RANDOM WALKS ON WEAKLY HYPERBOLIC GROUPS

JOSEPH MAHER, GIULIO TIOZZO

Abstract. Let G be a countable group which acts by isometries on a separable, but not necessarily proper, Gromov hyperbolic space X. We say the action of G is weakly hyperbolic if G contains two independent hyperbolic isometries. We show that a random walk on such G converges to the Gromov boundary almost surely. We apply the convergence result to show linear progress and linear growth of translation length, without any assumptions on the moments of the random walk.

If the action is acylindrical, and the random walk has finite entropy and finite logarithmic moment, we show that the Gromov boundary with the hitting measure is the Poisson boundary.

1. Introduction

We say a geodesic metric space X is *Gromov hyperbolic*, or *δ-hyperbolic*, if there is a number $\delta \geqslant 0$ for which every geodesic triangle in X satisfies the *δ-slim triangle* condition, i.e. any side is contained in a δ-neighbourhood of the other two sides. Throughout this paper we will assume that the space X is separable, i.e. it contains a countable dense set, but we will not assume that X is proper or locally compact. For example, any countable simplicial complex satisfies these conditions.

Let G be a countable group which acts by isometries on X. We say the action of G on X is *non-elementary* if G contains a pair of hyperbolic isometries with disjoint fixed point sets in the Gromov boundary. We say G is *weakly hyperbolic* if it admits a non-elementary action by isometries on some Gromov hyperbolic space X. In this case, a natural boundary for the group is given by the Gromov boundary ∂X of X, which however need not be compact.

Several widely studied group actions are weakly hyperbolic in this sense, in particular:

- Hyperbolic and relatively hyperbolic groups;
- Mapping class groups, acting on the curve complex;
- $\mathrm{Out}(F_n)$ acts on various Gromov hyperbolic simplicial complexes, for example the complex of free factors or the complex of free splittings;
- Right-angled Artin groups acting on their extension graphs;
- Finitely generated subgroups of the Cremona group.

2010 *Mathematics Subject Classification.* 60G50, 20F67, 57M60.

In particular, all acylindrically hyperbolic groups are weakly hyperbolic, see Section 1.2 for further discussion and more examples.

In this paper, we shall consider random walks on weakly hyperbolic groups, constructed by choosing products of random group elements. A probability distribution μ on G determines a random walk on G, by taking the product

$$w_n := g_1 g_2 \dots g_n$$

where the g_i are independent identically distributed elements of G, with distribution μ. A choice of basepoint $x_0 \in X$ determines an orbit map sending $g \mapsto gx_0$, and we can project the random walk on G to X by considering the sequence $(w_n x_0)_{n \in \mathbb{N}}$, which we call a *sample path*. The *support* of μ consists of all group elements g for which $\mu(g) > 0$. We say a measure μ on G is *non-elementary* if the semigroup generated by the support of μ contains a pair of hyperbolic isometries with disjoint fixed point sets.

1.1. **Results.** The first result we establish is that sample paths converge almost surely in the Gromov boundary:

Theorem 1.1. *Let G be a countable group of isometries of a separable Gromov hyperbolic space X, and let μ be a non-elementary probability distribution on G. Then, for any basepoint $x_0 \in X$, almost every sample path $(w_n x_0)_{n \in \mathbb{N}}$ converges to a point $\omega_+ \in \partial X$. The resulting hitting measure ν is non-atomic, and is the unique μ-stationary measure on ∂X.*

We use the convergence to the boundary result to show the following linear progress, or positive drift, result. Recall that a measure μ has finite first moment if $\int_G d_X(x_0, gx_0)\ d\mu(g) < \infty$.

Theorem 1.2. *Let G be a countable group of isometries of a separable Gromov hyperbolic space X, and let μ be a non-elementary probability distribution on G, and $x_0 \in X$ a basepoint. Then there is a constant $L > 0$ such that for almost every sample path we have*

$$\liminf_{n \to \infty} \frac{d_X(x_0, w_n x_0)}{n} \geqslant L > 0.$$

Furthermore, if μ has finite first moment, then the limit

$$\lim_{n \to \infty} \frac{d_X(x_0, w_n x_0)}{n} = L > 0$$

exists almost surely. Finally, if the support of μ is bounded in X, then there are constants $c < 1, K$ and $L > 0$ such that

$$\mathbb{P}(d_X(x_0, w_n x_0) \leqslant Ln) \leqslant Kc^n$$

for all n.

Note that the constants L, K, and c depend on the choice of the measure μ. Moreover, the first statement also implies that

$$\lim_{n \to \infty} \mathbb{P}(d_X(x_0, w_n x_0) \leqslant Ln) = 0$$

(for a possibly different constant L). If we assume that μ has finite first moment with respect to the distance function d_X, then we obtain the following geodesic tracking result.

Theorem 1.3. *Let G be a countable group which acts by isometries on a separable Gromov hyperbolic space X with basepoint x_0, and let μ be non-elementary probability distribution on G, with finite first moment. Then for almost every sample path $(w_n x_0)_{n\in\mathbb{N}}$ there is a quasigeodesic ray γ which tracks the sample path sublinearly, i.e.*

$$\lim_{n\to\infty} \frac{d_X(w_n x_0, \gamma)}{n} = 0, \text{ almost surely.}$$

If the support of μ is bounded in X, then in fact the tracking is logarithmic, i.e.

$$\limsup_{n\to\infty} \frac{d_X(w_n x_0, \gamma)}{\log n} < \infty, \text{ almost surely.}$$

Finally, we investigate the growth rate of translation length of group elements arising from the sample paths.

Theorem 1.4. *Let G be a countable group which acts by isometries on a separable Gromov hyperbolic space X, and let μ be a non-elementary probability distribution on G. Then the translation length $\tau(w_n)$ of the group element w_n grows at least linearly in n, i.e.*

$$\mathbb{P}(\tau(w_n) \leqslant Ln) \to 0 \text{ as } n \to \infty,$$

for some constant L strictly greater than zero.

If the support of μ is bounded in X, then there are constants $c < 1, K$ and $L > 0$, such that

$$\mathbb{P}(\tau(w_n) \leqslant Ln) \leqslant Kc^n$$

for all n.

Recall that the translation length $\tau(g)$ of an isometry g of X is defined as

$$\tau(g) := \lim_{n\to\infty} \tfrac{1}{n} d_X(x_0, g^n x_0).$$

As an element with non-zero translation length is a hyperbolic (= loxodromic) isometry, this shows that the probability that a random walk of length n gives rise to a hyperbolic isometry tends to one as n tends to infinity.

The Poisson boundary of acylindrically hyperbolic groups. A special class of weakly hyperbolic groups are the *acylindrically hyperbolic* groups. In this case, we show that we may identify the Gromov boundary $(\partial X, \nu)$ with the Poisson boundary.

Recall that a group G acts acylindrically on a Gromov hyperbolic space X, if for every $K \geqslant 0$ there are numbers R and N, which both depend on K, such that for any pair of points x and y in X, with $d_X(x, y) \geqslant R$, there are at most N group elements g in G such that $d_X(x, gx) \leqslant K$ and $d_X(y, gy) \leqslant K$.

Theorem 1.5. *Let G be a countable group of isometries which acts acylindrically on a separable Gromov hyperbolic space X, let μ be a non-elementary probability distribution on G with finite entropy and finite logarithmic moment, and let ν be the hitting measure on ∂X. Then $(\partial X, \nu)$ is the Poisson boundary of (G, μ).*

1.2. **Examples and discussion.**

Word hyperbolic groups. The simplest example of a weakly hyperbolic group is a (finitely generated) Gromov hyperbolic group acting on its Cayley graph, which by definition is a δ-hyperbolic space, and so any non-elementary Gromov hyperbolic group is weakly hyperbolic. Convergence of sample paths to the Gromov boundary in this case is due to Kaimanovich [Kai94], who also shows that the Gromov boundary may be identified with the Poisson boundary and the hitting measure is the unique μ-stationary measure. Note that in this case the space is locally compact, and the boundary is compact.

Relatively hyperbolic groups. The Cayley graph of a relatively hyperbolic group is δ-hyperbolic with respect to an infinite generating set, and so these groups are also weakly hyperbolic, but in this case the space on which the group acts need not be proper. In this case, convergence to the boundary was shown by Gautero and Mathéus [GM12], who also covered the case of groups acting on $\mathbb{R}$-trees. More recently, group actions on (locally infinite) trees have been considered by Malyutin and Svetlov [MS14].

There are then groups which are weakly hyperbolic, but not relatively hyperbolic, the two most important examples being the mapping class groups of surfaces, and $\mathrm{Out}(F_n)$.

Mapping class groups. The mapping class group $\mathrm{Mod}(S)$ of a surface S of genus g with p punctures acts on the curve complex $C(S)$, which is a locally infinite simplicial complex. As shown by Masur and Minsky [MM99], the curve complex is δ-hyperbolic, and moreover, the action is acylindrical, by work of Bowditch [Bow08], so we can apply our techniques to get convergence and the Poisson boundary.

Convergence to the boundary of the curve complex also follows from work of Kaimanovich and Masur [KM96] and Klarreich [Kla], using the action of $\mathrm{Mod}(S)$ on Teichmüller space (which is locally compact, but not hyperbolic). Indeed, Kaimanovich and Masur show that random walks on the mapping class group converge to points in Thurston's compactification of Teichmüller space $\mathcal{PMF}$, and then Klarreich (see also Hamenstädt [Ham06]) shows the relation between $\mathcal{PMF}$ and the boundary of the curve complex. Our approach does not use fine properties of Teichmüller geometry. A third approach is to consider the action of $\mathrm{Mod}(S)$ on Teichmüller space with the Weil-Petersson metric, which is a non-proper CAT(0) space. By work of Bestvina, Bromberg and Fujiwara [BBF10], this space has finite telescopic dimension, and one may then apply the results of Bader, Duchesne and Lécureux [BDL14].

We remark that $C(S)$ does not possess a CAT(0) metric, since it is homotopic to a wedge of spheres (Harer [Har86]); in fact, Kapovich and Leeb [KL96] showed that the mapping class group (of genus at least 3) does not act freely cocompactly on a CAT(0) space, though it is still open as to whether there is a proper CAT(0) space on which the mapping class group acts by isometries; Bridson [Bri12] showed that any such action must have elliptic or parabolic Dehn twists.

Out(F_n). The outer automorphism group of a non-abelian free group, $\mathrm{Out}(F_n)$, acts on a number of distinct Gromov hyperbolic spaces, as shown by Bestvina and Feighn [BF10, BF14] and Handel and Mosher [HM13], and so is weakly hyperbolic. Similarly to the case of $\mathrm{Mod}(S)$, convergence to the boundary also follows by considering the action of $\mathrm{Out}(F_n)$ on the (locally compact) outer space, as shown by Horbez [Hor14].

Right-angled Artin groups. A right-angled Artin group acts by simplicial isometries on its extension graph, which has infinite diameter as long as the group does not split as a non-trivial direct product, and is not quasi-isometric to $\mathbb{Z}$. Kim and Koberda showed that the extension graph is a (non-locally compact) quasi-tree [KK13], and in fact the action is acylindrical [KK14].

Finitely generated subgroups of the Cremona group. Manin [Man74] showed that the Cremona group acts faithfully by isometries on an infinite-dimensional hyperbolic space, known as the Picard-Manin space, which is not separable. However, any finite-generated subgroup preserves a totally geodesic closed subspace, which is separable, see for example Delzant and Py [DP12].

Acylindrically hyperbolic groups. The definition of an acylindrical group action is due to Sela [Sel97] for trees, and Bowditch [Bow08] for general metric spaces, see Osin [Osi13] for a discussion and several examples of acylindrical actions on hyperbolic spaces. As every acylindrically hyperbolic group is also weakly hyperbolic, this gives a number of additional examples of weakly hyperbolic groups which are not necessarily relatively hyperbolic; for example, all one relator groups with at least three generators, see [Osi13] for many other examples.

Isometries of CAT(0) spaces. Even though not all CAT(0) spaces are hyperbolic, the two theories overlap in many cases. For isometries of general CAT(0) spaces, Karlsson and Margulis [KM96] proved boundary convergence and identified the Poisson boundary. More recently, boundaries of CAT(0) cube complexes have been studied by Nevo and Sageev [NS13], and (not necessarily proper) CAT(0) spaces of finite telescopic dimension by Bader, Duchesne and Lécureux [BDL14].

Once we have proved convergence to the boundary, we apply this to show positive drift. In the locally compact case, positive drift results go back to Guivarc'h [Gui80]. In particular, when the space on which G acts is proper, positive drift follows from non-amenability of the group, but this need not be the case for non-proper spaces. In the curve complex case, linear progress is due to Maher [Mah10a].

We then show the sublinear tracking results, using work of Tiozzo [Tio12]. Sublinear tracking can be thought of as a generalization of Oseledec's multiplicative ergodic theorem [Ose68]. In our context, these results go back to Guivarc'h [Gui80], and are known for groups of isometries of CAT(0) spaces by Karlsson and Margulis [KM96], and for Teichmüller space by Duchin [Duc05]. Sublinear tracking on hyperbolic groups is due to Kaimanovich [Kai87, Kai94]; moreover, Karlsson and Ledrappier [KL06, KL07] proved a law of large numbers on general (proper) metric spaces using horofunctions.

Note that these results can be used to prove convergence to the boundary once one knows that the drift is positive. In the above-mentioned cases, the space is meant to be proper, so positive drift follows from non-amenability of the group, while a new argument is needed in general.

In this paper we give an argument for the non-proper weakly hyperbolic case, where sublinear tracking and positive drift *follow* from convergence to the boundary. Note that in our approach we use horofunctions, and indeed [KL06] can be used to simplify our proofs if one assumes positive drift. Recently (after the first version of this paper appeared), Mathieu and Sisto [MS14] provided a different argument for positive drift in the acylindrical case.

Logarithmic tracking was previously known for random walks on trees, due to Ledrappier [Led01], on hyperbolic groups, due to Blachère, Haïssinsky and Mathieu [BHM11], and on relatively hyperbolic groups, due to Sisto [Sis13].

Finally we show that the translation length grows linearly, which in particular shows that the probability that a random walk gives rise to a hyperbolic element tends to one. This generalizes earlier work of Rivin [Riv08], Kowalski [Kow08], Maher [Mah11] and Sisto [Sis11].

The methods in this paper build on previous work of Calegari and Maher [CM15], which showed convergence results with stronger conditions on X and μ.

1.3. **Outline of the argument.** To explain the argument in the proof of Theorem 1.1, we briefly remind the reader of the standard argument for convergence to the boundary for a random walk on a group G acting on a locally compact δ-hyperbolic space X. The argument ultimately goes back to Furstenberg [Fur63], who developed it for Lie groups.

Measures on the Gromov boundary. Let μ be the probability distribution on G generating the random walk. The first step is to find a μ-stationary measure ν on the Gromov boundary ∂X, and then apply the martingale convergence theorem to show that for almost every sample path $\omega = (w_n)_{n \in \mathbb{N}}$, the sequence of measures $(w_n \nu)_{n \in \mathbb{N}}$ converges to some measure ν_ω in $\mathcal{P}(\partial X)$, the space of probability measures on ∂X. One then uses geometric properties of the action of G on X to argue that ν_ω is a δ-measure δ_λ for some point $\lambda \in \partial X$, almost surely, and that the image of the sample path under the orbit map $(w_n x_0)_{n \in \mathbb{N}}$ converges to λ.

This argument uses local compactness in an essential way in the first step. For a locally compact hyperbolic space X, the Gromov boundary ∂X is compact, as is $X \cup \partial X$. The space of probability measures on ∂X is also compact, and so the existence of a μ-invariant measure on X just follows from taking weak limits. In the non-locally compact case, the Gromov boundary ∂X, and $X \cup \partial X$, need not be compact, as seen in the following example.

Example 1.6 (Countable wedge of rays)**.** *A ray is a half line $\mathbb{R}_+ = \{x \in \mathbb{R} : x \geqslant 0\}$, with basepoint 0. Let X be the wedge product of countably many rays. This space is a tree, and so is δ-hyperbolic, and is not locally compact at the basepoint. The Gromov boundary is homeomorphic to $\mathbb{N}$ with the discrete metric, and is not compact.*

The horofunction boundary. In order to address this issue, we shall consider the horofunction boundary of X, which was also initially developed by Gromov [BGS85], and has proved a useful tool in studying random walks, see for example Karlsson-Ledrappier [KL06, KL07] and Bjorklund [Bjö10]. We now give a brief description of this construction, giving full details in Section 3.

Let X be a metric space, and x_0 a basepoint. For each point x in X, one defines the horofunction ρ_x determined by x to be the function $\rho_x : X \to \mathbb{R}$

$$\rho_x(z) := d_X(x, z) - d_X(x, x_0).$$

This gives an embedding of X in the space $C(X)$ of (Lipschitz-) continuous functions on X, which we shall consider with the compact-open topology (we emphasize that we use uniform convergence on compact sets, not uniform convergence on bounded sets). With this topology, the closure of $\rho(X)$ in $C(X)$ is compact, even if X is not locally compact; it is called the *horofunction compactification* of X and denoted by $\overline{X}^h$. In particular, there is a μ-stationary measure ν on $\overline{X}^h$.

We now consider a basic but fundamental example in detail.

Example 1.7 ($\mathbb{R}$)**.** *Consider $X = \mathbb{R}$, with the usual metric. In this case the horofunction boundary $\overline{X}^h$ consists of $\rho(X)$ together with precisely two additional functions, namely $\rho_\infty(x) := -x$, and $\rho_{-\infty}(x) := x$.*

This example turns out to be very important in our case; indeed, if X is Gromov hyperbolic, then the restriction of an arbitrary horofunction to a geodesic is equal (up to a bounded additive error) to one of the horofunctions described above, i.e. ρ_x or $\rho_{\pm\infty}$.

In Example 1.6, the horofunction boundary equals the Gromov boundary as a set, but the topology is different: namely, any sequence of horofunctions

$(\rho_{x_n})_{n\in\mathbb{N}}$ corresponding to a sequence of points $(x_n)_{n\in\mathbb{N}}$ which leaves every compact set converges to the horofunction ρ_{x_0} associated to the basepoint x_0.

The Gromov boundary may be recovered from the horofunction boundary by identifying functions which differ by a bounded amount, but in general the horofunction boundary may be larger than the Gromov boundary, and is not a quasi-isometry invariant of the space.

Example 1.8. *Consider $X = \mathbb{Z}\times\mathbb{Z}/2\mathbb{Z}$, with the L^1-metric, $d_X((x,i),(y,j)) = |x-y| + |i-j|$. Then the sequences $\rho_{n,0}$ and $\rho_{n,1}$ have different values on $(0,1)$, and so converge to distinct horofunctions, and in fact in this case the horofunction boundary consists of the product of the Gromov boundary with $\mathbb{Z}/2\mathbb{Z}$.*

We shall distinguish two different types of horofunctions. We say a horofunction h is *finite* if $\inf_{x\in X} h(x) > -\infty$, and is *infinite* if $\inf h = -\infty$. This partitions $\overline{X}^h$ into two subsets: we shall write $\overline{X}^h_F$ for the set of finite horofunctions, and $\overline{X}^h_\infty$ for the set of infinite horofunctions.

The local minimum map. We shall now construct a map from the horofunction boundary to the Gromov boundary. Recall that the restriction of a horofunction h to a geodesic γ in X is coarsely equal to one of the standard horofunctions on $\mathbb{R}$: in particular, it has (coarsely) at most one local minimum on γ. Thus, if h is bounded below on γ, we can map h to the location where it attains its minimum, getting a map $\phi : \overline{X}^h_F \to X$. On the other hand, if the horofunction is not bounded below, then we can pick a sequence (x_n) of points for which the value of the horofunction tends to $-\infty$; it turns out that such a sequence converges to a unique point in the Gromov boundary, and the limit is independent of the choice of (x_n). Thus, we can extend ϕ to a map $\phi : \overline{X}^h \to X \cup \partial X$. We show that this map is continuous on $\overline{X}^h_\infty$ and G-equivariant, and that the stationary measure ν is supported on the infinite horofunctions $\overline{X}^h_\infty$. Therefore the stationary probability measure ν on $\overline{X}^h$ restricts to a probability measure on $\overline{X}^h_\infty$, and pushes forward to a μ-stationary probability measure $\widetilde{\nu}$ on ∂X. We may then complete the argument using with the geometric properties of the action of G on ∂X.

Plan of the paper. In Section 2, we review some useful material about Gromov hyperbolic spaces, and fix notation. In Section 3, we develop the properties of the horofunction boundary that we will use, including the local minimum map, and the behaviour of shadows. In these initial sections we give complete proofs in the non-proper case of certain statements that are already known in the proper case. In Section 4, we use the horofunction boundary to show that almost every sample path converges to the Gromov boundary. In Section 5, we use the convergence to the boundary result to show results on positive drift, sublinear tracking, and the growth rate of translation distance, and then finally in Section 6 we show that if the action of G is acylindrical, and μ has finite entropy, then the Gromov boundary with the hitting measure is the Poisson boundary for the random walk.

1.4. **Acknowledgements.** We would like to thank Jason Behrstock, Danny Calegari, Romain Dujardin, Camille Horbez, Vadim Kaimanovich, Anders Karlsson, Andrei Malyutin and Samuel Taylor for helpful conversations. We also thank the referee for the useful comments on the manuscript. The first author gratefully acknowledges the support of the Simons Foundation and PSC-CUNY. The authors would like to thank the referee for his helpful comments and suggestions.

2. Background on δ-hyperbolic spaces

Let X be a Gromov hyperbolic space, i.e. a geodesic metric space which satisfies the δ-slim triangles condition. We will not assume that X is *proper*, i.e. that closed balls are compact, but we will always assume that it is *separable*, i.e. that it contains a dense countable subset. We shall write d_X for the metric on X, and $B_X(x, r)$ for the closed ball of radius r about the point x in X. We shall now recall a few facts on the geometry of X.

2.1. **Notation.** We shall write $f(x) = O(\delta)$ to mean that the absolute value of the function f is bounded by a number which only depends on δ, though this need not be a linear multiple of δ. Similarly, we shall write $A = B + O(\delta)$ to mean that the difference between A and B is bounded by a constant, which depends only on δ.

2.2. **Coarse geometry.** Recall that the Gromov product in a metric space is defined to be

$$(x \cdot y)_{x_0} := \tfrac{1}{2}(d_X(x_0, x) + d_X(x_0, y) - d_X(x, y)).$$

In a δ-hyperbolic space, for all points x_0, x and y, the Gromov product $(x \cdot y)_{x_0}$ is equal to the distance from x_0 to a geodesic from x to y, up to an additive error of at most δ: if we write $[x, y]$ for a choice of geodesic from x to y, then

$$d_X(x_0, [x, y]) = (x \cdot y)_{x_0} + O(\delta), \tag{1}$$

see e.g. [BH99, III.H 1.19]. Moreover, for any three points $x, y, z \in X$ one has the following inequality

$$(x \cdot y)_{x_0} \geqslant \min\{(x \cdot z)_{x_0}, (y \cdot z)_{x_0}\} - O(\delta), \tag{2}$$

which we shall refer to as the triangle inequality for the Gromov product.

We now recall the definition of the Gromov boundary of X, which we shall write as ∂X. We say that a sequence $(x_n)_{n\in\mathbb{N}} \subseteq X$ is a *Gromov sequence* if $(x_m \cdot x_n)_{x_0}$ tends to infinity as $\min\{m, n\}$ tends to infinity. We say that two Gromov sequences $(x_n)_{n\in\mathbb{N}}$ and $(y_n)_{n\in\mathbb{N}}$ are equivalent if $(x_n \cdot y_n)_{x_0}$ tends to infinity as n tends to infinity. The *Gromov boundary* ∂X is defined as the set of equivalence classes of Gromov sequences.

We can extend the Gromov product to the boundary by

$$(x \cdot y)_{x_0} = \sup \liminf_{m,n \to \infty} (x_m \cdot y_n)_{x_0},$$

where the supremum is taken over all sequences $(x_m)_{m \in \mathbb{N}} \to x$ and $(y_n)_{n \in \mathbb{N}} \to y$. With this definition, the triangle inequality (2) also holds for any three points x, y, z in $X \cup \partial X$, but with a larger additive constant $O(\delta)$ (see e.g. [BH99, III.H Remark 3.17(4)]).

The Gromov product on the boundary may be used to define a complete metric on ∂X, see Bridson and Haefliger [BH99, III.H.3] for the proper case, and Väisälä [Väi05] for the non-proper case. Moreover, the space $X \cup \partial X$ can be equipped with a topology such that the relative topologies on both X and ∂X are equal respectively to the usual topology on X, and the above-mentioned metric topology on ∂X.

If $U \subset X$ we shall write $\overline{U}^{\delta}$ for the closure of U in $X \cup \partial X$. If X is proper, then ∂X is compact, but it need not be compact if X is not proper. However, a bounded set does not have limit points in the Gromov boundary, i.e. for $B_X(x_0, r) = \{x \in X : d_X(x_0, x) \leqslant r\}$ we have $\overline{B_X(x_0, r)}^{\delta} = B_X(x_0, r)$.

2.3. **Quasigeodesics.** Let I be a connected subset of $\mathbb{R}$, and let X be a metric space. A (Q, c)-quasigeodesic is a (not necessarily continuous) map $\gamma \colon I \to X$ such that for all s and t in I,

$$\frac{1}{Q} |t - s| - c \leqslant d_X(\gamma(s), \gamma(t)) \leqslant Q |t - s| + c.$$

If $I = \mathbb{R}$, then we will call the quasigeodesic γ a *bi-infinite quasigeodesic*. If $I = [0, \infty)$, then we shall call γ a *quasigeodesic ray based at* $\gamma(0)$. If the metric space is δ-hyperbolic, then quasigeodesics have the following stability property, which is often referred to as the Morse Lemma (see [BH99, III.H 1.7], [Väi05, Theorem 6.32]).

Proposition 2.1. *Let X be a δ-hyperbolic space. Given numbers Q and c, there is a number L such that for any two points x and y in $X \cup \partial X$, any two (Q, c)-quasigeodesics connecting x and y are contained in L-neighbourhoods of each other.*

We shall refer to a choice of constant L in Proposition 2.1 above as a Morse constant for the quasi-geodesic constants Q and c.

For any choice of basepoint x_0 and every point x in the boundary there is a quasigeodesic ray based at x_0 which converges to the point x, and any two points in the boundary are connected by a bi-infinite quasigeodesic. In fact, the quasigeodesics may be chosen to have quasigeodesic constants Q and c bounded above in terms of the hyperbolicity constant δ, independently of the choice of basepoint or boundary points, see e.g. Bonk and Schramm [BS00, Proposition 5.2] or Kapovich and Benakli [KB02, Remark 2.16]. By choosing Q and c sufficiently large, we may assume that we have chosen these constants so that at least one of the (Q, c)-quasigeodesics is continuous, see e.g. Bridson and Haefliger [BH99, III.H.1].

2.4. **Nearest point projection.** We will use the fact that in a δ-hyperbolic space nearest point projection onto a geodesic γ is coarsely well defined, i.e. there is a constant K_2, which only depends on δ, such that if p and q are nearest points on γ to x, then $d_X(p,q) \leqslant K_2$.

We will make use of the following *reverse triangle inequality.*

Proposition 2.2. *Let γ be a geodesic in X, $y \in X$ a point, and p a nearest point projection of y to γ. Then for any $z \in \gamma$ we have*

$$d_X(y,z) = d_X(y,p) + d_X(z,p) + O(\delta), \tag{3}$$

and furthermore, any geodesic from z to y passes within distance $O(\delta)$ of p.

Proof. The upper bound for $d_X(y,z)$ is immediate from the usual triangle inequality. To prove the lower bound, by the definition of nearest point projection,

$$d_X(y,p) = d_X(y,[p,z]).$$

Recall that by (1),

$$d_X(y,p) \leqslant (p \cdot z)_y + O(\delta),$$

and writing out the Gromov product, we get

$$d_X(y,p) \leqslant \frac{1}{2}\Big(d_X(y,p) + d_X(y,z) - d_X(p,z)\Big) + O(\delta),$$

which yields

$$d_X(y,z) \geqslant d_X(y,p) + d_X(p,z) - O(\delta),$$

as required. This implies that a path consisting of $[y,p] \cup [p,z]$ is a $(1,C)$-quasigeodesic for some $C = O(\delta)$, and so by stability of quasigeodesics in a δ-hyperbolic space, this path is contained in an $O(\delta)$-neighbourhood of any geodesic $[y,z]$ from y to z, so in particular, the distance from p to a $[y,z]$ is at most $O(\delta)$. $\square$

Finally we show that if two points x and y in X have nearest point projections p_x and p_y to a geodesic γ, and p_x and p_y are sufficiently far apart, then the path $[x,p_x] \cup [p_x,p_y] \cup [p_y,y]$ is a quasigeodesic, and in fact has the same length as a geodesic from x to y, up to an additive error depending only on δ.

Proposition 2.3. *Let γ be a geodesic in X, and let x and y be two points in X with nearest points p_x and p_y respectively on γ. Then if $d_X(p_x,p_y) \geqslant O(\delta)$, then*

$$d_X(x,y) = d_X(x,p_x) + d_X(p_x,p_y) + d_X(p_y,y) + O(\delta),$$

and furthermore, any geodesic from x to y passes within an $O(\delta)$-neighbourhood of both p_x and p_y.

This is well known, see e.g. [Mah10a, Proposition 3.4].

Given a point $x \in X$ and a number R, the *shadow* $S_{x_0}(x, R)$ is defined to be

$$S_{x_0}(x, R) := \{y \in X : (x \cdot y)_{x_0} \geqslant d_X(x_0, x) - R\}.$$

There are a number of similar definitions in the literature, and we emphasize that we define shadows to be subsets of X, rather than subsets of say $X \cup \partial X$ or ∂X. We allow arbitrary values of R; however, if $R < 0$, the shadow is empty, and if $R \geqslant d_X(x_0, x)$, the shadow consists of all of X. We will refer to the quantity $d_X(x_0, x) - R$ as the *distance parameter* of the shadow. For values of R between 0 and $d_X(x_0, x)$, the distance parameter is equal to the distance from x_0 to the shadow, up to an additive error depending only on δ.

By the triangle inequality for the Gromov product (2), for any two points y and z in the closure of a shadow $\overline{S_{x_0}(x, R)}^{\delta}$, with $R \geqslant 0$, there is a lower bound on their Gromov product

$$(y \cdot z)_{x_0} \geqslant d_X(x, x_0) - R + O(\delta). \tag{4}$$

We now show that the nearest point projection of a shadow $S_x(y, R)$ to a geodesic $[x, y]$ is contained in a bounded neighbourhood of the intersection of the shadow with $[x, y]$, and the same result holds for the complement of the shadow.

Proposition 2.4. *Let z be a point in the shadow $S_x(y, R)$, with $0 \leqslant R \leqslant d_X(x, y)$, let γ be a geodesic from x to y, and let p be a nearest point to z on γ. Then*

$$d_X(y, p) \leqslant R + O(\delta). \tag{5}$$

If $z \notin S_x(y, R)$, then

$$d_X(y, p) \geqslant R + O(\delta). \tag{6}$$

Proof. If z lies in the shadow $S_x(y, R)$,

$$(y \cdot z)_x \geqslant d_X(x, y) - R.$$

Using the definition of the Gromov product, we may rewrite this as

$$d_X(x, z) - d_X(z, y) \geqslant d_X(x, y) - 2R,$$

and then using (3), and the fact that x, p and y lie in that order on a common geodesic, gives

$$R + O(\delta) \geqslant d_X(p, y),$$

as required. If z does not lie in $S_x(y, R)$, then the same argument works, with the opposite inequality, □

As a consequence, the complement of a shadow is almost a shadow:

Corollary 2.5. *Let $0 \leqslant R \leqslant d_X(x, y)$. Then there is a number $C > 0$, which only depends on δ, such that the complement of the shadow $S_x(y, R)$, is contained in the shadow $S_y(x, \widetilde{R})$, where $\widetilde{R} = d_X(x, y) - R + C$.*

3. The horofunction boundary

Let (X, d_X) be a metric space. A function $f : X \to \mathbb{R}$ is called 1-*Lipschitz* if for each $x, y \in X$ we have

$$|f(x) - f(y)| \leqslant d_X(x, y).$$

Clearly, 1-Lipschitz functions are uniformly continuous. For each $x_0 \in X$, let us define

$$\mathrm{Lip}^1_{x_0}(X) := \{f \colon X \to \mathbb{R} \ : \ f \text{ is 1-Lipschitz, and } f(x_0) = 0\},$$

the space of 1-Lipschitz functions which vanish at x_0. We shall endow the space $\mathrm{Lip}^1_{x_0}(X)$ with the topology of pointwise convergence. Note that, since all elements of $\mathrm{Lip}^1_{x_0}(X)$ are uniformly continuous with the same modulus of continuity, this topology is equivalent to the topology of uniform convergence on compact sets, which is also equivalent to the compact-open topology as $\mathbb{R}$ is a metric space.

Proposition 3.1. *Let X be a separable metric space. Then for each $x_0 \in X$, the space $\mathrm{Lip}^1_{x_0}(X)$ is compact, Hausdorff and second countable (hence metrizable).*

Proof. Note that for any function $f \in \mathrm{Lip}^1_{x_0}(X)$ and each $z \in X$ we have

$$|f(z)| = |f(z) - f(x_0)| \leqslant d_X(x_0, z)$$

hence the space $\mathrm{Lip}^1_{x_0}(X)$ is a closed subspace of an infinite product of compact spaces, hence it is compact by Tychonoff's theorem. Let $C(X)$ be the space of real-valued continuous functions on X, with the compact-open topology. As $\mathbb{R}$ is Hausdorff, $C(X)$ is also Hausdorff, hence so is $\mathrm{Lip}^1_{x_0}(X)$. Since X is a separable metric space, it is second countable; moreover, as $\mathbb{R}$ is also second countable, $C(X)$ is second countable and so is $\mathrm{Lip}^1_{x_0}(X)$. $\square$

Let $x_0 \in X$ be a basepoint. We define the *horofunction* map ρ to be

$$\begin{aligned} \rho \colon X &\to C(X) \\ y &\mapsto \rho_y(z) := d_X(z, y) - d_X(x_0, y), \end{aligned}$$

where we write ρ_y for $\rho(y)$. Note that the horofunction map ρ depends on the choice of basepoint, but as we shall usually consider horofunctions defined from some fixed basepoint we omit this from the notation. In the few cases we consider horofunctions with different basepoints we will change notation on an ad hoc basis. For each y, the horofunction ρ_y is 1-Lipschitz, as

$$|\rho_y(z_1) - \rho_y(z_2)| = |d_X(z_1, y) - d_X(z_2, y)| \leqslant d_X(z_1, z_2), \tag{7}$$

and moreover $\rho_y(x_0) = 0$ for all y, hence ρ maps X into $\mathrm{Lip}^1_{x_0}(X, \mathbb{R})$.

Lemma 3.2. *The map $\rho \colon X \to C(X)$ defined above is continuous and injective.*

Proof. For any $y \in X$, the function $\rho_y(z) = d_X(z, y) - d_X(x_0, y)$ achieves a unique minimum at $z = y$, so the map ρ is injective on X. The map ρ is continuous, as for any $x, y, z \in X$ we have

$$\begin{aligned} |\rho_x(z) - \rho_y(z)| &= |d_X(z, x) - d_X(x_0, x) - d_X(z, y) + d_X(x_0, y)| \\ &\leqslant |d_X(z, x) - d_X(z, y)| + |d_X(x_0, y) - d_X(x_0, x)| \\ &\leqslant 2d_X(x, y) \end{aligned}$$

and so if $(y_n)_{n\in\mathbb{N}} \to y$ then $(\rho_{y_n})_{n\in\mathbb{N}} \to \rho_y$ uniformly on compact sets, in fact uniformly on all of X. □

Let us now define the fundamental object we are going to work with.

Definition 3.3. *Let X be a separable metric space with basepoint $x_0 \in X$. We define the* horofunction compactification $\overline{X}^h$ *to be the closure of $\rho(X)$ in* $\mathrm{Lip}^1_{x_0}(X)$. *We shall call the set $\partial X^h := \overline{X}^h \setminus X$ the* horofunction boundary *of X. Elements of $\overline{X}^h$ will be called* horofunctions.

Note that in the proper case, the space $\overline{X}^h$ contains X as an open, dense set. In the non-proper case, although the map ρ is injective on X, the image $\rho(X)$ need not be open in $\overline{X}^h$, so although $\overline{X}^h$ is compact, it is not a compactification of X in the standard sense.

Lemma 3.4. *Let G be a group of isometries of X. Then the action of G on X extends to a continuous action by homeomorphisms on $\overline{X}^h$, defined as*

$$g.h(z) := h(g^{-1}z) - h(g^{-1}x_0) \tag{8}$$

for each $g \in G$ and $h \in \overline{X}^h$.

Proof. The action of $g \in G$ on X translates into an action on $\rho(X)$, by defining $g.\rho_y := \rho_{gy}$ for each $g \in G$ and $y \in X$. Let us observe that

$$\begin{aligned} g.\rho_y(z) &= \rho_{gy}(z) \\ &= d_X(gy, z) - d_X(gy, x_0) \\ &= d_X(y, g^{-1}z) - d_X(y, g^{-1}x_0) \\ &= \rho_y(g^{-1}z) - \rho_y(g^{-1}x_0), \end{aligned}$$

thus we can define the action of $g \in G$ on each $h \in \overline{X}^h$ as in (8). It is immediate from the definition that if $h_n \to h$ pointwise then $g.h_n \to g.h$, hence g acts continuously on $\overline{X}^h$, and so acts by homeomorphisms, as G is a group. □

We shall write $\overline{U}^h$ for the closure of U in $\overline{X}^h$. We remark that as $B(x_0, r)$ need not be compact, a sequence of points contained in a bounded set may have images under ρ which converge to the horofunction boundary, i.e. $\overline{\rho(B(x_0, r))}^h$ may contain points in the horofunction boundary ∂X^h.

3.1. **Horofunctions in δ-hyperbolic spaces.** We now record some basic observations about the behaviour of horofunctions. We start by describing the restriction of a horofunction ρ_y to a geodesic γ in X. As we shall see, for any geodesic γ, the restriction of ρ_y to γ has a coarsely well defined local minimum a bounded distance away from the nearest point projection p of y to γ. Moreover, for any point $z \in \gamma$, the value of $\rho_y(z)$ is equal to $\rho_y(p) + d_X(p, z)$, up to bounded error depending only on δ. We now make this precise.

Proposition 3.5. *Let γ be a geodesic in X, $y \in X$, and let p be a nearest point projection of y to γ. Then the restriction of ρ_y to γ is given by*

$$\rho_y(z) = \rho_y(p) + d_X(z, p) + O(\delta), \tag{9}$$

for all $z \in \gamma$.

Proof. This follows from the reverse triangle inequality, Proposition 2.2, by adding $-d_X(y, x_0)$ to both sides. □

We now describe the restriction of an arbitrary horofunction h to a geodesic γ in X. An *orientation* for a geodesic γ is a strict total order on the points of γ, induced by a choice of unit speed parameterization (thus, each geodesic has exactly two orientations). We may then define the *signed distance function* along γ to be

$$d^+_\gamma(x, y) = \begin{cases} d_X(x, y) & \text{if } x \leqslant y \\ -d_X(x, y) & \text{if } x \geqslant y. \end{cases}$$

Proposition 3.6. *Let h be a horofunction in $\overline{X}^h$, and let γ be a geodesic in X. Then there is a point p on γ such that the restriction of h to γ is equal to exactly one of the following two functions, up to bounded additive error:*

either

$$h(x) = h(p) + d_X(p, x) + O(\delta), \qquad x \in \gamma; \tag{10}$$

or

$$h(x) = h(p) + d^+_\gamma(p, x) + O(\delta), \qquad x \in \gamma \tag{11}$$

for some choice of orientation on γ.

So for example, for geodesic rays starting at the basepoint x_0, the graphs of these functions are equal to one of the two graphs shown below, up to an error of $O(\delta)$.

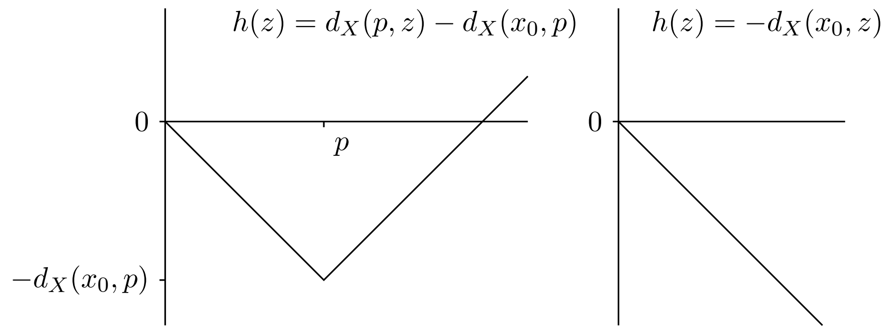

FIGURE 1. The behaviour of $h(z)$ along a geodesic ray starting at x_0.

Proof. Let h be a horofunction in $\overline{X}^h$, let $(\rho_{y_n})_{n\in\mathbb{N}}$ be sequence of horofunctions which converge to h, and let p_n be the nearest point projection of y_n to γ.

First, suppose that there is a subsequence of the $(\rho_{y_n})_{n\in\mathbb{N}}$ for which the projections p_n to γ are bounded, i.e. contained in a subinterval I of γ of finite length. We may pass to a further subsequence, which by abuse of notation we shall also call $(\rho_{y_n})_{n\in\mathbb{N}}$, such that the projections p_n converge to a point $p \in \gamma$, and in fact are all within distance 1 of p. Therefore, for this subsequence, $\rho_{y_n}(x) = \rho_{y_n}(p) + d_X(p, x) + O(\delta)$, by (9). As $\rho_{y_n} \to h$ pointwise, this implies that $h(x) = h(p) + d_X(p, x) + O(\delta)$ for each $x \in \gamma$, as required.

Now consider the case in which the nearest point projections p_n eventually exit every compact subinterval of γ. In this case it will be convenient to choose p to be a nearest point on γ to the basepoint x_0, and we may pass to a subsequence such that $p_n > p$ for all n, for some choice of orientation on γ.

For any $x \in \gamma$ all but finitely many p_n satisfy $p_n > x$. As p_n is a nearest point projection of y_n to γ, and p is a nearest point projection of x_0 to γ, we may use the reverse triangle inequality to rewrite $\rho_{y_n}(x)$ in terms of $d_X(x_0, p)$, $d_X(y_n, p_n)$, and distances between points on the geodesic γ. There are two cases, depending on whether $x > p$, or $x \leqslant p$, illustrated below in Figure 2.

The sign on each line segment in Figure 2 indicates the sign of the corresponding line segment in the approximation for $h_{y_n}(x)$. This shows that

$$\rho_{y_n}(x) = \begin{cases} -d_X(x_0, p) - d_X(p, x) + O(\delta), & \text{if } x > p, \\ -d_X(x_0, p) + d_X(p, x) + O(\delta), & \text{if } x \leqslant p. \end{cases}$$

As $(\rho_{y_n})_{n\in\mathbb{N}}$ converges to h, this implies the same coarse equalities for $h(x)$. As $h(p) = -d_X(x_0, p) + O(\delta)$, and using the definition of signed distance along γ, this gives (11) above. $\square$

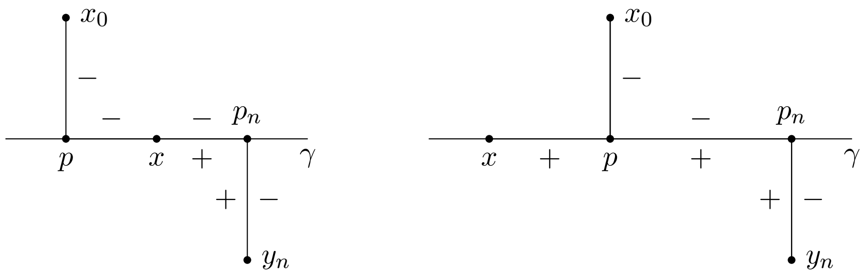

FIGURE 2. The horofunction ρ_{y_n} restricted to a geodesic γ.

We say a function $f(x)$ has no coarse local maxima if $f(x) = g(x) + O(\delta)$, where $g(x)$ has no local maxima. Similarly we say $f(x)$ has at most one coarse local minima if $f(x) = g(x) + O(\delta)$, where $g(x)$ has at most one local minima. So we have shown that any horofunction h restricted to a geodesic γ has no coarse local maxima on γ, and at most one coarse local minimum on γ.

3.2. **A partition of the horofunction boundary.** For any horofunction $h \in \overline{X}^h$ we may consider

$$\inf(h) = \inf_{y \in X} h(y),$$

which takes values in $[-\infty, 0]$. We may partition $\overline{X}^h$ into two sets depending on whether or not $\inf(h) = -\infty$.

Definition 3.7. *We shall denote* $\overline{X}^h_F$ *the set of* finite horofunctions

$$\overline{X}^h_F := \{h \in \overline{X}^h : \inf(h) > -\infty\},$$

and $\overline{X}^h_\infty$ *the set of* infinite horofunctions

$$\overline{X}^h_\infty := \{h \in \overline{X}^h : \inf(h) = -\infty\}.$$

Clearly, both $\overline{X}^h_F$ and $\overline{X}^h_\infty$ are invariant for the action of G. Note moreover that if a horofunction h is contained in $\rho(X)$, i.e $h = \rho_y$ for some $y \in X$, then

$$\inf(\rho_y) = -d_X(x_0, y) > -\infty,$$

and the infimum is achieved at the unique point $y \in X$, hence $\rho(X)$ is contained in the set $\overline{X}^h_F$ of finite horofunctions. More generally, by the same proof, if B is a bounded subset of X, then

$$\overline{\rho(B)}^h \subset \overline{X}^h_F.$$

Note however that the subset $\overline{X}^h_\infty \subset \overline{X}^h$ need not be compact, so there may be sequences of elements of $\overline{X}^h_\infty$ which do not have subsequences which converge in $\overline{X}^h_\infty$. If a sequence of horofunctions $(h_n)_{n \in \mathbb{N}}$ converges to h, this does not in general imply that $\inf h_n$ converges to $\inf h$. For example, in Example 1.6, consider the sequence $(\rho_{x_n})_{n\in\mathbb{N}}$, where x_n is the point distance n from 0 in the n-th ray. Then $\inf \rho_{x_n} = -n$, but $(\rho_{x_n})_{n\in\mathbb{N}}$ converges to ρ_{x_0}, for which $\inf \rho_{x_0} = 0$. In fact, in this example there are sequences in $\overline{X}^h_\infty$ which converge to horofunctions in $\overline{X}^h_F$. If we set h_n to be equal to the limit of $(\rho_{x_k})_{k\in\mathbb{N}}$, where the x_k are points distance k from 0 along the n-th ray, then $h_n \in \overline{X}^h_\infty$, but $(h_n)_{n\in\mathbb{N}}$ converges to ρ_{x_0}.

Lemma 3.8. *For each basepoint $x_0 \in X$, each horofunction $h \in \overline{X}^h$ and each pair of points $x, y \in X$ the following inequality holds:*

$$\min\{-h(x), -h(y)\} \leqslant (x \cdot y)_{x_0} + O(\delta). \tag{12}$$

Proof. Let $z \in X$. Then one has, by the triangle inequality

$$(x \cdot z)_{x_0} = \frac{d_X(x_0, x) + d_X(x_0, z) - d_X(x, z)}{2},$$

which implies

$$(x \cdot z)_{x_0} \geqslant d_X(x_0, z) - d_X(x, z),$$

and by definition, the right hand side is equal to $-\rho_z(x)$, which gives

$$(x \cdot z)_{x_0} \geqslant -\rho_z(x).$$

Now, by δ-hyperbolicity, combined with the previous estimate, one has

$$(x \cdot y)_{x_0} \geqslant \min\{(x \cdot z)_{x_0}, (y \cdot z)_{x_0}\} - O(\delta),$$

which we may rewrite as

$$(x \cdot y)_{x_0} \geqslant \min\{-\rho_z(x), -\rho_z(y)\} - O(\delta).$$

Since every horofunction is the pointwise limit of functions of type ρ_z, the claim follows. □

Definition 3.9. *We define a sequence $(x_n)_{n\in\mathbb{N}} \subseteq X$ to be* minimizing *for a horofunction h if $h(x_n) \to -\infty$ as $n \to \infty$.*

We shall now prove that every minimizing sequence is a Gromov sequence, hence it has a limit in the Gromov boundary.

Lemma 3.10. *Let $h \in \overline{X}^h_\infty$ an infinite horofunction, and (y_n) a sequence of points in X such that $h(y_n) \to -\infty$. Then the sequence (y_n) converges to a point in the Gromov boundary of X. Moreover, two minimizing sequences for the same horofunction converge to the same point in the Gromov boundary.*

Proof. By Lemma 3.8 one has

$$(y_n \cdot y_m)_{x_0} \geqslant \min\{-h(y_n), -h(y_m)\} - O(\delta) \to \infty$$

as $\min\{m, n\} \to \infty$, proving the first claim.

To prove uniqueness of the limit, suppose that there are two sequences (x_n) and (y_n) such that $h(x_n) \to -\infty$ and $h(y_n) \to -\infty$. Then, by Lemma 3.8, the Gromov product $(x_n \cdot y_n)_{x_0} \to \infty$, hence by definition the sequences (x_n) and (y_n) converge to the same point in the Gromov boundary ∂X. □

Note that if $h \in \overline{X}^h_\infty$, then there is actually a quasigeodesic sequence $(x_n)_{n\in\mathbb{N}}$ in X such that $h(x_n) \to -\infty$.

To see this, let $(y_n)_{n\in\mathbb{N}}$ be a sequence with $h(y_n) \to -\infty$: then by Lemma 3.10, $(y_n)_{n\in\mathbb{N}}$ converges to a point in the Gromov boundary. In particular, for any n there is an N_n such that $(y_{N_n} \cdot y_m)_{x_0} \geqslant n$ for all $m \geqslant N_n$, and so we may pass to a subsequence, which by abuse of notation we shall continue to call $(y_n)_{n\in\mathbb{N}}$, such that $(y_n \cdot y_m)_{x_0} \geqslant n$ for all $m \geqslant n$. For each n, choose a geodesic γ_n from x_0 to y_n, parameterized by arc length, and write $\gamma_n(t)$ for a point at distance t along this geodesic. Consider the sequence $x_n := \gamma_n(n)$. This sequence converges to the same boundary point as the original sequence $(y_n)_{n\in\mathbb{N}}$: indeed, as x_n lies on a geodesic from x_0 to y_n, this implies that $(x_n \cdot y_n)_{x_0} = n$, so $(x_n \cdot y_n)_{x_0} \to \infty$ as $n \to \infty$. Finally, we show that the sequence $(x_n)_{n\in\mathbb{N}}$ is quasigeodesic. As $(y_n \cdot y_m)_{x_0} \geqslant n$, for any pair x_m and x_n, with $m \geqslant n$, the two geodesics γ_n and γ_m $O(\delta)$-fellow travel for distance at least n,. This implies that the point x_n lies within $O(\delta)$ of the geodesic γ_m, and so $d_X(x_n, x_m) = m - n + O(\delta)$, which shows that the sequence $(x_n)_{n\in\mathbb{N}}$ is quasigeodesic.

Furthermore, we can recover the Gromov boundary from the horofunction boundary as follows. Define an equivalence relation on $\overline{X}^h$ by $h_1 \sim h_2$ if $\sup_{x\in X} |h_1(x) - h_2(x)|$ is finite. This collapses $\overline{X}^h_F$ to a single point, and the equivalence classes in $\overline{X}^h_\infty$ are precisely the point pre-images of the local minimum map $\phi\colon \overline{X}^h_\infty \to \partial X$, so the Gromov boundary is homeomorphic to $\overline{X}^h_\infty/_\sim$. However, we will not use this result so we omit the proof.

3.3. **The local minimum map.**

We now define a map $\phi\colon \overline{X}^h \to X \cup \partial X$, which may be thought of as the "local minimum" map, which sends a horofunction h to the location at which it attains its minimum. If the horofunction h does not attain a minimum in X, it turns out that it makes sense to think of the minimum value as lying in the Gromov boundary. We now make this precise.

Definition 3.11. *The* local minimum *map* $\phi\colon \overline{X}^h \to X \cup \partial X$ *is defined as follows.*

- *If* $h \in \overline{X}^h_F$, *i.e.* $\inf(h) > -\infty$, *then define*

$$\phi(h) := \{x \in X \ : \ h(x) \leqslant \inf h + 1\}$$

the set of points of X *where the value of* h *is close to its infimum;*

- *if $h \in \overline{X}^h_\infty$, i.e. $\inf(h) = -\infty$, then choose a sequence $(y_n)_{n\in\mathbb{N}}$ with $h(y_n) \to -\infty$, and set*
$$\phi(h) := \lim_{n\to\infty} y_n$$
to be the limit point of (y_n) in the Gromov boundary.

Lemma 3.12. *The local minimum map $\phi : \overline{X}^h \to X \cup \partial X$ is well-defined and G-equivariant.*

Proof. By Lemma 3.10, the map is well-defined on $\overline{X}^h_\infty$: indeed, every minimizing sequence for h converges in the Gromov boundary, and any two minimizing sequences yield the same limit.

To prove equivariance, let us first pick $h \in \overline{X}^h_F$, and $x \in \phi(h)$. Then for each $y \in X$ we have $h(x) \leqslant h(y) + 1$, thus
$$g.h(gx) = h(x) - h(g^{-1}x_0) \leqslant h(y) - h(g^{-1}x_0) + 1 = g.h(gy) + 1$$
for each $y \in X$, hence the value of $g.h$ at gx is close to its infimum hence $gx \in \phi(g.h)$. If instead $h \in \overline{X}^h_\infty$, then let (y_n) a minimizing sequence for h. Then by definition of the action one gets
$$g.h(gy_n) = h(y_n) - h(g^{-1}x_0) \to -\infty$$
hence (gy_n) is a minimizing sequence for $g.h$, so $\phi(g.h) = g.\phi(h)$ as required. □

Lemma 3.13. *There exists K, which depends only on δ, such that for each finite horofunction $h \in \overline{X}^h_F$ we have*
$$\operatorname{diam} \phi(h) \leqslant K.$$

Proof. Let $x, y \in \phi(h)$, for some $h \in \overline{X}^h_F$, and consider the restriction of h along a geodesic segment from x to y. By Proposition 3.6, the restriction has at most one coarse local minimum: hence, since x and y are coarse local minima of h, the distance between x and y is universally bounded in terms of δ. □

Proposition 3.14. *Let $(h_n)_{n\in\mathbb{N}}$ be a sequence of horofunctions which converges to some infinite horofunction $h \in \overline{X}^h_\infty$. Then $\phi(h_n)$ converges to $\phi(h)$ in the Gromov boundary. As a consequence, the local minimum map $\phi : \overline{X}^h_\infty \to \partial X$ is continuous.*

(If h_n is a finite horofunction, in the above statement we mean that $x_n \to \phi(h)$ for any choice of $x_n \in \phi(h_n)$.)

Proof. Let $N > 0$, and let (h_n) a sequence of horofunctions which converge to $h \in \overline{X}^h_\infty$. Let us pick a minimizing sequence $(x_m)_{m\in\mathbb{N}}$ for h, and for each n a sequence $(y_{m,n})_{m\in\mathbb{N}}$, such that $y_{m,n} \to \inf h_n \in \mathbb{R} \cup \{-\infty\}$ as $m \to \infty$, so that $\phi(h) = [x_m]$ and $\phi(h_n) = [y_{m,n}]$. The goal is to prove that $(\phi(h_n) \cdot \phi(h))_{x_0} \to \infty$ as $n \to \infty$.

Since (x_m) is a Gromov sequence and it is minimizing for h, there exists m_0 such that

$$h(x_{m_0}) \leqslant -N - 1$$

and for each $m, m' \geqslant m_0$ one has

$$(x_m \cdot x_{m'})_{x_0} \geqslant N + 1.$$

Since $h_n \to h$ pointwise, there exists n_0 such that

$$h_n(x_{m_0}) \leqslant -N$$

for each $n \geqslant n_0$. Now, since $(y_{m,n})$ is minimizing for h_n, there exists $m_1 = m_1(N, n)$ such that $m_1 \geqslant m_0$ and

$$h_n(y_{m,n}) \leqslant -N \qquad \text{for each } m \geqslant m_1$$

Hence, by Lemma 3.8 we have

$$(x_{m_0} \cdot y_{m,n})_{x_0} \geqslant \min\{-h_n(x_{m_0}), -h_n(y_{m,n})\} \geqslant N$$

and by property (2), for all $m, m' \geqslant m_1$

$$(x_{m'} \cdot y_{m,n})_{x_0} \geqslant \min\{(x_{m_0} \cdot y_{m,n})_{x_0}, (x_{m_0} \cdot x_{m'})_{x_0}\} - O(\delta) \geqslant N - \delta$$

thus $(\phi(h) \cdot \phi(h_n))_{x_0} = \sup \liminf_{m,m'} (x_{m'} \cdot y_{m,n})_{x_0} \geqslant N - O(\delta)$ for $n \geqslant n_0$, as claimed. □

Corollary 3.15. *The local minimum map* $\phi \colon \overline{X}^h_\infty \to \partial X$ *is surjective.*

Proof. Pick $\lambda \in \partial X$, let $\gamma : [0, \infty) \to X$ be a quasigeodesic ray converging to λ, and write γ_n for $\gamma(n)$. The sequence $(\gamma_n)_{n \in \mathbb{N}}$ of points of X converges to λ, and by compactness, the sequence of horofunctions ρ_{γ_n} has a subsequence $(\rho_{\gamma_{n_k}})_{k \in \mathbb{N}}$ which converges to some $h \in \overline{X}^h$. We now show that $h \in \overline{X}^h_\infty$. Consider $\rho_{\gamma_n}(\gamma_m)$ for $n > m$. The point γ_m lies a bounded distance from a geodesic from x_0 to γ_n, and so $\rho_{\gamma_n}(\gamma_m) = -d_X(x_0, \gamma_m) + O(\delta)$, for all $n \geqslant m$. As ρ_{γ_n} converges to h, this implies that $h(\gamma_m) = -d_X(x_0, \gamma_m) + O(\delta)$. As this holds for all m, this implies $\inf h = -\infty$, as required. Thus, by the Proposition, $\gamma_{n_k} \to \phi(h)$, hence by uniqueness of the limit $\phi(h) = \lambda$, as required. □

Note that $\inf_{x \in X} \rho(x)$ is not continuous in $\rho \in \overline{X}^h$, and we emphasize that if (h_n) converges to h in $\overline{X}^h_F$, then $(\phi(h_n))$ need not converge to $\phi(h)$. For instance, in the countable wedge of rays of Example 1.6, if x_n is the point on the branch X_n at distance n from the base point x_0, then $\rho_{x_n} \to \rho_{x_0}$ in the horofunction compactification, but the sequence $(x_n)_{n \in \mathbb{N}}$ does not converge in the Gromov boundary.

3.4. **Horofunctions and shadows.** We define the *depth* of the shadow $S = S_{x_0}(x, R)$ to be the quantity

$$dep(S) := 2R - d_X(x_0, x).$$

We now show that we may characterize points in a shadow in terms of the value of the corresponding horofunction at the basepoint, and the depth.

Lemma 3.16. *If $S = S_{x_0}(x, R)$ is a shadow, then $y \in S$ if and only if*

$$\rho_y(x) \leqslant dep(S).$$

Proof. By definition of shadow one has

$$(x \cdot y)_{x_0} \geqslant d_X(x_0, x) - R$$

hence by writing out the Gromov product

$$\tfrac{1}{2}(d_X(x_0, x) + d_X(x_0, y) - d_X(x, y)) \geqslant d_X(x_0, x) - R$$

and by simplifying we get

$$-\rho_y(x) \geqslant d_X(x_0, x) - 2R$$

which proves the claim. □

Corollary 3.17. *For any shadow $S = S_{x_0}(x, R)$, and for any $\epsilon > 0$, the closure of S in $\overline{X}^h$ satisfies*

$$\overline{S}^h \subseteq \{h \in \overline{X}^h : h(x) \leqslant dep(S)\} \subseteq \overline{S_{x_0}(x, R + \epsilon)}^h,$$

where $dep(S) = 2R - d_X(x_0, x)$.

It will also be useful to know how shadows are related to the topology of the Gromov boundary ∂X. We shall use the following property of shadows: there is a constant R_0, which only depends on the action of G on X, such that for any $g \in G$ the closure of the shadow $S_{x_0}(gx_0, R_0)$ in ∂X contains a non-empty open set. This follows from the Proposition below.

Proposition 3.18. *Let G be a group acting by isometries on a separable Gromov hyperbolic space X, such that G contains at least one hyperbolic isometry. Then there is a number $R_0 > 0$ such that for any $g \in G$ the set $\overline{S_{x_0}(gx_0, R_0)}^\delta$ contains a limit point of Gx_0 in its interior.*

This follows from the following result from Blachère, Haïssinsky and Mathieu [BHM11].

Proposition 3.19. [BHM11, Proposition 2.1] *For any $\epsilon > 0$ sufficiently small, and any $A > 0$, there are positive numbers C and R_0, such that for any $R > R_0$, and any $x \in X, y \in \partial X$ with $(x_0 \cdot y)_x \leqslant A$,*

$$B_\epsilon(y, \tfrac{1}{C} e^{\epsilon(R - d_X(x_0, x))}) \subset \overline{S_{x_0}(x, R)}^\delta \cap \partial X \subset B_\epsilon(y, C e^{\epsilon(R - d_X(x_0, x))}),$$

where $B_\epsilon(y, r)$ is the ball of radius r about y in the metric d_ϵ on ∂X.

Proof (of Proposition 3.18). Since G contains at least one hyperbolic isometry, then the limit set $\overline{Gx_0}^\delta$ contains at least two points in ∂X, which we call α^+ and α^-. Let α be a quasigeodesic from α^+ to α^-, and let p be a closest point on α to the basepoint x_0. Given a group element $g \in G$, consider the translate $g\alpha$, as illustrated in Figure 3 below.

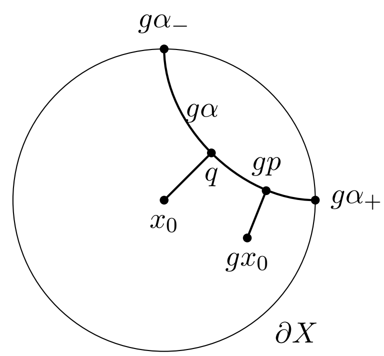

FIGURE 3. The translate of the quasigeodesic α under g.

For any element $g \in G$, let q be a nearest point on $g\alpha$ to x_0. Any quasigeodesic from x_0 to either $g\alpha_+$ or $g\alpha_-$ passes within distance $O(\delta)$ of q, so gx_0 lies within distance $d_X(x_0, \alpha) + O(\delta)$ of at least one of these quasigeodesics, which we may assume has endpoint α_+, up to relabeling.

Therefore, the product $(x_0 \cdot g\alpha^+)_{gx_0}$ is bounded above independently of $g \in G$, hence by Proposition 3.19, there is a number R_0, (depending only on δ and the choice of α) such that the closure $\overline{S_{x_0}(gx_0, R)}^\delta \cap \partial X$ contains an open set containing $g\alpha_+ \in \overline{Gx_0}^\delta$, for any $R \geqslant R_0$, as required. □

3.5. **Horofunctions and weak convexity of shadows.** *A priori*, shadows need not be convex, or even quasi-convex. However, we now show various results about horofunctions and nested shadows which we can think of as weak versions of convexity. For example, for any two points contained in a shadow $S_{x_0}(x, R)$, the geodesic connecting them is contained in $S_{x_0}(x, R+O(\delta))$. We start by showing that the value of a horofunction along a geodesic is bounded by its values on the endpoints, up to an additive error of $O(\delta)$.

Lemma 3.20. *Let X a δ-hyperbolic, geodesic metric space. Then there exists a constant $C \geqslant 0$, which depends only on δ, such that given any geodesic segment $[z_1, z_2]$ in X, with $z_1, z_2 \in X$, the following holds:*

$$\min\{\rho_{z_1}(x), \rho_{z_2}(x)\} - C \leqslant \rho_y(x) \leqslant \max\{\rho_{z_1}(x), \rho_{z_2}(x)\} + C \qquad (13)$$

for any $y \in [z_1, z_2]$ and any $x \in X$.

Proof. Let us first assume that x, x_0 belong to $[z_1, z_2]$. Up to swapping z_1 and z_2, we can assume $x \in [z_1, x_0]$; then we have the bound

$$\rho_{z_1}(x) = -d_X(x, x_0) \leqslant \rho_y(x) = d_X(x, y) - d_X(x_0, y) \leqslant d_X(x, x_0) = \rho_{z_2}(x)$$

which yields the claim. Now, in the general case let p_x be the closest point projection of x to $[z_1, z_2]$, and p_{x_0} the projection of x_0. By the definition of ρ_z,

$$\rho_z(x) = d_X(x, z) - d_X(x_0, z).$$

Then by the reverse triangle inequality (Proposition 2.2) we have for each $z \in X$,

$$\rho_z(x) = d_X(x, p_x) + d_X(p_x, z) - d_X(x_0, p_{x_0}) - d_X(p_{x_0}, z) + O(\delta),$$

hence for each $i = 1, 2$

$$\rho_y(x) - \rho_{z_i}(x) = \widetilde{\rho}_y(p_x) - \widetilde{\rho}_{z_i}(p_x) + O(\delta),$$

where $\widetilde{\rho}_z(x) := d_X(x, z) - d_X(p_{x_0}, z)$ denotes the horofunction based at p_{x_0}. Since now p_{x_0} and p_x lie on $[z_1, z_2]$, the claim follows by the previous case, and furthermore we may assume that the constant is positive. □

Lemma 3.20 implies the following weak convexity property of both shadows and their complements.

Corollary 3.21. *There exists a constant $C \geqslant 0$, which depends only on δ, such that for each shadow $S = S_{x_0}(x, R)$ the following hold:*

1. *if z_1 and z_2 belong to S, then the geodesic segment $[z_1, z_2]$ lies in $S_{x_0}(x, R + C)$;*
2. *if z_1 and z_2 do not belong to S, then the geodesic segment $[z_1, z_2]$ does not intersect $S_{x_0}(x, R - C)$.*

Moreover, for each (Q, c) which satisfy Proposition 2.1, the above statements still hold with "geodesic" replaced by "(Q, c)-quasi-geodesic", where this time C depends on δ, Q, and c.

Proof. If z_1, z_2 belong to S, then $\rho_{z_i}(x) \leqslant dep(S)$ for each $i = 1, 2$ by Lemma 3.16, hence Lemma 3.20 implies

$$\rho_y(x) \leqslant \max\{\rho_{z_1}(x), \rho_{z_2}(x)\} + C \leqslant dep(S) + C = dep(S')$$

where $S' = S_{x_0}(x, R + C/2)$, thus $y \in S'$ once again by Lemma 3.16. The proof of (2) is similar, using the left-hand side of equation (13). The extension to quasi-geodesic is immediate by the fellow-traveling property of Proposition 2.1. □

4. Convergence to the boundary

In this section we prove the following theorem.

Theorem 4.1. *Let G be a countable group of isometries of a geodesic, separable, δ-hyperbolic space X (not necessarily proper), and let μ be a non-elementary probability measure on G. Then for each $x_0 \in X$, almost every sample path $(w_n x_0)_{n\in\mathbb{N}}$ converges in $X \cup \partial X$ to a point of the Gromov boundary ∂X.*

The proof of the theorem takes several steps, and it exploits the action of G on the space of probability measures both on the horofunction compactification $\overline{X}^h$ and on the Gromov boundary ∂X. The strategy of the proof is the following:

(1) By compactness, there is a stationary measure ν on $\overline{X}^h$ (Lemma 4.3).
(2) The measure ν does not charge the finite part of the boundary: $\nu(\overline{X}^h_\infty) = 1$ (Proposition 4.4).
(3) By the martingale convergence theorem (Proposition 4.13), for almost every sample path the sequence of measures $(w_n\nu)_{n\in\mathbb{N}}$ converges to some measure ν_ω in $\mathcal{P}(\overline{X}^h)$.
(4) By pushing the sequence forward to ∂X, using the local minimum map ϕ, almost every sequence $(w_n\widetilde{\nu})_{n\in\mathbb{N}}$ converges to some measure $\phi_*\nu_\omega$ in $\mathcal{P}(\partial X)$ (Lemma 4.14).
(5) Almost every sample path $(w_n x_0)_{n\in\mathbb{N}}$ has a subsequence which converges to a point λ in the Gromov boundary ∂X (Proposition 4.7).
(6) Thus, by Lemma 4.15, almost every sample path $(w_n)_{n\in\mathbb{N}}$ has a subsequence $(w_{n_k})_{k\in\mathbb{N}}$ such that $(w_{n_k}\widetilde{\nu})_{k\in\mathbb{N}}$ converges to a delta-measure δ_λ, for some $\lambda \in \partial X$.
(7) Since the limit exists, almost every sequence $(w_n\widetilde{\nu})_{n\in\mathbb{N}}$ converges to a delta-measure δ_λ (Proposition 4.16).
(8) We prove that the fact that $(w_n\widetilde{\nu})_{n\in\mathbb{N}}$ converges to δ_λ implies that the sample path $(w_n x_0)_{n\in\mathbb{N}}$ converges to $\lambda \in \partial X$ (Proposition 4.18).

In the rest of the section we shall work out the details of the proof.

We remark that sample paths do not in general converge to points in the horofunction compactification.

Example 4.2. *Consider the nearest neighbour random walk on the Cayley graph of $F_2 \times \mathbb{Z}/2\mathbb{Z}$, with respect to the standard generating set $\langle a, b, c \mid [a, c], [b, c], c^2\rangle$, and with basepoint x_0 corresponding to the identity element. If $g \in F_2$, then $\rho_g(c) = 1$, and $\rho_{gc}(c) = -1$. As almost every sample path w_n hits each coset of F_2 infinitely often, sample paths do not converge in $\overline{X}^h$, almost surely.*

4.1. **Random walks.** We briefly review some background material on random walks and fix notation. Let μ be a probability distribution on G; the *step space* of the random walk generated by μ is the measure space $(G^{\mathbb{Z}}, \mu^{\mathbb{Z}})$, which is the countable infinite product of the measure spaces (G, μ). Each element of $G^{\mathbb{Z}}$ is a sequence $(g_n)_{n\in\mathbb{Z}}$, whose entries are the increments of our (bi-infinite) random walk. The shift map $T \colon G^{\mathbb{Z}} \to G^{\mathbb{Z}}$ sends $(g_n)_{n\in\mathbb{Z}}$ to $(g_{n-1})_{n\in\mathbb{Z}}$, and is measure preserving and ergodic.

We define the *location* of the random walk at time n, which we shall denote w_n to be

$$w_n = \begin{cases} g_0^{-1} g_{-1}^{-1} \cdots g_{n+1}^{-1} & \text{if } n \leqslant -1 \\ 1 & \text{if } n = 0 \\ g_1 g_2 \ldots g_n & \text{if } n \geqslant 1. \end{cases}$$

This gives a map $G^{\mathbb{Z}} \to G^{\mathbb{Z}}$, defined by $(g_n)_{n\in\mathbb{Z}} \mapsto (w_n)_{n\in Z}$. We shall denote the range of the location map as Ω, to distinguish it from the step space, and call $\mathbb{P}$ the pushforward to Ω of the product measure $\mu^{\mathbb{Z}}$ on the step space $G^{\mathbb{Z}}$. We shall refer to $(\Omega, \mathbb{P})$ as the *path space*, and elements $\omega \in \Omega$ as *sample paths* of the random walk. The shift map T acts on Ω by $T^k \colon (w_n)_{n\in\mathbb{Z}} \mapsto (w_k^{-1} w_{n+k})_{n\in\mathbb{Z}}$, and is measure preserving and ergodic.

4.2. **Stationary measures.** Let M be a metrizable topological space, and denote $\mathcal{P}(M)$ the space of Borel probability measures on M. The space $\mathcal{P}(M)$ is endowed with the weak-* topology, which is defined by saying $\nu_n \to \nu$ if for each continuous bounded function f on M one has $\nu_n(f) \to \nu(f)$.

If now G is a countable group which acts on M by homeomorphisms, we denote $g\nu$ the pushforward of $\nu \in \mathcal{P}(M)$ under the action of $g \in G$, i.e. $g\nu(U) = \nu(g^{-1}U)$, and define the *convolution operator* $\star \colon \mathcal{P}(G) \times \mathcal{P}(M) \to \mathcal{P}(M)$ as the average of the pushforwards:

$$\mu \star \nu := \sum_{g\in G} \mu(g) \; g\nu.$$

We say that a probability distribution ν on M is *μ-stationary* if $\mu \star \nu = \nu$, i.e. for each Borel set U we have

$$\nu(U) = \sum_{g\in G} \mu(g)\nu(g^{-1}U). \tag{14}$$

A space M equipped with a μ-stationary measure ν is called a *(G, μ)-space.* Now, if we fix a base point $x_0 \in M$, we shall write $\widetilde{\mu} \in \mathcal{P}(M)$ for the pushforward of μ under the orbit map, i.e. if $U \subset M$ then $\widetilde{\mu}(U) = \mu(\{g \in G : gx_0 \in U\})$. We shall write μ_n for the n-fold convolution of μ with itself on G, we shall write $\widetilde{\mu}_n$ for the pushforward of μ_n to M, and finally, we shall write $\overline{\mu}_n$ for the Cesàro averages of the pushforward measures, $\overline{\mu}_n := \frac{1}{n}(\widetilde{\mu} + \widetilde{\mu}_2 + \cdots + \widetilde{\mu}_n)$. Classical compactness arguments yield the following:

Lemma 4.3. *Let G be a countable group which acts by homeomorphisms on a compact metric space M, and let μ be a probability distribution on G. Then there exists a μ-stationary Borel probability measure ν on M.*

Proof. Since M is a compact metrizable space, then $\mathcal{P}(M)$ is compact in the weak-$*$ topology. Then any weak-$*$ limit point of the sequence $(\overline{\mu}_n)_{n\in\mathbb{N}}$ of the Cesàro averages is μ-stationary. An alternate proof follows from the Schauder-Tychonoff fixed point theorem. □

Applying the above arguments to $\overline{X}^h$ implies that there exists a μ-stationary measure ν on $\overline{X}^h$, i.e. $(\overline{X}^h, \nu)$ is a (G, μ)-space. We now show that the measure ν is supported on $\overline{X}^h_\infty$.

Proposition 4.4. *Let G be a countable group of isometries of a separable Gromov hyperbolic space X. Let μ be a non-elementary probability distribution on G, and let ν be a μ-stationary measure on $\overline{X}^h$. Then*

$$\nu(\overline{X}^h_F) = 0.$$

In order to show that some set Y has measure zero, the basic idea is to consider the translates gY of the set Y, and to consider the supremum of the measures of these sets. If we choose a translate gY with measure very close to the supremum, then by μ-stationarity, if hgY is another translate with $\mu(h) > 0$, then $\nu(hgY)$ will also be close to the supremum. If there are enough disjoint translates with ν-measures close to the supremum, then the total measure of ν is strictly greater than one, which contradicts the fact that ν is a probability measure. We now make this precise.

Lemma 4.5. *Let G be a countable group acting by homeomorphisms on a metric space M, let μ be a probability distribution on G, and let ν be a μ-stationary probability measure on M. Moreover, let us suppose that $Y \subset M$ has the property that there is a sequence of positive numbers $(\epsilon_n)_{n\in\mathbb{N}}$ such that for any translate fY of Y there is a sequence $(g_n)_{n\in\mathbb{N}}$ of group elements (which may depend on f), such that the translates $fY, g_1^{-1}fY, g_2^{-1}fY, \ldots$ are all disjoint, and for each g_n, there is an $m \geqslant 1$, such that $\mu_m(g_n) \geqslant \epsilon_n$. Then $\nu(Y) = 0$.*

The proof of this is a variation on [Mah10b, Lemma 3.5], but we provide a proof for the convenience of the reader.

Proof. Suppose that $s := \sup\{\nu(fY) : f \in G\} > 0$. Choose $N > 2/s$, let $\epsilon = \min\{\epsilon_i : 1 \leqslant i \leqslant N\}$, and let $\epsilon_s = \epsilon/N$. Finally, choose f such that the harmonic measure of fY is within ϵ_s of the supremum, i.e. $\nu(fY) \geqslant s - \epsilon_s$. By hypothesis, there is a sequence of group elements $g_1, \ldots, g_N$ such that the N translates $g_1^{-1}fY, \ldots, g_N^{-1}fY$, are all disjoint, and for each g_n there is an m such that with $\mu_m(g_n) \geqslant \epsilon$.

The harmonic measure ν is μ-stationary, and hence μ_m-stationary for any m, which implies

$$\nu(fY) = \sum_{h\in G} \mu_m(h)\nu(h^{-1}fY).$$

For any element $g \in G$ we may rewrite this as

$$\mu_m(g)\nu(g^{-1}fY) = \nu(fY) - \sum_{h \in G \setminus g} \mu_m(h)\nu(h^{-1}fY).$$

As we have chosen fY to have measure within ϵ_s of the supremum, this implies

$$\mu_m(g)\nu(g^{-1}fY) \geqslant s - \epsilon_s - \sum_{h \in G \setminus g} \mu_m(h)\nu(h^{-1}fY).$$

The harmonic measure of each translate of Y is at most the supremum s,

$$\mu_m(g)\nu(g^{-1}fY) \geqslant s - \epsilon_s - s \sum_{h \in G \setminus g} \mu_m(h),$$

and the sum of $\mu_m(h)$ over all $h \in G \setminus g$ is equal to $1 - \mu_m(g)$, which implies that

$$\mu_m(g)\nu(g^{-1}fY) \geqslant s - \epsilon_s - s(1 - \mu_m(g)).$$

Let us now suppose $g = g_n$ for some n. By hypothesis, there is an $m \geqslant 1$ such that $\mu_m(g) > 0$. For such an m we may divide by $\mu_m(g)$ to give the following estimate for the harmonic measure of the translate,

$$\nu(g^{-1}fY) \geqslant s - \epsilon_s/\mu_m(g).$$

By assumption, such an estimate holds for each of the N disjoint translates $g_i^{-1}fY$, and furthermore, $\mu_m(g_i) \geqslant \epsilon$ for each i. This implies that

$$\nu(\bigcup_{i=1}^{N} g_i^{-1}fY) = \nu(g_1^{-1}fY) + \cdots + \nu(g_N^{-1}fY) \geqslant N(s - \epsilon_s/\epsilon).$$

As we chose $N > 2/s$, and $\epsilon_s/\epsilon = 1/N$, this implies that the total measure $\nu(\bigcup g_i^{-1}fY)$ is greater than one, a contradiction. □

We now complete the proof of Proposition 4.4. Recall that the translation length $\tau(g)$ of an isometry g of X is defined to be

$$\tau(g) := \lim_{n \to \infty} \tfrac{1}{n} d_X(x_0, g^n x_0).$$

This definition is independent of the base point x_0, and $\tau(g) > 0$ if and only if g is a hyperbolic isometry, and furthermore $\tau(g^k) = k\tau(g)$.

Proof of Proposition 4.4. We shall apply Lemma 4.5 taking as Y the set of horofunctions whose local minimum lies in a given ball around the basepoint: precisely, $Y = \{h \in \overline{X}_F^h \ : \ \phi(h) \cap B(x_0, r) \neq \emptyset\}$, where ϕ is the local minimum map, and $B(x_0, r)$ is a ball of radius r in X. As μ is non-elementary, the semigroup generated by the support of μ contains hyperbolic isometries of arbitrarily large translation length. Choose such a hyperbolic isometry g with translation length $\tau(g)$ greater than $2r + K$, where K is the bound on the diameter of $\phi(h)$ from Lemma 3.13. Now, for any $f \in G$, the translates $g^{-n} f B(x_0, r)$ are all at least distance $\tau(g) - 2r > K$ apart, hence no $\phi(h)$ can intersect two of them, so the sets $g^{-n} f Y$ are all disjoint. As g lies in the semigroup generated by the support of μ, for each $n \geqslant 1$ there is an $m \geqslant 1$ such that $\mu_m(g^n) > 0$. Set $\epsilon_n = \mu_m(g^n)$, for some such m, then Lemma 4.5 implies that $\nu(Y) = 0$. As this holds for every r, this implies that $\nu(\overline{X}^h_F) = 0$, as required. □

The measure ν is therefore supported on $\overline{X}^h_\infty$, and as ϕ is continuous on $\overline{X}^h_\infty$, the measure ν pushes forward to a Borel probability measure

$$\widetilde{\nu} := \phi_* \nu$$

on the Gromov boundary ∂X. The measure $\widetilde{\nu}$ is a μ-stationary probability measure on ∂X, so $(\partial X, \widetilde{\nu})$ is a (G, μ)-space. We now show that $\widetilde{\nu}$ is non-atomic, which implies that ν is non-atomic as well. Recall that if the action of G on X is non-elementary, then G does not preserve any finite subset of the boundary ∂X.

Lemma 4.6. *Let G be a countable group which acts by isometries on a separable Gromov hyperbolic space X. Let μ be a non-elementary probability distribution on G, and let ν be a μ-stationary measure on $\overline{X}^h$, with pushforward $\widetilde{\nu}$ on ∂X under the local minimum map ϕ. Then the measure $\widetilde{\nu}$ is non-atomic (hence so is ν). Furthermore, any μ-stationary measure on ∂X is the pushforward of a μ-stationary measure on $\overline{X}^h_\infty$.*

Proof. Let Γ denote the group generated by the support of μ. We first observe that if there are atoms, then there must be an atom of maximal weight, as an infinite sequence of atoms $(b_n)_{n \in \mathbb{N}}$ of increasing weights has total measure greater than one. Let m be the maximal weight of any atom, and let A_m be the collection of atoms of weight m, which is a finite set. As $\widetilde{\nu}$ is μ-stationary, if $b \in A_m$, then

$$\widetilde{\nu}(b) = \sum_{g \in G} \mu(g) \widetilde{\nu}(g^{-1} b).$$

As no atom has weight greater than m, all elements of the orbit of b under Γ must have the same weight m, so A_m is a finite, Γ-invariant set, contradicting the fact that μ is non-elementary.

Finally, as the local minimum map $\phi \colon \overline{X}^h_\infty \to \partial X$ is surjective, the pushforward map $\phi_* \colon \mathcal{P}(\overline{X}^h_\infty) \to \mathcal{P}(\partial X)$ is also surjective, see e.g. [AB06, Theorem 15.14]. If λ is a μ-stationary measure on ∂X, then $\phi_*^{-1}(\lambda)$ is a non-empty, convex subspace of the space of measures on $\overline{X}^h_\infty$, which can also be seen as a subspace of the space $\mathcal{P}(\overline{X}^h)$ of probability measures on $\overline{X}^h$. Thus, the closure H of $\phi_*^{-1}(\lambda)$ in $\mathcal{P}(\overline{X}^h)$ is compact, convex, and invariant under convolution with μ, so by the Schauder-Tychonoff fixed point theorem there is a μ-stationary measure ν in $H \subseteq \mathcal{P}(\overline{X}^h)$. However, by Proposition 4.4, ν

vanishes on the set of finite horofunctions, hence it belongs to $\mathcal{P}(\overline{X}^h_\infty)$, and since ν is a limit of elements in $\phi_*^{-1}(\lambda)$ and ϕ_* is continuous, we also have $\phi_*\nu = \lambda$, as required. □

4.3. **Convergent subsequences.** The goal of this section is to prove the following step in the proof of Theorem 4.1:

Proposition 4.7. *Let G be a countable group of isometries of a separable Gromov hyperbolic space X, and let μ be a non-elementary probability distribution on G.*

Then, for $\mathbb{P}$-almost every sample path $(w_n)_{n\in\mathbb{N}}$ there is a subsequence of $(\rho_{w_n x_0})_{n\in\mathbb{N}}$ which converges to a horofunction in $\overline{X}^h_\infty$.

As a corollary, $\mathbb{P}$-almost every sample path $(w_n)_{n\in\mathbb{N}}$ has a subsequence $(w_{n_k})_{k\in\mathbb{N}}$ such that $(w_{n_k}x_0)_{k\in\mathbb{N}}$ converges to a point in the Gromov boundary ∂X.

Given a shadow $S = S_{x_0}(x, R)$, we define the *open shadow* $S^\circ = S^\circ_{x_0}(x, R)$ to be the subset of $\overline{X}^h$ given by

$$S^\circ_{x_0}(x, R) := \{h \in \overline{X}^h : h(x) < dep(S)\}.$$

As $\{x\}$ is compact and $(-\infty, dep(S))$ is open in $\mathbb{R}$, the set S° is an open subset of $\overline{X}^h$ contained in the interior of $\overline{S}^h$.

Lemma 4.8. *For each $T < 0$, the set $\overline{X}^h_\infty$ is contained in a countable collection of open shadows of depth $\leqslant T$.*

Proof. We have immediately from the definition

$$\overline{X}^h_\infty \subseteq \{h \in \overline{X}^h \mid \inf h < T\} = \bigcup_{x\in X} \{h \in \overline{X}^h \mid h(x) < T\}.$$

Now, by picking a countable dense subset $\{x_i\}_{i\in\mathbb{N}}$ of X and an enumeration $\{T_j\}_{j\in\mathbb{N}}$ of $(-\infty, T) \cap \mathbb{Q}$, we have

$$\overline{X}^h_\infty \subseteq \bigcup_{i,j\in\mathbb{N}} \{h \mid h(x_i) < T_j\},$$

as required. □

We shall now define a *descending shadow sequence* to be a sequence $\mathcal{S} = (\mathcal{S}_M)_{M\in\mathbb{N}}$, where each $\mathcal{S}_M$ is a finite collection of shadows, and each shadow $S \in \mathcal{S}_M$ has depth $dep(S) \leqslant -M$.

Given a descending shadow sequence $\mathcal{S}$, we shall introduce (for convenience of notation) an indexing of all its shadows, i.e. $\bigcup_M \mathcal{S}_M = \{S_1, S_2, \dots\}$. We say a M-tuple $I = (i_1, \dots, i_M)$ of positive integers is *an index set of depth M* for $\mathcal{S}$ if each $j = 1, \dots, M$, the shadow S_{i_j} is an element of $\mathcal{S}_j$ (hence, it has depth $\leqslant -j$). Given a descending shadow sequence $\mathcal{S}$, and an index set $I = (i_1, \ldots, i_M)$, we define the *cylinder* C_I to be the intersection of the open shadows $S^\circ_{i_j}$ corresponding to the indexed shadows S_{i_j}, i.e.

$$C_I := S^\circ_{i_1} \cap \cdots \cap S^\circ_{i_M},$$

so C_I is an open set in $\overline{X}^h$. Given a number $M \in \mathbb{N}$, let Σ_M be

$$\Sigma_M := \bigcup_{\substack{I \text{ index set} \\ \text{of depth } M}} C_I = \bigcup S^\circ_{i_1} \cap \cdots \cap S^\circ_{i_M},$$

where the union is taken over all index sets of depth M. The sets Σ_M form a nested sequence of open sets in $\overline{X}^h$, i.e.

$$\Sigma_1 \supseteq \Sigma_2 \supseteq \Sigma_3 \supseteq \cdots .$$

Finally, we observe that by Lemma 3.16, if $h \in \Sigma_M$ then $\inf h \leqslant -M$, so

$$\bigcap_{M \in \mathbb{N}} \Sigma_M \subseteq \overline{X}^h_\infty.$$

Lemma 4.9. *Let $(y_n)_{n\in\mathbb{N}}$ be a sequence of points of X, and let $(\mathcal{S}_M)_{M\in\mathbb{N}}$ be a descending shadow sequence. If $(\rho_{y_n})_{n\in\mathbb{N}}$ intersects Σ_M for each M, then there exists a subsequence $(y_{n_k})_{k\in\mathbb{N}}$ such that $(\rho_{y_{n_k}})_{k\in\mathbb{N}}$ converges to a horofunction in $\overline{X}^h_\infty$.*

Proof. Suppose the sequence $(\rho_{y_n})_{n\in\mathbb{N}}$ intersects each Σ_M. So for each $M \in \mathbb{N}$ there is an n_M such that $\rho_{y_{n_M}} \in \Sigma_M$. As $\bigcap \Sigma_M \subset \overline{X}^h_\infty$, each ρ_{y_n} may lie in only finitely many Σ_M, and so $n_M \to \infty$ as $M \to \infty$. The horofunction $\rho_{y_{n_k}}$ lies in Σ_k, which is a union of cylinders, so there is an index set $I_k = (i_1, \ldots, i_k)$ of depth k such that

$$\rho_{y_{n_k}} \in C_{I_k} = S^\circ_{i_1} \cap \cdots \cap S^\circ_{i_k}.$$

Recall that, by definition of descending shadow sequence and index set, there are only finitely many choices for the first entry i_1 in each index set I_k, so we may pass to a further subsequence in which i_1 is constant. Choose n_1 to be the first element of this subsequence, and relabel the remaining elements as $(n_k)_{k\in\mathbb{N}}$ for $k \geqslant 2$. Again, as there are only finitely many choices for the second entry i_2 in the index set I_k, we may pass to a further subsequence in which the indices i_2 are constant for all $k \geqslant 2$. Continuing by a standard diagonal procedure, we may construct a subsequence $(y_{n_k})_{k\in\mathbb{N}}$, and a sequence of indices $(i_k)_{k\in\mathbb{N}}$, such that

$$\rho_{y_{n_k}} \in S^\circ_{i_1} \cap \cdots \cap S^\circ_{i_k},$$

for each k.

By compactness, the sequence $(\rho_{y_{n_k}})_{k\in\mathbb{N}}$ has a limit point $h \in \overline{X}^h$. Then by Lemma 3.16, for each j and each $k \geqslant j$, if $S^\circ_{i_j} = S^\circ_{x_0}(x_j, R_j)$, we have $y_{n_k} \in S^\circ_{i_j}$, hence

$$\rho_{y_{n_k}}(x_j) \leqslant -j,$$

thus by passing to the limit as $k \to \infty$ we have

$$h(x_j) \leqslant -j \qquad \text{for all } j \geqslant 1,$$

which gives $h \in \overline{X}^h_\infty$. □

Lemma 4.10. *Let $\epsilon > 0$, and let ν be a μ-stationary measure on $\overline{X}^h$. Then there exists a finite descending shadow sequence $\mathcal{S} = (\mathcal{S}_M)_{M \in \mathbb{N}}$ such that for each M,*

$$\nu(\Sigma_M) \geqslant 1 - \epsilon.$$

Proof. Recall by Lemma 4.8 that for any $M \in \mathbb{N}$ there exists a countable collection of shadows of depth $\leqslant -M$ which covers $\overline{X}^h_\infty$, and $\overline{X}^h_\infty$ has full measure by Proposition 4.4. Thus, since probabilities are countably additive one can find a finite set $\mathcal{S}_M := \{S_{M,1}, \dots, S_{M,r_M}\}$ of shadows of depth $\leqslant -M$ such that the union $\overline{\mathcal{S}}_M := \bigcup_{i=1}^{r_M} S_{M,i}$ satisfies

$$\nu(\overline{\mathcal{S}}_M) \geqslant 1 - 2^{-M}\epsilon.$$

We may now set $\mathcal{S}$ to be the sequence $(\mathcal{S}_M)_{M \in \mathbb{N}}$, which is a descending shadow sequence. We now observe that $\nu(\Sigma_M) \geq 1 - \epsilon/2 - \cdots - \epsilon/2^M \geqslant 1 - \epsilon$ as required. □

We may now complete the proof of Proposition 4.7.

Proof (of Proposition 4.7). We shall show that the set Z of sequences $(w_n)_{n \in \mathbb{N}}$ in the path space $(\Omega, \mathbb{P})$ for which $(\rho_{w_n})_{n \in \mathbb{N}}$ does not have limit points in $\overline{X}^h_\infty$ has measure at most ϵ, for each $\epsilon > 0$. Fix $\epsilon > 0$, and let $(\mathcal{S}_M)_{M \in \mathbb{N}}$ be a descending shadow sequence constructed according to Lemma 4.10, using a measure ν which is a μ-stationary weak limit of the Cesàro averages $\overline{\mu}_n = \frac{1}{n}(\widetilde{\mu}_1 + \cdots + \widetilde{\mu}_n)$. Now, suppose that the sequence $(\rho_{w_n})_{n \in \mathbb{N}}$ does not have limit points in $\overline{X}^h_\infty$: then by Lemma 4.9 there exists an index M such that ρ_{w_n} does not belong to Σ_M for any n. Thus we have the inclusion

$$Z \subseteq \bigcup_M \bigcap_n \{(w_k)_{k \in \mathbb{N}} \; : \; \rho_{w_n} \notin \Sigma_M\},$$

so if we set $Y_M := \overline{X}^h \setminus \Sigma_M$, then

$$\mathbb{P}(Z) \leqslant \sup_M \inf_n \widetilde{\mu}_n(Y_M).$$

Then, by definition of the Cesàro averages, $\inf \overline{\mu}_n(Y_M) \geqslant \inf \widetilde{\mu}_n(Y_M)$, so this implies that

$$\mathbb{P}(Z) \leqslant \sup_M \inf_n \overline{\mu}_n(Y_M).$$

Furthermore as ν is a weak limit point of the Cesàro averages $\overline{\mu}_n$, and Y_M is closed, we have for each M,

$$\inf_n \overline{\mu}_n(Y_M) \leqslant \nu(Y_M),$$

hence by Lemma 4.10

$$\mathbb{P}(Z) \leqslant \sup_M \nu(Y_M) \leqslant \epsilon,$$

and the claim is proven. The corollary follows from Proposition 3.14. □

4.4. **The boundary action.** We now prove Theorem 1.1, convergence to the boundary. We start by showing that the action of G on $X \cup \partial X$ satisfies the following property (which need not hold for the action of G on $\overline{X}^h$): if the sequence $(g_n x_0)_{n\in\mathbb{N}}$ converges to a point in ∂X, then the sequence $(g_n y)_{n\in\mathbb{N}}$ converges to the same point, for any $y \in X$.

Lemma 4.11. *Let X be a Gromov hyperbolic space. If $x_0 \in X$ and the sequence $(g_n x_0)_{n\in\mathbb{N}}$ converges to a point λ in ∂X, then the sequence $(g_n y)_{n\in\mathbb{N}}$ converges to the same point λ, for any $y \in X$.*

Proof. Consider the Gromov product

$$(g_n x_0 \cdot g_n y)_{x_0} = \tfrac{1}{2}(d_X(x_0, g_n x_0) + d_X(x_0, g_n y) - d_X(g_n x_0, g_n y)).$$

By the triangle inequality $d_X(x_0, g_n y) \geqslant d_X(x_0, g_n x_0) - d_X(g_n x_0, g_n y)$, and as g_n is an isometry, $d_X(g_n x_0, g_n y) = d_X(x_0, y)$. This implies

$$(g_n x_0 \cdot g_n y)_{x_0} \geqslant d_X(x_0, g_n x_0) - d_X(x_0, y),$$

which tends to infinity as n tends to infinity, so $(g_n y)_{n\in\mathbb{N}}$ converges to the same limit point as $(g_n x_0)_{n\in\mathbb{N}}$. □

The following is a version of Kaimanovich [Kai00, Lemma 2.2] in the non-proper case.

Lemma 4.12. *Let G be a group of isometries of a Gromov hyperbolic space X. Let $(g_n)_{n\in\mathbb{N}}$ be a sequence in G such that $(g_n x_0)_{n\in\mathbb{N}} \to \lambda \in \partial X$. Then there is a subsequence $(g_{n_k})_{k\in\mathbb{N}}$ such that $(g_{n_k} x)_{k\in\mathbb{N}} \to \lambda$ for all but at most one point of $X \cup \partial X$.*

Proof. We show that there is a subsequence $(g_{n_k})_{k\in\mathbb{N}}$ for which there is at most one point $b \in \partial X$ such that $(g_{n_k} b)_{k\in\mathbb{N}} \not\to \lambda$. Suppose there is a point b_1 in ∂X such that $(g_n b_1)_{n\in\mathbb{N}} \not\to \lambda$. This implies there is an open set U_1 in $X \cup \partial X$ containing λ such that infinitely many $g_n b_1$ do not lie in U_1. Therefore, we may pass to a subsequence $(g_{n_k})_{k\in\mathbb{N}}$ such that $g_{n_k} b_1 \notin U_1$ for all k. Now suppose there is another point $b_2 \in \partial X$ such that $(g_{n_k} b_2)_{k\in\mathbb{N}} \not\to \lambda$. This implies there is an open set U_2 containing λ such that infinitely many $g_{n_k} b_2$ do not lie in U_2. As before, we may pass to a subsequence, which by abuse of notation we shall continue to call $(g_{n_k})_{k\in N}$, such that $(g_{n_k} b_2)_{k\in\mathbb{N}} \notin U_2$ for all n. As $U_1 \cap U_2$ is also an open neighbourhood of λ, it contains a shadow set of the form $S_1 = \overline{S_{x_0}(x, R)}^\delta$, with λ contained in the interior of the slightly smaller shadow $S_2 = \overline{S_{x_0}(x, R - C)}^\delta$, where C will be chosen as the weak convexity constant of Corollary 3.21 for (Q, c)-quasigeodesics. Let now γ be a (Q, c)-quasigeodesic from b_1 to b_2, and $y \in X$ a point on γ. By Corollary 3.21, since the endpoints $g_{n_k} b_1$ and $g_{n_k} b_2$ do not belong to S_1, then the point $g_{n_k} y$ does not belong to S_2. However, this is a contradiction, as by Lemma 4.11 we know that $g_{n_k} y \to \lambda$, hence it must eventually lie inside S_2. □

4.5. **Convergence of measures.** We will use the following result, which goes back to Furstenberg [Fur63, Corollary 3.1], see also Margulis [Mar91, Chapter VI].

Proposition 4.13. *Let M be a compact metric space on which the countable group G acts continuously, and ν a μ-stationary Borel probability measure on M. Then for $\mathbb{P}$-almost all sequences $(w_n)_{n \in \mathbb{N}}$ the limit*

$$\nu_\omega := \lim_{n \to \infty} g_1 g_2 \dots g_n \nu$$

exists in the space $\mathcal{P}(M)$ of probability measures on M.

To give a brief overview of the argument, one proves that, since the measure is stationary, for each continuous function $f \in C(M)$ the process

$$X_n := \int_X f(g_1 \dots g_n x) \ d\nu(x)$$

is a bounded martingale, hence converges almost surely; this defines a positive linear functional on the space $C(M)$, which is thus represented by a Borel measure.

Furthermore, if $w_n \nu \to \nu_\omega$ a.s., then we get the integral formula

$$\nu = \int_\Omega \nu_\omega \ d\mathbb{P}(\omega). \tag{15}$$

Lemma 4.14. *Let G be a countable group of isometries of a separable Gromov hyperbolic space X, let μ be a non-elementary probability distribution on G, and let $\widetilde{\nu}$ be a μ-stationary Borel measure on ∂X.*

Then for almost every sample path $\omega = (w_n)_{n \in \mathbb{N}}$, the sequence of measures $(w_n \widetilde{\nu})_{n \in \mathbb{N}}$ converges to a measure $\widetilde{\nu}_\omega \in \mathcal{P}(\partial X)$.

Proof. By Lemma 4.6, there is a μ-stationary probability measure ν on $\overline{X}^h_\infty$ such that $\widetilde{\nu}$ is the pushforward of ν. Applying Proposition 4.13 to the action of G on $\overline{X}^h$, the sequence $(w_n \nu)_{n \in \mathbb{N}}$ converges to a measure $\nu_\omega \in \mathcal{P}(\overline{X}^h)$, almost surely. Moreover, since ν vanishes on $\overline{X}^h_F$ and $\overline{X}^h_F$ is G-invariant, the measures $w_n \nu$ also vanish on $\overline{X}^h_F$ for each w_n; furthermore, by equation (15) the limit ν_ω also vanishes on $\overline{X}^h_F$ for almost every ω. Now, note that $\overline{X}^h_\infty$ is a countable intersection of open subsets of $\overline{X}^h$, hence it is a Borel subset of $\overline{X}^h$, so the weak-$*$ topology on $\mathcal{P}(\overline{X}^h_\infty)$ (arising from $C_b(\overline{X}^h_\infty)$) is the relativization of the weak-$*$ topology on $\mathcal{P}(\overline{X}^h)$, see e.g. [AB06, Theorem 15.4]. Since $w_n \nu \to \nu_\omega$ a.s. in $\mathcal{P}(\overline{X}^h)$ and both $w_n \nu$ and ν_ω belong to $\mathcal{P}(\overline{X}^h_\infty)$, this implies that a.s. $w_n \nu \to \nu_\omega$ in the weak-$*$ topology of $\mathcal{P}(\overline{X}^h_\infty)$. Finally, since ϕ is continuous, the pushforward map $\mathcal{P}(\overline{X}^h_\infty) \to \mathcal{P}(\partial X)$ is continuous, hence $w_n \widetilde{\nu} = \phi_* w_n \nu \to \phi_* \nu_\omega = \widetilde{\nu}_\omega$ as claimed. □

We wish to show that for $\mathbb{P}$-almost all sequences $(w_n)_{n\in\mathbb{N}}$, the measure $\widetilde{\nu}_\omega$ is a δ-measure, and as the limit exists, it suffices to show this for any subsequence $(w_{n_k})_{k\in\mathbb{N}}$. We have already shown, by Proposition 4.7, that almost every sequence $(w_n)_{n\in\mathbb{N}}$ in $(\Omega, \mathbb{P})$ has a subsequence $(w_{n_k})_{k\in\mathbb{N}}$ such that $(w_{n_k}x_0)_{k\in\mathbb{N}}$ converges to a point in ∂X, so it suffices to show that if a sequence $(g_n)_{n\in\mathbb{N}}$ has the property that $(g_n x_0)_{n\in\mathbb{N}}$ converges to $\lambda \in \partial X$, then the measures $(g_n\widetilde{\nu})_{n\in\mathbb{N}}$ converge to δ_λ.

Lemma 4.15. *Let G be a non-elementary countable group of isometries of a separable Gromov hyperbolic space X, and let $(g_n)_{n\in\mathbb{N}}$ be a sequence in G such that $g_n x_0 \to \lambda \in \partial X$. Then for any non-atomic probability measure $\widetilde{\nu}$ on ∂X there is a subsequence $(g_{n_k})_{k\in\mathbb{N}}$ such that the translations $(g_{n_k}\widetilde{\nu})_{k\in\mathbb{N}}$ converge in the weak-$*$ topology to a delta-measure δ_λ, supported on λ.*

Proof. By Lemma 4.12, there exists a subsequence $(n_k)_{k\in\mathbb{N}}$ and a point $x \in \partial X$ such that for all $b \in \partial X, b \neq x$, one has $g_{n_k} b \to \lambda$. Since the measure $\widetilde{\nu}$ is non-atomic, $\widetilde{\nu}(\{x\}) = 0$ and the claim follows by the dominated convergence theorem. □

Proposition 4.16. *Let G be a countable group of isometries of a separable Gromov hyperbolic space X, let μ be a non-elementary probability distribution on G, and let $\widetilde{\nu}$ be a μ-stationary probability measure on ∂X.*

Then for almost every sample path $\omega = (w_n)_{n\in\mathbb{N}}$, there is a boundary point $\lambda(\omega) \in \partial X$ such that the sequence of measures $(w_n\widetilde{\nu})_{n\in\mathbb{N}}$ converges in $\mathcal{P}(\partial X)$ to a delta-measure $\delta_{\lambda(\omega)}$.

Proof. By Proposition 4.7, almost every sample path ω has a subsequence which converges to some point $\lambda(\omega) \in \partial X$. Thus, by Lemma 4.15, there exists a subsequence $(w_{n_k})_{k\in\mathbb{N}}$ such that $(w_{n_k}\widetilde{\nu})_{k\in\mathbb{N}} \to \delta_{\lambda(\omega)}$. By Lemma 4.14, the sequence $(w_n\widetilde{\nu})_{n\in\mathbb{N}}$ has a limit, hence the limit must coincide with $\delta_{\lambda(\omega)}$. □

4.6. **Convergence to the boundary: end of proof.** We have shown that $\mathbb{P}$-almost every sequence $(w_n\widetilde{\nu})_{n\in\mathbb{N}}$ converges to a δ-measure on ∂X. Finally, we now show that this implies that $\mathbb{P}$-almost every sequence $(w_n x_0)_{n\in\mathbb{N}}$ converges to some point in the Gromov boundary ∂X.

We start by showing that there are two shadows with disjoint closures and positive measure.

Lemma 4.17. *Let X be a separable, Gromov hyperbolic space, and $\widetilde{\nu}$ a non-atomic probability measure on ∂X. Then there exist two shadows S_1, S_2 in X such that their closures $U_i := \overline{S_i}^\delta$ in ∂X are disjoint, and both have positive $\widetilde{\nu}$-measure.*

Proof. As ∂X is a separable metric space, the support of $\widetilde{\nu}$ is a non-empty closed set, and, since $\widetilde{\nu}$ is non-atomic, it contains at least two points λ_1 and λ_2. Now, for each λ_i we can choose a shadow S_i in X such that λ_i is contained in the interior of the closure $U_i = \overline{S_i}^\delta$, hence $\widetilde{\nu}(U_i) > 0$ for each

$i = 1, 2$, and such that the distance parameter of each S_i is much larger than $(\lambda_1 \cdot \lambda_2)_{x_0}$, so that U_1 and U_2 are disjoint. □

The next proposition completes the proof of Theorem 4.1.

Proposition 4.18. *Let G be a countable group of isometries of a separable Gromov hyperbolic space X, let μ be a non-elementary probability distribution on G, and let $\widetilde{\nu}$ be a μ-stationary measure on ∂X.*

Suppose that the sequence $(w_n \widetilde{\nu})_{n \in \mathbb{N}}$ converges to a delta-measure δ_λ on ∂X, for some $\lambda \in \partial X$. Then $(w_n x_0)_{n \in \mathbb{N}}$ converges to λ in $X \cup \partial X$.

Proof. Let $U_1 = \overline{S_1}^\delta$ and $U_2 = \overline{S_2}^\delta$ as in Lemma 4.17, applied to the stationary measure $\widetilde{\nu}$. Set $\epsilon = \min\{\widetilde{\nu}(U_1), \widetilde{\nu}(U_2)\}$; as U_1 and U_2 both have positive $\widetilde{\nu}$-measure, ϵ is strictly greater than zero. As the sequence of measures $(w_n \widetilde{\nu})_{n \in \mathbb{N}}$ converges to the delta-measure δ_λ, for any shadow set $V = \overline{S_{x_0}(x, R)}^\delta$ containing λ in its interior, there is an N such that

$$w_n \widetilde{\nu}(V) \geqslant 1 - \epsilon/2 \qquad \text{for all } n \geqslant N.$$

As $w_n \widetilde{\nu}(w_n U_1) = \widetilde{\nu}(U_1) \geqslant \epsilon$, the set $w_n U_1$ intersects V, hence $w_n S_1$ intersects the (slightly larger) shadow $S_0 = S_{x_0}(x, R + O(\delta))$, and similarly $w_n S_2$ intersects S_0. For each $i = 1, 2$, let us pick $y_i \in S_i \cap w_n^{-1} S_0$ and denote $x_i = w_n y_i \in S_0$.

By disjointness of U_1 and U_2, there exists a constant C such that

$$(y_1 \cdot y_2)_{x_0} \leqslant C$$

for each $y_1 \in U_1$ and $y_2 \in U_2$. Moreover, since G acts by isometries, we get

$$(x_1 \cdot x_2)_{w_n x_0} = (y_1 \cdot y_2)_{x_0} \leqslant C$$

hence we can bound the distance from $w_n x_0$ to the geodesic $[x_1, x_2]$ as

$$d_X(w_n x_0, [x_1, x_2]) = (x_1 \cdot x_2)_{w_n x_0} + O(\delta) \leqslant C + O(\delta);$$

note that the constant on the right-hand side depends only on U_1, U_2 and δ, and not on n.

As x_1 and x_2 both lie in S_0, by weak convexity (Corollary 3.21), a geodesic $[x_1, x_2]$ connecting them lies in a shadow $S_{x_0}(x, R + O(\delta))$, and as $w_n x_0$ is a bounded distance from $[x_1, x_2]$, this implies that $w_n x_0$ lies in the slightly larger shadow $S_0^+ = S_{x_0}(x, R + C + O(\delta))$, for all $n \geqslant N$.

As this holds for all shadow sets V containing λ in their interiors, this implies that $(w_n x_0)_{n \in \mathbb{N}}$ converges to λ, as required. □

Theorem 4.1 implies the convergence statement in Theorem 1.1, and so it remains to show that the hitting measure $\widetilde{\nu}$ is the unique μ-stationary measure on ∂X, and the convolution measures $(\widetilde{\mu}_n)$ converge weakly to $\widetilde{\nu}$.

By Proposition 4.16, for any μ-stationary measure $\widetilde{\nu}$ on ∂X, for $\mathbb{P}$-almost every sample path ω, the sequence of measures $(w_n \widetilde{\nu})_{n \in \mathbb{N}}$ converges to $\delta_{\lambda(\omega)}$, where $\lambda(\omega)$ is the limit point of the sample path $(w_n x_0)_{n \in \mathbb{N}}$, hence it only depends on ω, not on $\widetilde{\nu}$. Thus, uniqueness follows from the integral formula (15). Furthermore,

$$\widetilde{\mu}_n = \int_\Omega w_n \delta_{x_0} \, d\mathbb{P}(\omega).$$

We may take the limit as n tends to infinity, and by the integral formula (15), the distribution of the limit points is given by ν, so $\widetilde{\mu}_n$ weakly converges to $\widetilde{\nu}$.

5. Applications

In this section we use convergence to the boundary to show the results on positive drift, sublinear tracking and translation length.

We will no longer use measures on the horofunction boundary, and so we shall from now on simply denote by ν the hitting measure on the Gromov boundary ∂X. Moreover, given $S \subseteq X$, the symbol $\overline{S}$ from now on will always mean the closure of S in the space $X \cup \partial X$.

5.1. **Hitting measures of shadows.** We start by showing that the measure of a shadow tends to zero as the distance parameter of the shadow tends to infinity. In order to simplify notation, we shall denote

$$Sh(x_0, r) := \{S_{x_0}(gx_0, R) \ : \ g \in G, \ d_X(x_0, gx_0) - R \geqslant r\}$$

the set of shadows based at x_0, with centers on the orbit Gx_0 and with distance parameter $\geqslant r$.

Proposition 5.1. *Let G be a countable group of isometries of a separable Gromov hyperbolic space X. Let μ be a non-elementary probability distribution on G, and let ν be the hitting measure on ∂X. Then we have*

$$\lim_{r \to \infty} \sup_{S \in Sh(x_0, r)} \nu(\overline{S}) = 0.$$

This result also holds for the reflected measures $\check{\mu}_n$ and $\check{\nu}$ determined by $\check{\mu}(g) = \mu(g^{-1})$, as if μ satisfies the hypotheses of Proposition 5.1, then so does $\check{\mu}$.

Proof. Propositions 3.18 and 3.19 show that a shadow centered at gx_0 of distance parameter r is contained in a ball of radius $Ce^{-\epsilon r}$ in the metric d_ϵ on ∂X, where C is independent of r, and as ν is non-atomic. It therefore suffices to show that for any non-atomic probability measure ν on a metric space (M, d),

$$\sup_{x \in M} \nu(B(x, r)) \to 0 \text{ as } r \to 0.$$

Suppose not, then there is an $\epsilon > 0$ and a sequence of metric balls $B(x_n, r_n)$ with $r_n \to 0$ and $\nu(B(x_n, r_n)) \geqslant \epsilon > 0$, for all n. Consider

$$U_n = \bigcup \{B(x, r) \mid r \leqslant 1/n \text{ and } \nu(B(x, r)) \geqslant \epsilon\},$$

and let $U = \bigcap U_n$. The U_n form a nested decreasing family of sets, i.e. $U_m \subseteq U_n$ for all $m \geqslant n$, and furthermore, $\nu(U_n) \geqslant \epsilon$ for all n. This implies that $\nu(U) \geqslant \epsilon$, and so in particular U is non-empty. For any point $x \in U$, $\nu(B(x, r)) \geqslant \epsilon$ for all $r > 0$, and this contradicts the fact that ν is non-atomic. □

If μ has bounded range in X then the argument from [Mah12, Lemma 2.10] shows that the ν- and μ_n-measures of a shadow $\overline{S_{x_0}(gx_0, R)}$ decay exponentially in the distance parameter, i.e. there are positive constants K and $c < 1$ such that

$$\nu(\overline{S_{x_0}(gx_0, R)}) \leqslant Kc^{d_X(x_0, gx_0) - R}. \tag{16}$$

For U a subset of X, let $H^+_x(U)$ denote the probability that a random walk starting at x ever hits U in forward time, i.e.

$$H^+_x(U) := \mathbb{P}(\exists\ n \geqslant 0\ :\ w_n x \in U).$$

Similarly, let $H^-_x(U)$ be the probability that a random walk starting at x ever hits U in reverse time, i.e. the probability that $w_n x$ lies in U for some $n \leqslant 0$.

Proposition 5.2. *Let G be a countable group which acts by isometries on a separable Gromov hyperbolic space X, and μ a non-elementary probability distribution on G. Then*

$$\sup_{S \in Sh(x_0, r)} H^+_{x_0}(S) \to 0 \qquad \textit{as } r \to \infty.$$

This immediately implies the same result with $H^+_{x_0}$ replaced by $H^-_{x_0}$, by replacing μ with the reflected measure $\check{\mu}(g) = \mu(g^{-1})$.

Proof. Suppose a sample path starting at x_0 hits a shadow $S_1 = S_{x_0}(x, R)$ with $x = hx_0$ in forward time, at gx_0 say. Let γ be a geodesic from x_0 to x, and let p be a nearest point on γ to gx_0, as illustrated below in Figure 4.

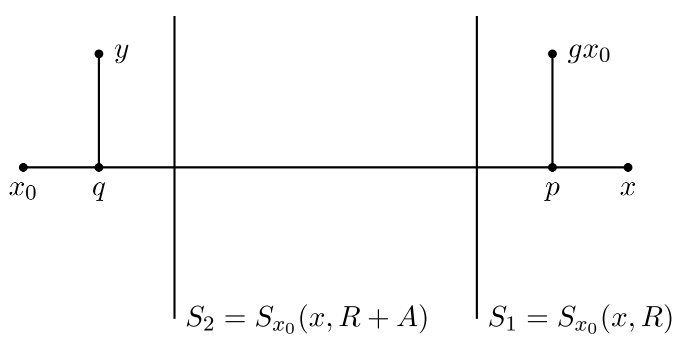

FIGURE 4. Nested shadows

Consider the shadow $S_2 = S_{x_0}(x, R + A)$, for some fixed $A > 0$ which we will choose later. The main idea is that if the random walk ever hits S_1, it will likely converge inside the closure of S_2, and the probability of that happening is small if the distance parameter of S_2 is large.

To make the idea precise, let y be a point in the complement of S_2, and let q be the nearest point projection of y to γ. By Proposition 2.4, p is within distance $O(\delta)$ of S_1, and q is within distance $O(\delta)$ of the complement of S_2, so the distance between p and q is at least $A+O(\delta)$. Now using Proposition 2.3, if $A \geqslant O(\delta)$, then the Gromov product satisfies

$$(x_0 \cdot y)_{gx_0} = d_X(gx_0, p) + d_X(p, q) + O(\delta) \geqslant A + O(\delta).$$

Therefore, the complement of S_2 is contained in a shadow $S_3 = S_{gx_0}(x_0, R')$ (where $R' = d_X(gx_0, x_0) - A + O(\delta)$) of distance parameter $A + O(\delta)$.

Fix some positive $\epsilon < 1$; then, since the measure of shadows tends to zero as the distance parameter tends to infinity (Proposition 5.1), there is a number A_0 sufficiently large such that $\nu(\overline{S}) \leqslant \epsilon$ for all shadows $S \in Sh(x_0, A_0)$ with distance parameter larger than A_0. As a consequence, if we choose A such that $A + O(\delta) \geqslant A_0$ in the above construction, we have

$$\nu(g^{-1}\overline{S_3}) = \nu(\overline{S_{x_0}(g^{-1}x_0, R')}) \leqslant \epsilon$$

hence, since we proved $\partial X \setminus \overline{S_2} \subseteq \overline{X \setminus S_2} \subseteq \overline{S_3}$,

$$\nu(g^{-1}\overline{S_2}) \geqslant 1 - \epsilon. \tag{17}$$

Now, by the Markov property of the random walk, the conditional probability of ending up in $\overline{S_2}$ after hitting an element $gx_0 \in S_1$ at time k is given by

$$\mathbb{P}(\lim_{n\to\infty} w_n x_0 \in \overline{S_2} \mid w_k = g) = \mathbb{P}(\lim_{n\to\infty} w_n x_0 \in g^{-1}\overline{S_2}) = \nu(g^{-1}\overline{S_2})$$

for each k and g, and such probability is large by equation (17). This implies the following lower bound on the probability of ending up in $\overline{S_2}$: if we denote by $A_{k,g}$ the event "$w_k = g$ and $w_k x_0 \notin S_1$ for $0 \leqslant h \leqslant k-1$" we have

$$\begin{aligned}\mathbb{P}(\lim_{n\to\infty} w_n x_0 \in \overline{S_2}) \geqslant \mathbb{P}(\lim_{n\to\infty} w_n x_0 \in \overline{S_2} \text{ and } \exists\, n \;:\; w_n x_0 \in S_1) = \\ = \sum_{\substack{k\in\mathbb{N}\\ gx_0\in S_1}} \mathbb{P}(A_{k,g})\mathbb{P}(\lim_{n\to\infty} w_n x_0 \in \overline{S}_2 \mid w_k = g) \geqslant \\ \geqslant \mathbb{P}(\exists\, n \;:\; w_n x_0 \in S_1)(1-\epsilon)\end{aligned}$$

which, by recalling the definitions of ν and $H^+_{x_0}$, becomes

$$H^+_{x_0}(S_1) \leqslant \frac{1}{1-\epsilon}\nu(\overline{S_2}). \tag{18}$$

Now, as the distance parameter of S_1 tends to ∞, so does the distance parameter of S_2, hence $\nu(\overline{S_2}) \to 0$ by Proposition 5.1, and by the above equation $H^+_{x_0}(S_1)$ tends to 0, as required. □

As $H^+_{x_0}(S_1)$ is an upper bound for $\mu_n(S_1)$ for any n, equation (18) implies the following corollary.

Corollary 5.3. *Let G be a countable group which acts by isometries on a separable Gromov hyperbolic space X, and let μ be a non-elementary probability distribution on G. Then there is a function $f(r)$, with $f(r) \to 0$ as $r \to \infty$ such that for all n one has*

$$\sup_{S \in Sh(x_0, r)} \mu_n(S) \leqslant f(r). \tag{19}$$

As the reflected random walk also satisfies the hypotheses of Corollary 5.3, we obtain a similar result for $\check{\mu}_n$, though possibly for a different function f.

Proposition 5.4. *Let G be a non-elementary, countable group acting by isometries on a separable Gromov hyperbolic space X, and let μ be a non-elementary probability distribution on G. Then there is a number R_0 such that for any group elements g and h in the semigroup generated by the support of μ, the closure of the shadow $S_{hx_0}(gx_0, R_0)$ has positive hitting measure for the random walk determined by μ.*

Proof. Let us first assume $h = 1$. By Proposition 3.18, there is a constant R_0 such that every shadow $S_1 = S_{x_0}(gx_0, R_0)$ contains a limit point λ of Gx_0 in the interior of its closure. We may now follow the same argument as in Proposition 5.2. Choose a shadow $S_2 = S_{x_0}(\widetilde{g}x_0, R_0)$ containing λ such that $d_X(S_2, X \setminus S_1)$ is at least $A + O(\delta)$ (note that here the roles of S_1 and S_2 are reversed with respect to Proposition 5.2, as $S_2 \subseteq S_1$). In particular, this implies that for any point $y \in S_2$, the complement of S_1 is contained in a shadow $S_y(x_0, R)$ with distance parameter at least A. By Proposition 5.1 the measure of shadows tends to zero as the distance parameter tends to infinity, so given a positive number $\epsilon < 1$ we may choose A sufficiently large so that $\nu(\overline{S}) \leqslant \epsilon$ for all shadows $S \in Sh(x_0, A)$.

Let $(g_n)_{n \in \mathbb{N}}$ be a sequence in the semigroup generated by the support of μ such that $(g_n x_0)_{n \in \mathbb{N}}$ converges to λ. Let g_n be an element of the sequence with $g_n x_0 \in S_2$, and let k be such that $\mu_k(g_n) > 0$. Now using the Markov property of the random walk, the conditional probability of converging to the closure of S_1, having hit S_2, is at least $1 - \epsilon$, and so

$$\nu(\overline{S_1}) \geqslant (1 - \epsilon)\mu_k(g_n),$$

which is positive, as required. Now, the case $h \neq 1$ can be reduced to the previous one; indeed, given h in the support of the random walk generated by μ, there exists by hypothesis an n such that $\mu_n(h) > 0$, which implies

$$\nu(\overline{S_{hx_0}(gx_0, R_0)}) \geqslant \mathbb{P}(w_n = h)\ \mathbb{P}(\lim_m w_m x_0 \in \overline{S_{hx_0}(gx_0, R_0)} \ \mid \ w_n = h)$$

which by the Markov property of the walk equals

$$\mu_n(h)\ \mathbb{P}(\lim_m w_m x_0 \in \overline{S_{x_0}(h^{-1}gx_0, R_0)})$$

and this is positive by the previous case. $\square$

5.2. **Positive drift.** In this section we prove Theorem 1.2, i.e. that the sample paths $(w_n x_0)_{n\in\mathbb{N}}$ of the random walk have positive drift in X.

It will be convenient to consider the k-step random walk $(w_{kn})_{n\in\mathbb{N}}$, and introduce the notation $x_i := w_{ki}x_0$ for each i. Let $\chi_i^k \colon \Omega \to \mathbb{R}$ be the random variable given by the distance in X traveled by the sample path from time $k(i-1)$ to time ki, i.e.

$$\chi_i^k(\omega) := d_X(w_{k(i-1)}x_0, w_{ki}x_0) = d_X(x_{i-1}, x_i).$$

For fixed k, the χ_i^k are independent identically distributed random variables with common distribution χ_1^k.

Given a number R, we say a subsegment $[x_i, x_{i+1}]$ of the sample path is *persistent* if the following three conditions are satisfied:

$$d_X(x_i, x_{i+1}) \geqslant 2R + 2C + C_0 \tag{20}$$

$$x_n \in S_{x_{i+1}}(x_i, R) \text{ for all } n \leqslant i \tag{21}$$

$$x_n \in S_{x_i}(x_{i+1}, R) \text{ for all } n \geqslant i+1 \tag{22}$$

The constant C in (20) is the weak convexity constant from Corollary 3.21, while C_0 will only depend on δ and will be chosen later. A persistent subsegment is illustrated in Figure 5 below.

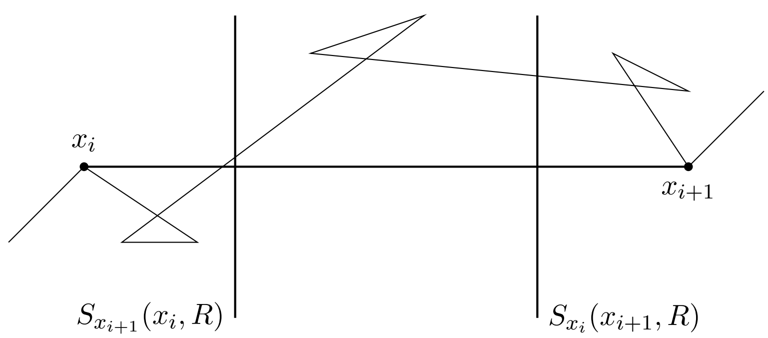

FIGURE 5. A persistent segment of the sample path.

Choose an ϵ, with $0 < \epsilon < \frac{1}{3}$. We now show that given such a choice of ϵ, we may choose both R and k sufficiently large such that for any i each of the three conditions holds with probability at least $1 - \epsilon$.

The probability that (21) holds is the same as the probability that $w_{kn}x_0$ never hits the complement of the shadow $S_{x_{i+1}}(x_i, R)$ for any $n \leqslant i$. As the complement of this shadow is contained in a shadow

$$S_i = S_{x_i}(x_{i+1}, R_i)$$

where $R_i = d_X(x_i, x_{i+1}) - R + O(\delta)$, the probability that (21) holds is at least

$$1 - \mathbb{P}(\exists\, n \leqslant ki \;:\; w_n x_0 \in S_i)$$

which equals by the Markov property

$$1 - H^-_{x_0}(w_{ki}^{-1} S_i). \tag{23}$$

The distance parameter of $w_{ki}^{-1} S_i$, which equals the distance parameter of S_i, is $R + O(\delta)$; hence, by Proposition 5.2, we may choose R sufficiently large such that (23) is at least $1 - \epsilon$.

A similar argument shows that the probability that (22) holds is at least

$$1 - H^+_{x_0}(w_{k(i+1)}^{-1} S_{x_{i+1}}(x_i, R_i)) \tag{24}$$

and again we may choose R sufficiently large such that (24) is at least $1 - \epsilon$.

Finally, the probability that (20) holds is $\mathbb{P}(\chi_i^k \geqslant 2R+2C+C_0) = \mathbb{P}(\chi_1^k \geqslant 2R + 2C + C_0)$, since χ_i^k and χ_1^k have the same law. We have shown that almost every sample path converges to a point in the Gromov boundary, so in particular, sample paths are transient on bounded sets. This implies that for any R and ϵ, there is a sufficiently large k, depending on R and ϵ, such that

$$\mathbb{P}(\chi_1^k \leqslant 2R + 2C + C_0) < \epsilon$$

as required. Therefore, for the choice of ϵ, R and k described above, the probability that each condition holds individually is at least $1 - \epsilon$. The three conditions need not be independent, but the probability that all three hold simultaneously is at least $\eta := 1 - 3\epsilon$, which is positive as $\epsilon < \frac{1}{3}$.

Thus, if we define for each i the random variable $Y_i^k : \Omega \to \mathbb{R}$

$$Y_i^k(\omega) := \begin{cases} 1 \text{ if } [x_i, x_{i+1}] \text{ is persistent} \\ 0 \text{ otherwise.} \end{cases}$$

we get that the Y_i^k are identically distributed (but not independent), with finite expectation since they are bounded; moreover, for each i

$$\mathbb{E}(Y_i^k) \geqslant \eta > 0. \tag{25}$$

We now show that the number of persistent segments lying between x_0 and $w_{kn} x_0$ gives a lower bound on the distance $d_X(x_0, w_{kn} x_0)$.

Let γ be a geodesic from x_0 to $x_n = w_{kn} x_0$, and suppose that $[x_i, x_{i+1}]$ is a persistent subsegment of the sample path. By (21), x_0 lies in $S_{x_{i+1}}(x_i, R)$ for $i \geqslant 0$, and by (22), x_n lies in $S_{x_i}(x_{i+1}, R)$ for $n \geqslant i+1$. As the two shadows are at least distance $2C+C_0-O(\delta)$ apart, the geodesic γ has a subsegment γ_i of length at least $C_0 - O(\delta) \geqslant C_0/2$ which fellow travels with $[x_i, x_{i+1}]$, and which is disjoint from both $S_{x_{i+1}}(x_i, R+C)$ and $S_{x_i}(x_{i+1}, R+C)$. Now let $[x_j, x_{j+1}]$ be a different persistent subsegment. The same argument as above shows that there is a subsegment γ_j of γ of length at least $C_0/2$ which fellow travels with $[x_j, x_{j+1}]$. We now show that γ_i and γ_j are disjoint subsegments of γ. Up to relabeling, we may assume that $i < j$. Then both x_j and x_{j+1} lie in $S_{x_i}(x_{i+1}, R)$, and so by weak convexity, Corollary 3.21, any geodesic connecting them lies in $S_{x_i}(x_{i+1}, R+C)$, and so in particular γ_i and γ_j are disjoint subsegments of γ. Therefore the distance $d_X(x_0, w_{kn} x_0)$ is at least $C_0/2$ times the number of persistent subsegments between x_0 and $w_{kn} x_0$.

We will now apply Kingman's subadditive ergodic theorem, [Kin68], using the following version from [Woe00, Theorem 8.10]:

Theorem 5.5. *Let $(\Omega, \mathbb{P})$ be a probability space and $U : \Omega \to \Omega$ a measure preserving transformation. If W_n is a subadditive sequence of non-negative real-valued random variables on Ω, that is, $W_{n+m} \leqslant W_n + W_m \circ U^n$ for all $m, n \in \mathbb{N}$, and W_1 has finite first moment, then there is a U-invariant random variable W_∞ such that*

$$\lim_{n\to\infty} \tfrac{1}{n} W_n = W_\infty$$

$\mathbb{P}$-almost surely, and in $L^1(\Omega, \mathbb{P})$.

In order to apply the theorem, let us define for each n the variable

$$Z_n^k := \sum_{i=0}^{n-1} Y_i^k = \#\{0 \leqslant i \leqslant n-1 \ : \ [x_i, x_{i+1}] \text{ is persistent}\}$$

which gives the number of persistent subsegments along a given sample path from x_0 to $x_n = w_{kn}x_0$. The random variables $(Z_n^k)_{n\in\mathbb{N}}$ are non-negative and have finite expectation, since $Z_n^k \leqslant n$ for each n, and the sequence is subadditive by the Markov property. Moreover, as expectation is additive, we get from equation (25)

$$\mathbb{E}(Z_n^k) = \sum_{i=0}^{n-1} \mathbb{E}(Y_i^k) \geqslant n\eta$$

with $\eta > 0$. We now apply Theorem 5.5 taking as Ω the step space of the k^{th}-step random walk, U the shift map, and the Z_n^k as random variables (for fixed k); we get that the sequence $(\frac{1}{n} Z_n^k)_{n\in\mathbb{N}}$ converges almost surely and in L^1 to some random variable Z_∞^k; moreover, since U is ergodic, Z_∞^k must be constant almost everywhere, thus there exists a constant $A \geqslant 0$ such that

$$\frac{1}{n} Z_n^k \to A$$

in L^1; finally, since $\mathbb{E}(Z_\infty^k) = \lim_n \mathbb{E}(\frac{1}{n} Z_n^k) \geqslant \eta > 0$, we have that $A > 0$. Thus, since Z_n^k is a lower bound for the distance $d_X(x_0, w_{kn}x_0)$, we get almost surely for the k^{th}-step random walk

$$\liminf_{n\to\infty} \frac{d_X(x_0, w_{kn}x_0)}{kn} \geqslant \frac{C_0}{2k} \liminf_{n\to\infty} \frac{1}{n} Z_n^k = \frac{AC_0}{2k} > 0$$

which proves the first part of Theorem 1.2, where we make no assumptions on the moments of μ.

For the second part of Theorem 1.2, we assume that μ has finite first moment with respect to the distance function d_X. In this case, we can apply Kingman's Theorem directly to $d_X(x_0, w_{kn}x_0)$, and we know that the limiting value L is positive, by the previous case.

Finally, if the support of μ is bounded in X, then the arguments from [Mah12] apply directly.

5.3. **Geodesic tracking.** We will now prove Theorem 1.3, using the following sublinearity result from Tiozzo [Tio12].

Lemma 5.6. *Let* $f\colon \Omega \to \mathbb{R}$ *be a non-negative measurable function,* $T : \Omega \to \Omega$ *an ergodic, measure preserving transformation, and suppose that*

$$g(\omega) = f(T\omega) - f(\omega) \text{ lies in } L^1(\Omega, \mathbb{P}). \tag{26}$$

Then

$$\lim_{n\to\infty} \frac{f(T^n\omega)}{n} = 0$$

for almost all $\omega \in \Omega$.

In order to apply it in this case, let us note that there are constants Q and c, depending only on δ, such that any two distinct points in ∂X are connected by a (Q,c)-quasigeodesic. We shall write $\Gamma(x,y)$ for the set of (Q,c)-quasigeodesics connecting x and y. We then define $f : \Omega \to \mathbb{R}$ as

$$f(\omega) := \sup\{d_X(x_0, \gamma) : \gamma \in \Gamma(\omega_-, \omega_+)\}.$$

As ν and $\check{\nu}$ are non-atomic, $\nu \times \check{\nu}$ gives measure zero to the diagonal in $\partial X \times \partial X$, so $\Gamma(\omega_-, \omega_+)$ is non-empty $(\nu \times \check{\nu})$-almost surely, and the function $f(\omega)$ is well-defined $\mathbb{P}$-almost surely.

Then by the triangle inequality, $|f(T\omega) - f(\omega)| \leqslant d_X(x_0, w_1x_0)$, and so $g(\omega)$ lies in $L^1(\Omega, \mathbb{P})$. Thus it follows from Lemma 5.6 (where T is the shift map on the step space) that sample paths track quasigeodesics sublinearly, i.e.

$$\frac{d_X(w_nx_0, \gamma(\omega))}{n} \to 0, \text{ as } n \to \infty, \text{ almost surely,}$$

proving the first part of Theorem 1.3.

To prove the second part, we need to show that, if μ has bounded support in X, then the tracking is in fact logarithmic. We apply the argument from Blachère, Haïssinsky and Mathieu [BHM11, Section 3], combined with our exponential decay of shadows. We now give the details for the convenience of the reader.

We first show that the distribution of distances from the locations of the sample path to the quasigeodesic satisfies an exponential decay property.

Proposition 5.7. *Let* G *be a countable group of isometries of a separable Gromov hyperbolic space* X*, and let* μ *be a non-elementary probability distribution on* G*. There are positive constants* K *and* $c < 1$*, which depend on* μ*, such that*

$$\mathbb{P}(d_X(w_nx_0, \gamma(\omega)) \geqslant D) \leqslant Kc^D,$$

where γ *is a quasigeodesic ray from* x_0 *to the limit point in* ∂X *of* $(w_nx_0)_{n\in\mathbb{N}}$.

Proof. By the definition of $\gamma(\omega)$,

$$d_X(w_nx_0, \gamma(\omega)) = d_X(w_nx_0, [x_0, \lambda(\omega)]),$$

where $\lambda(\omega)$ is the limit point of $(w_m x_0)_{n\in\mathbb{N}}$ in ∂X, which exists for $\mathbb{P}$-almost every ω. Applying the isometry w_n^{-1} gives

$$d_X(w_n x_0, \gamma(\omega)) = d_X(x_0, [w_n^{-1}x_0, w_n^{-1}\lambda(\omega)]).$$

Recall from (1), that the Gromov product $(x \cdot y)_{x_0}$ may be estimated up to an error of $O(\delta)$ in terms of the distance $d_X(x_0, [x, y])$, and a similar estimate holds if one of x or y is a point in ∂X, and $[x, y]$ is a quasigeodesic connecting them. This implies that

$$d_X(w_n x_0, \gamma(\omega)) = (w_n^{-1}x_0 \cdot w_n^{-1}\lambda(\omega))_{x_0} + O(\delta)$$

so, the condition

$$d_X(w_n x_0, \gamma(\omega)) \geqslant D$$

implies

$$(w_n^{-1}x_0 \cdot w_n^{-1}\lambda(\omega))_{x_0} \geqslant D + O(\delta)$$

which, by definition of shadow, implies the condition

$$w_n^{-1}\lambda(\omega) \in \overline{S_{x_0}(w_n^{-1}x_0, R)},$$

where the parameter R is given by

$$R = d_X(x_0, w_n^{-1}x_0) - D + O(\delta).$$

The boundary point $w_n^{-1}\lambda(\omega)$ only depends on the increments of the random walk of index greater than n, so $w_n^{-1}\lambda(\omega)$ and w_n are independent. Furthermore, the distribution of $w_n^{-1}\lambda(\omega)$ is equal to ν. Therefore

$$\mathbb{P}(d_X(w_n x_0, \gamma(\omega)) \geqslant D) \leqslant \nu\left(\overline{S_{x_0}(w_n^{-1}x_0, R)}\right),$$

and as μ has bounded range in X, we may use the exponential decay estimate for shadows (16), which gives

$$\mathbb{P}(d_X(w_n x_0, \gamma(\omega)) \geqslant D) \leqslant Kc^D,$$

as required. □

It follows immediately from the proposition above that there is a constant $\kappa > 0$ such that

$$\mathbb{P}(d_X(w_n x_0, \gamma(\omega)) \geqslant \kappa \log n) \leqslant \frac{1}{n^2}.$$

The logarithmic tracking result,

$$\limsup \frac{d_X(w_n x_0, \gamma(\omega))}{\log n} < \infty, \text{ almost surely,}$$

then follows from the Borel-Cantelli lemma. This completes the proof of Theorem 1.3.

5.4. **Translation length.** We briefly review some results about the translation length of isometries, see e.g. Bridson and Haefliger [BH99] or Fujiwara [Fuj08].

We start by observing that the translation length of an isometry g may be estimated in terms of the distance it moves the basepoint x_0, together with the Gromov product of gx_0 and $g^{-1}x_0$.

Proposition 5.8. *There exists a constant $C_0 > 0$, which depends only on δ, such that the following holds. For any isometry g of a δ-hyperbolic space X, if g satisfies the inequality*

$$d_X(x_0, gx_0) \geqslant 2(gx_0 \cdot g^{-1}x_0)_{x_0} + C_0, \tag{27}$$

then the translation length of g is

$$\tau(g) = d_X(x_0, gx_0) - 2(g^{-1}x_0 \cdot gx_0)_{x_0} + O(\delta). \tag{28}$$

This is well known, but we provide a proof in the appendix for the convenience of the reader.

In order to complete the proof of Theorem 1.4, we shall now estimate the probability that the translation length is small for a sample path of length n. To apply the estimate for translation length (28) we need a lower bound for $d_X(x_0, w_nx_0)$, which is given by positive drift, and an upper bound for the Gromov product $(w_n^{-1}x_0 \cdot w_nx_0)_{x_0}$, which we now obtain.

Let $m = \lceil n/2 \rceil$; we shall introduce the notation $u_m := w_m^{-1}w_n = g_{m+1}g_{m+2}\cdots g_n$, and we may think of w_mx_0 as an approximate midpoint of the sample path from x_0 to w_nx_0, and of $u_m^{-1}x_0 = w_n^{-1}w_mx_0$ as an approximate midpoint of the inverse sample path from x_0 to $w_n^{-1}x_0$. Note that for each m, the G-valued processes $w_m = g_1g_2\ldots g_m$ and $u_m := g_{m+1}g_{m+2}\cdots g_n$ are independent.

Because of this independence, and the fact that the hitting measures are non-atomic, it is easy to prove the following upper bound on the Gromov product $(u_m^{-1}x_0 \cdot w_mx_0)_{x_0}$.

Lemma 5.9. *Let G be a countable group of isometries of a separable Gromov hyperbolic space X, and let μ be a non-elementary probability distribution on G. If $l : \mathbb{N} \to \mathbb{N}$ is any function such that $l(n) \to \infty$ as $n \to \infty$, then we have*

$$\mathbb{P}\left((u_m^{-1}x_0 \cdot w_mx_0)_{x_0} \leqslant l(n)\right) \to 1, \text{ as } n \to \infty,$$

for all n, where $m = \lceil n/2 \rceil$.

Proof. By the definition of shadows,

$$\mathbb{P}\left((u_m^{-1}x_0 \cdot w_mx_0)_{x_0} \leqslant l(n)\right) = \mathbb{P}(u_m^{-1}x_0 \notin S_{x_0}(w_mx_0, R)),$$

where $R = d_X(x_0, w_mx_0) - l(n)$. As w_m and u_m^{-1} are independent and the distribution of u_m^{-1} is $\check{\mu}_{n-m}$,

$$\mathbb{P}\left((u_m^{-1}x_0 \cdot w_mx_0)_{x_0} \leqslant l(n)\right) = 1 - \sum_{g \in G} \check{\mu}_{n-m}(S_{x_0}(gx_0, R))\mu_m(g)$$

Now, since the distance parameter of the shadows on the right-hand side is $l(n)$, using the estimate (19) gives

$$\mathbb{P}\left((u_m^{-1}x_0 \cdot w_m x_0)_{x_0} \leqslant l(n)\right) \geqslant 1 - f(l(n))$$

which tends to 1 as $n \to \infty$. □

We will now use the fact that if the Gromov products $(w_m x_0 \cdot w_n x_0)_{x_0}$ and $(u_m^{-1}x_0 \cdot w_n^{-1}x_0)_{x_0}$ are large, and the Gromov product $(u_m^{-1}x_0 \cdot w_m x_0)_{x_0}$ is small, then the two Gromov products $(w_n^{-1}x_0 \cdot w_n x_0)_{x_0}$ and $(w_n^{-1}x_0 \cdot w_m x_0)_{x_0}$ are equal, up to bounded additive error depending only on δ. This follows from the following lemma, which is a standard exercise in coarse geometry. We omit the proof, but the appropriate approximate tree is illustrated in Figure 6, with the points labeled according to our application.

Lemma 5.10. *For any four points a, b, c and d in a Gromov hyperbolic space X, if there is a number A such that $(a \cdot b)_{x_0} \geqslant A$, $(c \cdot d)_{x_0} \geqslant A$ and $(a \cdot c)_{x_0} \leqslant A - O(\delta)$ then $(a \cdot c)_{x_0} = (b \cdot d)_{x_0} + O(\delta)$.*

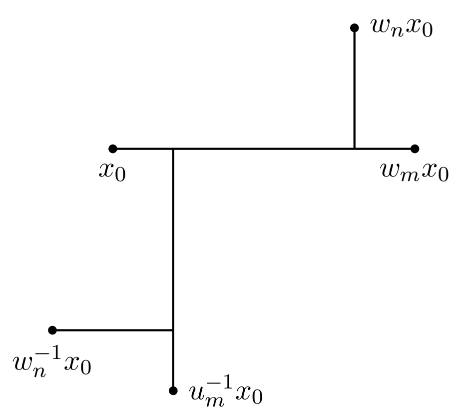

FIGURE 6. Estimating the Gromov product.

We now observe that with high probability, the Gromov product $(w_m x_0 \cdot w_n x_0)_{x_0}$ is large. Recall that by linear progress there exists $L > 0$ such that

$$\mathbb{P}\left(d_X(x_0, w_n x_0) \geqslant Ln\right) \to 1. \tag{29}$$

Lemma 5.11. *Let G be a countable group of isometries of a separable Gromov hyperbolic space X, and let μ be a non-elementary probability distribution on G, and L as in eq.* (29)*. Then for any $l < L/2$ we have*

$$\mathbb{P}\left((w_m x_0 \cdot w_n x_0)_{x_0} \geqslant ln\right) \to 1, \text{ as } n \to \infty,$$

for all n, where $m = \lceil n/2 \rceil$.

Proof. Note that by definition of shadows, we have the equality

$$\mathbb{P}\left((w_m x_0 \cdot w_n x_0)_{x_0} \geqslant ln\right) = \mathbb{P}\left(w_n x_0 \in S_{x_0}(w_m x_0, R)\right),$$

where $R = d_X(x_0, w_m x_0) - ln$. Applying the isometry w_m^{-1} yields (recall $u_m = w_m^{-1} w_n$)

$$\mathbb{P}\left((w_m x_0 \cdot w_n x_0)_{x_0} \geqslant ln\right) = \mathbb{P}\left(u_m x_0 \in S_{w_m^{-1} x_0}(x_0, R)\right).$$

As the complement of a shadow is approximately a shadow (Corollary 2.5),

$$\mathbb{P}\left((w_m x_0 \cdot w_n x_0)_{x_0} \geqslant ln\right) \geqslant \mathbb{P}\left(u_m x_0 \notin S_{x_0}(w_m^{-1} x_0, \widetilde{R})\right)$$

with $\widetilde{R} = ln + O(\delta)$, and by conditioning with $w_m = g$ we get

$$\mathbb{P}\left((w_m x_0 \cdot w_n x_0)_{x_0} \geqslant ln\right) \geqslant \sum_{g \in G} \mathbb{P}\left(u_m x_0 \notin S_{x_0}(w_m^{-1} x_0, \widetilde{R}) \;\middle|\; w_m = g\right) \mathbb{P}\left(w_m = g\right).$$

As u_m and w_m are independent, and the distribution of u_m is μ_{n-m} we have

$$\mathbb{P}\left((w_m x_0 \cdot w_n x_0)_{x_0} \geqslant ln\right) \geqslant \sum_{g \in G} \left(1 - \mu_{n-m}\left(S_{x_0}(g^{-1} x_0, \widetilde{R})\right)\right) \mu_m(g).$$

Now, if we restrict to the set of g such that $d_X(x_0, g x_0) \geqslant Ln/2$, then the distance parameter of the shadow is $d_X(x_0, g x_0) - ln + O(\delta) \geqslant \epsilon n + O(\delta)$ for $\epsilon = \frac{L}{2} - l > 0$, hence by the estimate for μ_{n-m} in terms of the distance parameter, we get

$$\mathbb{P}\left((w_m x_0 \cdot w_n x_0)_{x_0} \geqslant ln\right) \geqslant \left(1 - f\left(\epsilon n + O(\delta)\right)\right) \mathbb{P}\left(d(x_0, w_m x_0) \geqslant Ln/2\right).$$

The result now follows by positive drift. □

The same argument applied to $w_n^{-1} x_0$, which has approximate midpoint $u_m^{-1} x_0$, shows that

$$\mathbb{P}\left((u_m^{-1} x_0 \cdot w_n^{-1} x_0)_{x_0} \geqslant ln\right) \to 1 \text{ as } n \to \infty$$

for any $l < L/2$. Now using Lemma 5.10, together with the lower bounds on the Gromov products of $(w_m x_0 \cdot w_n x_0)_{x_0}$ and $(u_m^{-1} x_0 \cdot w_n^{-1} x_0)_{x_0}$ from Lemma 5.11, and the upper bound on the Gromov product $(u_m^{-1} x_0 \cdot w_m x_0)_{x_0}$ from Lemma 5.9 implies

$$\mathbb{P}\left((w_n^{-1} x_0 \cdot w_n x_0)_{x_0} \leqslant l(n)\right) \to 1 \text{ as } n \to \infty.$$

for any function $l : \mathbb{N} \to \mathbb{N}$ such that $l(n) \to \infty$ as $n \to \infty$. Applying this to the estimate for translation length (28), shows that the probability that

$$\tau(w_n) \geqslant \tfrac{1}{2} Ln + O(\delta)$$

tends to 1 as $n \to \infty$, as required.

If μ has bounded support in X, then this happens exponentially fast, by [Mah12].

6. The Poisson boundary for acylindrically hyperbolic groups

In this section we prove Theorem 1.5, i.e. we show that if the action of G on X is acylindrical and μ has finite entropy and finite logarithmic moment, then in fact the Gromov boundary with the hitting measure is the Poisson boundary.

We shall assume from now on that G is a non-elementary, countable group of isometries of a separable Gromov hyperbolic space X, and μ a non-elementary probability measure on G. Recall that the entropy of μ is $H(\mu) := -\sum_{g \in G} \mu(g) \log \mu(g)$, and μ is said to have *finite entropy* if $H(\mu) < \infty$. The measure μ is said to have *finite logarithmic moment* if

$$\sum_{g \in G} \mu(g) |\log d_X(x_0, gx_0)| < \infty.$$

Let us recall the definition of acylindrical action, which is due to Sela [Sel97] for trees, and Bowditch [Bow08] for general metric spaces.

Definition 6.1. *We say a group G acts* acylindrically *on a Gromov hyperbolic space X, if for every $K \geqslant 0$ there are numbers $R = R(K)$ and $N = N(K)$ such that for any pair of points x and y in X, with $d_X(x, y) \geqslant R$, there are at most N group elements g in G such that $d_X(x, gx) \leqslant K$ and $d_X(y, gy) \leqslant K$.*

For a discussion and several examples of acylindrical actions on hyperbolic spaces, see [Osi13].

The proof will use Kaimanovich's strip criterion from [Kai00]. Briefly, the criterion uses the existence of "strips", that is subsets of G which are associated to each pair of boundary points in a G-equivariant way.

In order to apply the criterion, however, one also needs to control the number of elements in the strips; in fact, we will show that for each strip the number of elements whose images in X lie in a ball of radius r can grow at most linearly in r. In a proper space, one may often choose the strips to consist of all geodesics connecting the endpoints of the sample path, but in our case, this usually gives infinitely many points in a ball of finite radius. Instead, we observe that by recurrence, the sample path returns close to a geodesic connecting its endpoints for a positive density of times $n \in \mathbb{Z}$. Using this it can be shown that there are infinitely many pairs of locations $w_n x_0$ and $w_{n+m} x_0$, where the sample path has gone a definite distance along the geodesic in bounded time. In fact, we may choose a suitable group element v, and look at all group elements g whose orbit points gx_0 are close to a geodesic γ, such that both gx_0 and gvx_0 are close to γ. We shall call the collection of such group elements *bounded geometry* elements, and we will choose our strips to consist of these elements. We will use acylindricality to show that this set is locally finite, and in fact the intersection of its image in X with $B_X(x_0, r)$ grows at most linearly with r. Let us now make this precise.

6.1. **Bounded geometry points.** Let $v \in G$ be a group element, K, R two constants, and $\alpha, \beta \in \partial X$ two boundary points.

We say that a group element g has (K, R, v)-*bounded geometry* with respect to the pair of boundary points α, $\beta \in \partial X$ if the three following conditions hold:

(1) $d_X(gx_0, gvx_0) \geqslant R$;
(2) α belongs to the interior of the closure (in $X \cup \partial X$) of $S_{gvx_0}(gx_0, K)$;
(3) β belongs to the interior of the closure of $S_{gx_0}(gvx_0, K)$.

This is illustrated in Figure 7 below. We shall write $\mathcal{O}(\alpha, \beta)$ for the set of bounded geometry elements determined by α and β (or $\mathcal{O}_{K,R,v}(\alpha, \beta)$ if we want to explicitly keep track of the constants). This definition is G-equivariant, i.e. $g\mathcal{O}(\alpha, \beta) = \mathcal{O}(g\alpha, g\beta)$ for any $g \in G$. We will refer to the image of a bounded geometry element in X under the orbit map as a *bounded geometry point*.

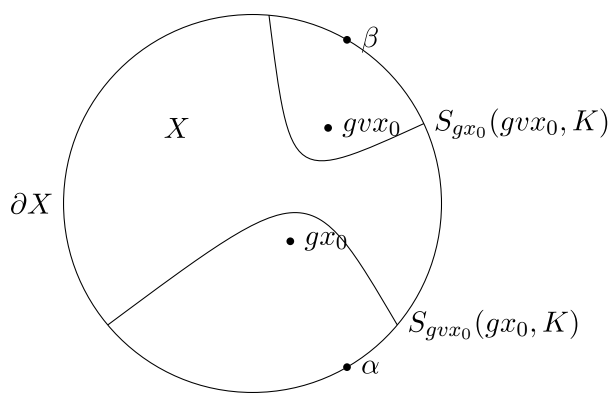

FIGURE 7. A bounded geometry point gx_0 in $\mathcal{O}_{K,R,v}(\alpha, \beta)$.

We say a set of group elements $\mathcal{O}$ is *locally finite* if the set $\mathcal{O}x_0 \cap B_X(x, r)$ is finite for all $x \in X$ and all $r \geqslant 0$, and that $\mathcal{O}$ has *linear growth* if there is a constant C such that for all $r \geqslant 0$

$$|\mathcal{O}x_0 \cap B_X(x_0, r)| \leqslant Cr.$$

We now show that the set of bounded geometry elements has linear growth.

Proposition 6.2. *There exists K_0 such that for any $K \geqslant K_0$, there exists R_0, such that for any $R \geqslant R_0$, there exists a constant C such that we have the estimate*

$$|B_X(x_0, r) \cap \mathcal{O}_{K,R,v}(\alpha, \beta)x_0| \leqslant Cr.$$

for any $\alpha, \beta \in \partial X$, any $r > 0$ and any group element $v \in G$.

In order to prove the proposition, let us start by proving that the number of bounded geometry elements in a ball $B_X(x, 4K)$ is bounded in terms of K.

Lemma 6.3. *There exists K_0 such that, for any $K \geqslant K_0$, there exists R_0 such that for any $R \geqslant R_0$ and any group element $v \in G$, we have the estimate*

$$|B_X(x, 4K) \cap \mathcal{O}_{K,R,v}(\alpha, \beta)x_0| \leqslant N(22K)$$

for any pair of boundary points α and β, and any point $x \in X$.

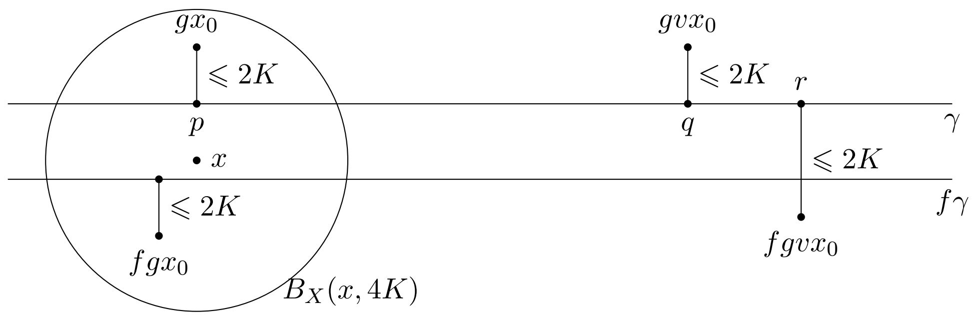

FIGURE 8. Bounded geometry points in $B_X(x, 4K)$.

Proof. Recall from Section 2.3 that we have chosen Q and c to be two numbers such that every pair of points in the Gromov boundary ∂X are connected by a continuous (Q, c)-quasigeodesic. We shall choose L to be a Morse constant for the (Q, c)-quasigeodesics, i.e. for any pair of points x and y in a (Q, c)-quasigeodesic γ, the segment of γ between x and y is contained in an L-neighbourhood of any geodesic connecting x and y.

Let g be a bounded geometry element, and write S_1 for $S_{gvx_0}(gx_0, K)$ and S_2 for $S_{gx_0}(gvx_0, K)$. As g has bounded geometry, each (Q, c)-quasigeodesic from α to β passes within distance $L+K+O(\delta)$ of both gx_0 and gvx_0. Therefore we may choose K_0 to be sufficiently large such that for any $K \geqslant K_0$ the distance from any bounded geometry point gx_0 to a (Q, c)-quasigeodesic connecting α and β is at most $2K$. Furthermore, we will choose K_0 to be larger than the quasigeodesic constants Q and c, and also larger than the Morse constant L.

Given K, we will choose R_0 sufficiently large such that for any (Q, c)-quasigeodesic γ with one endpoint in the interior of $\overline{S_1}$, and the other endpoint in the interior of $\overline{S_2}$, the nearest point projections of S_1 and S_2 onto γ are disjoint. Furthermore, we shall choose R_0 to be at least the acylindricality constant $R(22K)$.

Let gx_0 and $g'x_0 \in B_X(x, 4K) \cap \mathcal{O}_{K,R,v}(\alpha, \beta)x_0$ be two bounded geometry elements with respect to the same boundary points (α, β), and the same element v. We may write $g' = fg$, for some group element $f \in G$. This is illustrated in Figure 8 above. The isometry f moves the point gx_0 distance at most $8K$.

Now let γ be a (Q, c)-quasigeodesic joining α to β. By construction, both γ and $f\gamma$ have endpoints in fS_1 and fS_2, hence they both pass within distance $2K$ from both fgx_0 and $fgvx_0$. Let us now consider $y := gvx_0$; we now show that the isometry f also moves the point y a bounded distance, which yields the claim by definition of an acylindrical action.

Let p be a closest point on γ to gx_0, let q be a closest point on γ to gvx_0, and let r be a closest point on γ to $fgvx_0$. As the shadows $\overline{S_1}$ and $\overline{S_2}$ have disjoint nearest point projections to γ, the order in which the three points appear on γ is either (p, q, r) or (p, r, q). As the distances from p to r and p to q are both approximately $d_X(x_0, vx_0)$, this implies that $d_X(q, r)$ is small; we now make this precise.

Claim 6.4. *$d_X(q, r) \leqslant 18K$.*

Proof (of claim). The point p is within distance $2K$ of gx_0, and the point q is within distance $2K$ of gvx_0, so

$$|d_X(p, q) - d_X(gx_0, gvx_0)| \leqslant 4K.$$

Similarly, the point p is within distance $10K$ of fgx_0, and the point r is within distance $2K$ of $fgvx_0$, so

$$|d_X(p, r) - d_X(fgx_0, fgvx_0)| \leqslant 12K.$$

Combining these two estimates and using $d_X(gx_0, gvx_0) = d_X(fgx_0, fgvx_0)$ gives

$$|d_X(p, q) - d_X(p, r)| \leqslant 16K. \tag{30}$$

First consider the case in which the order is (p, q, r). As γ is a (Q, c)-quasigeodesic, the point q lies within distance L of a geodesic from p to r: if we denote q' the projection of q onto this geodesic segment, then we have

$$d_X(q', r) = d_X(p, r) - d_X(p, q')$$

which, using the fact that q' is within distance L from q, gives

$$d_X(q, r) \leqslant d_X(p, r) - d_X(p, q) + 2L$$

and now combining this with the estimate (30) one gets

$$d_X(q, r) \leqslant 16K + 2L \leqslant 18K.$$

If the points lie on γ in the order (p, r, q), the proof is analogous. □

As q is close to gvx_0 and r is close to $fgvx_0$, we obtain the following estimate

$$d_X(gvx_0, fgvx_0) \leqslant 22K.$$

Therefore, f moves each of gx_0 and gvx_0 distance at most $22K$, and so by acylindricality there are at most $N(22K)$ possible choices for f, as we have chosen $R_0 \geqslant R(22K)$, as required. □

Proof of Proposition 6.2. Let γ be a (Q, c)-quasigeodesic connecting α and β. We shall choose the number K_0 to be the same as the number K_0 from Lemma 6.3. Then K is sufficiently large such that every element of $\mathcal{O}(\alpha, \beta)$ has an image in X which lies within distance at most $2K$ of γ, and any pair of points $\gamma(n)$ and $\gamma(n+1)$ on the quasigeodesic are distance at most $2K$ apart. Therefore $\mathcal{O}(\alpha, \beta)x_0$ is covered by balls of the form $B_X(\gamma(n), 4K)$ for $n \in \mathbb{N}$, and the claim follows from applying Lemma 6.3 to each of these balls. □

6.2. **Recurrence and the strip criterion.** Given a bi-infinite sample path $(w_n)_{n\in\mathbb{Z}}$, we shall define the forward and backward limit points to be

$$\lambda_+(\omega) := \lim_{n\to\infty} w_n x_0, \text{ and } \lambda_-(\omega) := \lim_{n\to\infty} w_{-n} x_0.$$

As the forward and backward random walks converge to the Gromov boundary, these limit points are defined for $\mathbb{P}$-almost all ω, and the joint distribution of the pair $(\lambda_+(\omega), \lambda_-(\omega))$ is $\nu \times \check{\nu}$.

For any bi-infinite sequence $\omega \in \Omega = G^{\mathbb{Z}}$, we define $\mathcal{O}(\omega)$ to be the set $\mathcal{O}(\lambda_+(\omega), \lambda_-(\omega))$ of bounded geometry elements determined by the limit point $\lambda_+(\omega)$ of the forward random walk and the limit point $\lambda_-(\omega)$ of the backward random walk.

Finally, we show that we can choose K, R, and v such that the set of bounded geometry elements is non-empty and locally finite for $\nu \times \check{\nu}$-almost all $(\alpha, \beta) \in \partial X \times \partial X$.

Proposition 6.5. *There are constants K, R and a group element $v \in G$ such that the set*

$$\mathcal{O}_{K,R,v}(\alpha, \beta)$$

of bounded geometry elements has linear growth and is non-empty (in fact, infinite) for $\nu \times \check{\nu}$-almost all pairs (α, β).

Proof. By Proposition 5.4, we can choose K large enough so that for any group element v, the closure of the shadow $S = S_{x_0}(vx_0, K)$ has positive ν-measure, and the closure of the shadow $S' = S_{vx_0}(x_0, K)$ has positive $\check{\nu}$-measure. Thus, the probability that the group identity element 1 lies in $\mathcal{O}(\omega)$ is positive, because

$$\mathbb{P}\left(\omega \ : \ 1 \in \mathcal{O}(\omega)\right) = \nu(\overline{S})\check{\nu}(\overline{S'}) = p > 0.$$

Consider the probability that the location of the random walk w_n lies in $\mathcal{O}(\omega)$, i.e.

$$\mathbb{P}\left(\omega \ : \ w_n \in \mathcal{O}(\lambda_+(\omega), \lambda_-(\omega))\right).$$

By G-equivariance, this is equal to

$$\mathbb{P}\left(\omega \ : \ 1 \in \mathcal{O}(w_n^{-1}\lambda_+(\omega), w_n^{-1}\lambda_-(\omega))\right),$$

and by definition of the shift map this is equal to

$$\mathbb{P}\left(\omega \ : \ 1 \in \mathcal{O}(T^n\omega)\right).$$

As the shift map preserves the measure $\mathbb{P}$, this is equal to

$$\mathbb{P}\left(\omega \ : \ 1 \in \mathcal{O}(\omega)\right) = p > 0.$$

Therefore the events $\{\omega \in \Omega \ : \ w_n \in \mathcal{O}(\omega)\}$ occur with the same positive probability p, though they are not independent. By ergodicity of the shift map, the proportion of locations $\{w_1, \ldots w_N\}$ satisfying $w_n \in \mathcal{O}(\omega)$ converges to p as N tends to infinity. As $(w_n x_0)_{n \in \mathbb{N}}$ converges to the boundary $\mathbb{P}$-almost surely, $\mathcal{O}(\omega)$ contains infinitely many elements $\mathbb{P}$-almost surely. □

We remind the reader of Kaimanovich's strip criterion from [Kai00, Theorem 6.4]. We shall write $B_G(1, r)$ for all group elements whose image in X under the orbit map is distance at most r from the basepoint x_0, i.e.

$$B_G(1, r) = \{g \in G \mid d_X(x_0, gx_0) \leqslant r\}.$$

Theorem 6.6. *Let μ be a probability measure with finite entropy on G, and let $(\partial X, \nu)$ and $(\partial X, \check{\nu})$ be μ- and $\check{\mu}$-boundaries, respectively. If there exists a measurable G-equivariant map S assigning to almost every pair of points $(\alpha, \beta) \in \partial X \times \partial X$ a non-empty "strip" $S(\alpha, \beta) \subset G$, such that for all g*

$$\frac{1}{n} \log |S(\alpha, \beta) g \cap B_G(1, d_X(x_0, w_n x_0))| \to 0 \qquad \text{as } n \to \infty,$$

for $(\nu \times \check{\nu})$-almost every $(\alpha, \beta) \in \partial X \times \partial X$, then $(\partial X, \nu)$ and $(\partial X, \check{\nu})$ are the Poisson boundaries of the random walks (G, μ) and $(G, \check{\mu})$, respectively.

In order to prove Theorem 1.5, we define the strip $S(\alpha, \beta)$ as the set $\mathcal{O}_{K,R,v}(\alpha, \beta)$ of bounded geometry elements. By right multiplication by g^{-1}, the set

$$S(\alpha, \beta) g \cap B_G(1, d_X(x_0, w_n x_0))$$

has the same cardinality as

$$S(\alpha, \beta) \cap B_G(1, d_X(x_0, w_n x_0)) g^{-1}.$$

Furthermore,

$$B_G(1, d_X(x_0, w_n x_0)) g^{-1} \subset B_G(1, d_X(x_0, w_n x_0) + d_X(x_0, g x_0)),$$

and so

$$|S(\alpha, \beta) g \cap B_G(1, d_X(x_0, w_n x_0))| \leqslant |S(\alpha, \beta) \cap B_G(1, d_X(x_0, w_n x_0) + d_X(x_0, g x_0))| .$$

Proposition 6.5 shows that there are suitable choices of K, R and v such that the sets of bounded geometry elements are non-empty almost surely and have linear growth, so there is a number K such that

$$|S(\alpha, \beta) g \cap B_G(1, d_X(x_0, w_n x_0))| \leqslant K(d_X(x_0, w_n x_0) + d_X(x_0, g x_0)).$$

Therefore, it suffices to show that almost surely $\log d_X(x_0, w_n x_0)/n \to 0$ as $n \to \infty$, and this follows from the fact that μ has finite logarithmic moment, as we now briefly explain. Finite logarithmic moment implies that $\log d_X(x_0, g_n x_0)/n \to 0$ almost surely, and so for any $\epsilon > 0$ there exists a $C > 1$ such that we have $d_X(x_0, g_n x_0) \leqslant Ce^{\epsilon n}$ for all n. By the triangle inequality

$$d_X(x_0, w_n x_0) \leqslant d_X(x_0, g_1 x_0) + \cdots + d_X(x_0, g_n x_0),$$

and so

$$\log d_X(x_0, w_n x_0) \leqslant \log C + \log n + \epsilon n.$$

As this holds for all $\epsilon > 0$, this implies that $\log d_X(x_0, w_n x_0)/n \to 0$ as $n \to \infty$, as required.

Finally, the statement that the map S is measurable means that for any $g \in G$, the set

$$\{(\alpha, \beta) \in \partial X \times \partial X \mid g \in S(\alpha, \beta)\}$$

is a Borel set; this holds, since by definition, $g \in S(\alpha, \beta)$ if and only if (α, β) belongs to the product of the closures of two shadows, which is closed, hence Borel. This completes the proof of Theorem 1.5.

Appendix A. Estimating translation length

In this section we provide a proof of Proposition 5.8, which estimates the translation length in terms of the distance an isometry moves the basepoint, and the Gromov product.

Proposition 5.8. *There exists a constant $C_0 > 0$, which depends only on δ, such that the following holds. For any isometry g of a δ-hyperbolic space X, if g satisfies the inequality*

$$d_X(x_0, gx_0) \geqslant 2(gx_0 \cdot g^{-1}x_0)_{x_0} + C_0, \tag{31}$$

then the translation length of g is

$$\tau(g) = d_X(x_0, gx_0) - 2(g^{-1}x_0 \cdot gx_0)_{x_0} + O(\delta). \tag{32}$$

Proof. We start by showing that if γ is a geodesic segment from x_0 to $g^n x_0$, then $g^k x_0$ is contained in a bounded neighbourhood of γ, for all $0 \leqslant k \leqslant n$. We shall write x_k for $g^k x_0$. Note that, since the action is isometric, we have for each k

$$d_X(x_{k+1}, x_{k+2}) = d_X(x_0, gx_0), \text{ and } (x_k \cdot x_{k+2})_{x_{k+1}} = (gx_0 \cdot g^{-1}x_0)_{x_0}. \tag{33}$$

Claim A.1. *Let γ be a geodesic from x_0 to x_n. Then*

$$d_X(x_k, \gamma) \leqslant (gx_0 \cdot g^{-1}x_0)_{x_0} + O(\delta), \tag{34}$$

for all $0 \leqslant k \leqslant n$.

Proof (of claim). Let p_k be a nearest point on γ to x_k, and let x_k be an element of $\{x_k\}_{k=0}^n$ furthest from γ. Consider the quadrilateral formed by $x_{k-1}, x_{k+1}, p_{k+1}$ and p_{k-1}, as illustrated below in Figure 9.

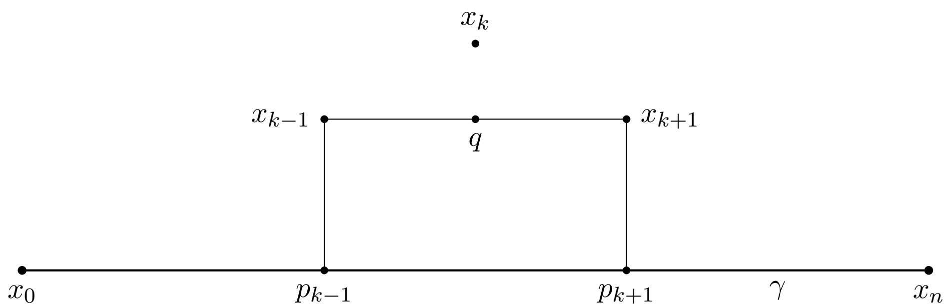

FIGURE 9. The points $x_k = g^k x_0$ lie in a bounded neighbourhood of γ.

Let q be a nearest point to x_k on a geodesic segment from x_{k-1} to x_{k+1}. By the estimate for the Gromov product in terms of distance to a geodesic (1),

$$(x_{k-1} \cdot x_{k+1})_{x_k} = d_X(x_k, [x_{k-1}, x_{k+1}]) + O(\delta),$$

and as g is an isometry, this implies that

$$d_X(x_k, q) = (gx_0 \cdot g^{-1}x_0)_{x_0} + O(\delta).$$

By thin triangles, the point q lies within distance 2δ of at least one of the other three sides of the quadrilateral. Suppose q lies within 2δ of a geodesic from x_{k-1} to p_{k-1}.

Since x_k is the furthest point from the geodesic γ, we have $d_X(x_k, \gamma) \geqslant d_X(x_{k-1}, \gamma)$, and since q lies in a 2δ-neighbourhood of the geodesic from x_{k-1} to p_{k-1}, we have

$$d_X(x_{k-1}, q) \leqslant d_X(x_k, q) + O(\delta). \tag{35}$$

On the other hand, if we now assume (31) and apply the reverse triangle inequality (3), we get

$$d_X(x_k, x_{k-1}) = d_X(x_k, q) + d_X(x_{k-1}, q) + O(\delta)$$
$$d_X(x_k, x_{k-1}) \geqslant 2d_X(x_k, q) + C_0 + O(\delta)$$

hence

$$d_X(x_{k-1}, q) \geqslant d_X(x_k, q) + C_0 + O(\delta),$$

which contradicts (35) if C_0 is large enough (depending only on δ). The same argument applies if q lies within 2δ of a geodesic from x_{k+1} to p_{k+1}. Therefore, q lies within 2δ of γ, and so $d_X(x_k, \gamma) \leqslant (gx_0 \cdot g^{-1}x_0)_{x_0} + O(\delta)$, as required. □

Consider a pair of adjacent points x_k and x_{k+1}. Combining (31), (34), and the triangle inequality, gives

$$d_X(p_k, p_{k+1}) \geqslant d_X(x_0, gx_0) - 2(gx_0 \cdot g^{-1}x_0)_{x_0} + O(\delta). \tag{36}$$

Consider a triple of consecutive points, x_k, x_{k+1} and x_{k+2}. If their corresponding nearest point projections p_k, p_{k+1} and p_{k+2} to γ do not lie in the same order, then using Proposition 2.3 repeatedly one gets, if p_{k+2} lies in between p_k and p_{k+1}, the equality

$$(x_k \cdot x_{k+2})_{x_{k+1}} = d_X(x_{k+1}, x_{k+2}) - d_X(x_{k+2}, p_{k+2}) + O(\delta)$$

which, using (33) and (34), implies

$$d_X(x_0, gx_0) - 2(gx_0 \cdot g^{-1}x_0)_{x_0} = O(\delta)$$

which contradicts (31) if C_0 is large enough. The case where p_k lies between p_{k+1} and p_{k+2} is completely analogous, therefore the p_k are monotonically ordered on γ, and so by (36)

$$d_X(p_0, p_k) \geqslant k(d_X(x_0, gx_0) - 2(gx_0 \cdot g^{-1}x_0)_{x_0} + O(\delta)),$$

which implies, by Proposition 2.3, a similar bound for $d_X(x_0, x_k)$, and so in fact $(x_k)_{k\in\mathbb{N}}$ is quasi-geodesic, with $\tau(g) \geqslant d_X(x_0, gx_0) - 2(gx_0 \cdot g^{-1}x_0)_{x_0} + O(\delta)$.

The upper bound on $\tau(g)$ follows from the triangle inequality; indeed, for each $y \in X$ one has

$$\tau(g) \leqslant d_X(y, gy)$$

and the desired bound follows by taking as y the midpoint of the geodesic segment between x_0 and gx_0, completing the proof of Proposition 5.8. □

References

[AB06] Charalambos D. Aliprantis and Kim C. Border, *Infinite dimensional analysis*, 3rd ed., Springer, Berlin, 2006.

[BDL14] Uri Bader, Bruno Duchesne, and Jean Lécureux, *Furstenberg maps for CAT(0) targets of finite telescopic dimension* (2014), available at `arXiv:1404.3187`.

[BGS85] Werner Ballmann, Mikhael Gromov, and Viktor Schroeder, *Manifolds of nonpositive curvature*, Progress in Mathematics, vol. 61, Birkhäuser Boston, Inc., Boston, MA, 1985.

[BBF10] Mladen Bestvina, Kenneth Bromberg, and Koji Fujiwara, *Constructing group actions on quasi-trees and applications to mapping class groups* (2010), available at `arXiv:1006.1939`.

[BF10] Mladen Bestvina and Mark Feighn, *A hyperbolic* Out(F_n)*-complex*, Groups Geom. Dyn. **4** (2010), no. 1, 31–58.

[BF14] Mladen Bestvina and Mark Feighn, *Hyperbolicity of the complex of free factors*, Adv. Math. **256** (2014), 104–155.

[Bjö10] Michael Björklund, *Central limit theorems for Gromov hyperbolic groups*, J. Theoret. Probab. **23** (2010), no. 3, 871–887.

[BHM11] Sébastien Blachère, Peter Haïssinsky, and Pierre Mathieu, *Harmonic measures versus quasiconformal measures for hyperbolic groups*, Ann. Sci. Éc. Norm. Supér. (4) **44** (2011), no. 4, 683–721.

[BS00] M. Bonk and O. Schramm, *Embeddings of Gromov hyperbolic spaces*, Geom. Funct. Anal. **10** (2000), no. 2, 266–306.

[Bow08] Brian H. Bowditch, *Tight geodesics in the curve complex*, Invent. Math. **171** (2008), no. 2, 281–300.

[Bri12] Martin R. Bridson, *On the dimension of CAT(0) spaces where mapping class groups act*, J. Reine Angew. Math. **673** (2012), 55–68.

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

[CM15] Danny Calegari and Joseph Maher, *Statistics and compression of scl*, Ergodic Theory Dynam. Systems **35** (2015), no. 1, 64–110.

[DP12] Thomas Delzant and Pierre Py, *Kähler groups, real hyperbolic spaces and the Cremona group*, Compos. Math. **148** (2012), no. 1, 153–184.

[Duc05] Moon Duchin, *Geodesics track random walks in Teichmüller space*, ProQuest LLC, Ann Arbor, MI, 2005. Thesis (Ph.D.)–The University of Chicago.

[Fuj08] Koji Fujiwara, *Subgroups generated by two pseudo-Anosov elements in a mapping class group. I. Uniform exponential growth*, Groups of diffeomorphisms, Adv. Stud. Pure Math., vol. 52, Math. Soc. Japan, Tokyo, 2008, pp. 283–296.

[Fur63] Harry Furstenberg, *Noncommuting random products*, Trans. Amer. Math. Soc. **108** (1963), 377–428.

[GM12] François Gautero and Frédéric Mathéus, *Poisson boundary of groups acting on $\mathbb{R}$-trees*, Israel J. Math. **191** (2012), no. 2, 585–646.

[Gui80] Y. Guivarc'h, *Sur la loi des grands nombres et le rayon spectral d'une marche aléatoire*, Conference on Random Walks (Kleebach, 1979), Astérisque, vol. 74, Soc. Math. France, Paris, 1980, pp. 47–98, 3 (French, with English summary).

[Har86] John L. Harer, *The virtual cohomological dimension of the mapping class group of an orientable surface*, Invent. Math. **84** (1986), no. 1, 157–176.

[Ham06] Ursula Hamenstädt, *Train tracks and the Gromov boundary of the complex of curves*, Spaces of Kleinian groups, London Math. Soc. Lecture Note Ser., vol. 329, Cambridge Univ. Press, Cambridge, 2006, pp. 187–207.

[HM13] Michael Handel and Lee Mosher, *The free splitting complex of a free group, I: hyperbolicity*, Geom. Topol. **17** (2013), no. 3, 1581–1672.

[Hor14] Camille Horbez, *The Poisson boundary of Out(F_N)* (2014), available at `arXiv:1405.7938`.

[Kai87] V. A. Kaimanovich, *Lyapunov exponents, symmetric spaces and a multiplicative ergodic theorem for semisimple Lie groups*, Zap. Nauchn. Sem. Leningrad. Otdel. Mat. Inst. Steklov. (LOMI) **164** (1987), no. Differentsialnaya Geom. Gruppy Li i Mekh. IX, 29–46, 196–197 (Russian, with English summary); English transl., J. Soviet Math. **47** (1989), no. 2, 2387–2398.

[Kai94] Vadim A. Kaimanovich, *The Poisson boundary of hyperbolic groups*, C. R. Acad. Sci. Paris Sér. I Math. **318** (1994), no. 1, 59–64 (English, with English and French summaries).

[Kai00] Vadim A. Kaimanovich, *The Poisson formula for groups with hyperbolic properties*, Ann. of Math. (2) **152** (2000), no. 3, 659–692.

[KM96] Vadim A. Kaimanovich and Howard A. Masur, *The Poisson boundary of the mapping class group*, Invent. Math. **125** (1996), no. 2, 221–264.

[KB02] Ilya Kapovich and Nadia Benakli, *Boundaries of hyperbolic groups*, Combinatorial and geometric group theory (New York, 2000/Hoboken, NJ, 2001), Contemp. Math., vol. 296, Amer. Math. Soc., Providence, RI, 2002, pp. 39–93.

[KL96] Michael Kapovich and Bernhard Leeb, *Actions of discrete groups on nonpositively curved spaces*, Math. Ann. **306** (1996), no. 2, 341–352.

[KL06] Anders Karlsson and François Ledrappier, *On laws of large numbers for random walks*, Ann. Probab. **34** (2006), no. 5, 1693–1706.

[KL07] Anders Karlsson and François Ledrappier, *Linear drift and Poisson boundary for random walks*, Pure Appl. Math. Q. **3** (2007), no. 4, Special Issue: In honor of Grigory Margulis., 1027–1036.

[KK13] Sang-hyun Kim and Thomas Koberda, *Embedability between right-angled Artin groups*, Geom. Topol. **17** (2013), no. 1, 493–530.

[KK14] Sang-Hyun Kim and Thomas Koberda, *The geometry of the curve graph of a right-angled Artin group*, Internat. J. Algebra Comput. **24** (2014), no. 2, 121–169.

[Kin68] J. F. C. Kingman, *The ergodic theory of subadditive stochastic processes*, J. Roy. Statist. Soc. Ser. B **30** (1968), 499–510.

[Kla] E. Klarreich, *The boundary at infinity of the curve complex and the relative Teichmüller space.*

[Kow08] E. Kowalski, *The large sieve and its applications*, Cambridge Tracts in Mathematics, vol. 175, Cambridge University Press, Cambridge, 2008. Arithmetic geometry, random walks and discrete groups.

[Led01] François Ledrappier, *Some asymptotic properties of random walks on free groups*, Topics in probability and Lie groups: boundary theory, CRM Proc. Lecture Notes, vol. 28, Amer. Math. Soc., Providence, RI, 2001, pp. 117–152.

[Mah10a] Joseph Maher, *Linear progress in the complex of curves*, Trans. Amer. Math. Soc. **362** (2010), no. 6, 2963–2991.

[Mah10b] Joseph Maher, *Random Heegaard splittings*, J. Topol. **3** (2010), no. 4, 997–1025.

[Mah11] Joseph Maher, *Random walks on the mapping class group*, Duke Math. J. **156** (2011), no. 3, 429–468.

[Mah12] Joseph Maher, *Exponential decay in the mapping class group*, J. Lond. Math. Soc. (2) **86** (2012), no. 2, 366–386.

[MS14] A. Malyutin and P. Svetlov, *Poisson–Furstenberg boundaries of fundamental groups of closed 3-manifolds* (2014), available at `arXiv:1403.2135`.

[Man74] Yu. I. Manin, *Cubic forms: algebra, geometry, arithmetic*, North-Holland Publishing Co., Amsterdam-London; American Elsevier Publishing Co., New York, 1974. Translated from the Russian by M. Hazewinkel; North-Holland Mathematical Library, Vol. 4.

[Mar91] G. A. Margulis, *Discrete subgroups of semisimple Lie groups*, Ergebnisse der Mathematik und ihrer Grenzgebiete (3) [Results in Mathematics and Related Areas (3)], vol. 17, Springer-Verlag, Berlin, 1991.

[MM99] Howard A. Masur and Yair N. Minsky, *Geometry of the complex of curves. I. Hyperbolicity*, Invent. Math. **138** (1999), no. 1, 103–149.

[MS14] Pierre Mathieu and Alessandro Sisto, *Deviation inequalities for random walks* (2014), available at `arXiv:1411.7865`.

[NS13] Amos Nevo and Michah Sageev, *The Poisson boundary of* CAT(0) *cube complex groups*, Groups Geom. Dyn. **7** (2013), no. 3, 653–695.

[Ose68] V. I. Oseledec, *A multiplicative ergodic theorem. Characteristic Ljapunov, exponents of dynamical systems*, Trudy Moskov. Mat. Obšč. **19** (1968), 179–210 (Russian).

[Osi13] Dennis Osin, *Acylindrically hyperbolic groups* (2013), available at `arXiv:1304.1246`.

[Riv08] Igor Rivin, *Walks on groups, counting reducible matrices, polynomials, and surface and free group automorphisms*, Duke Math. J. **142** (2008), no. 2, 353–379.

[Sel97] Z. Sela, *Acylindrical accessibility for groups*, Invent. Math. **129** (1997), no. 3, 527–565.

[Sis11] Alessandro Sisto, *Contracting elements and random walks* (2011), available at `arXiv:1112.2666`.

[Sis13] Alessandro Sisto, *Tracking rates for random walks* (2013), available at `arXiv:1305.5472`.

[Tio12] Giulio Tiozzo, *Sublinear deviation between geodesics and sample paths* (2012). to appear on Duke Math. J.

[Väi05] Jussi Väisälä, *Gromov hyperbolic spaces*, Expo. Math. **23** (2005), no. 3, 187–231.

[Woe00] Wolfgang Woess, *Random walks on infinite graphs and groups*, Cambridge Tracts in Mathematics, vol. 138, Cambridge University Press, Cambridge, 2000.

Joseph Maher
CUNY College of Staten Island and CUNY Graduate Center
joseph.maher@csi.cuny.edu

Giulio Tiozzo
Yale University
giulio.tiozzo@yale.edu