $x \in X$, the previous map induces a homeomorphism

$$\phi_x : \mathscr{R}_x(X) \longrightarrow c\partial X. \tag{6}$$

This gives us a one-to-one correspondence between the set of unitary dynamical transportation plans starting at δ_x and the set $\mathscr{P}_1(c\partial X)$.

Given $\xi \in \partial \mathscr{W}_2(X)$, consider a unit ray (μ_t) in $\mathscr{W}_2(X)$ belonging to ξ and $\mu_\infty \in \mathscr{P}_1(c\partial X)$ its asymptotic measure. We claim that ${\phi_x^{-1}}_{\#}\mu_\infty$ is the unique ray starting at δ_x and belonging to ξ. Indeed, ${\phi_x^{-1}}_{\#}\mu_\infty$ and (μ_t) have the same asymptotic measure, thus they are asymptotic thanks to the asymptotic formula (Theorem 4.2). The asymptotic formula also implies that two asymptotic rays (starting at δ_x) have the same asymptotic measure, thus they are equal thanks to the first part of the lemma. ☐

Lemma 6.4. — *Let $(\mu_t), (\sigma_t)$ be two unitary geodesics (possibly rays) in $\mathscr{W}_2(X)$ starting at δ_x. Then, the comparison angle $\widetilde{\angle}_{\mu_s\delta_x\sigma_t}$ at δ_x of the triangle $\Delta(\delta_x, \mu_s, \sigma_t)$ is a nondecreasing function of s and t. Consequently, the map $t \longrightarrow \mathrm{W}(\mu_t, \sigma_t)/t$ is a nondecreasing function as well.*

Proof. — We set $d_m, d_s \leqslant +\infty$ the length of (μ_t) and (σ_t) respectively and μ, σ the corresponding optimal dynamical plans. Thanks to the lifting lemma, we set $\Theta \in \Gamma(\mu, \sigma)$ a dynamical plan such that, for given $s \leqslant d_m$ and $t \leqslant d_s$, $(e_s, e_t)_{\#}\Theta$ is an optimal plan. By definition of the Wasserstein distance, we get, for any $s' \leqslant s$ and $t' \leqslant t$, the following estimate

$$\mathrm{W}^2(\mu_{s'}, \sigma_{t'}) \leqslant \int d^2(\gamma(s'), \gamma'(t'))\, \Theta(d\gamma, d\gamma').$$

Now, the fact that X is nonpositively curved yields

$$\begin{aligned} d^2(\gamma(s'), \gamma'(t')) \quad \leqslant \quad & \frac{s'^2}{s^2} d^2(x, \gamma(s)) + \frac{t'^2}{t^2} d^2(x, \gamma'(t)) \\ & - 2\frac{s't'}{st} d(x, \gamma(s)) d(x, \gamma'(t)) \cos \widetilde{\angle}_{\gamma(s)x\gamma'(t)}. \end{aligned}$$

where $\widetilde{\angle}_{\gamma(s)x\gamma'(t)}$ is the comparison angle at x (here, we use the fact that the initial measure is a Dirac mass). By integrating this inequality against Θ, we get

$$\begin{aligned}\mathrm{W}^2(\mu_{s'},\sigma_{t'}) \leqslant{} & s'^2+t'^2 \\ & -2\frac{s't'}{st}\int d(x,\gamma(s))d(x,\gamma'(t))\cos\widetilde{\angle}_{\gamma(s)x\gamma'(t)}\,\Theta(d\gamma,d\gamma').\end{aligned}$$

We conclude by noticing that the inequality above is an equality when $s=s'$ and $t=t'$, so we get

$$\mathrm{W}^2(\mu_{s'},\sigma_{t'}) \leqslant s'^2+t'^2-2s't'\cos\widetilde{\angle}_{\mu_s\delta_x\sigma_t}$$

which is equivalent to the property $\widetilde{\angle}_{\mu_{s'}\delta_x\sigma_{t'}} \leqslant \widetilde{\angle}_{\mu_s\delta_x\sigma_t}$. The remaining statement follows readily. □

Lemma 6.5. — *Let (μ_t) be a unitary geodesic, possibly a ray, starting at δ_y and $\delta_x \neq \delta_y$. For any $\theta \in \mathscr{W}_2(X)$ such that $\theta \neq \delta_x, \delta_y$, the comparison angle at θ satisfies*

$$\cos\widetilde{\angle}_{\delta_x\theta\delta_y} = \frac{1}{\mathrm{W}(\delta_x,\theta)\,\mathrm{W}(\delta_y,\theta)}\int d(x,z)d(y,z)\cos\widetilde{\angle}_{xzy}\,\theta(dz).$$

Moreover, given two nonnegative numbers $s \neq t$, the following inequality holds for $0<t<T$

$$\widetilde{\angle}_{\mu_0\mu_t\delta_x}+\widetilde{\angle}_{\mu_T\mu_t\delta_x} \geqslant \pi.$$

Proof. — For any $z \in X$, the following equality holds

$$d^2(x,y)=d^2(x,z)+d^2(y,z)-2d(x,z)d(y,z)\cos\widetilde{\angle}_{xzy}.$$

By integrating this inequality against θ, we get the first statement by definition of the comparison angle. Let μ be the unique optimal dynamical coupling that induces (μ_t). The first step of the proof is to get the equality below:

$$\begin{aligned}\cos\widetilde{\angle}_{\mu_T\mu_t\delta_x} ={} & \\ & \frac{1}{\mathrm{W}(\mu_T,\mu_t)\,\mathrm{W}(\mu_t,\delta_x)}\int d(\gamma(t),\gamma(T))d(\gamma(t),x)\cos\widetilde{\angle}_{x\gamma(t)\gamma(T)}\,\mu(d\gamma)\end{aligned} \tag{7}$$

For $\gamma \in \operatorname{supp}\mu$, the following equality holds

$$\begin{aligned}d^2(\gamma(T),x)={} & d^2(\gamma(t),x)+d^2(\gamma(t),\gamma(T)) \\ & -2d(\gamma(t),x)d(\gamma(t),\gamma(T))\cos\widetilde{\angle}_{x\gamma(t)\gamma(T)}.\end{aligned}$$

By integrating this equality against μ, we get (7). Now, using that X is nonpositively curved, we have $\widetilde{\angle}_{x\gamma(t)\gamma(T)} + \widetilde{\angle}_{x\gamma(t)y} \geqslant \pi$. This gives

$$\begin{aligned}
&\cos \widetilde{\angle}_{\mu_T \mu_t \delta_x} \\
&\qquad \leqslant \tfrac{-1}{\mathrm{W}(\mu_T,\mu_t)\,\mathrm{W}(\mu_t,\delta_x)} \int d(\gamma(t),\gamma(T))d(\gamma(t),x)\cos\widetilde{\angle}_{x\gamma(t)y}\,\mu(d\gamma) \\
&\qquad \leqslant \tfrac{-1}{(T-t)\,\mathrm{W}(\mu_t,\delta_x)} \int \tfrac{T-t}{t} d(\gamma(t),y)d(\gamma(t),x)\cos\widetilde{\angle}_{x\gamma(t)y}\,\mu(d\gamma) \\
&\qquad \leqslant \tfrac{-1}{\mathrm{W}(\mu_t,\delta_y)\,\mathrm{W}(\mu_t,\delta_x)} \int d(\gamma(t),y)d(\gamma(t),x)\cos\widetilde{\angle}_{x\gamma(t)y}\,\mu(d\gamma) \\
&\qquad \leqslant -\cos\widetilde{\angle}_{\mu_0\mu_t\delta_x}
\end{aligned}$$

where the last inequality follows from the first statement and the result is proved. □

As a consequence, we get the following result.

Proposition 6.6. — *Given $\varepsilon > 0$, $a > 0$, and $R > 0$, there exists a constant $T = T(\varepsilon, a, R) > 0$ such that the followings holds: for any $x, y \in X$ such that $d(x, y) = a$ and a unitary geodesic (possibly a ray) (μ_t) of length greater than T and starting at δ_y, if (σ^s_t) is the unitary geodesic from δ_x to μ_s then*

$$\mathrm{W}(\sigma^s_R, \sigma^{s'}_R) < \varepsilon$$

for any $s' > s > T$.

In particular, if (μ_t) is a ray and s goes to infinity, $(\sigma^s)_{s\geqslant 0}$ converges uniformly on compact subsets to the unitary ray $\mu_{\delta_x,\xi}$ where ξ is the asymptote class of (μ_t) (see Figure 3).

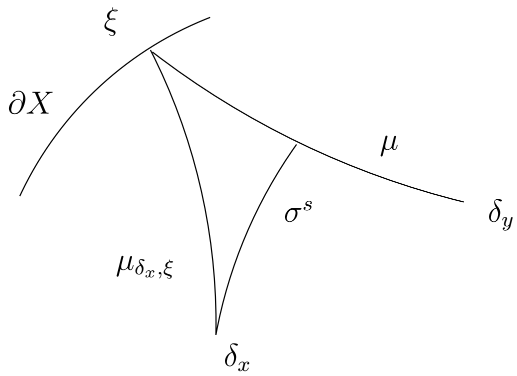

FIGURE 3. Uniform convergence of σ^s on compact subsets.

Proof. — Under these assumptions, the comparison angle $\widetilde{\angle}_{\delta_y \mu_s \delta_x}$ is arbitrary small provided s is sufficiently large. Consequently, thanks to Lemma 6.5, $\widetilde{\angle}_{\mu_{s'} \mu_s \delta_x}$ is close to π; therefore the comparison angle $\widetilde{\angle}_{\mu_s \delta_x \mu_{s'}}$ is small. This gives the first part of the result since

$$\widetilde{\angle}_{\sigma_R^s \delta_x \sigma_R^{s'}} \leqslant \widetilde{\angle}_{\mu_s \delta_x \mu_{s'}}$$

thanks to Lemma 6.4. Using Lemma 6.4 again, it only remains to prove the pointwise convergence of $(\sigma^s)_{s \geqslant 0}$ to $\mu_{\delta_x,\xi}$. Thanks to the asymptotic formula, there exists $C > 0$ such that

$$\mathrm{W}((\mu_{\delta_x,\xi})_t, \mu_t) \leqslant C \tag{8}$$

for any nonnegative number t. Finally, we conclude by using s' sufficiently large and the bound

$$\mathrm{W}((\mu_{\delta_x,\xi})_R, \sigma_R^s) \leqslant \mathrm{W}((\mu_{\delta_x,\xi})_R, \sigma_R^{s'}) + \mathrm{W}(\sigma_R^s, \sigma_R^{s'})$$

where the same reasoning as above and (8) show that the first term on the right-hand side is small provided s' is large. □

Now, we can prove that the topology above is well-defined and does not depend on the choice of the base point δ_x. This the content of the lemma below.

Lemma 6.7. — *Given two positive numbers R, ε and $y \in X$, $\xi \in U(x, \eta, R, \varepsilon) \cap \partial \mathscr{W}_2(X)$, there exists $S, \varepsilon' > 0$ such that*

$$U(y, \xi, S, \varepsilon') \subset U(x, \eta, R, \varepsilon).$$

Proof. — We set $\alpha = \varepsilon - \mathrm{W}((\mu_{\delta_x,\eta})_R, (\mu_{\delta_x,\xi})_R) > 0$. Let $\theta \in U(y, \xi, S, \varepsilon')$ and Θ (respectively Ξ) be the unitary geodesic $\mu_{\delta_y,\theta}$ (respectively the unitary ray $\mu_{\delta_y,\xi}$). We have

$$\begin{aligned}\mathrm{W}((\mu_{\delta_x,\theta})_R, (\mu_{\delta_x,\eta})_R) \leqslant\ & \mathrm{W}((\mu_{\delta_x,\theta})_R, (\mu_{\delta_x,\Theta_S})_R) + \mathrm{W}((\mu_{\delta_x,\Theta_S})_R, (\mu_{\delta_x,\Xi_S})_R) \\ & + \mathrm{W}((\mu_{\delta_x,\Xi_S})_R, (\mu_{\delta_x,\xi})_R) + \mathrm{W}((\mu_{\delta_x,\xi})_R, (\mu_{\delta_x,\eta})_R)\end{aligned}$$

The first and the third term on the right-hand side are smaller than $\alpha/3$ for large S thanks to Proposition 6.6 while the second term is smaller than $\alpha/3$ for large S and small ε' thanks to lemma 6.4. □

7. The boundary of $\mathscr{W}_2(X)$ viewed as a set of measures

To state the main result of this section, we first need to introduce a definition.

Definition 7.1. — We set

$$\begin{array}{rccc} Am: & \mathscr{R}_{\delta_x,1}(\mathscr{W}_2(X)) & \longrightarrow & \mathscr{P}_1(c\partial X) \\ & (\mu_t) & \longmapsto & \mu_\infty \end{array}$$

the map that sends a unitary ray starting at δ_x to its asymptotic measure.

The main result of this part is the following theorem.

Theorem 7.2. — *The map $Am : \mathscr{R}_{\delta_x,1}(\mathscr{W}_2(X)) \longrightarrow \mathscr{P}_1(c\partial X)$ induces a homeomorphism from $\partial\mathscr{W}_2(X)$ onto $\mathscr{P}_1(c\partial X)$.*

Note that a straightforward consequence of the result above is

Corollary 7.3. — *Let d be a metric on ∂X that induces the cone topology on ∂X and d_C the cone metric induced by d on $c\partial X$ (see (2) for a definition). Let us denote by $\mathscr{W}_2(c\partial X)$ the quadratic Wasserstein space over the Polish space $(c\partial X, d_C)$. Then, $\partial\mathscr{W}_2(X)$ is homeomorphic to the subset of probability measures with unitary speed in $\mathscr{W}_2(c\partial X)$.*

Remark 7.4. — In particular we get the more symmetric result that $c\partial\mathscr{W}_2(X)$ is homeomorphic to $\mathscr{W}_2(c\partial X)$.

The rest of this part is devoted to the proof of the theorem above. Recall that we have proved in Lemma 6.3 that both Am and the map $\widetilde{Am} : \partial\mathscr{W}_2(X) \to \mathscr{P}_1(c\partial X)$ it induces are bijective.

The proof of Theorem 7.2 is in two steps. First, we prove that the map Am is a homeomorphism. Then, we use this fact to prove that $\widetilde{Am}$ is a homeomorphism as well.

We start the proof with a definition.

Definition 7.5. — Let $x \in X$. We denote by

$$ODT_x \quad = \quad \Big\{\mu \in \mathscr{P}(\mathscr{R}_x(X));\int s^2(\gamma)\mu(d\gamma) = 1\Big\}$$

the set of unitary dynamical transport plans endowed with the weak topology. We also set

$$\begin{array}{rccc} (e_t)_{\geqslant 0\#}: & ODT_x & \longrightarrow & \mathscr{R}_{\delta_x,1}(\mathscr{W}_2(X)) \\ & \mu & \longmapsto & (\mu_t) \end{array}$$

and

$$\begin{array}{rccc} e_{\infty\#}: & ODT_x & \longrightarrow & \mathscr{P}_1(c\partial X) \\ & \mu & \longmapsto & \phi_{x\#}\mu \end{array}$$

where ϕ_x is defined in (6).

Thanks to Lemma 6.3, we have the following commutative diagram where all the maps are one-to-one.

$$\begin{array}{ccc} ODT_x & \xrightarrow{\quad e_{\infty\#} \quad} & \mathscr{P}_1(c\partial X) \\ {\scriptstyle (e_t)_{t\geqslant 0\#}} \downarrow & \nearrow_{Am} & \\ \mathscr{R}_{\delta_x,1}(\mathscr{W}_2(X)) & & \end{array}$$

We first prove that

Lemma 7.6. — *The map $e_{\infty\#}$ is a homeomorphism onto $\mathscr{P}_1(c\partial X)$.*

Proof. — The map $\phi_x : \mathscr{R}_x(X) \longrightarrow c\partial X$ is a homeomorphism. Therefore it induces a homeomorphism between ODT_x and $\mathscr{P}_1(c\partial X)$ when endowed with the weak topology. □

Lemma 7.7. — *The map $(e_t)_{\geqslant 0\#}$ is a continuous map.*

Proof. — Since the spaces we consider are metrizable, we just have to prove the sequential continuity. Consequently, we are given a sequence $(\mu_n)_{n\in\mathbb{N}}$ such that $\mu_n \rightharpoonup \mu$ in ODT_x. Now, since $e_t : \mathscr{R}_x(X) \longrightarrow X$ is a continuous map, we get that $e_{t\#}\mu_n \rightharpoonup e_{t\#}\mu$ in $\mathscr{P}(X)$. By definition of ODT_x, we have

$$\int s^2(\gamma)\,\mu_n(d\gamma) = \int s^2(\gamma)\,\mu(d\gamma) = 1.$$

Since $\int s^2(\gamma)\,\mu_n(d\gamma) = \int d^2(x,\gamma(1))\,\mu_n(d\gamma)$, the equality above implies the convergence of the second moment. Namely, we have

$$\int d^2(x,\gamma(t))\,\mu_n(d\gamma) = t^2 \int d^2(x,\gamma(1))\,\mu_n(d\gamma) = \int d^2(x,\gamma(t))\,\mu(d\gamma).$$

This implies the convergence of $e_{t\#}\mu_n$ to $e_{t\#}\mu$ with respect to the Wasserstein distance (see for instance [**Vil09**], Theorem 6.9). Thus, we have proved the pointwise convergence of rays. Now, since $t \longrightarrow \mathrm{W}(e_{t\#}\mu_n, e_{t\#}\mu)$ is nondecreasing as proved in Lemma 6.4, we get the result. □

We end the first part of the proof with the following lemma.

Lemma 7.8. — *The map $(e_t)_{\geqslant 0\#}$ is a homeomorphism.*

Proof. — Since the topology of both $\mathscr{R}_{\delta_x,1}(\mathscr{W}_2(X))$ and ODT_x is induced by a metric, it is sufficient to prove that $(e_t)_{\geqslant 0\#}$ is a proper map. Moreover, we just have to prove sequential compactness. We set K a compact subset of $\mathscr{R}_{\delta_x,1}(\mathscr{W}_2(X))$. Let $(\mu_n)_{n\in\mathbb{N}} \in (e_t)_{\geqslant 0\#}^{-1}(K)$. We first notice that $(\mu_n)_{n\in\mathbb{N}}$ is tight. Indeed, by assumption on K, the sequence $(e_{1\#}\mu_n)_{n\in\mathbb{N}}$ is tight in $\mathscr{P}(X)$. Therefore, by arguing as in the end of the proof of Proposition 3.2, we obtain the claim. Consequently, since $\mathscr{R}_x(X)$ is a Polish space, we can apply Prokhorov's theorem to get a converging subsequence $(\mu_{n_k})_{k\in\mathbb{N}}$ to $\widetilde{\mu}$. It remains to prove that $\widetilde{\mu} \in ODT_x$, namely that $\int s^2(\gamma)\,\widetilde{\mu}(d\gamma) = 1$. Since K is compact, we can also assume without loss of generality that $(e_{t\#}\mu_{n_k}) \longrightarrow (\bar{\mu}_t)$ in $\mathscr{R}_{\delta_x,1}(\mathscr{W}_2(X))$. Moreover, since $e_{t\#}\mu_{n_k} \rightharpoonup e_{t\#}\widetilde{\mu}$ for any t, we get $(e_t)_{\geqslant 0\#}(\widetilde{\mu}) = (\bar{\mu}_t)$. Therefore, $e_{1\#}\mu_{n_k} \longrightarrow e_{1\#}\widetilde{\mu}$ in $\mathscr{W}_2(X)$. This implies the convergence of the second moment $\int d^2(x,\gamma(1))\,\mu_{n_k}(d\gamma) = \int s^2(\gamma)\,\mu_{n_k}(d\gamma) = 1$ (see for instance [**Vil09**], Theorem 6.9) and the result is proved. □

We are now in position to prove Theorem 7.2. We set $p_{\partial\mathscr{W}_2}$ the canonical projection on $\partial\mathscr{W}_2(X)$. We have the following commutative diagram.

$$\begin{array}{ccc} \mathscr{R}_{\delta_x,1}(\mathscr{W}_2(X)) & \xrightarrow{\quad Am \quad} & \mathscr{P}_1(c\partial X) \\ {\scriptstyle p_{\partial\mathscr{W}_2}}\Big\downarrow & \nearrow_{\widetilde{Am}} & \\ \partial\mathscr{W}_2(X) & & \end{array}$$

We have seen at the beginning of the proof that all the maps above are one-to-one. To conclude, it remains to prove that $p^{-1}_{\partial\mathscr{W}_2}$ is a continuous map. Since $\partial\mathscr{W}_2(X)$ is first-countable (see Remark 6.2), it is sufficient to prove sequential continuity. To this aim, let $\xi_n \longrightarrow \xi$ in $\partial\mathscr{W}_2(X)$ and $(e_{t\#}\mu_n), (\mu_t) \in \mathscr{R}_{\delta_x,1}(\mathscr{W}_2(X))$ such that $p_{\partial\mathscr{W}_2}((e_{t\#}\mu_n)) = \xi_n$ and $p_{\partial\mathscr{W}_2}((\mu_t)) = \xi$. Recall that under these assumptions, the map $t \longrightarrow \mathrm{W}(e_{t\#}\mu_n, \mu_t)$ is nondecreasing (see Lemma 6.4), thus we just have to show the pointwise convergence of $(e_{t\#}\mu_n)$. This pointwise convergence follows readily from the definition of the cone topology on $\overline{\mathscr{W}_2(X)}$.

References

[Bal95] W. Ballmann – *Lectures on spaces of nonpositive curvature*, DMV Seminar, vol. 25, Birkhäuser Verlag, Basel, 1995, With an appendix by Misha Brin.

[BH99] M. R. Bridson & A. Haefliger – *Metric spaces of non-positive curvature*, Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 319, Springer-Verlag, Berlin, 1999.

[Bil99] P. Billingsley – *Convergence of probability measures*, second ed., Wiley Series in Probability and Statistics: Probability and Statistics, John Wiley & Sons Inc., New York, 1999, A Wiley-Interscience Publication.

[BK10] J. Bertrand & B. Kloeckner – "A geometric study of wasserstein spaces: Hadamard spaces", arXiv:1010.0590v1, 2010.

[Bou97] M. Bourdon – "Immeubles hyperboliques, dimension conforme et rigidité de Mostow", *Geom. Funct. Anal.* **7** (1997), no. 2, p. 245–268.

[BP99] M. Bourdon & H. Pajot – "Poincaré inequalities and quasiconformal structure on the boundary of some hyperbolic buildings", *Proc. Amer. Math. Soc.* **127** (1999), no. 8, p. 2315–2324.

[Bus55] H. Busemann – *The geometry of geodesics*, Academic Press Inc., New York, N. Y., 1955.

[EO73] P. Eberlein & B. O'Neill – "Visibility manifolds", *Pacific J. Math.* **46** (1973), p. 45–109.

[Klo10] B. Kloeckner – "A geometric study of Wasserstein spaces: Euclidean spaces", *Ann. Sc. Norm. Super. Pisa Cl. Sci. (5)* **9** (2010), no. 2, p. 297–323.

[Lot08] J. Lott – "Some geometric calculations on Wasserstein space", *Comm. Math. Phys.* **277** (2008), no. 2, p. 423–437.

[LV09] J. Lott & C. Villani – "Ricci curvature for metric-measure spaces via optimal transport", *Ann. of Math. (2)* **169** (2009), no. 3, p. 903–991.

[McC97] R. J. McCann – "A convexity principle for interacting gases", *Adv. Math.* **128** (1997), no. 1, p. 153–179.

[Sch70] L. Schwartz – "Systèmes projectifs de mesures et théorème de Prokhorov", Séminaire Laurent Schwartz 1969–1970: Applications radonifiantes, Exp. No. 1, Centre de Math., École Polytech., Paris, 1970, p. 5.

[ST09] W. Schachermayer & J. Teichmann – "Characterization of optimal transport plans for the Monge-Kantorovich problem", *Proc. Amer. Math. Soc.* **137** (2009), no. 2, p. 519–529.

[Stu06] K.-T. Sturm – "On the geometry of metric measure spaces. I, II", *Acta Math.* **196** (2006), no. 1, p. 65–131, 133–177.

[TY12] A. Takatsu & T. Yokota – "Cone structure of L^2-Wasserstein spaces", *J. Topol. Anal.* **4** (2012), no. 2, p. 237–253, arXiv:0812.2752.

[Vil03] C. Villani – *Topics in optimal transportation*, Graduate Studies in Mathematics, vol. 58, American Mathematical Society, Providence, RI, 2003.

[Vil09] ———, *Optimal transport*, Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 338, Springer-Verlag, Berlin, 2009, Old and new.

Jérôme Bertrand, Institut de Mathématiques, Université Paul Sabatier, 118 route de Narbonne, F31062 Cedex 9 Toulouse, France • *E-mail :* `bertrand@math.univ-toulouse.fr`

Benoît Kloeckner, Université de Grenoble I, Institut Fourier, CNRS UMR 5582, BP 74, 38402 Saint Martin d'Hères cedex, France *E-mail :* `benoit.kloeckner@ujf-grenoble.fr`