

# RANDOM WALKS ON BAUMSLAG–SOLITAR GROUPS

BY

JOHANNES CUNO\*

*Département de mathématiques et applications, École normale supérieure  
CNRS, PSL Research University, 45 rue d’Ulm, 75005 Paris, France  
e-mail: johannes.cuno@ens.fr*

AND

ECATERINA SAVA-HUSS\*\*

*Institute of Discrete Mathematics, Graz University of Technology  
Steyrergasse 30 / III, 8010 Graz, Austria  
e-mail: sava-huss@tugraz.at*

ABSTRACT

We consider random walks on non-amenable Baumslag–Solitar groups  $BS(p, q)$  and describe their Poisson–Furstenberg boundary. The latter is a probabilistic model for the long-time behaviour of the random walk. In our situation, we identify it in terms of the space of ends of the Bass–Serre tree and the real line using Kaimanovich’s strip criterion.

## 1. Introduction

For any two non-zero integers  $p$  and  $q$  the Baumslag–Solitar group  $BS(p, q)$  is given by the presentation  $BS(p, q) = \langle a, b : ab^p = b^q a \rangle$ . These groups were introduced by Baumslag and Solitar in [BS62], who identified  $BS(2, 3)$  as the first example of a two-generator one-relator non-Hopfian group. We consider

---

\* Research supported by the Austrian Science Fund (FWF): W1230-N13 and P24028-N18, the Canada Research Chairs Program, and the European Research Council (ERC): No 725773 “GroIsRan”.

\*\* Research supported by the Austrian Science Fund (FWF): J3575-N26.

Received January 21, 2016 and in revised form November 15, 2017

random walks on  $\text{BS}(p, q)$ . Each of them is driven by a probability measure  $\mu$  whose support generates  $\text{BS}(p, q)$  as a semigroup. The random walk starts at the identity element and proceeds with independent  $\mu$ -distributed increments  $X_1, X_2, \dots$  being multiplied from the right to the current state.

The Poisson–Furstenberg boundary was introduced by Furstenberg in [Fur63] and [Fur71]. It is a probabilistic model for the long-time behaviour of the random walk and simultaneously provides a way to represent all bounded harmonic functions on the state space. In [Kai91, Theorem 5.1], Kaimanovich considered random walks on  $\text{BS}(1, 2)$ . Under the assumption of finite first moment, he identified their Poisson–Furstenberg boundary geometrically. In particular, he showed that the latter is trivial if the random walk has no vertical drift, i.e., the expected exponent sum of the increments with respect to the generator  $a$  is equal to zero.

For random walks on non-amenable groups the situation is different. As long as they are irreducible, their Poisson–Furstenberg boundary can never be trivial. This motivates the present paper, in which we study random walks on non-amenable Baumslag–Solitar groups. It is organised as follows. In Section 2, we discuss some algebraic and geometric properties of Baumslag–Solitar groups  $\text{BS}(p, q)$  with  $1 \leq p < q$ . We explain how these groups can be understood through their projections to the Bass–Serre tree  $T$  and the hyperbolic plane  $\mathbb{H}$ . Afterwards, we recall the construction of the space of ends  $\partial T$  and the hyperbolic boundary  $\partial\mathbb{H}$ , which contains the real line  $\mathbb{R}$  as a subset. These spaces shall later be used to associate a geometric boundary to  $\text{BS}(p, q)$ . In Section 3, we turn to random walks on groups. We outline some results about the Poisson–Furstenberg boundary and then state Kaimanovich’s strip criterion, which is an important tool to identify this boundary geometrically.

In Section 4, we study random walks on  $\text{BS}(p, q)$  with finite first moment. We consider the pointwise projections of the random walk to  $T$  and  $\mathbb{H}$ . If the random walk has negative vertical drift, then the projection to  $\mathbb{H}$  converges almost surely to a random element in  $\mathbb{R}$ . For the projection to  $T$ , we do not need to distinguish between different vertical drifts; as soon as  $1 < p < q$ , it converges almost surely to a random element in  $\partial T$ . We thus endow  $\partial T$  (or even  $\partial T \times \mathbb{R}$ ) with the Borel  $\sigma$ -algebra  $\mathcal{B}_{\partial T}$  (or  $\mathcal{B}_{\partial T \times \mathbb{R}}$ ) and the hitting measure  $\nu_{\partial T}$  (or  $\nu_{\partial T \times \mathbb{R}}$ ). Finally, Kaimanovich’s strip criterion shows that the resulting probability space is isomorphic to the Poisson–Furstenberg boundary.

Up to and including Section 4.1, we assume that the two non-zero integers  $p$  and  $q$  satisfy  $1 \leq p < q$ . Then, we restrict ourselves to the non-amenable subcase  $1 < p < q$ . In the appendix, we explain how to obtain similar results for the remaining non-amenable cases  $1 < p < -q$  and  $1 < p = |q|$ . Our main result is the following.

**THEOREM 1.1** (“identification theorem”): *Let  $Z = (Z_0, Z_1, \dots)$  be a random walk on a non-amenable Baumslag–Solitar group  $G = \text{BS}(p, q)$  with  $1 < p < q$  and increments  $X_1, X_2, \dots$  of finite first moment. Depending on the vertical drift  $\delta$ , we distinguish three cases:*

- (1) *If  $\delta > 0$ , then the Poisson–Furstenberg boundary is isomorphic to  $(\partial T, \mathcal{B}_{\partial T}, \nu_{\partial T})$  endowed with the boundary map  $\text{bnd}_{\partial T} : \Omega \rightarrow \partial T$ .*
- (2) *If  $\delta < 0$ , then the Poisson–Furstenberg boundary is isomorphic to  $(\partial T \times \mathbb{R}, \mathcal{B}_{\partial T \times \mathbb{R}}, \nu_{\partial T \times \mathbb{R}})$  endowed with  $\text{bnd}_{\partial T \times \mathbb{R}} : \Omega \rightarrow \partial T \times \mathbb{R}$ .*
- (3) *If  $\delta = 0$  and  $\ln(A_{X_1})$  has finite second moment and there is an  $\varepsilon > 0$  such that  $\ln(1 + |B_{X_1}|)$  has finite  $(2 + \varepsilon)$ -th moment, then the Poisson–Furstenberg boundary is isomorphic to  $(\partial T, \mathcal{B}_{\partial T}, \nu_{\partial T})$  endowed with  $\text{bnd}_{\partial T} : \Omega \rightarrow \partial T$ .*

Note that the driftless case is a little more subtle and requires additional assumptions on the moments. Here, the terms  $A_g$  and  $B_g$  denote the imaginary and real part of the projection of an element  $g \in G$  to the hyperbolic plane  $\mathbb{H}$ . The two assumptions are certainly satisfied if  $X_1$  has finite  $(2 + \varepsilon)$ -th moment. Further details can be found at the beginning of Section 4.1.

**ACKNOWLEDGEMENTS.** We would like to thank Wolfgang Woess for suggesting this problem to us and supporting us with references and ideas while the research was carried out. Moreover, we are grateful to Vadim Kaimanovich and the anonymous reviewer, both of whom made valuable suggestions that led to a substantial improvement of the present paper.

## 2. Baumslag–Solitar groups

**2.1. AMENABILITY OF BAUMSLAG–SOLITAR GROUPS.** The structure of Baumslag–Solitar groups can be studied by means of HNN extensions. Indeed,  $\text{BS}(p, q)$  is precisely the HNN extension  $\mathbb{Z} *_\varphi$  with isomorphism  $\varphi : p\mathbb{Z} \rightarrow q\mathbb{Z}$  given by  $\varphi(p) := q$ . This allows us to use the respective machinery, such as

Britton's lemma, see [Bri63], which implies that a freely reduced non-empty word  $w$  over the letters  $a$  and  $b$  and their formal inverses can only represent the identity element  $1 \in \text{BS}(p, q)$  if it contains  $ab^r a^{-1}$  with  $p \mid r$  or  $a^{-1}b^r a$  with  $q \mid r$  as a subword. In particular, if neither  $|p| = 1$  nor  $|q| = 1$ , then the elements  $x := a$  and  $y := bab^{-1}$  generate a non-abelian free subgroup and  $\text{BS}(p, q)$  is non-amenable. On the other hand, if  $|p| = 1$  or  $|q| = 1$ , a simple calculation shows that the normal subgroup  $\langle\langle b \rangle\rangle \trianglelefteq \text{BS}(p, q)$  is abelian with quotient isomorphic to  $\mathbb{Z}$ . In this case,  $\text{BS}(p, q)$  is solvable and therefore amenable. As we will address briefly in Section 3.4, the distinction between these two cases is of importance when working with random walks.

**2.2. PROJECTION TO THE BASS–SERRE TREE.** Assume first that  $1 \leq p < q$ . The Cayley graph  $\Gamma$  of the group  $G := \text{BS}(p, q)$  with respect to the standard generators  $a$  and  $b$  is the directed multigraph with vertex set  $G$ , edge set  $G \times \{a, b\}$ , source function  $s : G \times \{a, b\} \rightarrow G$  given by  $s(g, x) := g$ , and target function  $t : G \times \{a, b\} \rightarrow G$  given by  $t(g, x) := gx$ . Every directed multigraph can be converted into a simple graph by ignoring the direction and the multiplicity of the edges and deleting the loops. For the purpose of this paper, it is sufficient to think of  $\Gamma$  as a simple graph, and we shall tacitly do so.

*Definition 2.1* (“levels”): Let  $\lambda : \{a, b\} \rightarrow \mathbb{Z}$  be the map given by  $\lambda(a) := 1$  and  $\lambda(b) := 0$ . It follows from von Dyck's theorem, see, e.g., [CZ93, §1.1.3], that this map can be uniquely extended to a group homomorphism  $\lambda : G \rightarrow \mathbb{Z}$ . We think of it as a level function.

Consider the illustration of  $\Gamma$  in Figure 1. Intuitively speaking, we may look at it from the side to see the associated Bass–Serre tree. Formally, let

$$B := \langle b \rangle \leq G$$

and let  $T$  be the graph with vertex set  $G/B = \{gB : g \in G\}$  and edge set  $\{\{gB, gaB\} : g \in G\}$ . This graph is actually a tree; it is connected and, by Britton's lemma, it does not contain any cycle. We use the symbol  $\pi_T$  to denote the canonical projection to the cosets, i.e., the map  $\pi_T : G \rightarrow G/B$  given by  $\pi_T(g) := gB$ .

*Remark 2.2:* Since  $\lambda(b) = 0$ , the level function is well-defined on the vertices of  $T$ . It is not hard to see that every vertex of  $T$  has exactly  $p + q$  neighbours;  $p$  of them are one level below and  $q$  of them are one level above the vertex.

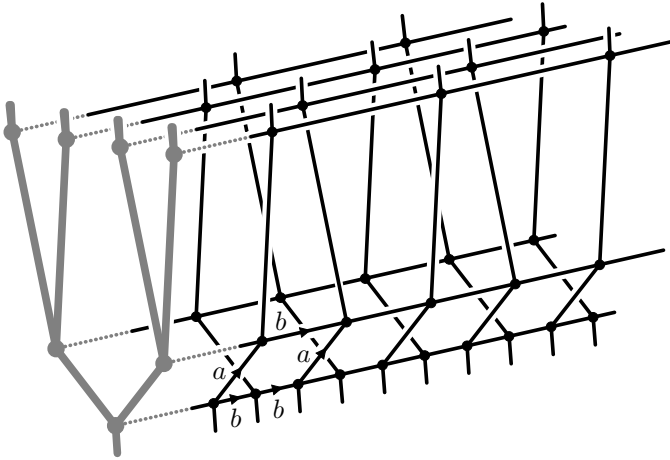


Figure 1. The Cayley graph  $\Gamma$  of  $BS(1,2)$  with respect to the standard generators  $a$  and  $b$ .

2.3. PROJECTION TO THE HYPERBOLIC PLANE. The second projection captures the information that is obtained by looking at  $\Gamma$  from the front. It is convenient to describe it in terms of the hyperbolic plane. So, let  $\mathbb{H}$  be the hyperbolic plane as per the Poincaré half-plane model, i.e.,  $\mathbb{H} := \{z \in \mathbb{C} : \text{Im}(z) > 0\}$ , endowed with the standard metric

$$d_{\mathbb{H}}(z_1, z_2) := \ln \left( \frac{|z_1 - \bar{z}_2| + |z_1 - z_2|}{|z_1 - \bar{z}_2| - |z_1 - z_2|} \right) = \text{arcosh} \left( 1 + \frac{|z_1 - z_2|^2}{2 \cdot \text{Im}(z_1) \text{Im}(z_2)} \right).$$

The isometry group  $\text{Isom}(\mathbb{H})$  consists of all maps  $\varphi : \mathbb{H} \rightarrow \mathbb{H}$  of the form

$$\underbrace{\varphi(z) = \frac{\alpha z + \beta}{\gamma z + \delta}}_{\textcircled{1}} \quad \text{or} \quad \underbrace{\varphi(z) = \frac{\alpha \cdot (-\bar{z}) + \beta}{\gamma \cdot (-\bar{z}) + \delta}}_{\textcircled{2}} \quad \text{with } \alpha, \beta, \gamma, \delta \in \mathbb{R}, \alpha\delta - \beta\gamma > 0,$$

see, e.g., [Bea83, Theorem 7.4.1]. For the time being, we shall only work with the orientation-preserving isometries  $\textcircled{1}$ . The orientation-reversing ones  $\textcircled{2}$  will later be of relevance, in Section A.1 of the appendix. Let  $\pi_{\text{Isom}(\mathbb{H})} : \{a, b\} \rightarrow \text{Isom}(\mathbb{H})$  be the map given by  $\pi_{\text{Isom}(\mathbb{H})}(a) := (z \mapsto \frac{a}{p} \cdot z)$  and  $\pi_{\text{Isom}(\mathbb{H})}(b) := (z \mapsto z + 1)$ . As in Definition 2.1, it follows from von Dyck’s theorem that this map can be uniquely extended to a group homomorphism  $\pi_{\text{Isom}(\mathbb{H})} : G \rightarrow \text{Isom}(\mathbb{H})$ . Now, we define  $\pi_{\mathbb{H}} : G \rightarrow \mathbb{H}$  by  $\pi_{\mathbb{H}}(g) := \pi_{\text{Isom}(\mathbb{H})}(g)(i)$ .

LEMMA 2.3: *For every  $g \in G$  the point  $\pi_{\mathbb{H}}(ga) \in \mathbb{H}$  is above the point  $\pi_{\mathbb{H}}(g) \in \mathbb{H}$ ; the two points have the same real part and their distance in the hyperbolic plane is  $\ell_a := \ln(\frac{a}{p})$ . Similarly, the point  $\pi_{\mathbb{H}}(gb) \in \mathbb{H}$  is to the right of the point  $\pi_{\mathbb{H}}(g) \in \mathbb{H}$ ; the two points have the same imaginary part and their distance in the hyperbolic plane is  $\ell_b := \ln(\frac{3+\sqrt{5}}{2})$ .*

*Proof.* This is clear for  $g = 1$ . Now, pick an arbitrary element  $g \in G$ . The points  $\pi_{\mathbb{H}}(ga) \in \mathbb{H}$  and  $\pi_{\mathbb{H}}(g) \in \mathbb{H}$  are obtained by evaluating  $\pi_{\text{Isom}(\mathbb{H})}(g)$  at  $\pi_{\mathbb{H}}(a) \in \mathbb{H}$  and  $\pi_{\mathbb{H}}(1) \in \mathbb{H}$ . Since  $g$  can be written as a product over  $a^{\pm 1}$  and  $b^{\pm 1}$ , its image  $\pi_{\text{Isom}(\mathbb{H})}(g)$  is the respective composition of  $\pi_{\text{Isom}(\mathbb{H})}(a^{\pm 1})$  and  $\pi_{\text{Isom}(\mathbb{H})}(b^{\pm 1})$ . Being dilations and translations, the latter preserve the relative position of any two points in  $\mathbb{H}$ , and so does  $\pi_{\text{Isom}(\mathbb{H})}(g)$ . The same argument works for the second assertion, which completes the proof. ■

Here and throughout the present paper, we use the symbol  $\mathbb{N}_0$  to denote the non-negative integers and the symbol  $\mathbb{N}$  to denote the strictly positive ones.

*Definition 2.4* (“path”, “reduced path”): Given a simple graph with vertex set  $V$ , we consider finite paths  $v : \{0, 1, \dots, n\} \rightarrow V$ , infinite paths  $v : \mathbb{N}_0 \rightarrow V$ , and doubly infinite paths  $v : \mathbb{Z} \rightarrow V$ . In any case, being a path means that for every possible choice of  $k$  the vertices  $v(k)$  and  $v(k + 1)$  are adjacent. A path is reduced if for every possible choice of  $k$  the vertices  $v(k)$  and  $v(k + 2)$  are distinct.

*Remark 2.5* (“discrete hyperbolic plane”): One way to “recover hyperbolic structures within the Cayley graph  $\Gamma$  is the following. Fix an ascending doubly infinite path  $v : \mathbb{Z} \rightarrow G/B$  in the tree  $T$ . Ascending refers to the level function constructed in Definition 2.1 and Remark 2.2, and it means that for every  $k \in \mathbb{Z}$  the vertex  $v(k)$  is located below the following vertex  $v(k + 1)$ . Let  $G_v$  be the full  $\pi_T$ -preimage of the path, i.e., the set consisting of all  $g \in G$  such that  $\pi_T(g)$  is contained in  $v(\mathbb{Z})$ . The subgraph  $\Gamma_v \leq \Gamma$  spanned by  $G_v$ , see ① in Figure 2, is connected so that the graph distance  $d_{\Gamma_v}$  is a metric. This subgraph is sometimes referred to as discrete hyperbolic plane or plane of bricks, which makes particular sense in light of the fact that the restriction  $\pi_{\mathbb{H}}|_{G_v} : G_v \rightarrow \mathbb{H}$  is a quasi-isometry, even a bi-Lipschitz map, between the graph  $\Gamma_v$  endowed with the graph distance  $d_{\Gamma_v}$  and the hyperbolic plane  $\mathbb{H}$  endowed with the standard metric  $d_{\mathbb{H}}$ .

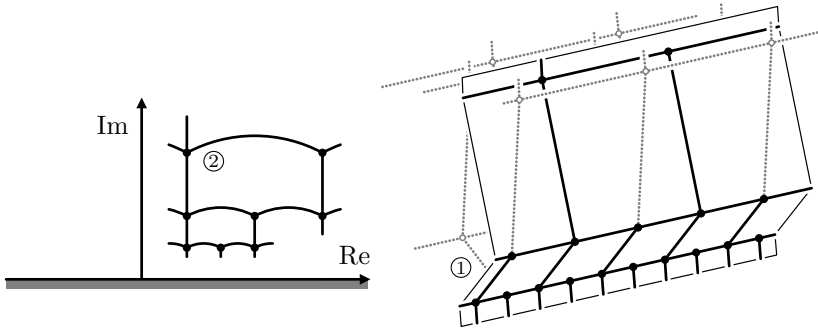


Figure 2. A part of a discrete hyperbolic plane  $\Gamma_v$  (right) and its projection to  $\mathbb{H}$  (left).

2.4. COMPACTIFICATIONS. Both the tree  $T$  and the hyperbolic plane  $\mathbb{H}$  have a natural compactification. In case of  $T$ , it is the **end compactification**, which can be constructed as follows. Fix a base point, say  $B \in G/B$ , and consider the set  $\widehat{T}$  of all reduced paths that start in  $B$ , be they finite or infinite. The endpoint map yields a one-to-one correspondence between the finite paths and the vertices  $G/B$ . We may therefore think of  $G/B$  as a subset of  $\widehat{T}$ . The set  $\widehat{T}$  can be endowed with the metric

$$d_{\widehat{T}}(x, y) := \begin{cases} 2^{-|x \wedge y|} & \text{if } x \neq y, \\ 0 & \text{if } x = y, \end{cases}$$

where  $|x \wedge y|$  denotes the number of edges the two paths run together until they separate, see ① in Figure 3. In other words,  $|x \wedge y|$  is the maximal number  $k \in \mathbb{N}_0$  such that  $x$  and  $y$  are both defined at  $k$  and the vertices  $x(k)$  and  $y(k)$  agree. Hence, the later the paths separate the closer they are. The set  $\widehat{T}$  endowed with the metric  $d_{\widehat{T}}$  is a compact metric space that contains  $G/B$  as a discrete and dense subset. The complement of  $G/B$  is the set of infinite paths, it is denoted by  $\partial T$  and called the **space of ends**.

In case of  $\mathbb{H}$ , we temporarily switch to the Poincaré disc model. Instead of working in the half-plane  $\mathbb{H} = \{z \in \mathbb{C} : \text{Im}(z) > 0\}$ , we consider the open unit disc  $\mathbb{D} := \{z \in \mathbb{C} : |z| < 1\}$ . The Cayley transform  $W : \mathbb{H} \leftrightarrow \mathbb{D}$  given by

$$W(z) := \frac{z - i}{z + i}$$

is one possibility to convert between the two models. The hyperbolic topology on  $\mathbb{D}$  is the one induced by the Cayley transform. It agrees with the standard topology on  $\mathbb{D}$  so that, topologically speaking, the hyperbolic plane in the Poincaré disc model is just a subspace of the complex plane  $\mathbb{C}$ . We may therefore compactify it by taking the closed unit disc  $\widehat{\mathbb{D}} := \{z \in \mathbb{C} : |z| \leq 1\}$ , see Figure 3. In order to translate this compactification back to the Poincaré half-plane model, we first extend both the domain and the codomain of the Cayley transform so that we obtain a bijection  $W : \mathbb{H} \cup \mathbb{R} \cup \{\infty\} \leftrightarrow \widehat{\mathbb{D}}$ , and then apply its inverse. The space

$$\widehat{\mathbb{H}} := \mathbb{H} \cup \mathbb{R} \cup \{\infty\}$$

is our compactification. It is, once again, endowed with the induced topology, and thus a compact space that contains  $\mathbb{H}$  as a dense subset. The complement of  $\mathbb{H}$  is the union  $\mathbb{R} \cup \{\infty\}$ , it is denoted by  $\partial\mathbb{H}$  and called the **hyperbolic boundary**. Having introduced the hyperbolic boundary this way, the following lemma gives us a helpful criterion for convergence. Its proof is elementary and we leave it to the reader.

LEMMA 2.6: *A sequence  $(x_0, x_1, \dots)$  in  $\mathbb{H}$  converges to  $r \in \mathbb{R} = \partial\mathbb{H} \setminus \{\infty\}$  if and only if it does with respect to the standard topology on the complex plane  $\mathbb{C}$ . The sequence converges to  $\infty \in \partial\mathbb{H}$  if and only if the absolute values  $|x_n|$  tend to  $\infty$ .*

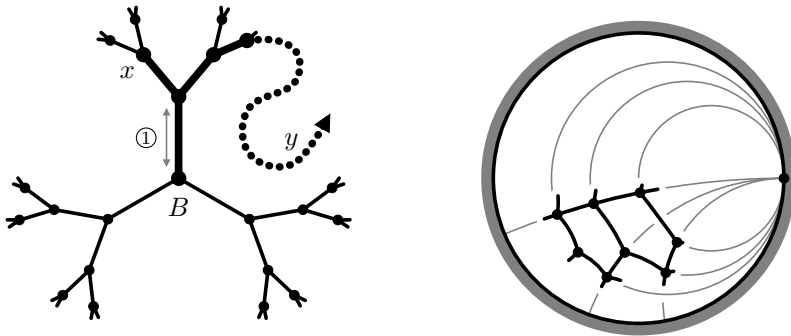


Figure 3. The space of ends (left) and the hyperbolic boundary in the Poincaré disc model (right).



### 3. Random walks on groups

3.1. PRELIMINARIES. The aim of the present paper is to study random walks. Given a countable state space  $X$ , an initial probability measure  $\vartheta : X \rightarrow [0, 1]$ , and transition probabilities  $p : X \times X \rightarrow [0, 1]$ , we are interested in the Markov chain  $Z = (Z_0, Z_1, \dots)$  that starts according to  $\vartheta$  and proceeds according to  $p$ . Formally, we construct the probability space  $(\Omega, \mathcal{A}, \mathbb{P})$ , where  $\Omega := \{(x_0, x_1, \dots) : x_n \in X\}$  is the set of trajectories,  $\mathcal{A}$  is the product  $\sigma$ -algebra, and  $\mathbb{P}$  is the probability measure induced by  $\vartheta$  and  $p$ . The projections  $Z_n : \Omega \rightarrow X$  given by  $Z_n(x_0, x_1, \dots) := x_n$  then become random variables that constitute the Markov chain. We shall use the term **random walk** instead of Markov chain.

Assume now that  $X$  is a countable group  $G$ . In this situation, we may adapt the transition probabilities  $p : G \times G \rightarrow [0, 1]$  to the group structure. More precisely, let  $\mu : G \rightarrow [0, 1]$  be a probability measure on  $G$  and consider the random walk given by the following data. The initial probability measure  $\vartheta : G \rightarrow [0, 1]$  puts all mass on the identity element  $1 \in G$  and the transition probabilities  $p : G \times G \rightarrow [0, 1]$  are given by  $p(g, h) := \mu(g^{-1}h)$ . We could also have said  $p(g, gx) := \mu(x)$ , which leads to a handy interpretation. The random walk starts at the identity element and has independent  $\mu$ -distributed increments being multiplied from the right to the current state. Hence,  $Z_0 = 1$  a.s. (= almost surely) and for every  $n \in \mathbb{N}$ ,  $Z_n = X_1 \cdots X_n$ , where  $X_1, X_2, \dots$  is a sequence of independent  $\mu$ -distributed random variables.

Throughout the present paper, we assume that the support

$$\text{supp}(\mu) = \{g \in G : \mu(g) > 0\}$$

generates  $G$  as a semigroup. In other words, the random walk is irreducible; any two states can be reached from each other with positive probability.

Given a probability space, e.g.,  $(\Omega, \mathcal{A}, \mathbb{P})$  described above, and a real valued random variable  $X : \Omega \rightarrow \mathbb{R}$ , the latter has finite first moment if  $\int |X| \, d\mathbb{P}$  is finite. In this case, both  $\int X^+ \, d\mathbb{P}$  and  $\int X^- \, d\mathbb{P}$  are finite and we can define the expectation  $\mathbb{E}(X) := \int X^+ \, d\mathbb{P} - \int X^- \, d\mathbb{P}$ . Of course, the difference would still make sense if only one of the two integrals was finite. But this case is not of relevance for us and when writing  $\mathbb{E}(X)$  we implicitly mean that  $\mathbb{E}(X)$  is a real number. More generally, given any non-negative  $k \in \mathbb{R}$ , the random variable  $X : \Omega \rightarrow \mathbb{R}$  has finite  $k$ -th moment if  $\int |X|^k \, d\mathbb{P}$  is finite.

*Definition 3.1* (“word metric”): If  $G$  is a finitely generated group and  $S \subseteq G$  is a finite generating set, then the word metric  $d_S : G \times G \rightarrow \mathbb{N}_0$  is the distance in the respective Cayley graph. In other words,

$$d_S(g, h) := \min\{n \in \mathbb{N}_0 : \exists s_1, \dots, s_n \in S \text{ and } \varepsilon_1, \dots, \varepsilon_n \in \{1, -1\} \\ \text{such that } g^{-1}h = s_1^{\varepsilon_1} \cdots s_n^{\varepsilon_n}\}.$$

*Definition 3.2* (“finite  $k$ -th moment” for  $G$ -valued random variables): Let  $G$  be a finitely generated group and  $S \subseteq G$  be a finite generating set. A random variable  $X : \Omega \rightarrow G$  has finite  $k$ -th moment if the image  $d_S(1, X) : \Omega \rightarrow \mathbb{N}_0$  has finite  $k$ -th moment in the classical sense, i.e., if

$$\int d_S(1, X)^k d\mathbb{P} < \infty.$$

It is well-known that this property does not depend on the choice of  $S$ .

**3.2. LEBESGUE–ROHLIN SPACES.** In order to define the Poisson–Furstenberg boundary, we need to ensure that we are working with Lebesgue–Rohlin spaces, which are also known as standard probability spaces. For definitions, basic examples, and fundamental results, we refer to [Roh52], [Hae73], and [Rud90]. Two short summaries can also be found in [CFS82, Appendix 1] and [KKR04, Appendix].

*Example 3.3:* A Polish space is a topological space that is separable and completely metrisable. All Polish spaces endowed with their Borel  $\sigma$ -algebra  $\mathcal{B}$  and a Borel measure  $\mu$  become, after completion, examples of Lebesgue–Rohlin spaces, see [Roh52, §2.7] and [Hae73, p. 248, Example 1].

In light of Example 3.3, we observe that the space of trajectories  $\Omega$  introduced in Section 3.1 is the product  $X^{\mathbb{N}_0}$  and can therefore be endowed with the product topology. The resulting space is actually Polish; see, e.g., [Wil70, Theorem 24.11]. Since its Borel  $\sigma$ -algebra agrees with the product  $\sigma$ -algebra  $\mathcal{A}$ , the completion of  $(\Omega, \mathcal{A}, \mathbb{P})$  is a Lebesgue–Rohlin space. Let us assume that, as soon as a measurable space is endowed with a probability measure, we are working with its completion. We may therefore say that  $(\Omega, \mathcal{A}, \mathbb{P})$  is a Lebesgue–Rohlin space.

As it is customary, we always deal with Lebesgue–Rohlin spaces mod 0, i.e., up to subsets of measure 0, even when it is not indicated explicitly.

**3.3. POISSON–FURSTENBERG BOUNDARY.** Several equivalent definitions of the Poisson–Furstenberg boundary are given in [KV83]. Since we are interested in the long-time behaviour of the trajectories  $x = (x_0, x_1, \dots) \in \Omega$ , we identify those pairs of trajectories whose tails sooner or later behave identically. More precisely, we define an equivalence relation  $\sim$  on  $\Omega$  by setting

$$x \sim y : \iff \exists t_1, t_2 \in \mathbb{N}_0 \text{ such that } \forall n \in \mathbb{N}_0 \text{ the equation } x_{t_1+n} = y_{t_2+n} \text{ holds.}$$

Consider the partition  $\zeta$  of  $\Omega$  into the equivalence classes mod  $\sim$ . It induces a sub- $\sigma$ -algebra  $\mathcal{A}(\zeta) \subseteq \mathcal{A}$  consisting of all those measurable sets  $A \in \mathcal{A}$  that are, mod 0, unions of elements in  $\zeta$ . Being a complete sub- $\sigma$ -algebra,  $\mathcal{A}(\zeta)$  corresponds to a measurable partition  $\zeta_1$  of  $\Omega$ . This partition is unique, up to equivalence mod 0. It has the properties that the induced sub- $\sigma$ -algebra  $\mathcal{A}(\zeta_1) \subseteq \mathcal{A}$  coincides with  $\mathcal{A}(\zeta)$  and that the quotient of  $(\Omega, \mathcal{A}, \mathbb{P})$  by  $\zeta_1$  is a Lebesgue–Rohlin space.

The latter is called the **Poisson–Furstenberg boundary** and we denote it by  $(B, \mathcal{B}, \nu)$ . It is naturally endowed with a **boundary map**  $\text{bnd} : \Omega \rightarrow B$  assigning to every trajectory  $x \in \Omega$  the element in  $\zeta_1$  that contains it. The boundary map is a measurable and measure-preserving map between Lebesgue–Rohlin spaces; such a map is called a **homomorphism**.

Here, we consider irreducible random walks on countable groups  $G$ . In this situation, the measurable  $G$ -action on  $\Omega$  given by  $g(x_0, x_1, \dots) := (gx_0, gx_1, \dots)$  induces a measurable  $G$ -action on  $B$  so that the measure  $\nu$  is  **$\mu$ -stationary**, i.e.,

$$\nu(A) = \sum_{g \in G} \mu(g) \cdot \nu(g^{-1}A),$$

and **quasi-invariant**, i.e., the  $G$ -action maps null sets  $A$  to null sets  $gA$ .

**3.4. SOME RESULTS ABOUT THE POISSON–FURSTENBERG BOUNDARY.** Given such a random walk, it is a challenging problem to decide whether the Poisson–Furstenberg boundary is trivial or not. In the latter case, one may wonder how to identify it geometrically. We shall only address a few results; a survey was given by Erschler in [Ers10]. As always, we assume that the random walk is irreducible.

If  $G$  is abelian, then the Poisson–Furstenberg boundary is trivial, see [Bla55] and [CD60]. The same holds true for all groups of polynomial growth, and for groups of subexponential growth endowed with a probability measure  $\mu$  with finite first moment. For the special case of probability measures with finite

support see [Ave74] and for the general case see, e.g., [KW02, Theorem 5.3] and [Ers04, §4]. Moreover, it was shown in [Ers04] that the assumption of finite first moment cannot be dropped. If  $G$  is amenable, then there is at least one symmetric probability measure  $\mu$  such that the Poisson–Furstenberg boundary is trivial, see the conjecture in [Fur73, §9]. The proof was announced in [VK79, Theorem 4] and given in [Ros81] and [KV83].

For random walks on the Baumslag–Solitar group  $G = BS(1, 2)$  with finite first moment, one can be more specific. Here, the Poisson–Furstenberg boundary is isomorphic to  $\mathbb{R}$  for  $\delta < 0$ , trivial for  $\delta = 0$ , and isomorphic to  $\mathbb{Q}_2$  for  $\delta > 0$ , see [Kai91, Theorem 5.1]. Further results about random walks on rational affinities are given in [Bro06]. If  $G$  is non-amenable, then the Poisson–Furstenberg boundary can never be trivial, see [Fur73, §9] and [KV83, §4.2]. This holds in particular for random walks on non-amenable Baumslag–Solitar groups, even when  $\delta = 0$ .

*Remark 3.4:* There are striking similarities between solvable Baumslag–Solitar groups and lamplighter groups. For example, while  $BS(1, 2)$  can be described as the semidirect product  $\mathbb{Z}[\frac{1}{2}] \rtimes \mathbb{Z}$ , where  $1 \in \mathbb{Z}$  acts by doubling, the lamplighter group  $\mathbb{Z}_2 \wr \mathbb{Z}$  is defined as  $(\bigoplus_{i \in \mathbb{Z}} \mathbb{Z}_2) \rtimes \mathbb{Z}$ , where  $1 \in \mathbb{Z}$  acts by shifting the index by 1. For results on the Poisson–Furstenberg boundary of random walks on lamplighter groups, see [VK79], [KV83], [LP15], and also [Sav10].

**3.5. KAIMANOVICH’S STRIP CRITERION.** Kaimanovich’s strip criterion, which we recall below, is a tool for identifying the Poisson–Furstenberg boundary geometrically. We state it as a theorem and then briefly discuss the notions appearing in the statement. For the proof, we refer to [Kai00, §6.4].

**THEOREM 3.5 (“strip criterion”):** *Let  $Z = (Z_0, Z_1, \dots)$  be a random walk on a countable group  $G$  driven by a probability measure  $\mu$  with finite entropy  $H(\mu)$ . Moreover, let  $(B_-, \mathcal{B}_-, \nu_-)$  and  $(B_+, \mathcal{B}_+, \nu_+)$  be  $\check{\mu}$ - and  $\mu$ -boundaries, respectively. If there exist a gauge  $\mathcal{G} = (\mathcal{G}_1, \mathcal{G}_2, \dots)$  on  $G$  with associated gauge function  $|\cdot| = |\cdot|_{\mathcal{G}}$  and a measurable  $G$ -equivariant map  $\mathcal{S}$  assigning to pairs of points  $(b_-, b_+) \in B_- \times B_+$  non-empty strips  $\mathcal{S}(b_-, b_+) \subseteq G$  such that for every  $g \in G$  and  $\nu_- \otimes \nu_+$ -almost every  $(b_-, b_+) \in B_- \times B_+$*

$$\frac{1}{n} \cdot \ln(\text{card}(\mathcal{S}(b_-, b_+)g \cap \mathcal{G}_{|Z_n|})) \xrightarrow{n \rightarrow \infty} 0 \quad \text{in probability,}$$

*then the  $\mu$ -boundary  $(B_+, \mathcal{B}_+, \nu_+)$  is maximal.*

*Remark 3.6:* The proof shows that it is not even necessary to verify the convergence for every  $g \in G$ . It suffices to consider the special case  $g = 1$  as long as we can ensure that a random strip contains the identity element  $1 \in G$  with positive probability, i.e., that  $\nu_- \otimes \nu_+ \{(b_-, b_+) \in B_- \times B_+ : 1 \in \mathcal{S}(b_-, b_+)\} > 0$ .

The **entropy** of the probability measure  $\mu$  is given by

$$H(\mu) := \sum_{g \in G} -\log_2(\mu(g)) \cdot \mu(g).$$

Here, as usual, one defines  $-\log_2(0) \cdot 0 := 0$ . The assumption of finite entropy will be no issue for us because Baumslag–Solitar groups are finitely generated and the increments of the random walks under consideration will all have finite first moment. This implies that their probability measures  $\mu$  have finite entropy, as stated in the following lemma. Its proof is elementary. For an idea see, e.g., [GPS94, Theorem 4.1].

**LEMMA 3.7:** *Let  $G$  be a finitely generated group and let  $\mu : G \rightarrow [0, 1]$  be a probability measure. If a  $\mu$ -distributed random variable  $X : \Omega \rightarrow G$  has finite first moment, then  $\mu$  has finite entropy.*

Kaimanovich defines a  **$\mu$ -boundary** as the quotient of the Poisson–Furstenberg boundary with respect to some  $G$ -invariant measurable partition; see, e.g., [Kai00, §1.5]. The Poisson–Furstenberg boundary is therefore itself a  $\mu$ -boundary, the maximal one. Moreover, every Lebesgue–Rohlin space  $(B_+, \mathcal{B}_+, \nu_+)$  endowed with a measurable  $G$ -action and a homomorphism  $\text{bnd}_+ : \Omega \rightarrow B_+$  that is ①  $\sim$ -invariant and ②  $G$ -equivariant is a  $\mu$ -boundary. At this point, recall that the random walk is irreducible, whence the properties ① and ② already imply that the measure  $\nu_+$  is  $\mu$ -stationary and quasi-invariant.

While  $\mu$  is the probability measure that drives the random walk, the symbol  $\check{\mu}$  denotes the reflected probability measure, which is given by  $\check{\mu}(g) := \mu(g^{-1})$ . Accordingly, a  **$\check{\mu}$ -boundary** is a Lebesgue–Rohlin space  $(B_-, \mathcal{B}_-, \nu_-)$  that satisfies the requirements of a  $\mu$ -boundary when replacing  $\mu$  by  $\check{\mu}$ .

A **gauge**  $\mathcal{G}$  is an exhaustion  $\mathcal{G} = (\mathcal{G}_1, \mathcal{G}_2, \dots)$  of the group  $G$ , i.e., a sequence of subsets  $\mathcal{G}_k \subseteq G$  which is increasing  $\mathcal{G}_1 \subseteq \mathcal{G}_2 \subseteq \dots$  and whose union  $\mathcal{G}_1 \cup \mathcal{G}_2 \cup \dots$  is the whole group  $G$ . Given a gauge  $\mathcal{G}$  and an element  $g \in G$ , we may ask for the minimal index  $k \in \mathbb{N}$  with the property that  $g \in \mathcal{G}_k$ . This index is the value of the associated gauge function  $|\cdot| = |\cdot|_{\mathcal{G}}$  at  $g$ .

*Remark 3.8:* Kaimanovich distinguishes between various kinds of gauges, see [Kai00]. For example, a gauge  $\mathcal{G}$  is subadditive if any two group elements  $g_1, g_2 \in G$  satisfy  $|g_1 g_2| \leq |g_1| + |g_2|$  and it is temperate if all gauge sets  $\mathcal{G}_k$  are finite and grow at most exponentially. Even though these two properties do play a crucial role in the corollaries to the strip criterion given in [Kai00, §6.5], they are not required in the strip criterion itself. And, in fact, not all of our gauges will have these two properties.

The power set  $\{0, 1\}^G$  is naturally endowed with the product  $\sigma$ -algebra, which enables us to talk about measurability of the map  $\mathcal{S} : B_- \times B_+ \rightarrow \{0, 1\}^G$ . More precisely, the product  $\sigma$ -algebra on  $\{0, 1\}^G$  is generated by the coordinate projections. Since the set  $\{0, 1\}$  consists of only two elements, the  $\sigma$ -algebra is already generated by the preimages of  $1 \in \{0, 1\}$ . In order to show that  $\mathcal{S}$  is measurable, it thus suffices to verify that for every  $g \in G$  the set of all  $(b_-, b_+) \in B_- \times B_+$  whose strip  $\mathcal{S}(b_-, b_+) \subseteq G$  contains the element  $g \in G$  is measurable. As soon as we know that  $\mathcal{S}$  is  $G$ -equivariant, it even suffices to verify measurability for  $g = 1$ , which will be immediate for the strips under consideration.

## 4. Identification of the Poisson–Furstenberg boundary

4.1. CONVERGENCE TO THE BOUNDARY OF THE HYPERBOLIC PLANE. Let us now return to  $G = \text{BS}(p, q)$  with  $1 \leq p < q$  and consider a random walk  $Z = (Z_0, Z_1, \dots)$  on  $G$ . When working with the projection  $\pi_{\mathbb{H}} : G \rightarrow \mathbb{H}$ , we may analyse the imaginary part  $\text{Im}(\pi_{\mathbb{H}}(g))$  and the real part  $\text{Re}(\pi_{\mathbb{H}}(g))$  separately, and it is convenient to abbreviate the former by  $A_g$  and the latter by  $B_g$ . Occasionally, we do not assume that  $X_1$  has some finite moment but impose this assumption on the images  $\ln(A_{X_1})$  and  $\ln(1 + |B_{X_1}|)$ . The following lemma relates the two situations.

LEMMA 4.1: *If  $X_1$  has finite  $k$ -th moment, then  $\ln(A_{X_1})$  and  $\ln(1 + |B_{X_1}|)$  have finite  $k$ -th moment, too.*

*Remark 4.2:* It follows from the definition of  $\pi_{\mathbb{H}}$  that for every  $g \in G$  the equation  $A_g = q^{\lambda(g)} p^{-\lambda(g)}$  holds. Taking the logarithm on both sides yields the equation  $\ln(A_g) = \ln\left(\frac{q}{p}\right) \cdot \lambda(g)$ . So, instead of thinking of  $\ln(A_g)$  we may think of a multiple of  $\lambda(g)$ .

*Proof of Lemma 4.1.* Let  $S := \{a, b\} \subseteq G$  be the standard generating set. Then

$$\begin{aligned} \int |\ln(A_{X_1})|^k \, d\mathbb{P} &= \left( \ln \left( \frac{q}{p} \right) \right)^k \cdot \int |\lambda(X_1)|^k \, d\mathbb{P} \\ &\leq \left( \ln \left( \frac{q}{p} \right) \right)^k \cdot \underbrace{\int d_S(1, X_1)^k \, d\mathbb{P}}_{< \infty} < \infty. \end{aligned}$$

This proves the first assertion. For the second one, Lemma 2.3 implies that the distance  $d_{\mathbb{H}}(\pi_{\mathbb{H}}(1), \pi_{\mathbb{H}}(g))$  is at most  $\max\{\ell_a, \ell_b\} \cdot d_S(1, g)$ . This allows us to estimate  $\ln(1 + |B_g|)$  by a multiple of  $d_S(1, g)$ . Indeed,

$$\begin{aligned} \ln(1 + |B_g|) &\leq \ln \left( 1 + \frac{1}{2} \cdot |B_g|^2 + \sqrt{\left( 1 + \frac{1}{2} \cdot |B_g|^2 \right)^2 - 1} \right) \\ &= \operatorname{arcosh} \left( 1 + \frac{1}{2} \cdot |B_g|^2 \right) \\ &= d_{\mathbb{H}}(i, i + B_g) \leq d_{\mathbb{H}}(i, A_g \cdot i + B_g) + d_{\mathbb{H}}(A_g \cdot i + B_g, i + B_g) \\ &= d_{\mathbb{H}}(\pi_{\mathbb{H}}(1), \pi_{\mathbb{H}}(g)) + |\ln(A_g)| \\ &\leq \max\{\ell_a, \ell_b\} \cdot d_S(1, g) + \ln \left( \frac{q}{p} \right) \cdot |\lambda(g)| \\ &\leq \max\{\ell_a, \ell_b\} \cdot d_S(1, g) + \ln \left( \frac{q}{p} \right) \cdot d_S(1, g). \end{aligned}$$

Therefore,

$$\int \ln(1 + |B_{X_1}|)^k \, d\mathbb{P} \leq \left( \max\{\ell_a, \ell_b\} + \ln \left( \frac{q}{p} \right) \right)^k \cdot \underbrace{\int d_S(1, X_1)^k \, d\mathbb{P}}_{< \infty} < \infty,$$

which proves the second assertion.  $\blacksquare$

*Definition 4.3* (“vertical drift”): If  $\ln(A_{X_1})$  has finite first moment, then  $\lambda(X_1)$  has finite first moment and we can define the expectation  $\mathbb{E}(\lambda(X_1))$ . We call the latter the vertical drift and denote it by  $\delta$ .

The following lemmas concern the behaviour of the projections  $\pi_{\mathbb{H}}(Z_n)$ . They seem to be well-known and we do not claim originality. But, for the sake of completeness, we give rigorous proofs.

**LEMMA 4.4:** *Assume that  $\ln(A_{X_1})$  has finite first moment. If  $Z$  has positive vertical drift  $\delta > 0$ , then the projections  $\pi_{\mathbb{H}}(Z_n)$  converge a.s. to  $\infty \in \partial\mathbb{H}$ .*

*Proof.* By the strong law of large numbers,

$$\frac{\lambda(Z_n)}{n} = \frac{\lambda(X_1) + \dots + \lambda(X_n)}{n} \xrightarrow[n \rightarrow \infty]{\text{a.s.}} \mathbb{E}(\lambda(X_1)) = \delta > 0.$$

Therefore, the projections  $\lambda(Z_n)$  tend a.s. to  $\infty$ . By Remark 4.2, so do the imaginary parts  $A_{Z_n}$  and the absolute values  $|\pi_{\mathbb{H}}(Z_n)|$ . Now, Lemma 2.6 completes the proof. ■

LEMMA 4.5: *Assume that both  $\ln(A_{X_1})$  and  $\ln(1 + |B_{X_1}|)$  have finite first moment. If  $Z$  has negative vertical drift  $\delta < 0$ , then the projections  $\pi_{\mathbb{H}}(Z_n)$  converge a.s. to a random element  $r \in \mathbb{R} = \partial\mathbb{H} \setminus \{\infty\}$ .*

*Proof.* The proof of Lemma 4.4 can be adapted to show that the imaginary parts  $A_{Z_n}$  converge a.s. to 0, whence we only need to understand the behaviour of the real parts  $B_{Z_n}$ . By the construction of the group homomorphism  $\pi_{\text{Isom}(\mathbb{H})} : G \rightarrow \text{Isom}(\mathbb{H})$ , each isometry  $\pi_{\text{Isom}(\mathbb{H})}(g)$  with  $g \in G$  is of the form  $z \mapsto \alpha z + \beta$  with  $\alpha, \beta \in \mathbb{R}$  and  $\alpha > 0$ . So, the equation  $\pi_{\mathbb{H}}(g) = A_g \cdot i + B_g$  yields  $\pi_{\text{Isom}(\mathbb{H})}(g)(z) = A_g \cdot z + B_g$  and, in light of the multiplication

$$(\pi_{\text{Isom}(\mathbb{H})}(g_1) \circ \pi_{\text{Isom}(\mathbb{H})}(g_2))(z) = A_{g_1} \cdot A_{g_2} \cdot z + A_{g_1} \cdot B_{g_2} + B_{g_1},$$

we obtain

$$\begin{aligned} \pi_{\mathbb{H}}(Z_n) &= \pi_{\text{Isom}(\mathbb{H})}(Z_n)(i) = \pi_{\text{Isom}(\mathbb{H})}(X_1 \cdots X_n)(i) \\ &= (\pi_{\text{Isom}(\mathbb{H})}(X_1) \circ \dots \circ \pi_{\text{Isom}(\mathbb{H})}(X_n))(i) \\ &= A_{X_1} \cdots A_{X_n} \cdot i + \sum_{k=1}^n A_{X_1} \cdots A_{X_{k-1}} \cdot B_{X_k}. \end{aligned}$$

Therefore, the real parts  $B_{Z_n}$  are the partial sums of the series  $\sum_{k=1}^{\infty} C_k$  with  $C_k := A_{X_1} \cdots A_{X_{k-1}} \cdot B_{X_k}$ . In order to verify a.s. convergence of this series, we apply Cauchy’s root test,

$$\begin{aligned} |C_k|^{\frac{1}{k}} &\leq \exp\left(\ln\left(\frac{q}{p}\right) \cdot \underbrace{\frac{\lambda(X_1) + \dots + \lambda(X_{k-1})}{k-1}}_{\rightarrow \mathbb{E}(\lambda(X_1)) = \delta < 0 \text{ a.s.}} \cdot \underbrace{\frac{k-1}{k}}_{\rightarrow 1}\right) \cdot \exp\left(\underbrace{\frac{\ln(1 + |B_{X_k}|)}{k}}_{\rightarrow 0 \text{ a.s.}}\right) \\ &\xrightarrow[n \rightarrow \infty]{\text{a.s.}} \left(\frac{q}{p}\right)^{\delta} < 1. \end{aligned}$$

The convergence claimed in the first factor follows from the strong law of large numbers, the one claimed in the second factor from the Borel–Cantelli lemma. Indeed, let  $Q_k := \frac{1}{k} \cdot \ln(1 + |B_{X_k}|)$ . In order to show that  $Q_k$  converges a.s. to 0,



recall that  $\ln(1 + |B_{X_1}|)$  has finite first moment. For every  $\varepsilon > 0$  we may thus estimate

$$\sum_{k=1}^{\infty} \mathbb{P}(Q_k > \varepsilon) \leq \sum_{k=1}^{\infty} \mathbb{P}\left(\left\lceil \frac{\ln(1 + |B_{X_1}|)}{\varepsilon} \right\rceil \geq k\right) = \mathbb{E}\left(\left\lceil \frac{\ln(1 + |B_{X_1}|)}{\varepsilon} \right\rceil\right).$$

Now, the Borel–Cantelli lemma yields

$$\mathbb{P}(\exists \text{ infinitely many } k \in \mathbb{N} \text{ such that } Q_k > \varepsilon) = 0,$$

from where we may conclude that  $Q_k$  converges a.s. to 0, as claimed above. Therefore,  $\limsup_{k \rightarrow \infty} |C_k|^{\frac{1}{k}} < 1$  a.s., whence  $\sum_{k=1}^{\infty} C_k$  converges a.s. to a random element  $r \in \mathbb{R}$ . ■

What remains is the driftless case. An answer was given by Brofferio in [Bro03, Theorem 1]. It says that if  $\ln(A_{X_1})$  and  $\ln(1 + |B_{X_1}|)$  have finite first moment, then the projections  $\pi_{\mathbb{H}}(Z_n)$  converge a.s. to  $\infty \in \partial\mathbb{H}$ . For us, a result of slightly different flavour will be of relevance.

LEMMA 4.6: *Assume that  $\ln(A_{X_1})$  has finite second moment and there is an  $\varepsilon > 0$  such that  $\ln(1 + |B_{X_1}|)$  has finite  $(2 + \varepsilon)$ -th moment. If  $Z$  has no vertical drift, i.e.,  $\delta = 0$ , then the projections  $\pi_{\mathbb{H}}(Z_n)$  have sublinear speed, i.e.,*

$$\frac{d_{\mathbb{H}}(\pi_{\mathbb{H}}(Z_0), \pi_{\mathbb{H}}(Z_n))}{n} \xrightarrow[n \rightarrow \infty]{\text{a.s.}} 0.$$

The proof is based on ideas that go back to Élie [Éli82, Lemme 5.49] and have also been used in [CKW94, Proposition 4b]. We first adapt these ideas to our situation in Lemma 4.7 and then deduce Lemma 4.6. By assumption, there is no vertical drift so that the pointwise projection  $\lambda(Z) = (\lambda(Z_0), \lambda(Z_1), \dots)$  to  $\mathbb{Z}$  is recurrent. In particular, we know that there exists a.s. a strictly increasing sequence  $\tau(0), \tau(1), \dots$  given by  $\tau(0) := 0$  and

$$\tau(n) := \inf \{k \in \mathbb{N} : k > \tau(n-1) \text{ and } \lambda(Z_k) > \lambda(Z_{\tau(n-1)})\}$$

for all  $n \in \mathbb{N}$ . We call  $\tau(n)$  the  $n$ -th **ladder time**, see Figure 4 for an illustration, and write  $\tau := \tau(1)$  for short.

LEMMA 4.7: *Under the same assumptions as in Lemma 4.6, the random variable*

$$\ln \left( 1 + \sum_{k=1}^{\tau} |B_{X_k}| \right)$$

*has finite first moment.*

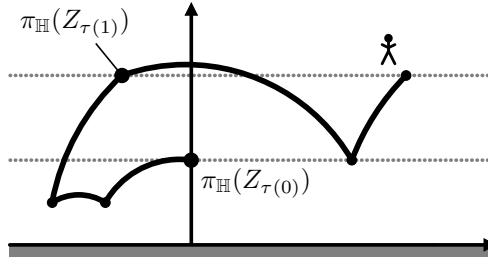


Figure 4. The first ladder times  $\tau(0)$  and  $\tau(1)$ .

*Proof.* We adapt the proof of [Éli82, Lemme 5.49] to our situation. Pick an  $\varepsilon > 0$  that satisfies the requirements of Lemmas 4.6 and 4.7 and let

$$\beta := \frac{1}{2 + \varepsilon}.$$

Since  $\ln(A_{X_1})$  has finite second moment, we know that  $\lambda(X_1)$  has finite second moment and therefore  $\mathbb{P}(\tau > k)$  is asymptotically equivalent to  $\text{const} \cdot k^{-\frac{1}{2}}$  with a strictly positive constant; see [Éli82, §5.44] referring to [Fel71, p. 415]. In other words, the quotient of  $\mathbb{P}(\tau > k)$  and  $\text{const} \cdot k^{-\frac{1}{2}}$  converges to 1. Thus

$$\int \tau^\beta \, d\mathbb{P} \leq \int \lceil \tau^\beta \rceil \, d\mathbb{P} = \sum_{k=1}^{\infty} \mathbb{P}(\lceil \tau^\beta \rceil \geq k) = \sum_{k=0}^{\infty} \underbrace{\mathbb{P}(\tau > k^{\frac{1}{\beta}})}_{\sim \text{const} \cdot k^{-(1+\frac{\varepsilon}{2})}}.$$

In particular, there is a  $k_0 \in \mathbb{N}$  such that for all  $k \geq k_0$  the inequality

$$\mathbb{P}(\tau > k^{\frac{1}{\beta}}) < k^{-(1+\frac{\varepsilon}{4})}$$

holds. Since  $\sum_{k=k_0}^{\infty} k^{-(1+\frac{\varepsilon}{4})}$  is finite, we know that  $\int \tau^\beta \, d\mathbb{P}$  is finite. By construction, the increments  $\tau(1) - \tau(0), \tau(2) - \tau(1), \dots$  are i.i.d. (= independent and identically distributed), whence the fact that  $0 < \beta < 1$ , which implies that  $(x + y)^\beta \leq x^\beta + y^\beta$ , and the strong law of large numbers yield

$$\begin{aligned}
 (*) \quad \frac{\tau(n)^\beta}{n} &\leq \frac{(\tau(1) - \tau(0))^\beta + \dots + (\tau(n) - \tau(n-1))^\beta}{n} \xrightarrow[\text{a.s.}]{n \rightarrow \infty} \mathbb{E}(\tau^\beta), \\
 &\implies \limsup_{n \rightarrow \infty} \frac{\tau(n)^\beta}{n} < \infty \quad \text{a.s.}
 \end{aligned}$$

Now, we are prepared for the main argument. The sums

$$\sum_{k=\tau(0)+1}^{\tau(1)} |B_{X_k}|, \quad \sum_{k=\tau(1)+1}^{\tau(2)} |B_{X_k}|, \dots$$

are i.i.d., they are non-negative and not a.s. equal to zero. Hence, in view of [Éli82, Lemme 5.23], the following equivalence holds:

$$\int \ln \left( 1 + \sum_{k=1}^{\tau} |B_{X_k}| \right) d\mathbb{P} < \infty \iff \underbrace{\limsup_{n \rightarrow \infty} \left( \sum_{k=\tau(n-1)+1}^{\tau(n)} |B_{X_k}| \right)^{\frac{1}{n}}}_{=: K} < \infty \quad \text{a.s.}$$

It thus suffices to verify the right-hand side. In order to do so, we would like to estimate

$$\begin{aligned} K &\leq \limsup_{n \rightarrow \infty} \exp \left( \frac{\ln(1 + \sum_{k=1}^{\tau(n)} |B_{X_k}|)}{n} \right) \\ &\leq \exp \left( \underbrace{\limsup_{n \rightarrow \infty} \frac{\ln(1 + \sum_{k=1}^{\tau(n)} |B_{X_k}|)}{\tau(n)^\beta}_{=: L}} \cdot \underbrace{\limsup_{n \rightarrow \infty} \frac{\tau(n)^\beta}{n}}_{< \infty \text{ a.s. } (*)} \right). \end{aligned}$$

A priori, it might be the case that  $L$  is infinite and the second factor in the rightmost term is 0, in which case the product would not make any sense. We claim that  $L$  is a.s. finite, which does not only legitimate the above estimate but also completes the proof. Indeed, observe that

$$\begin{aligned} L &\leq \limsup_{n \rightarrow \infty} \frac{\ln(1 + \tau(n) \cdot \max\{|B_{X_k}| : 1 \leq k \leq \tau(n)\})}{\tau(n)^\beta} \\ &\leq \underbrace{\limsup_{n \rightarrow \infty} \frac{\ln(\tau(n))}{\tau(n)^\beta}}_{=0} + \limsup_{n \rightarrow \infty} \frac{\ln(1 + \max\{|B_{X_k}| : 1 \leq k \leq \tau(n)\})}{\tau(n)^\beta} \\ &= \limsup_{n \rightarrow \infty} \left( \frac{\max\{\ln(1 + |B_{X_k}|)^{\frac{1}{\beta}} : 1 \leq k \leq \tau(n)\}}{\tau(n)} \right)^\beta \\ &\leq \limsup_{n \rightarrow \infty} \left( \underbrace{\frac{\sum_{k=1}^{\tau(n)} \ln(1 + |B_{X_k}|)^{\frac{1}{\beta}}}{\tau(n)}}_{=: M_n} \right)^\beta. \end{aligned}$$

Now, recall that  $\frac{1}{\beta} = 2 + \varepsilon$ . By the strong law of large numbers,  $M_n$  converges a.s. to  $\mathbb{E}(\ln(1 + |B_{X_1}|)^{\frac{1}{\beta}})$ . This implies that  $\limsup_{n \rightarrow \infty} M_n^\beta$  is a.s. finite, and so is  $L$ .    ■

*Proof of Lemma 4.6.* Recall from the proof of Lemma 4.7, that  $\mathbb{P}(\tau > k)$  is asymptotically equivalent to  $\text{const} \cdot k^{-\frac{1}{2}}$  with a strictly positive constant. In particular, there is a  $k_0 \in \mathbb{N}$  such that for all  $k \geq k_0$  the inequality  $\mathbb{P}(\tau > k) > k^{-1}$  holds, whence

$$\int \tau \, d\mathbb{P} = \sum_{k=1}^{\infty} \mathbb{P}(\tau \geq k) = \sum_{k=0}^{\infty} \mathbb{P}(\tau > k) \geq \sum_{k=k_0}^{\infty} k^{-1} = \infty.$$

Since  $\tau(1) - \tau(0), \tau(2) - \tau(1), \dots$  are i.i.d. and non-negative, we may deduce from the strong law of large numbers by truncating the random variables (see, e.g., [Rou14, p. 309, Lemma 6]) that

$$\frac{\tau(n)}{n} = \frac{(\tau(1) - \tau(0)) + (\tau(2) - \tau(1)) + \dots + (\tau(n) - \tau(n-1))}{n} \xrightarrow[n \rightarrow \infty]{\text{a.s.}} \infty$$

and

$$(**) \quad \frac{n}{\tau(n)} \xrightarrow[n \rightarrow \infty]{\text{a.s.}} 0.$$

This convergence can be used to estimate the distance between  $\pi_{\mathbb{H}}(Z_0)$  and  $\pi_{\mathbb{H}}(Z_n)$  from above. Indeed, for every  $n \in \mathbb{N}_0$  let  $m = m(n) \in \mathbb{N}_0$  be the unique element with

$$\tau(m) \leq n < \tau(m+1).$$

It exists a.s. because the ladder times

$$0 = \tau(0) < \tau(1) < \dots$$

do. Now, observe that

$$\begin{aligned} \frac{d_{\mathbb{H}}(\pi_{\mathbb{H}}(Z_0), \pi_{\mathbb{H}}(Z_n))}{n} &\leq \underbrace{\frac{d_{\mathbb{H}}(i, A_{Z_{\tau(m)}} \cdot i)}{n}}_{\textcircled{1}} \\ &+ \underbrace{\frac{d_{\mathbb{H}}(A_{Z_{\tau(m)}} \cdot i, A_{Z_{\tau(m)}} \cdot i + B_{Z_n})}{n}}_{\textcircled{2}} \\ &+ \underbrace{\frac{d_{\mathbb{H}}(A_{Z_{\tau(m)}} \cdot i + B_{Z_n}, A_{Z_n} \cdot i + B_{Z_n})}{n}}_{\textcircled{3}}. \end{aligned}$$

The meaning of the three summands is illustrated in Figure 5. We will consider them separately and show that each of them converges a.s. to 0. For  $\textcircled{1}$  and  $\textcircled{3}$ ,

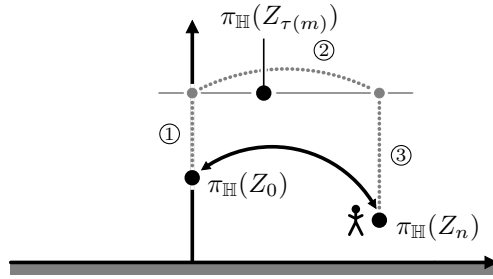


Figure 5. Estimation of the distance between  $\pi_{\mathbb{H}}(Z_0)$  and  $\pi_{\mathbb{H}}(Z_n)$ .

this is straightforward. Indeed,

$$\begin{aligned} \textcircled{1} = \frac{|\ln(A_{Z_{\tau(m)}})|}{n} &\leq \frac{|\ln(A_{Z_{\tau(m)}})|}{\tau(m)} = \ln\left(\frac{q}{p}\right) \cdot \left| \frac{\lambda(X_1) + \dots + \lambda(X_{\tau(m)})}{\tau(m)} \right| \\ &\xrightarrow[\text{a.s.}]{n \rightarrow \infty} \ln\left(\frac{q}{p}\right) \cdot |\mathbb{E}(\lambda(X_1))| = \ln\left(\frac{q}{p}\right) \cdot |\delta| = 0 \end{aligned}$$

and similarly

$$\begin{aligned} \textcircled{3} &\leq \frac{d_{\mathbb{H}}(A_{Z_{\tau(m)}} \cdot i, i)}{n} + \frac{d_{\mathbb{H}}(i, A_{Z_n} \cdot i)}{n} = \textcircled{1} + \frac{|\ln(A_{Z_n})|}{n} \\ &= \textcircled{1} + \ln\left(\frac{q}{p}\right) \cdot \left| \frac{\lambda(X_1) + \dots + \lambda(X_n)}{n} \right| \xrightarrow[\text{a.s.}]{n \rightarrow \infty} 0. \end{aligned}$$

For  $\textcircled{2}$ , recall from the proof of Lemma 4.5 that

$$B_{Z_n} = \sum_{k=1}^n A_{X_1} \cdots A_{X_{k-1}} \cdot B_{X_k}$$

and observe that for all  $\ell, m \in \mathbb{N}_0$  with  $\tau(m) \leq \ell \leq \tau(m+1)$  the following holds:

$$\begin{aligned} \frac{|B_{Z_\ell} - B_{Z_{\tau(m)}}|}{A_{Z_{\tau(m)}}} &\leq \frac{A_{X_1} \cdots A_{X_{\tau(m)}} \cdot \sum_{k=\tau(m)+1}^{\ell} A_{X_{\tau(m)+1}} \cdots A_{X_{k-1}} \cdot |B_{X_k}|}{A_{X_1} \cdots A_{X_{\tau(m)}}} \\ \textcircled{\dagger} \quad &= \sum_{k=\tau(m)+1}^{\ell} \underbrace{A_{X_{\tau(m)+1}} \cdots A_{X_{k-1}}}_{\leq 1} \cdot |B_{X_k}| \\ &\leq \sum_{k=\tau(m)+1}^{\ell} |B_{X_k}| \quad \text{a.s.} \end{aligned}$$

Hence, using that  $A_{Z_{\tau(0)}} < A_{Z_{\tau(1)}} < \dots < A_{Z_{\tau(m)}}$ , we obtain

$$\begin{aligned} \textcircled{2} &= \frac{1}{n} \cdot \operatorname{arcosh} \left( 1 + \frac{1}{2} \cdot \left( \frac{|B_{Z_n}|}{A_{Z_{\tau(m)}}} \right)^2 \right) \\ &= \frac{1}{n} \cdot \ln \left( 1 + \frac{1}{2} \cdot \left( \frac{|B_{Z_n}|}{A_{Z_{\tau(m)}}} \right)^2 + \sqrt{\left( 1 + \frac{1}{2} \cdot \left( \frac{|B_{Z_n}|}{A_{Z_{\tau(m)}}} \right)^2 - 1 \right)} \right) \\ &\leq \frac{1}{n} \cdot \left( \ln(2) + \ln \left( 1 + \left( \frac{|B_{Z_n}|}{A_{Z_{\tau(m)}}} \right)^2 \right) \right) \\ &\leq \frac{1}{n} \cdot \left( \ln(2) + 2 \cdot \ln \left( 1 + \frac{|B_{Z_n}|}{A_{Z_{\tau(m)}}} \right) \right) \\ &\leq \frac{1}{n} \cdot \left( \ln(2) + 2 \cdot \ln \left( 1 + \frac{|B_{Z_{\tau(1)}} - B_{Z_{\tau(0)}}|}{A_{Z_{\tau(0)}}} + \frac{|B_{Z_{\tau(2)}} - B_{Z_{\tau(1)}}|}{A_{Z_{\tau(1)}}} \right. \right. \\ &\qquad \qquad \qquad \left. \left. + \dots + \frac{|B_{Z_{\tau(m)}} - B_{Z_{\tau(m-1)}}|}{A_{Z_{\tau(m-1)}}} + \frac{|B_{Z_n} - B_{Z_{\tau(m)}}|}{A_{Z_{\tau(m)}}} \right) \right), \end{aligned}$$

which allows us to use that  $\tau(m) \leq n < \tau(m + 1)$  together with  $(\dagger)$ , and then by applying  $(**)$  we obtain

$$\begin{aligned} \dots &\leq \frac{1}{n} \cdot \left( \ln(2) + 2 \cdot \ln \left( 1 + \sum_{k=1}^n |B_{X_k}| \right) \right) \\ &\leq \frac{1}{n} \cdot \left( \ln(2) + 2 \cdot \ln \left( 1 + \sum_{k=1}^{\tau(m+1)} |B_{X_k}| \right) \right) \\ &\leq \underbrace{\frac{\ln(2)}{n}}_{\rightarrow 0} + 2 \cdot \underbrace{\frac{\ln(1 + \sum_{k=\tau(0)+1}^{\tau(1)} |B_{X_k}|) + \dots + \ln(1 + \sum_{k=\tau(m)+1}^{\tau(m+1)} |B_{X_k}|)}{m+1}}_{\rightarrow \mathbb{E}(\ln(1 + \sum_{k=1}^{\tau} |B_{X_k}|)) \text{ a.s. by Lemma 4.7}} \\ &\qquad \cdot \underbrace{\frac{m+1}{\tau(m)}}_{\rightarrow 0 \text{ a.s.}} \cdot \underbrace{\frac{\tau(m)}{n}}_{\leq 1 \text{ a.s.}} \xrightarrow[n \rightarrow \infty]{\text{a.s.}} 0. \quad \blacksquare \end{aligned}$$

4.2. CONVERGENCE TO THE SPACE OF ENDS OF THE BASS–SERRE TREE. Even though the projections  $\pi_T(Z_n)$  do not need to satisfy the Markov property, we are still able to show that they converge a.s. to a random end by applying a result of Cartwright and Soardi, [CS89, p. 820, Theorem], which is based on a technique developed by Furstenberg in [Fur63] and [Fur71]. The authors consider a random walk  $\Phi = (\Phi_0, \Phi_1, \dots)$  on the automorphism group of a locally

finite and infinite tree and prove under a mild assumption on the probability measure that the sequence of vertices obtained by evaluating each automorphism  $\Phi_n$  at a fixed vertex  $v$  converges a.s. to a random end. Their assumption is that the random walk is driven by a regular Borel probability measure whose support is not contained in any amenable subgroup. However, the proof of [CS89, p. 820, Theorem] shows that it suffices to assume that the support is not contained in any amenable closed subgroup. Given that  $1 < p < q$ , this result can be immediately applied to our setting.

LEMMA 4.8: *Let  $1 < p < q$ . Then, the projections  $\pi_T(Z_n)$  converge a.s. to a random end  $\xi \in \partial T$ .*

*Proof.* Since the group  $G$  acts on the tree  $T$ , we may consider the group homomorphism  $\varphi : G \rightarrow \text{Aut}(T)$  associated to this action. The automorphism group  $\text{Aut}(T)$  is endowed with the topology of pointwise convergence. Since  $G$  is discrete,  $\varphi$  is certainly measurable. The pointwise images

$$(\varphi(Z_0), \varphi(Z_1), \dots)$$

constitute a random walk on  $\text{Aut}(T)$  that satisfies the assumption of [CS89, p. 820, Theorem]. Indeed, the random walk on  $\text{Aut}(T)$  is driven by the push-forward Borel probability measure  $\varphi_*(\mu)$ . Because  $\text{Aut}(T)$  is a locally compact Hausdorff space with a countable base (see, e.g., [Woe91, §2]) every Borel probability measure on  $\text{Aut}(T)$  is regular (see, e.g., [Coh13, Proposition 7.2.3]) and so is  $\varphi_*(\mu)$ . It remains to show that the support of the latter is not contained in any amenable closed subgroup.

Observe that the support of  $\varphi_*(\mu)$  generates the subgroup

$$\varphi(G) \leq \text{Aut}(T).$$

Since  $\varphi(G)$  acts transitively on  $T$  and does not fix an end, every closed subgroup that contains  $\varphi(G)$  has these two properties as well and is therefore not amenable; see [Neb88, Theorem 2]. In other words, the support of  $\varphi_*(\mu)$  is not contained in any amenable closed subgroup of  $\text{Aut}(T)$ . Now, [CS89, p. 820, Theorem] yields that the sequence obtained by evaluating each automorphism  $\varphi(Z_n)$  at a fixed vertex  $v$  converges a.s. to a random end  $\xi \in \partial T$ . Since

$$\pi_T(Z_n) = Z_n B = \varphi(Z_n)(B),$$

setting  $v := B$  completes the proof.    ■

4.3. CONSTRUCTION OF  $\mu$ -BOUNDARIES. Resuming Sections 4.1 and 4.2, we may formulate the following theorem.

**THEOREM 4.9** (“convergence theorem”): *Let  $Z = (Z_0, Z_1, \dots)$  be a random walk on a non-amenable Baumslag–Solitar group  $G = \text{BS}(p, q)$  with  $1 < p < q$  and increments  $X_1, X_2, \dots$  of finite first moment. Then, the projections  $\pi_T(Z_n)$  converge a.s. to a random end  $\xi \in \partial T$ . Moreover, depending on the vertical drift  $\delta$ , we distinguish three cases:*

- (1) *If  $\delta > 0$ , then the projections  $\pi_{\mathbb{H}}(Z_n)$  converge a.s. to  $\infty \in \partial \mathbb{H}$ .*
- (2) *If  $\delta < 0$ , then the projections  $\pi_{\mathbb{H}}(Z_n)$  converge a.s. to a random element  $r \in \mathbb{R} = \partial \mathbb{H} \setminus \{\infty\}$ .*
- (3) *If  $\delta = 0$  and  $\ln(A_{X_1})$  has finite second moment and there is an  $\varepsilon > 0$  such that  $\ln(1 + |B_{X_1}|)$  has finite  $(2 + \varepsilon)$ -th moment, then projections  $\pi_{\mathbb{H}}(Z_n)$  have sublinear speed.*

So, let us assume that  $1 < p < q$  and that the increments  $X_1, X_2, \dots$  have finite first moment. We may therefore consider the map  $\text{bnd}_{\partial T} : \Omega \rightarrow \partial T$  and, in the special case that  $\delta < 0$ , also the map  $\text{bnd}_{\mathbb{R}} : \Omega \rightarrow \mathbb{R}$ , defined almost everywhere, assigning to a trajectory  $\omega = (x_0, x_1, \dots) \in \Omega$  the limit

$$\text{bnd}_{\partial T}(\omega) := \lim_{n \rightarrow \infty} \pi_T(x_n) \in \partial T \quad \text{and} \quad \text{bnd}_{\mathbb{R}}(\omega) := \lim_{n \rightarrow \infty} \pi_{\mathbb{H}}(x_n) \in \mathbb{R}.$$

The topological spaces  $\partial T$  and  $\mathbb{R}$  are endowed with their Borel  $\sigma$ -algebras  $\mathcal{B}_{\partial T}$  and  $\mathcal{B}_{\mathbb{R}}$ . Even though the maps  $\text{bnd}_{\partial T}$  and  $\text{bnd}_{\mathbb{R}}$  are only defined almost everywhere, they are measurable in the sense that the preimages of measurable sets are measurable. Given  $\text{bnd}_{\partial T}$  and  $\text{bnd}_{\mathbb{R}}$ , we may construct their product map  $\text{bnd}_{\partial T \times \mathbb{R}} : \Omega \rightarrow \partial T \times \mathbb{R}$ . It is measurable with respect to the product  $\sigma$ -algebra  $\mathcal{B}_{\partial T} \otimes \mathcal{B}_{\mathbb{R}}$ . Because both  $\partial T$  and  $\mathbb{R}$  are metrisable and separable topological spaces, it is not hard to see that the product  $\sigma$ -algebra  $\mathcal{B}_{\partial T} \otimes \mathcal{B}_{\mathbb{R}}$  agrees with the Borel  $\sigma$ -algebra  $\mathcal{B}_{\partial T \times \mathbb{R}}$ ; see, e.g., [Bil99, Appendix M.10]. The pushforward probability measures

$$\nu_{\partial T} := (\text{bnd}_{\partial T})_*(\mathbb{P}) \quad \text{and} \quad \nu_{\partial T \times \mathbb{R}} := (\text{bnd}_{\partial T \times \mathbb{R}})_*(\mathbb{P})$$

on the respective measurable spaces are called the **hitting measures**. Since  $\partial T$  and  $\mathbb{R}$ , and therefore also  $\partial T \times \mathbb{R}$ , are Polish spaces, Example 3.3 implies that  $(\partial T, \mathcal{B}_{\partial T}, \nu_{\partial T})$  and  $(\partial T \times \mathbb{R}, \mathcal{B}_{\partial T \times \mathbb{R}}, \nu_{\partial T \times \mathbb{R}})$  are Lebesgue–Rohlin spaces. The maps  $\text{bnd}_{\partial T}$  and  $\text{bnd}_{\partial T \times \mathbb{R}}$  are homomorphisms and, by construction, they are  $\textcircled{1}$   $\sim$ -invariant.



Each of the topological spaces  $\partial T$  and  $\mathbb{R}$  is endowed with a continuous  $G$ -action. The one on  $\partial T$  is induced by the left-multiplication  $g(hB) := (gh)B$  on  $T$ . More precisely, recall that ends are infinite reduced paths that start in  $B$ . The pointwise left-multiplication maps every such path  $\xi \in \partial T$  to some other path that need not start in  $B$  anymore. The end  $g\xi \in \partial T$  is obtained by connecting  $B$  to the initial vertex of this path and reducing the concatenation if necessary. The  $G$ -action on  $\mathbb{R}$  is induced by the isometric  $G$ -action on  $\mathbb{H}$  that we addressed in Section 2.3. In light of the representation of the elements of  $\text{Isom}(\mathbb{H})$  as rational functions, we can also evaluate them on the boundary  $\partial\mathbb{H}$  and finally observe that the isometries associated to the elements of  $G$  leave the subset  $\mathbb{R} \subseteq \partial\mathbb{H}$  invariant. The  $G$ -actions on  $\partial T$  and  $\mathbb{R}$  induce a componentwise  $G$ -action on the product  $\partial T \times \mathbb{R}$ , which is also continuous.

All three  $G$ -actions are measurable with respect to the Borel  $\sigma$ -algebras and, since they map null sets  $A$  to null sets  $gA$ , they remain measurable when we proceed to the completions. In particular, the spaces  $(\partial T, \mathcal{B}_{\partial T}, \nu_{\partial T})$  and  $(\partial T \times \mathbb{R}, \mathcal{B}_{\partial T \times \mathbb{R}}, \nu_{\partial T \times \mathbb{R}})$  are endowed with measurable  $G$ -actions and, by construction, the maps  $\text{bnd}_{\partial T}$  and  $\text{bnd}_{\partial T \times \mathbb{R}}$  are  $\textcircled{2}$   $G$ -equivariant. We have thus derived the following lemma.

LEMMA 4.10: *For any vertical drift  $\delta$ , in particular for  $\delta \geq 0$ , the Lebesgue–Rohlin space  $(\partial T, \mathcal{B}_{\partial T}, \nu_{\partial T})$  endowed with the homomorphism  $\text{bnd}_{\partial T} : \Omega \rightarrow \partial T$  is a  $\mu$ -boundary. If  $\delta < 0$ , then  $(\partial T \times \mathbb{R}, \mathcal{B}_{\partial T \times \mathbb{R}}, \nu_{\partial T \times \mathbb{R}})$  endowed with  $\text{bnd}_{\partial T \times \mathbb{R}} : \Omega \rightarrow \partial T \times \mathbb{R}$  is also a  $\mu$ -boundary.*

Before we will use Kaimanovich’s strip criterion to show that the above  $\mu$ -boundaries are maximal, we analyse the hitting measures. This requires a preliminary observation.

LEMMA 4.11: *The  $G$ -actions on  $\partial T$  and  $\mathbb{R}$ , as well as the componentwise  $G$ -action on the product  $\partial T \times \mathbb{R}$ , are topologically minimal, i.e., each orbit is dense. Because all three spaces are infinite and Hausdorff, this implies that each orbit is infinite.*

*Proof.* Consider the  $G$ -action on  $\partial T$ . Choose an end  $\xi \in \partial T$  and a non-empty open subset  $S \subseteq \partial T$ . We shall construct an element  $g \in G$  such that  $g\xi \in S$ . Because  $S$  is non-empty and open, all ends with a certain finite initial piece belong to  $S$ . In other words, there is an element  $h \in G$  such that all ends that start in  $B$  and traverse the vertex  $hB$  are contained in  $S$ . If we set  $g := h \in G$ ,

then the end  $g\xi \in \partial T$  will have the correct finite initial piece unless cancellation takes place. In the latter case, we set  $g := hb \in G$  instead. Since  $|p| \neq 1$  and  $|q| \neq 1$ , cancellation will take place in at most one of the two cases, which proves the first assertion.

Next, consider the  $G$ -action on  $\mathbb{R}$ , an element  $r \in \mathbb{R}$ , and a non-empty open subset  $S \subseteq \mathbb{R}$ . Because  $S$  is non-empty and open, there are  $s \in \mathbb{R}$  and  $\varepsilon > 0$  such that the interval  $(s - \varepsilon, s + \varepsilon)$  is contained in  $S$ . We assume that  $1 < p < q$ , so we can find integers  $k_1, k_2 \in \mathbb{Z}$  with  $k_1 < 0$  such that  $q^{k_1}p^{-k_1} < \varepsilon$  and the elements  $q^{k_1}p^{-k_1}(r + k_2)$  and  $q^{k_1}p^{-k_1}(r + k_2 + 1)$  are both contained in  $(s - \varepsilon, s + \varepsilon)$ . Therefore, we set  $g := a^{k_1}b^{k_2} \in G$  to obtain

$$gr = a^{k_1}b^{k_2}r = a^{k_1}(r + k_2) = q^{k_1}p^{-k_1}(r + k_2) \in (s - \varepsilon, s + \varepsilon) \subseteq S.$$

Finally, consider the  $G$ -action on the product  $\partial T \times \mathbb{R}$ , an element  $(\xi, r) \in \partial T \times \mathbb{R}$ , and a non-empty open subset  $S \subseteq \partial T \times \mathbb{R}$ . Because  $S$  is non-empty and open, there are non-empty and open subsets  $S_1 \subseteq \partial T$  and  $S_2 \subseteq \mathbb{R}$  such that  $S_1 \times S_2$  is contained in  $S$ . We shall now construct an element  $g \in G$  such that both  $g\xi \in S_1$  and  $gr \in S_2$ . Look at the tree component first. We already know that there is an element  $h \in G$  such that all ends that start in  $B$  and traverse the vertex  $hB$  are contained in  $S_1$ . Let  $k_0 \in \{0, 1\}$ , whichever ensures that the reduced path from  $B$  to the vertex  $hb^{k_0}a^{-1}B$  traverses the vertex  $hB$ . Now, look at the real component. We can find integers  $k_1, k_2 \in \mathbb{Z}$  with  $k_1 < 0$  such that the elements  $a^{k_1}b^{k_2}r$  and  $a^{k_1}b^{k_2+1}r$  are both contained in  $(hb^{k_0})^{-1}S_2$ . Back to the tree component, we choose  $k_3 \in \{0, 1\}$  such that the end  $hb^{k_0}a^{k_1}b^{k_2+k_3}\xi$  traverses the vertex  $hb^{k_0}a^{k_1}B$ . Then, by construction, it also traverses the vertex  $hB$ . We set  $g := hb^{k_0}a^{k_1}b^{k_2+k_3} \in G$  to obtain  $g\xi \in S_1$  and  $gr \in S_2$ . ■

Given Lemma 4.11 and the  $\mu$ -stationarity and quasi-invariance of the hitting measures, it is well-known that the latter are non-atomic and have full support. Indeed, if there were atoms, then we could choose an atom  $\xi$  of maximal mass. Because the respective hitting measure  $\nu$  is  $\mu$ -stationary, the value  $\nu(\xi)$  is a convex combination of all values  $\nu(g^{-1}\xi)$  with  $g \in \text{supp}(\mu)$ . Therefore, each  $\nu(g^{-1}\xi)$  must be equal to  $\nu(\xi)$ . Iteration of this procedure yields that the equality does not only hold for every  $g \in \text{supp}(\mu)$  but also for every  $g$  in the semigroup generated by  $\text{supp}(\mu)$ , i.e., for every  $g \in G$ . Since the orbit  $G\xi$  is infinite, this contradicts the finiteness of the hitting measure  $\nu$ . Concerning the

assertion of full support, if there was a non-empty open null set  $S$ , then the topological minimality of the  $G$ -action would imply that the translates  $gS$  with  $g \in G$  form a countable covering of the whole space with null sets, which is a contradiction. For further details see, e.g., [Woe89, Lemma 3.4] and [MNS17, Lemma 2.2 and 2.3].

4.4. PROOF OF THE MAIN RESULT. We are now ready to prove our main result, Theorem 1.1 announced in the introduction. It identifies the Poisson–Furstenberg boundary of random walks  $Z = (Z_0, Z_1, \dots)$  on non-amenable Baumslag–Solitar groups

$$G = \text{BS}(p, q)$$

with  $1 < p < q$  and increments  $X_1, X_2, \dots$  of finite first moment.

*Proof of Theorem 1.1.* We seek to apply the strip criterion, Theorem 3.5. By Lemma 3.7, the probability measure  $\mu$  driving the random walk has finite entropy. By Lemma 4.10, the Lebesgue–Rohlin space  $(\partial T, \mathcal{B}_{\partial T}, \nu_{\partial T})$  is a  $\mu$ -boundary. If  $Z$  has negative vertical drift  $\delta < 0$ , then  $(\partial T \times \mathbb{R}, \mathcal{B}_{\partial T \times \mathbb{R}}, \nu_{\partial T \times \mathbb{R}})$  is also a  $\mu$ -boundary. Let us consider the case  $\delta < 0$  first. We thus take the  $\mu$ -boundary  $(\partial T \times \mathbb{R}, \mathcal{B}_{\partial T \times \mathbb{R}}, \nu_{\partial T \times \mathbb{R}})$  and the  $\check{\mu}$ -boundary  $(\partial T, \mathcal{B}_{\partial T}, \check{\nu}_{\partial T})$ . Here,  $\check{\nu}_{\partial T}$  denotes the hitting measure of the pointwise projection of the random walk  $\check{Z} = (\check{Z}_0, \check{Z}_1, \dots)$  driven by the reflected probability measure  $\check{\mu}$  to the tree  $T$ .

Next, we define gauges and strips. Let  $S := \{a, b\} \subseteq G$  be the standard generating set and define gauges  $\mathcal{G}_k := \{g \in G : d_S(1, g) \leq k\}$ , i.e., the gauges exhaust the group  $G$  with balls centred at the identity element  $1 \in G$  and the gauge function  $|\cdot| = |\cdot|_{\mathcal{G}}$  is essentially the distance to 1 with respect to the word metric  $d_S$ .

Since the hitting measures are non-atomic, see Section 4.3, we know that  $\check{\nu}_{\partial T} \otimes \nu_{\partial T \times \mathbb{R}}$ -almost every pair of points  $(\xi_-, (\xi_+, r_+)) \in \partial T \times (\partial T \times \mathbb{R})$  has distinct ends  $\xi_-, \xi_+ \in \partial T$ . In this situation, we may connect  $\xi_-$  and  $\xi_+$  by a doubly infinite reduced path  $v : \mathbb{Z} \rightarrow T$  and define the strip  $\mathcal{S}(\xi_-, (\xi_+, r_+))$  as follows. It consists of all  $g \in G$  in the full  $\pi_T$ -preimage of the path, i.e., the image  $\pi_T(g)$  is contained in  $v(\mathbb{Z})$ , with the additional property that the real part  $\text{Re}(\pi_{\mathbb{H}}(g))$  has minimal distance to  $r_+ \in \mathbb{R}$  among all real parts  $\text{Re}(\pi_{\mathbb{H}}(h))$  with  $h \in gB$ , see the left-hand side of Figure 6. To all remaining pairs we assign the whole of  $G$  as a strip. This way, the map  $\mathcal{S}$  becomes measurable and  $G$ -equivariant. Since the hitting measures have full support, see Section 4.3

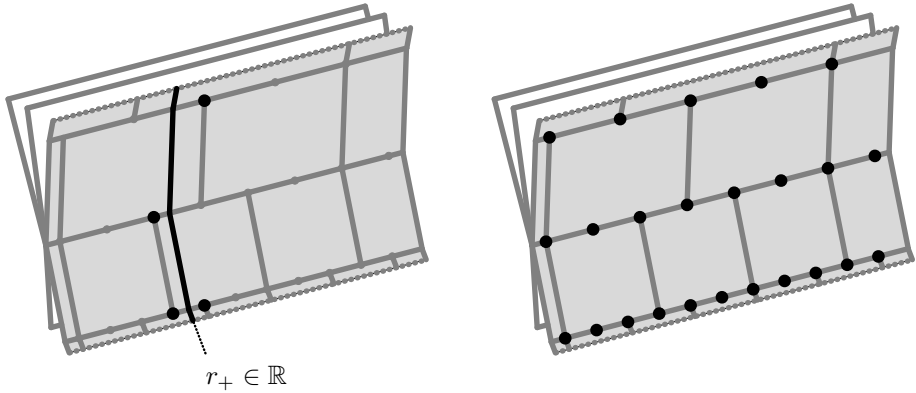


Figure 6. Strips for the cases  $\delta \neq 0$  (left) and  $\delta = 0$  (right).

again, it is not hard to see that a random strip contains the identity element  $1 \in G$  with positive probability, i.e., the map  $\mathcal{S}$  satisfies the inequality of Remark 3.6. So, it suffices to verify the following convergence for an arbitrary pair  $(\xi_-, (\xi_+, r_+)) \in \partial T \times (\partial T \times \mathbb{R})$  with distinct ends  $\xi_-, \xi_+ \in \partial T$ ,

$$\frac{1}{n} \cdot \ln(\text{card}(\mathcal{S}(\xi_-, (\xi_+, r_+)) \cap \mathcal{G}_{|Z_n|})) \xrightarrow[n \rightarrow \infty]{\text{a.s.}} 0.$$

The strip  $\mathcal{S}(\xi_-, (\xi_+, r_+))$  intersects the gauge  $\mathcal{G}_{|Z_n|}$  in at most  $2 \cdot |Z_n| + 1$  many cosets from  $G/B$  and each of them contains at most two elements of the strip. Therefore,

$$\begin{aligned} \frac{1}{n} \cdot \ln(\text{card}(\mathcal{S}(\xi_-, (\xi_+, r_+)) \cap \mathcal{G}_{|Z_n|})) &\leq \frac{\ln((2 \cdot |Z_n| + 1) \cdot 2)}{n} \\ &= \frac{\ln((2 \cdot d_S(1, Z_n) + 1) \cdot 2)}{n} \xrightarrow[n \rightarrow \infty]{\text{a.s.}} 0. \end{aligned}$$

In the final step above, we used that  $X_1$  has finite first moment. Indeed,

$$\begin{aligned} \frac{1}{n} \cdot d_S(1, Z_n) &= \frac{1}{n} \cdot d_S(1, X_1 \cdots X_n) \\ &\leq \frac{1}{n} \cdot \sum_{k=1}^n d_S(1, X_k) \xrightarrow[n \rightarrow \infty]{\text{a.s.}} \mathbb{E}(d_S(1, X_1)), \end{aligned}$$

from where we may first conclude that the sequence  $\frac{1}{n} \cdot d_S(1, Z_n)$  is a.s. bounded and second that the sequence  $\frac{1}{n} \cdot \ln((2 \cdot d_S(1, Z_n) + 1) \cdot 2)$  converges a.s. to 0.

So, we can finally apply the strip criterion and obtain that

$$(\partial T \times \mathbb{R}, \mathcal{B}_{\partial T \times \mathbb{R}}, \nu_{\partial T \times \mathbb{R}})$$

is isomorphic to the Poisson–Furstenberg boundary. Vice versa, if  $Z$  has positive vertical drift  $\delta > 0$ , then the same argument yields that  $(\partial T, \mathcal{B}_{\partial T}, \nu_{\partial T})$  is isomorphic to the Poisson–Furstenberg boundary.

It remains to consider the driftless case, i.e.,  $\delta = 0$ . Then, both  $\mu$  and  $\check{\mu}$  are driftless and there is no natural candidate for a real number that determines the horizontal position of the strip. But the fact that the projections  $\pi_{\mathbb{H}}(Z_n)$  have sublinear speed, see Lemma 4.6, allows us to solve this issue. More precisely, take the  $\mu$ -boundary  $(\partial T, \mathcal{B}_{\partial T}, \nu_{\partial T})$  and the  $\check{\mu}$ -boundary  $(\partial T, \mathcal{B}_{\partial T}, \check{\nu}_{\partial T})$  and define gauges

$$\mathcal{G}_k := \{g \in G : d_T(\pi_T(1), \pi_T(g)) \leq k^2 \text{ and } d_{\mathbb{H}}(\pi_{\mathbb{H}}(1), \pi_{\mathbb{H}}(g)) \leq k\}.$$

Again,  $\check{\nu}_{\partial T} \otimes \nu_{\partial T}$ -almost every pair of points  $(\xi_-, \xi_+) \in \partial T \times \partial T$  has distinct ends  $\xi_-, \xi_+ \in \partial T$ , which we may connect by a doubly infinite reduced path  $v : \mathbb{Z} \rightarrow T$ . Let  $\mathcal{S}(\xi_-, \xi_+)$  be the full  $\pi_T$ -preimage of the path, i.e., the set consisting of all  $g \in G$  such that  $\pi_T(g)$  is contained in  $v(\mathbb{Z})$ ; see the right-hand side of Figure 6. Again, to all remaining pairs we assign the whole of  $G$  as a strip. This way, the map  $\mathcal{S}$  becomes measurable,  $G$ -equivariant, and satisfies the inequality of Remark 3.6. Now, pick an arbitrary pair  $(\xi_-, \xi_+) \in \partial T \times \partial T$  with distinct ends  $\xi_-, \xi_+ \in \partial T$ . We claim that

$$\begin{aligned} \frac{1}{n} \cdot \ln(\text{card}(\mathcal{S}(\xi_-, \xi_+) \cap \mathcal{G}_{|Z_n|})) &\leq \frac{\ln((2 \cdot |Z_n|^2 + 1) \cdot \exp(|Z_n| + 2))}{n} \\ &= \underbrace{\frac{\ln(2 \cdot |Z_n|^2 + 1)}{n}}_{\textcircled{1}} + \underbrace{\frac{|Z_n| + 2}{n}}_{\textcircled{2}}. \end{aligned}$$

Indeed, the inequality holds for a similar reason as above. The strip  $\mathcal{S}(\xi_-, \xi_+)$  intersects the gauge  $\mathcal{G}_{|Z_n|}$  in at most  $2 \cdot |Z_n|^2 + 1$  many cosets from  $G/B$ . Slightly more involved is the observation that each of them contains at most  $\exp(|Z_n| + 2)$  many elements of the gauge. Fix a coset  $gB$ . The projections  $\pi_{\mathbb{H}}(h)$  of the elements  $h \in gB$  are located on a horizontal line  $L \subseteq \mathbb{H}$  with imaginary part  $y := \text{Im}(\pi_{\mathbb{H}}(g))$ . One necessary condition for such an element  $h \in gB$  to be contained in the gauge  $\mathcal{G}_{|Z_n|}$  is that the projection  $\pi_{\mathbb{H}}(h)$  is contained in the closed disc  $D := \{z \in \mathbb{H} : d_{\mathbb{H}}(i, z) \leq |Z_n|\} \subseteq \mathbb{H}$ . If  $L \cap D$  is empty, then the coset  $gB$  does not contain any element of the gauge and we

are done. Otherwise, there is a unique  $x \in \mathbb{R}$  with  $x \geq 0$  such that  $L \cap D$  is the horizontal line between  $z_1 := -x + iy$  and  $z_2 := x + iy$ ; see Figure 7. The projections  $\pi_{\mathbb{H}}(h)$  with  $h \in gB$  have the property that the real parts  $\operatorname{Re}(\pi_{\mathbb{H}}(h))$  and  $\operatorname{Re}(\pi_{\mathbb{H}}(hb))$  differ exactly by  $y$ . So,  $L \cap D$  contains at most  $1 + \frac{2x}{y}$  many of them. Let us now estimate  $1 + \frac{2x}{y}$  in terms of  $|Z_n|$ . Since  $z_1$  and  $z_2$  are both contained in  $D$ , their distance is at most  $2 \cdot |Z_n|$ . Therefore,

$$\begin{aligned} 2 \cdot |Z_n| &\geq d_{\mathbb{H}}(z_1, z_2) = \operatorname{arcosh} \left( 1 + \frac{|z_2 - z_1|^2}{2 \cdot \operatorname{Im}(z_1) \operatorname{Im}(z_2)} \right) \\ &= \operatorname{arcosh} \left( 1 + \frac{2x^2}{y^2} \right) \geq \ln \left( 1 + \frac{2x^2}{y^2} \right). \end{aligned}$$

And, in particular,

$$\begin{aligned} \exp(2 \cdot |Z_n|) > \frac{2x^2}{y^2}, &\iff \exp \left( |Z_n| + \frac{1}{2} \cdot \ln(2) \right) > \frac{2x}{y}, \\ &\implies \exp(|Z_n| + 2) > 1 + \frac{2x}{y}. \end{aligned}$$

So, the coset  $gB$  contains strictly fewer than  $\exp(|Z_n| + 2)$  elements of the gauge. We will now show that both summands ① and ② converge a.s. to 0, which will complete the proof. Let us first observe that

$$\begin{aligned} (\ddagger) \quad |Z_n| - 1 &\leq \max\{d_{\mathbb{H}}(\pi_{\mathbb{H}}(1), \pi_{\mathbb{H}}(Z_n)), \sqrt{d_T(\pi_T(1), \pi_T(Z_n))}\} \\ &\leq \max\{d_{\mathbb{H}}(\pi_{\mathbb{H}}(1), \pi_{\mathbb{H}}(Z_n)), \sqrt{d_S(1, Z_n)}\}. \end{aligned}$$

Concerning ①, we deduce from  $(\ddagger)$  and Lemma 2.3 that

$$|Z_n| \leq \max\{\ell_a, \ell_b, 1\} \cdot d_S(1, Z_n) + 1$$

and finally obtain by the same argument as above

$$\textcircled{1} = \frac{\ln(2 \cdot |Z_n|^2 + 1)}{n} \leq \frac{\ln(2 \cdot (\max\{\ell_a, \ell_b, 1\} \cdot d_S(1, Z_n) + 1)^2 + 1)}{n} \xrightarrow[\text{a.s.}]{n \rightarrow \infty} 0.$$

On the other hand, concerning ②, we apply  $(\ddagger)$  and Lemma 4.6 to obtain

$$\textcircled{2} = \frac{|Z_n| + 2}{n} \leq \frac{\max\{d_{\mathbb{H}}(\pi_{\mathbb{H}}(1), \pi_{\mathbb{H}}(Z_n)), \sqrt{d_S(1, Z_n)}\} + 3}{n} \xrightarrow[\text{a.s.}]{n \rightarrow \infty} 0. \quad \blacksquare$$

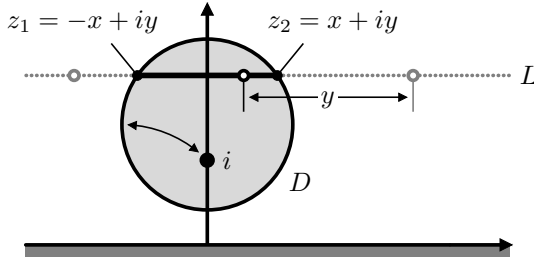


Figure 7. The horizontal line  $L$ , the closed disc  $D$ , and their intersection  $L \cap D$ .

**Appendix: The remaining non-amenable cases**

Recall from Section 2.1 that a Baumslag–Solitar group  $BS(p, q)$  is non-amenable if and only if neither  $|p| = 1$  nor  $|q| = 1$ . Until now, we have only identified the Poisson–Furstenberg boundary for random walks on non-amenable Baumslag–Solitar groups  $BS(p, q)$  with  $1 < p < q$ . Replacing one of the two generators by its inverse yields

$$BS(p, q) \cong BS(q, p) \quad \text{and} \quad BS(p, q) \cong BS(-p, -q).$$

So, in order to cover the remaining non-amenable cases, it suffices to consider  $1 < p < -q$  and  $1 < p = |q|$ . Below, we explain how to adjust our methods to obtain similar results for these cases.

A.1. ACTION BY SUITABLE ISOMETRIES ON THE HYPERBOLIC PLANE. Assume that  $G = BS(p, q)$  with  $1 < p < -q$ . The definition of the tree  $T$  and the level function  $\lambda$  remain the same and even Remark 2.2 can be adapted, replacing  $q$  by  $|q|$ . Recall that, in Section 2.3, we first constructed the group homomorphism  $\pi_{\text{Isom}(\mathbb{H})} : G \rightarrow \text{Isom}(\mathbb{H})$  and then used it to define the projection  $\pi_{\mathbb{H}} : G \rightarrow \mathbb{H}$ . This is precisely what we are going to do again; but, this time, with another isometry  $\pi_{\text{Isom}(\mathbb{H})}(a)$ . Let  $\pi_{\text{Isom}(\mathbb{H})} : \{a, b\} \rightarrow \text{Isom}(\mathbb{H})$  be the map given by

$$\pi_{\text{Isom}(\mathbb{H})}(a) := \left( z \mapsto \frac{|q|}{p} \cdot (-\bar{z}) \right) \quad \text{and} \quad \pi_{\text{Isom}(\mathbb{H})}(b) := (z \mapsto z + 1).$$

It follows from von Dyck’s theorem that this map can be uniquely extended to a group homomorphism  $\pi_{\text{Isom}(\mathbb{H})} : G \rightarrow \text{Isom}(\mathbb{H})$ . Now, we define  $\pi_{\mathbb{H}} : G \rightarrow \mathbb{H}$  by

$$\pi_{\mathbb{H}}(g) := \pi_{\text{Isom}(\mathbb{H})}(g)(i).$$

LEMMA A.1: *For every  $g \in G$  the point  $\pi_{\mathbb{H}}(ga) \in \mathbb{H}$  is above the point  $\pi_{\mathbb{H}}(g) \in \mathbb{H}$ ; the two points have the same real part and their distance in the hyperbolic plane is  $\ell_a := \ln(\frac{|q|}{p})$ . The point  $\pi_{\mathbb{H}}(gb) \in \mathbb{H}$  is either to the right or to the left of the point  $\pi_{\mathbb{H}}(g) \in \mathbb{H}$  depending on whether the level  $\lambda(g)$  is even or odd; in any case, the two points have the same imaginary part and their distance in the hyperbolic plane is  $\ell_b := \ln(\frac{3+\sqrt{5}}{2})$ .*

*Proof sketch.* The proof is similar to the one of Lemma 2.3, and we shall only address the differences. The points  $\pi_{\mathbb{H}}(gb) \in \mathbb{H}$  and  $\pi_{\mathbb{H}}(g) \in \mathbb{H}$  are obtained by evaluating  $\pi_{\text{Isom}(\mathbb{H})}(g)$  at  $\pi_{\mathbb{H}}(b) \in \mathbb{H}$  and  $\pi_{\mathbb{H}}(1) \in \mathbb{H}$ . Again, the image  $\pi_{\text{Isom}(\mathbb{H})}(g)$  is a composition of  $\pi_{\text{Isom}(\mathbb{H})}(a^{\pm 1})$  and  $\pi_{\text{Isom}(\mathbb{H})}(b^{\pm 1})$ . While  $\pi_{\text{Isom}(\mathbb{H})}(b^{\pm 1})$  are translations, each occurrence of  $\pi_{\text{Isom}(\mathbb{H})}(a^{\pm 1})$  yields both a dilation and a reflection at the imaginary axis. This implies that the point  $\pi_{\mathbb{H}}(gb) \in \mathbb{H}$  is to the right of the point  $\pi_{\mathbb{H}}(g)$  if and only if the number of occurrences of  $\pi_{\text{Isom}(\mathbb{H})}(a^{\pm 1})$  is even, which is the case if and only if  $\lambda(g)$  is even. ■

Using this projection, and replacing  $q$  by  $|q|$  wherever it is necessary, we may repeat most of the arguments from Section 4. For example, the definitions of the imaginary part  $A_g$  and the real part  $B_g$  now yield the equation  $\ln(A_g) = \ln(\frac{|q|}{p}) \cdot \lambda(g)$ . In order to identify the Poisson–Furstenberg boundary geometrically, we first showed that the pointwise projections of the random walk to  $\mathbb{H}$  and  $T$  converge a.s. to random elements in the respective boundaries.

While the proof of Lemma 4.4 for  $\delta > 0$  can be adapted, the one of Lemma 4.5 for  $\delta < 0$  requires some additional work. We have to show that the real parts  $B_{Z_n}$  converge a.s. to a random element  $r \in \mathbb{R}$ . In the original proof, we used that  $A_{Z_n} = A_{X_1} \cdots A_{X_n}$  and  $B_{Z_n} = \sum_{k=1}^n C_k$  with  $C_k := A_{X_1} \cdots A_{X_{k-1}} \cdot B_{X_k}$ . The first formula for  $A_{Z_n}$  remains true. However, the second one for  $B_{Z_n}$  does not because we are now in a situation where not only the scaling but also the direction of the next horizontal increment depends on the current level. Instead, we obtain that

$$B_{Z_n} = \sum_{k=1}^n C_k$$

with  $C_k := \varepsilon_{X_1} \cdot A_{X_1} \cdots \varepsilon_{X_{k-1}} \cdot A_{X_{k-1}} \cdot B_{X_k}$  where  $\varepsilon_g := 1$  if  $\lambda(g)$  is even and  $\varepsilon_g := -1$  if  $\lambda(g)$  is odd. This allows us to apply Cauchy’s root test as in the proof of Lemma 4.5. Moreover, the proofs of Lemmas 4.6 and 4.7 for  $\delta = 0$  can also be adapted because the estimates are not in terms of the actual horizontal increments but only of their absolute values.



Concerning the boundary  $\partial T$ , it suffices to observe that the proof of Lemma 4.8 only requires the property that the subgroup  $\varphi(G) \leq \text{Aut}(T)$  acts transitively on  $T$  and does not fix an end, which is always the case unless  $|p| = 1$  or  $|q| = 1$ . Therefore, it still shows that the projections  $\pi_T(Z_n)$  converge a.s. to a random end in  $\partial T$ . As in Lemma 4.11, we can show that the  $G$ -actions on  $\partial T$  and  $\partial T \times \mathbb{R}$  are topologically minimal, whence the hitting measures  $\nu_{\partial T}$  and  $\nu_{\partial T \times \mathbb{R}}$  are non-atomic and have full support. This allows us to adapt the proof of Theorem 1.1 and to obtain the following version of the identification theorem.

**THEOREM A.2** (“identification theorem” for  $1 < p < -q$ ): *Let  $Z = (Z_0, Z_1, \dots)$  be a random walk on a non-amenable Baumslag–Solitar group  $G = \text{BS}(p, q)$  with  $1 < p < -q$  and increments  $X_1, X_2, \dots$  of finite first moment. Depending on the vertical drift  $\delta$ , we distinguish three cases:*

- (1) *If  $\delta > 0$ , then the Poisson–Furstenberg boundary is isomorphic to  $(\partial T, \mathcal{B}_{\partial T}, \nu_{\partial T})$  endowed with the boundary map  $\text{bnd}_{\partial T} : \Omega \rightarrow \partial T$ .*
- (2) *If  $\delta < 0$ , then the Poisson–Furstenberg boundary is isomorphic to  $(\partial T \times \mathbb{R}, \mathcal{B}_{\partial T \times \mathbb{R}}, \nu_{\partial T \times \mathbb{R}})$  endowed with  $\text{bnd}_{\partial T \times \mathbb{R}} : \Omega \rightarrow \partial T \times \mathbb{R}$ .*
- (3) *If  $\delta = 0$  and  $\ln(A_{X_1})$  has finite second moment and there is an  $\varepsilon > 0$  such that  $\ln(1 + |B_{X_1}|)$  has finite  $(2 + \varepsilon)$ -th moment, then the Poisson–Furstenberg boundary is isomorphic to  $(\partial T, \mathcal{B}_{\partial T}, \nu_{\partial T})$  endowed with  $\text{bnd}_{\partial T} : \Omega \rightarrow \partial T$ .*

**A.2. ACTION BY ISOMETRIES ON THE EUCLIDEAN PLANE.** Let us now assume that  $G = \text{BS}(p, q)$  with  $1 < p = |q|$ . Again, the definition of the tree  $T$  and the level function  $\lambda$  remain the same and Remark 2.2 can be adapted. However, the situation differs fundamentally from the ones discussed so far because each brick, see ① and ② in Figure 2, would now have equally many edges on its upper and lower level. Therefore, we shall use the Euclidean plane  $\mathbb{R}^2$  instead of the hyperbolic plane  $\mathbb{H}$ . In order to construct a projection  $\pi_{\mathbb{R}^2} : G \rightarrow \mathbb{R}^2$ , consider the map  $\pi_{\text{Isom}(\mathbb{R}^2)} : \{a, b\} \rightarrow \text{Isom}(\mathbb{R}^2)$  given by

$$\pi_{\text{Isom}(\mathbb{R}^2)}(a) := \begin{cases} ((x, y) \mapsto (x, y + 1)) & \text{if } q > 0 \\ ((x, y) \mapsto (-x, y + 1)) & \text{if } q < 0 \end{cases}$$

and

$$\pi_{\text{Isom}(\mathbb{R}^2)}(b) := ((x, y) \mapsto (x + 1, y)).$$

In both cases,  $q > 0$  and  $q < 0$ , we may apply von Dyck’s theorem to extend the map uniquely to a group homomorphism  $\pi_{\text{Isom}(\mathbb{R}^2)} : G \rightarrow \text{Isom}(\mathbb{R}^2)$ . Now, we define  $\pi_{\mathbb{R}^2} : G \rightarrow \mathbb{R}^2$  by  $\pi_{\mathbb{R}^2}(g) := \pi_{\text{Isom}(\mathbb{R}^2)}(g)(0, 0)$ . Note that, instead of the discrete hyperbolic plane, we obtain a discrete Euclidean plane  $\Gamma_v$ .

We want to show that, as soon as the projections converge to a random element in  $\partial T$ , independently of the vertical drift, the Poisson–Furstenberg boundary is isomorphic to  $(\partial T, \mathcal{B}_{\partial T}, \nu_{\partial T})$ . In particular, we do not need to introduce any boundary to capture the behaviour of the projections  $\pi_{\mathbb{R}^2}(Z_n)$ .

Even though the action of the group  $G$  on the tree  $T$  is not faithful anymore, the proof of Lemma 4.8 still shows that the projections  $\pi_T(Z_n)$  converge a.s. to a random end in  $\partial T$ . As in the first assertion of Lemma 4.11, we can show that the  $G$ -action on  $\partial T$  is topologically minimal, whence the hitting measure  $\nu_{\partial T}$  is non-atomic and has full support. This finally allows us to prove the following version of the identification theorem.

**THEOREM A.3** (“identification theorem” for  $1 < p = |q|$ ): *Let  $Z = (Z_0, Z_1, \dots)$  be a random walk on a non-amenable Baumslag–Solitar group  $G = \text{BS}(p, q)$  with  $1 < p = |q|$  and increments  $X_1, X_2, \dots$  of finite first moment. Then, the Poisson–Furstenberg boundary is isomorphic to  $(\partial T, \mathcal{B}_{\partial T}, \nu_{\partial T})$  endowed with the boundary map  $\text{bnd}_{\partial T} : \Omega \rightarrow \partial T$ .*

*Proof sketch.* As in the proof of Theorem 1.1, we take the  $\mu$ -boundary  $(\partial T, \mathcal{B}_{\partial T}, \nu_{\partial T})$  and the  $\check{\mu}$ -boundary  $(\partial T, \mathcal{B}_{\partial T}, \check{\nu}_{\partial T})$ . Then, we define gauges

$$\mathcal{G}_k := \{g \in G : d_T(\pi_T(1), \pi_T(g)) \leq k \text{ and } d_{\mathbb{R}^2}(\pi_{\mathbb{R}^2}(1), \pi_{\mathbb{R}^2}(g)) \leq k\}.$$

Again,  $\check{\nu}_{\partial T} \otimes \nu_{\partial T}$ -almost every pair of points  $(\xi_-, \xi_+) \in \partial T \times \partial T$  has distinct ends  $\xi_-, \xi_+ \in \partial T$ , which we may connect by a doubly infinite reduced path  $v : \mathbb{Z} \rightarrow T$ . Let  $\mathcal{S}(\xi_-, \xi_+)$  be the full  $\pi_T$ -preimage of the path, i.e., the set consisting of all  $g \in G$  such that  $\pi_T(g)$  is contained in  $v(\mathbb{Z})$ . To all remaining pairs we assign the whole of  $G$  as a strip. This way, the map  $\mathcal{S}$  becomes measurable,  $G$ -equivariant, and satisfies the inequality of Remark 3.6. Now, pick an arbitrary pair  $(\xi_-, \xi_+) \in \partial T \times \partial T$  with distinct ends  $\xi_-, \xi_+ \in \partial T$ . We claim that

$$\frac{1}{n} \cdot \ln(\text{card}(\mathcal{S}(\xi_-, \xi_+) \cap \mathcal{G}_{|Z_n|})) \leq \frac{\ln((2 \cdot |Z_n| + 1) \cdot (2 \cdot |Z_n| + 1))}{n}.$$

Indeed, the strip  $\mathcal{S}(\xi_-, \xi_+)$  intersects the gauge  $\mathcal{G}_{|Z_n|}$  in at most  $2 \cdot |Z_n| + 1$  many cosets from  $G/B$ , and each of them contains at most  $2 \cdot |Z_n| + 1$  many

elements of the gauge. Now, it suffices to consider the standard generating set  $S := \{a, b\} \subseteq G$  and to observe that  $|Z_n| \leq d_S(1, Z_n) + 1$ . Then, as in the proof of Theorem 1.1, we may use the fact that  $\frac{1}{n} \cdot d_S(1, Z_n)$  is a.s. bounded and conclude that

$$\dots = \frac{\ln((2 \cdot |Z_n| + 1)^2)}{n} \leq \frac{\ln((2 \cdot d_S(1, Z_n) + 3)^2)}{n} \xrightarrow[\text{a.s.}]{n \rightarrow \infty} 0,$$

which allows us to apply the strip criterion.    ■

## References

- [Ave74] A. Avez, *Théorème de Choquet–Deny pour les groupes à croissance non exponentielle*, Comptes Rendus Hebdomadaires des Séances de l’Académie des Sciences. Série A. Mathématique **279** (1974), 25–28.
- [Bea83] A. F. Beardon, *The Geometry of Discrete Groups*, Graduate Texts in Mathematics, Vol. 91, Springer-Verlag, New York, 1983.
- [Bil99] P. Billingsley, *Convergence of Probability Measures*, Wiley Series in Probability and Statistics: Probability and Statistics, John Wiley & Sons, New York, 1999.
- [Bla55] D. Blackwell, *On transient Markov processes with a countable number of states and stationary transition probabilities*, Annals of Mathematical Statistics **26** (1955), 654–658.
- [Bri63] J. L. Britton, *The word problem*, Annals of Mathematics **77** (1963), 16–32.
- [Bro03] S. Brofferio, *How a centred random walk on the affine group goes to infinity*, Annales de l’Institut Henri Poincaré. Probabilités et Statistiques **39** (2003), 371–384.
- [Bro06] S. Brofferio, *The Poisson boundary of random rational affinities*, Université de Grenoble. Annales de l’Institut Fourier **56** (2006), 499–515.
- [BS62] G. Baumslag and D. Solitar, *Some two-generator one-relator non-Hopfian groups*, Bulletin of the American Mathematical Society **68** (1962), 199–201.
- [CD60] G. Choquet and J. Deny, *Sur l’équation de convolution  $\mu = \mu * \sigma$* , Comptes Rendus Hebdomadaires des Séances de l’Académie des Sciences **250** (1960), 799–801.
- [CFS82] I. P. Cornfeld, S. V. Fomin and Ya. G. Sinai, *Ergodic Theory*, Grundlehren der mathematischen Wissenschaften, Vol. 245, Springer-Verlag, New York, 1982.
- [CKW94] D. I. Cartwright, V. A. Kaimanovich and W. Woess, *Random walks on the affine group of local fields and of homogeneous trees*, Université de Grenoble. Annales de l’Institut Fourier **44** (1994), 1243–1288.
- [Coh13] D. L. Cohn, *Measure Theory*, Birkhäuser Advanced Texts: Basler Lehrbücher, Birkhäuser, Springer-Verlag, New York, 2013.
- [CS89] D. I. Cartwright and P. M. Soardi, *Convergence to ends for random walks on the automorphism group of a tree*, Proceedings of the American Mathematical Society **107** (1989), 817–823.
- [CZ93] D. J. Collins and H. Zieschang, *Combinatorial group theory and fundamental groups*, in *Algebra VII*, Encyclopaedia of Mathematical Sciences, Vol. 58, Springer-Verlag, Berlin, 1993, pp. 1–166, 233–240.

- [Éli82] L. Élie, *Comportement asymptotique du noyau potentiel sur les groupes de Lie*, Annales Scientifiques de l'École Normale Supérieure **15** (1982), 257–364.
- [Ers04] A. Erschler, *Boundary behavior for groups of subexponential growth*, Annals of Mathematics **160** (2004), 1183–1210.
- [Ers10] A. Erschler, *Poisson–Furstenberg boundaries, large-scale geometry and growth of groups*, in *Proceedings of the International Congress of Mathematicians, Vol. II*, Hindustan Book Agency, New Delhi, 2010, pp. 681–704.
- [Fel71] W. Feller, *An Introduction to Probability Theory and its Applications*, Vol. II, John Wiley & Sons, New York–London–Sydney, 1971.
- [Fur63] H. Furstenberg, *A Poisson formula for semi-simple Lie groups*, Annals of Mathematics **77** (1963), 335–386.
- [Fur71] H. Furstenberg, *Random walks and discrete subgroups of Lie groups*, in *Advances in Probability and Related Topics, Vol. 1*, Dekker, New York, 1971, pp. 1–63.
- [Fur73] H. Furstenberg, *Boundary theory and stochastic processes on homogeneous spaces*, in *Harmonic Analysis on Homogeneous Spaces*, Proceedings of Symposia in Pure Mathematics, Vol.26, American Mathematical Society, Providence, RI, 1973, pp. 193–229.
- [GPS94] S. W. Golomb, R. E. Peile and R. A. Scholtz, *Basic Concepts in Information Theory and Coding*, Applications of Communications Theory, Plenum Press, New York, 1994.
- [Hae73] J. Haezendonck, *Abstract Lebesgue–Rohlin spaces*, Bulletin de la Société Mathématique de Belgique **25** (1973), 243–258.
- [Kai91] V. A. Kaimanovich, *Poisson boundaries of random walks on discrete solvable groups*, in *Probability Measures on Groups X (Oberwolfach, 1990)*, Plenum, New York, 1991, pp. 205–238.
- [Kai00] V. A. Kaimanovich, *The Poisson formula for groups with hyperbolic properties*, Annals of Mathematics **152** (2000), 659–692.
- [KKR04] V. A. Kaimanovich, Y. Kifer and B.-Z. Rubshtein, *Boundaries and harmonic functions for random walks with random transition probabilities*, Journal of Theoretical Probability **17** (2004), 605–646.
- [KV83] V. A. Kaimanovich and A. M. Vershik, *Random walks on discrete groups: boundary and entropy*, Annals of Probability **11** (1983), 457–490.
- [KW02] V. A. Kaimanovich and W. Woess, *Boundary and entropy of space homogeneous Markov chains*, Annals of Probability **30** (2002), 323–363.
- [LP15] R. Lyons and Y. Peres, *Poisson boundaries of lamplighter groups: proof of the Kaimanovich–Vershik conjecture*, <http://arxiv.org/abs/1508.01845>.
- [MNS17] A. Malyutin, T. Nagnibeda and D. Serbin, *Boundaries of  $\mathbb{Z}^n$ -free groups*, in *Groups, Graphs and Random Walks*, London Mathematical Society Lecture Note Series, Vol. 436, Cambridge University Press, Cambridge, 2017, pp. 355–390.
- [Neb88] C. Nebbia, *Amenability and Kunze–Stein property for groups acting on a tree*, Pacific Journal of Mathematics **135** (1988), 371–380.
- [Roh52] V. A. Rohlin, *On the fundamental ideas of measure theory*, American Mathematical Society Translations **71** (1952).

- [Ros81] J. Rosenblatt, *Ergodic and mixing random walks on locally compact groups*, *Mathematische Annalen* **257** (1981), 31–42.
- [Rou14] G. G. Roussas, *An Introduction to Measure-Theoretic Probability*, Elsevier/Academic Press, Amsterdam, 2014.
- [Rud90] D. J. Rudolph, *Fundamentals of Measurable Dynamics*, Oxford Science Publications, The Clarendon Press, Oxford University Press, New York, 1990.
- [Sav10] E. Sava, *A note on the Poisson boundary of lamplighter random walks*, *Monatshefte für Mathematik* **159** (2010), 379–396.
- [VK79] A. M. Vershik and V. A. Kaimanovich, *Random walks on groups: boundary, entropy, uniform distribution*, *Soviet Mathematics. Doklady* **20** (1979), 1170–1173.
- [Wil70] S. Willard, *General Topology*, Addison–Wesley Publishing, Reading, MA–London–Don Mills, ON, 1970.
- [Woe89] W. Woess, *Boundaries of random walks on graphs and groups with infinitely many ends*, *Israel Journal of Mathematics* **68** (1989), 271–301.
- [Woe91] W. Woess, *Topological groups and infinite graphs*, *Discrete Mathematics* **95** (1991), 373–384.